



**HAL**  
open science

# Contribution à la perception augmentée de scènes dynamiques : schémas temps réels d'assimilation de données pour la mécanique du solide et des structures

Adrien Goeller

► **To cite this version:**

Adrien Goeller. Contribution à la perception augmentée de scènes dynamiques : schémas temps réels d'assimilation de données pour la mécanique du solide et des structures. Autre. Université Paris Saclay (COMUE), 2018. Français. NNT : 2018SACLC013 . tel-01753169

**HAL Id: tel-01753169**

**<https://theses.hal.science/tel-01753169>**

Submitted on 29 Mar 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Contribution à la perception augmentée de scènes dynamiques : schémas temps réels d'assimilation de données pour la mécanique du solide et des structures

Thèse de doctorat de l'Université Paris-Saclay  
préparée à CentraleSupélec

École doctorale n°579 : sciences mécaniques et énergétiques,  
matériaux et géosciences (SMEMAG)  
Spécialité de doctorat : mécanique des solides

Thèse présentée et soutenue à Saint-Ouen, le 19 janvier 2018, par

**Adrien Goeller**

Composition du Jury :

Emmanuel FOLTETE Professeur des universités, ENSMM – FEMTO-ST	Rapporteur
Xavier CHIEMENTIN Maître de Conférences HDR, Université de Reims Champagne-Ardenne – GRESPI	Rapporteur
Pierre VACHER Professeur des universités, Université Savoie Mont-Blanc – Laboratoire SYMME	Président
Ludovic CHAMOIN Professeur des Universités, ENS Cachan – LMT	Examineur
Jean-Luc DION Professeur des Universités, Institut Supérieur de Mécanique de Paris – Laboratoire QUARTZ	Directeur de thèse
Thierry SORIANO Professeur des Universités, SEATECH – Université de Toulon – Laboratoire COSMER	Co-Directeur de thèse
Ronan LE BRETON Maître de Conférences, INSA Rennes – LGCGM	Encadrant
Bernard ROUX, Directeur Général, VANNIER-KINOPTIK	Invité

# Table des matières

<b>Table des matières</b>	<b>I</b>
<b>Liste des tableaux</b>	<b>V</b>
<b>Table des figures</b>	<b>VII</b>
<b>Remerciements</b>	<b>1</b>
<b>Introduction générale</b>	<b>5</b>
<b>1 Problématiques de l'étude</b>	<b>9</b>
I Contexte de la CIFRE . . . . .	11
I.A. Présentation du projet . . . . .	11
I.B. Acteurs . . . . .	12
I.C. Objectifs de la thèse . . . . .	12
II Cadre technologique de l'étude . . . . .	14
II.A. Introduction à la vidéo rapide . . . . .	15
II.B. Contexte particulier d'une vidéo rapide . . . . .	22
III Conclusion : apports proposés . . . . .	23
Bibliographie du chapitre 1 . . . . .	25
<b>2 Analyse de vidéos par approche eulérienne basée sur un modèle cinématique élémentaire</b>	<b>27</b>
I Introduction du contexte vidéo et présentation de deux descriptions d'une même scène vidéo . . . . .	30
I.A. Introduction à la vidéo numérique . . . . .	30
I.B. Représentation eulérienne et lagrangienne : proposition d'un formalisme pour décrire une scène vidéo . . . . .	33
II Estimer le mouvement des pixels et interpoler leur trajectoire . . . . .	36
II.A. Cadre de l'interpolation vidéo temps réel : quels modèles pour quelles informations ? . . . . .	36
II.B. Application à travers la philosophie des normes de compression MPEG . . . . .	42
II.C. L'interpolation pour modifier la fréquence temporelle . . . . .	46
III Caractériser un objet dans une séquence d'images : un problème mal posé en 2D . . . . .	47
III.A. Cas particulier de dispositifs matériels ajoutant des informations à l'image 2D . . . . .	47
III.B. Détecter et suivre . . . . .	48
III.C. Extraire et mesurer . . . . .	59
III.D. Conclusion et remarque sur une simplification par le passage de la 2D à la 3D . . . . .	61

IV	Une nouvelle méthode pour l'interpolation et la compression à l'aide de la composition cinématique du mouvement . . . . .	62
IV.A.	Lien entre l'approche eulérienne et lagrangienne . . . . .	63
IV.B.	Utilisation de l'interpolation petite échelle pour des grands mouvements . . . . .	65
V	Formalisation de la chaîne complète . . . . .	67
V.A.	Détection et extraction . . . . .	68
V.B.	Caractérisation des objets . . . . .	70
V.C.	Reconstruction et rendu amélioré . . . . .	71
V.D.	Applications . . . . .	72
VI	Conclusion . . . . .	77
	Bibliographie du chapitre 2 . . . . .	79
<b>3</b>	<b>Analyse de vidéos par approche lagrangienne basée sur un modèle dynamique physique</b>	<b>85</b>
I	Apports d'un modèle dynamique . . . . .	88
I.A.	Définition des modèles physiques . . . . .	88
II	L'assimilation de données : une famille de méthodes pour combiner modèle et expérience	90
II.A.	Principes de l'assimilation . . . . .	90
II.B.	Applications du cadre de l'assimilation de données dans d'autres domaines . . . . .	99
III	Courte Kalmanographie pour les expérimentateurs à vocations non linéaires . . . . .	102
III.A.	Contexte bayésien . . . . .	102
III.B.	Approximer le filtrage bayésien . . . . .	103
III.C.	Adaptation du filtre pour traiter des modèles non-linéaires . . . . .	107
III.D.	Quelques versions du filtre . . . . .	114
III.E.	Schémas de filtrage . . . . .	125
III.F.	Performances des filtres de Kalman . . . . .	128
III.G.	Conclusion . . . . .	130
IV	Proposition d'un schéma d'assimilation de vidéos rapides sur des modèles dynamiques complexes et réduits . . . . .	132
IV.A.	Identification et extrapolation avec un modèle réduit : une solution pour concilier la donnée image et la rapidité . . . . .	132
IV.B.	Application à la dynamique multicorps : la règle glissante . . . . .	136
V	Conclusion . . . . .	149
	Bibliographie du chapitre 3 . . . . .	150
<b>4</b>	<b>Application à l'analyse modale : une méthode temps réel pour l'identification de modes non linéaires</b>	<b>157</b>
I	Contexte de l'analyse modale . . . . .	160
I.A.	Introduction : un cadre favorable . . . . .	160
I.B.	Analyse modale . . . . .	161
I.C.	Apports et limites de la mesure par caméra rapide . . . . .	175
I.D.	Conclusion . . . . .	176
II	Construction d'une fonction d'évolution pour l'assimilation séquentielle modale . . . . .	177
II.A.	Contexte particulier de l'analyse modale . . . . .	177
II.B.	La démodulation comme modélisation d'un signal vibratoire . . . . .	178
III	Un premier schéma d'assimilation utilisant un modèle théorique modal de la structure : le Kinematic multiSensors Assimilation Modal Identification (K-SAMI) . . . . .	186
IV	Formalisation de la chaîne complète . . . . .	187

## Table des matières

IV.A.	Méthodes de vision . . . . .	187
IV.B.	Formalisation pour un cas 2D . . . . .	190
IV.C.	Formalisation pour un cas 3D . . . . .	191
IV.D.	Fusion d'informations : utilisation de capteurs complémentaires . . . . .	193
IV.E.	Fonctionnement en prédiction, en lissage, en interpolateur (cf partie 3) . . . . .	194
V	Schéma d'assimilation développé sans connaissances sur la structure : Blind multi-Sensors Assimilation Modal Identification (B-SAMI) . . . . .	195
VI	Application au cas d'une poutre encastrée libre . . . . .	197
VI.A.	Cadre expérimental . . . . .	197
VI.B.	Validation du schéma Kinematic-SAMI . . . . .	200
VI.C.	Validation du schéma Blind-SAMI . . . . .	209
VI.D.	Conclusion sur l'application des SAMI . . . . .	213
VII	Modifier la perception du système réel à l'aide de sa physique : proposition d'un autre algorithme de <i>motion magnification</i> . . . . .	215
VII.A.	Modification d'une vidéo à partir de l'espace d'observation du K-SAMI . . . . .	215
VII.B.	Application sur le cas réel d'un ski . . . . .	216
VIII	Conclusion et perspectives . . . . .	218
	Bibliographie du chapitre 4 . . . . .	220
	<b>Conclusion générale</b> . . . . .	<b>225</b>
	<b>Publications associées à ces travaux</b> . . . . .	<b>229</b>
	<b>Annexes</b> . . . . .	<b>231</b>
A.	Application dans le cadre de l'interpolation vidéo . . . . .	231
1.	Trajectoire et échantillonnage . . . . .	231
2.	Modèle d'évolution de la trajectoire . . . . .	233
3.	Cas d'application sur une simulation de balle rebondissante . . . . .	234
B.	Modèle de la Règle à 2 points de contacts . . . . .	237
1.	Bilan des actions mécaniques . . . . .	237
2.	Cinématique du système . . . . .	238
3.	Cinétique . . . . .	238
4.	Dynamique . . . . .	239
5.	Principe Fondamental de la Dynamique . . . . .	239
C.	Comparaison des performances des méthodes de vision . . . . .	241
D.	Résultats animés sur le cas test réel de la poutre encastrée libre pour l'identification avec le schéma Kinematic-SAMI : applications à l'interpolation et l'extrapolation spatiale . . . . .	245
E.	Résultats animés sur le cas test réel de la poutre encastrée libre pour l'identification avec le schéma Blind-SAMI . . . . .	249
	<b>Bibliographie générale</b> . . . . .	<b>252</b>



# Liste des tableaux

1.1	Paramètres réglables et intrinsèques au capteur utilisé . . . . .	18
1.2	Débits de données suivants différents paramètres . . . . .	19
1.3	Différentes connectiques d'interface . . . . .	20
3.1	Synthèse des différentes erreurs entrant en jeu durant le processus de l'assimilation de données . . . . .	93
3.2	Utilisation des filtres de Kalman en fonction de la position des données par rapport à l'état courant du système . . . . .	105
3.3	Paramétrage choisi pour les deux scénarios tests présentés . . . . .	138
3.4	Propriétés de la caméra utilisée . . . . .	142
4.1	Différents domaines fréquentiels . . . . .	160
4.2	Tableau présentant un état de l'art récent d'utilisation de caméras rapides dans des contextes vibratoires (avec l'aimable permission de Xavier Régat) . . . . .	164
4.3	Tableau synthétisant les avantages et inconvénients des différents capteurs détaillés dans l'état de l'art . . . . .	165
4.4	Propriétés de la caméra utilisée . . . . .	197
4.5	Fréquences théoriques de la poutre encastree libre simulée . . . . .	198
4.6	Propriétés de la poutre encastree libre . . . . .	199
4.7	Analyse modale : comparaison entre une analyse de Fourier et notre méthode . . . . .	204
4.8	Résolution des capteurs utilisés . . . . .	206
9	Fréquences propres extraites des spectres . . . . .	242



# Table des figures

1.1	Illustration des objectifs attendus de l'utilisation de méthodes d'assimilation de données.	14
1.2	Schéma d'une caméra rapide . . . . .	15
1.3	Schéma d'une caméra rapide . . . . .	16
1.4	Fonctionnement d'un capteur pour les technologies CMOS et CCD . . . . .	17
1.5	Deux exemples d'architecture de stockage de vidéos rapides pendant l'acquisition. . .	21
1.6	Vidéo rapide d'une formule 1 des mers : l'objet rapide de la scène est le bateau, la mer qui n'a pas la même dynamique peut être considérée comme immobile. . . . .	22
2.1	Résumé graphique présentant la démarche et les résultats du chapitre 2 . . . . .	29
2.2	Deux manières différentes de voir le signal vidéo . . . . .	30
2.3	Schéma de Bayer utilisé sur les capteurs noir et blanc afin de permettre de capturer les couleurs. On réduit de moitié la résolution du capteur . . . . .	32
2.4	Quelques représentations de modèles colorimétriques . . . . .	33
2.5	Représentation Eulérienne et Lagrangienne pour une règle encastrée libre avec des marqueurs extraits par des méthodes de vision . . . . .	35
2.6	Pyramide gaussienne en rouge avec une fenêtre d'analyse spatiale en jaune : la pyramide permet d'utiliser la même fenêtre pour traiter des échelles différentes. . . . .	39
2.7	Schéma montrant la reconstruction d'image à l'aide du champ de mouvement des différents blocs de l'image précédente . . . . .	40
2.8	Succession des images dans une vidéo : seules les images jaunes I existent dans le flux compressé, puis sont reconstruites les images P en vert et enfin les images B en bleu . . . . .	44
2.9	Schéma mettant en lumière les zones de l'espace colorimétrique HSL où la distinction colorimétrique est plus difficile . . . . .	51
2.10	Différentes manières d'analyser l'image, d'un point de vue de la densité et d'un point de vue de localisation et concentration des informations. Figure extraite de Krig (2014) page 192. . . . .	56
2.11	Exemple de détection MSER sur les mires qui seront utilisées dans le chapitre 4 . . .	57
2.12	Figure présentant la segmentation automatique d'une scène d'un homme sur fond fixe. Chaque zone colorée représente un objet qui ne correspond pas à un objet réel : c'est ce type d'objet que nous appellerons objet vidéo. . . . .	63
2.13	Schéma résumant notre méthode de décomposition de mouvement permettant l'extension de l'estimation de mouvement pour de grands déplacements . . . . .	65
2.14	Schéma de la structure algorithmique choisie pour l'implémentation de notre méthode. . . . .	67

## Table des figures

2.15	Représentation des différentes manières de détecter un même objet. Dans notre cas, seules les trois dernières descriptions conviennent pour faire correspondre objet vidéo et objet mécanique. . . . .	68
2.16	Vidéo originale du dé. Les vidéos suivantes se serviront de certaines images de cette vidéo initiale et les interpolations seront comparées à cette référence. . . . .	72
2.17	Représentation de la trajectoire du centre de gravité du dé (en bas à gauche) et de son ombre (en haut à droite) . . . . .	73
2.18	Représentation des différences entre les images originales (canal rouge), les images reconstruites à partir du flot optique de Färneback (canal vert) et celles reconstruites à partir du flot optique de Brox (canal bleu) . . . . .	73
2.19	Vidéo présentant les masques interpolés des objets superposés à la vidéo originale. Cette interpolation est faite à partir d'une image sur quatre de la vidéo originale. . . . .	74
2.20	Vidéo présentant la vidéo interpolée avec notre méthode. . . . .	74
2.21	Vidéo originale « Balle de Tennis ». . . . .	75
2.22	Représentation de la trajectoire du centre de gravité de la balle de Tennis . . . . .	76
2.23	Représentation des différences entre les images originales (canal rouge), les images reconstruites à partir du flot optique de Färneback (canal vert) et celles reconstruites à partir du flot optique de Brox (canal bleu) . . . . .	77
2.24	Vidéo superposant le masque d'une amélioration de rendu de facteur 10 avec la vidéo originale : on multiplie par 10 les images dans la nouvelle séquence. . . . .	77
2.25	Vidéo synthétisée proposant une amélioration de rendu d'un facteur 10. . . . .	77
3.1	Résumé graphique présentant la démarche et les résultats du chapitre 3 . . . . .	87
3.2	Schéma synthétisant une assimilation de données . . . . .	92
3.3	Illustration synthétisant les différents espaces, variables et fonctions ainsi que les erreurs introduites en assimilation de données . . . . .	93
3.4	Schéma comparant les étapes d'assimilation de données avec le formalisme stochastique . . . . .	98
3.5	Schéma animé exposant les différents fonctionnements d'un filtre <i>EKF</i> et d'un filtre <i>UKF</i> . . . . .	112
3.6	Schéma de la structure d'un filtre dual. . . . .	126
3.7	Vision chronologique des articles principaux abordés dans la section. . . . .	131
3.8	Modèle de la règle glissant sur la table . . . . .	136
3.9	Vidéo du scénario 1 . . . . .	138
3.10	Vidéo du scénario 2 . . . . .	138
3.11	Vidéo du scénario 1 . . . . .	139
3.12	Vidéo du scénario 2 . . . . .	139
3.13	Évolution des paramètres pendant l'assimilation des différentes simulations . . . . .	140
3.14	Évolution des paramètres pendant l'assimilation des différentes simulations . . . . .	140
3.15	Évolution des paramètres pendant l'assimilation des différentes simulations . . . . .	141
3.16	Présentation de la disposition de l'essai d'estimation d'une règle glissante et de l'affichage de l'estimation . . . . .	141
3.17	Vidéo du scénario 1 avec affichage des forces de frottement en réalité augmentée et de l'extrapolation de la position finale en vert. . . . .	142
3.18	Vidéo du scénario 2 avec affichage des forces de frottement en réalité augmentée et de l'extrapolation de la position finale en vert. . . . .	142
3.19	Évolution des variables d'état sur l'essai 1 . . . . .	143
3.20	Évolution des variables d'état sur l'essai 2 . . . . .	143
3.21	Évolution des variables d'état sur l'essai 1 . . . . .	144

## Table des figures

3.22	Évolution des variables d'état sur l'essai 2 . . . . .	144
3.23	Vidéo du scénario 1 avec une assimilation sur des observations aléatoires de fréquence moyenne équivalent à un facteur 20 . . . . .	145
3.24	Vidéo de l'essai 1 assimilé et interpolé d'un facteur 10 . . . . .	145
3.25	Vidéo de l'essai 1 assimilé et interpolé d'un facteur 20 . . . . .	145
3.26	Vidéo de l'essai 2 assimilé et interpolé d'un facteur 10 . . . . .	146
3.27	Vidéo de l'essai 2 assimilé et interpolé d'un facteur 20 . . . . .	146
3.28	Évolution des erreurs de position et d'angle sur l'essai 1 . . . . .	146
3.29	Évolution des paramètres sur l'essai 1 . . . . .	147
3.30	Évolution des paramètres sur l'essai 2 . . . . .	147
4.1	Résumé graphique présentant la démarche et les résultats du chapitre 4 . . . . .	159
4.2	Figure expliquant la transformée de Hilbert tirée de Feldman (2011a) p13 . . . . .	182
4.3	Les fameuses mires à damier connues pour leur utilisation en crash-test et les mires utilisées chez Vannier-Kinoptik . . . . .	188
4.4	Proposition d'un nouveau design de mires comprenant à la fois des hautes et des basses fréquences : les mires HBF. . . . .	189
4.5	Schéma présentant l'assimilation de données associée à des mesures vidéo pour des scènes 2D . . . . .	191
4.6	Schéma présentant les espaces présentés pour l'assimilation 3D . . . . .	193
4.7	Photographie du montage expérimental pour le réglage . . . . .	200
4.8	Exemple d'acquisition de la poutre sollicitée au choc. À noter que les points rouges au centre de chaque mire sont issus du suivi par seuillage . . . . .	200
4.9	Résultats de l'identification en fréquence et amplitude pour la simulation non-linéaire d'une poutre encastree-libre . . . . .	201
4.10	Résultats de l'identification en fréquence et en amplitude pour la simulation non-linéaire d'une poutre encastree-libre : RMSE . . . . .	202
4.11	Analyse Spectrale des amplitudes identifiées des 3 différents modes . . . . .	202
4.12	Résultats de l'identification pour le cas test réel d'une poutre encastree-libre . . . . .	203
4.13	Gain de l'analyse de Fourier pour chaque point de mesure sur le cas test réel d'une poutre encastree libre . . . . .	204
4.14	Diagramme Q-Q pour vérification de la nature gaussienne du résidu de l'identification pour différents nombres de modes identifiés . . . . .	205
4.15	Deux manières différentes de présenter les erreurs : le RMSE en fonction du temps et la fonction de densité de probabilité centrée en fonction du nombre de modes identifiés	205
4.16	Résidu spatial et modes normés identifiés avec l'affichage des modes théoriques . . . . .	206
4.17	Fusion de plusieurs sources de données avec les réglages sur la covariance . . . . .	207
4.18	Configuration de l'observation et résultat temporel du RMSE pour l'extrapolation spatiale . . . . .	207
4.19	Configuration de l'observation et résultat temporel du RMSE pour l'interpolation . . . . .	208
4.20	Configuration de l'observation et résultat temporel du RMSE pour l'interpolation temporelle d'un facteur 5 . . . . .	209
4.21	Explication pour les matrices de MAC exposées dans la suite . . . . .	210
4.22	Résultats de l'estimation des paramètres modaux par le schéma Blind-SAMI et comparaison avec les vraies valeurs de la simulation . . . . .	211
4.23	Résultats de l'identification en fréquence et amplitude pour la simulation non-linéaire d'une poutre encastree-libre : RMSE . . . . .	211
4.24	Caractérisation de l'erreur pour le schéma B-SAMI . . . . .	212

## Table des figures

4.25	Comparaison entre les estimations de K-SAMI et de B-SAMI . . . . .	212
4.26	Résultats de l'estimation des paramètres modaux par le schéma B-SAMI sur le cas test réel de la poutre encastrée libre . . . . .	213
4.27	MAC et déformées pour deux instants de l'estimation . . . . .	214
4.28	Une nouvelle méthode de magnification basée sur de l'identification modale et sur de la compensation de mouvements . . . . .	217
4.29	Image 160 de la séquence du ski avec une excitation de type choc (canal rouge) - Magnification du 1er mode uniquement (canal vert) - Magnification du second mode (canal bleu) en gardant la contribution du premier mode . . . . .	217
30	Schéma montrant les erreurs occasionnées par la discrétisation pour une perturbation de trajectoire . . . . .	233
31	Vidéo présentant la vidéo interpolée avec notre méthode. . . . .	234
32	Comparaison des vitesses aux points interpolés . . . . .	235
33	Comparaison de l'interpolation de positions par spline et de celle linéaire de l'accélération avec la trajectoire réelle de la balle . . . . .	236
34	Modèle de la règle glissant sur la table . . . . .	237
35	Image issue de la séquence de la poutre munie de mires HBF traitée par seuillage : le centre déterminé est en rouge. . . . .	241
36	Spectre avant, pendant et à la fin du choc pour comparer les mesures par extraction vidéo et les mesures de position . . . . .	242
37	Comparaison des mesures par caméra et par capteur de position avant et après un choc : valeurs et erreurs en fonction du temps . . . . .	243
38	Schéma explicatif sur la provenance des erreurs de mesures entre les capteurs . . . . .	244
39	Résultats dynamiques de la comparaison entre déformées modales extrapolées estimées et déformées modales réelles pour des observations partielles sur l'extrémité de la règle sur le cas test réel de la poutre encastrée libre pour les trois premiers modes . . . . .	245
40	Résultats dynamiques de la comparaison entre déformées modales interpolées estimées et déformées modales réelles pour des observations équiréparties sur la règle sur le cas test réel de la poutre encastrée libre pour les trois premiers modes . . . . .	246
41	Résultats dynamiques de la comparaison entre déformées modales interpolées estimées et déformées modales réelles pour une interpolation temporelle d'un facteur 10 sur le cas test réel de la poutre encastrée libre pour les trois premiers modes . . . . .	247
42	Résultats dynamiques de l'évolution du critère de MAC et de l'identification des déformées modales correspondantes sur le cas test réel de la règle encastrée libre pour les trois premiers modes . . . . .	249

# Remerciements

Lorsqu'on vient à regarder en arrière, il est particulièrement difficile de retranscrire la complexité des choix qui ont été faits. Éclairés par l'actuel présent, les faits se teignent irrémédiablement d'une couleur rétroactive. Dans un océan de perception où l'empirisme règne, décrire les pavés qui ont composé le chemin se heurte à la réalité de l'interprétation. Comment restaurer ces choix ? En faisant confiance au reflet des choses par les personnes qui ont bordé le chemin et qui en forme comme un kaléidoscope à la fois rationnel et émotionnel.

Je tiens à remercier très chaleureusement Emmanuel Foltête et Xavier Chimentin pour avoir accepté de rapporter mon travail, c'est un privilège d'avoir bénéficié de leur attention et de leur bienveillance. Tous mes remerciements vont également à Pierre Vacher, président du jury et Ludovic Chamoin, qui ont accepté avec enthousiasme d'examiner ces travaux.

J'ai eu la chance d'avoir un jury composé de personnes que j'ai lues, et des personnes que j'admire : c'est un grand honneur pour moi d'avoir été examiné par ces experts.

On ne devient pas docteur tout seul.

Chaque thèse est une aventure unique, un chemin à travers les hypothèses et les difficultés qu'il est difficile de comparer. Pourtant j'ai le sentiment d'avoir eu un encadrement exceptionnel. Une thèse, c'est un sujet, mais c'est surtout une aventure humaine. Cette aventure a été belle, et je le dois énormément à mon directeur de thèse, Jean-Luc Dion. C'est un directeur en qui j'ai eu une entière confiance et qui me l'a rendue. Tu m'as appris à franchir des montagnes avec un mot qui te caractérise si bien : ton dynamisme ! Merci pour ta générosité dans l'encadrement, ta ténacité et ta vision de la recherche : investie, galvanisante et pragmatique.

L'aventure m'a amené à rencontrer Ronan Le Breton dans mon encadrement quotidien. Au-delà de toute forme de technique, Ronan m'a encadré avec minutie, générosité et empathie. Ce sont des qualités humaines dont j'espère avoir réussi à m'imprégner, essentielles pour le travail en équipe. Malgré son arrivée à mi-parcours, Ronan a été un encadrant à part entière dans ces travaux. Merci également à Thierry Soriano pour avoir accepté d'encadrer ce travail depuis Toulon et d'avoir eu cette difficile tâche de donner un avis extérieur.

Cet encadrement a été très complet par sa complémentarité et cette relation de confiance, un véritable fil rouge sous-jacent de ces travaux.

Pour m'avoir fait confiance, j'aimerais tout particulièrement remercier Bernard Roux, dirigeant de l'entreprise centenaire Vannier-Kinoptik. Bernard m'a montré une vision humaniste et transparente de l'entreprise et du management, doublée d'une éloquence et d'une gestion du risque rare qui forcent l'engouement.

Alors non, on ne devient pas docteur tout seul.

C'est aussi le reflet de mon environnement professionnel. Côté entreprise, je tiens à remercier Pascal Lorette et Luc Fesquet, les deux piliers de Vannier-Kinoptik pour m'avoir accueilli et transmis un peu de cette entreprise de bientôt 130 ans ! Au laboratoire, dans l'équipe VAST bien sûr, réunie dans ses nouveaux locaux, merci à Nicolas Peyret, Franck Renaud, Olivia Penas, Jean-Baptiste Casimir, Imad Tawfiq, à mes voisins acousticiens Stéphane Job et Benoît Nennig, à mes plus récents voisins PRAG Alain Stricher, Sylvain Courtois, Nicolas Tijoux, Alexandre Mares et Pierre Vinter, et les nouveaux venus Martin Ghienne et Stefania Lo Feudo. Que ce soit autour d'un café, pour de l'enseignement ou dans le cadre d'un projet, c'est aussi toutes ces discussions, ces idées, et ces conseils qui ont ponctués mon travail et même invisibles, ils ont eu leur importance.

Un grand merci à Gwendal Cumunel et Xavier Régat pour ces discussions autour de la mesure de champ par vidéo, mais également à Gaël Chevallier que je n'ai pas eu la chance de connaître à Supméca mais à Besançon. S'il y a une personne qui n'a pas participé à ces travaux mais qui m'a montré avec bienveillance quelques-unes des facettes de l'expertise industrielle, il s'agit de Mickaël Bray dont la subtilité et les traits d'esprits me manquent déjà.

J'ai découvert ce joli château au sommet de la colline face à La Défense qu'est Supméca et je m'y suis bien senti, avant tout parce qu'on m'y a bien installé. Un grand merci aux secrétaires Christel, Véronique et Amel qui déchargent avec le sourire les chercheurs d'un travail chronophage. Merci à Mireille et aux pôles techniques qui ont toujours une solution, je pense notamment à Arkadiusz, Christophe et Christophe (Make great with 3D video) et Abdel. Merci à Adrien et Antoine de la communication. Les autres équipes ne sont pas en reste, j'ai pu découvrir Tony Da Silva, Muriel Quillien, Jean-Philippe Créte et Julien Fortes Da Cruz, mais également Philippe Serré et Jean-Yves Choley dont je remercie la bienveillance. Merci à Christophe Massacrier, à Marianne Séchet, et aux PLACIS Boys : Alexis, Antoine et Elmehdi.

L'ambiance n'aurait pas été la même sans les doctorants bien entendu ! Je commencerai dans l'ordre par Etienne et Antoine<sup>1</sup>, pour avoir instauré quelques pauses salvatrices dans des après-midis à l'horizon morose. J'aimerais remercier également Kamil, Marco, Hadrien et Adrien mais aussi Karl, Georges et Romain pour tous ces petits moments de vie ô combien importants partagés au labo, ou ailleurs ! Bonne chance à Emna et Reza. Aux bons souvenirs de quelques voyages dont Los Angeles, je pense à Kevin et Pauline de FEMTO que je n'ai pas encore cités et les autres aventuriers qui se reconnaîtront. Merci aussi à l'équipe des JJCAB qui ont réussi le pari d'organiser cette conférence sur deux sites non sans humour et bonne humeur. Je pense à Aro, Arthur, Sylvain et Christophe du CNAM. Un clin d'oeil à Antoine du LAUM et Anne du GIPSA. Merci également à Pierre Margerit et Xavier Amandolese pour ces enrichissantes sessions jazzistiques, qui ne vont pas s'arrêter là !

On ne devient vraiment pas docteur tout seul.

Le lecteur arrivant jusqu'à cette ligne a déjà eu un aperçu de cet environnement professionnel qui m'a porté pendant ces trois années.

Sur le chemin de la vie, il y a ces personnes qui laissent une empreinte fugace mais tellement influente sur la trajectoire. Je commencerai chronologiquement, en adressant mes remerciements éternels à Mme Maria Cristian qui m'a transmis avec sagesse et générosité tant de valeurs à travers sa sensibilité musicale, son équilibre et sa passion. Un équilibre et une passion que j'ai retrouvés quelques années plus tard dans la classe de Mme Claire Pigeot que je remercie tout autant. Je souhaite également remercier Mme Lafon pour m'avoir redonné du cœur à l'ouvrage, à une époque où je n'en n'avais plus.

---

1. qui n'est pas un doctorant mais un autre PLACIS boys.

## Remerciements

On n'a pas souvent l'occasion dans la vie professionnelle de remercier l'ensemble des personnes qui gravitent autour nous, qui vous donnent l'énergie, l'initiative, et avant-tout le sourire! J'en profite! Un grand merci à mon entourage parisien, Mathieu & Laura mes voisins du destin, Vincent & Alexis pour les rendez-vous ESPCI, Antoine mon infatigable compagnon de volley, la team PVC du mercredi et Flo et Tony les volleyeurs de Grenoble, mais aussi Alexandra, Nicolas et Laura.

Un immense merci aux Groomjes pour ces innombrables aventures musicales, cette complicité mutuelle, la confiance musicale que vous m'accordez, et cette volonté de perfectionnisme, malgré la distance non négligeable Paris Grenoble Clermont. J'aimerais remercier Thomas la flûte de caractère, Audrey la voix à la volonté de fer (parce que le dire c'est bien, mais le fer, c'est mieux!), Quentin le beatbox précheur du Groomje, Paul le sax « Macéo Approved » et Valentin l'afro batteur, Thibaut la basse brillante et modeste, Antoine la trompette avec Miles sur les talons, Max l'alto sax, Cyril la trompette fraîchement doctorante (bonne chance!), Rémy la guitare à l'ADN mi-signal processing mi-blues (un peu plus blues quand même), Carl le crooner et Grégoire l'ingé son/lumière qui connaît mieux le set que les musiciens! Je n'oublie pas non plus Manu, de passage au saxophone et expert de la caméra, qui m'a aidé à la fois pour m'engager dans ce sujet de thèse comme pour le clôturer. Un grand merci aussi aux deux feux follets musicaux strasbourgeois Léo & Chloé qui vous donnent toujours la force d'aller plus loin dans vos projets. Une pensée à Anne, Raph, et Mido. Toutes ces personnes m'ont énormément apporté dans la vie comme dans la musique, et je ne serai pas le même sans eux.

Les voyages forment la jeunesse, que pourrait-on dire des compagnons de voyage alors! Je pense à Kevin, le docteur/producteur, Guillaume, Aubry, Grand & Loulou, Alicia & Greg, Bouli, Pretox et Titou, le Britano-Birman-Gwada-Limousin. Un remerciement tout particulier à Clara, Claire et Claire (en italien *Chiara, Chiara è Chiara*) qui sont là à tous les moments importants. Merci également à Brigitte, Agathe, Lise et Antoine pour leur amitié.

Enfin, j'aimerais remercier la *famiglia* qui m'a toujours soutenu. À ma grand-mère qui voulait un médecin dans la famille qui aura finalement son docteur. À mon grand-père qui m'a montré qu'avec de l'esprit et beaucoup de cœur on peut devenir quelqu'un d'exception. Le destin aura voulu que cette thèse aurait été presque à domicile pour lui.

Merci à Daniel, Anne-Marie, Jean-Pierre et Antoinette, Antonin, Marina, Guillaume et Elene, sans oublier Michel, Gabriel et Eden.

Je ne sais que dire pour remercier mes parents de m'avoir mis sur ce chemin pas forcément de la science, mais de la curiosité et de la passion. Peut-être qu'au-delà de la connaissance qui pousse la technologie, c'est avec la subtilité de leur pédagogie que les vocations se sont formées. Un grand merci à mon frère Charles qui m'a donné avec classe et expertise de nombreux coups de pouce pour balayer ou confirmer des idées d'un revers de main.

Enfin, qu'aurais-je pu faire sans Pauline pour partager chaque jour de cette aventure, qui a su écouter avec patience mes idées et mes questions en éclairant toujours de son reflet d'autres chemins inexplorés...



# Introduction générale

En une décennie la vidéo numérique s'est développée dans les applications grand public et industrielles comme un outil de tous les instants. Son potentiel est immense. Dans les systèmes biologiques, la vision est un des sens apportant le plus d'informations sur son environnement. Cependant il serait restrictif de ne parler que de l'œil, c'est-à-dire la partie du système qui effectue l'acquisition. En réalité, le système visuel doit également prendre en compte les parties du cerveau où l'information est distribuée, dont le cortex visuel fait partie. L'ensemble des mécanismes et des fonctions des différentes zones actives pour la vision ne sont pas encore élucidés. Il apparaît que l'interaction de ces zones permettent les traitements de détection bas niveau (orientation, contours), haut niveau (formes, mouvement), de reconstruction 3D, ou d'interprétation des images en fonction des sens et des connaissances. Un système de vision prend tout son sens non pas à l'acquisition d'une vidéo mais lors de l'extraction d'informations de celle-ci. Dans cette course numérique pour maximiser la perception il est primordial de maîtriser toute la chaîne de mesure, du phénomène physique du capteur jusqu'à l'extraction d'informations de l'image. Ce manuscrit s'intéresse exclusivement à des séquences particulières, prises à l'aide de caméras rapides, permettant de faire ressortir plus d'informations de la dynamique d'une scène filmée. Extraire des informations pertinentes de ces séquences est une tâche d'autant plus complexe que les observations obtenues sont une projection des informations en trois dimensions sur un capteur en deux dimensions. Différentes stratégies sont développées pour permettre de pallier cette projection. Outre une solution matérielle qui suggère l'ajout d'autres capteurs à la scène, l'état de l'art de la vision par ordinateur propose toute une série de méthodes permettant de segmenter une séquence vidéo, d'isoler les objets d'intérêt et de les mesurer. Ces méthodes demandent de poser des hypothèses supplémentaires sur la scène comme de faibles mouvements entre deux images ou des objets d'intérêts ayant une dynamique plus rapide que le fond. Il est alors utile de distinguer les objets considérés en vision, c'est-à-dire des groupes de pixels, des objets réels filmés. Cette distinction limite beaucoup la perception des systèmes de vision par ordinateur, qui n'ont pas toujours les informations suffisantes pour analyser en continu les séquences vidéos.

Les travaux proposés dans ce manuscrit consistent à concevoir des méthodes proposant d'enrichir l'acquisition de séquences issues de caméras rapides. Ils s'inscrivent dans un projet fixant un cadre et des objectifs. Ainsi, ils devront proposer des méthodes permettant de faire de l'identification de paramètres physiques invisibles, d'offrir une prédiction sur une scène et de pouvoir interpoler deux images successives. Par ailleurs, le potentiel de ces objectifs pour la compression ou la réalité augmentée peut être important, à condition de proposer une reconstruction de vidéos synthétiques à partir de l'acquisition initiale et des informations extraites.

L'utilisation de caméras rapides étant souvent spécifique à de l'analyse de mouvements, du contrôle en ligne ou du diagnostic, il est particulièrement intéressant de relier ces informations à des modèles mécaniques pour pouvoir accéder à des paramètres invisibles mais influençant suffisamment

la dynamique pour être identifiés. Aussi, compte-tenu du contexte applicatif il semblait intéressant de chercher à aller plus loin dans l'utilisation de modèles dans les vidéos. L'évolution technologique tendant vers les flux vidéos, il est nécessaire de concevoir des méthodes permettant à un système de vision d'analyser les informations dans un temps réduit après l'acquisition, idéalement en suivant un schéma temps réel. Cette notion de schéma temps réel est importante : il s'agit de schémas dont la structure et le coût algorithmique permettent d'envisager une optimisation vers une utilisation temps réel. Allier les observations d'une vidéo avec la connaissance d'un modèle présente un intérêt en vidéo rapide puisque les scénarios des scènes filmées sont généralement connus à l'avance.

Ce manuscrit se focalise sur l'utilisation de modèles dynamiques et physiques en schémas temps réels avec des données issues de caméras rapides. Trois axes de recherches sont présentés.

- Les outils du domaine de la vision utilisent des modèles pour caractériser les mouvements intervenant dans une séquence vidéo. Aussi l'interpolation d'images est un problème déjà résolu dans le cadre particulier de petits mouvements entre deux images consécutives. Nos problématiques pouvant précisément amener à des déplacements importants entre deux positions, ces méthodes ne sont donc pas appropriées dans notre cas. De plus les modèles utilisés étaient de nature cinématique, aucun paramètre physique n'est introduit, ce qui limite les possibilités du processus d'identification. Cependant la nature générique de ces méthodes présente un intérêt et nos objectifs peuvent être atteints partiellement de cette manière à condition de réussir à correctement estimer les mouvements d'objets en schéma temps réel.
- Afin de pouvoir introduire des modèles physiques plus complexes décrivant la scène filmée, nous avons choisi de nous placer dans le cadre des méthodes d'assimilation de données. Cette famille de méthodes regroupe des algorithmes permettant d'assimiler des données provenant de capteurs de natures différentes avec un modèle physique. Afin de satisfaire l'exigence du schéma temps réel, nous nous intéresserons aux méthodes séquentielles, connues également sous le nom de filtres de Kalman. Ces filtres offrent la possibilité de relier des observations issues de caméras rapides avec des modèles mécaniques à condition de concevoir un modèle adéquat et d'utiliser une chaîne de traitement vidéo adaptée. La correspondance entre les observations vidéos et le modèle souffre toutefois d'une différence de description puisque la vidéo observe la scène à travers une grille de pixels fixes tandis que les modèles mécaniques considèrent l'évolution d'un objet décrit par un nombre de paramètres réduits.
- Deux cadres d'application sont choisis pour illustrer les concepts développés. Dans le cas de la dynamique de grands mouvements, la prédiction et l'identification présentent un intérêt immédiat, tandis que l'interpolation permet de ralentir la séquence tout en respectant la physique de la scène. D'un certain point de vue, l'analyse modale présente des problématiques en adéquation avec les concepts développés dans ces travaux. La recherche de techniques de mesures sans contact permettant de caractériser le comportement modal de structures dans des cas non-stationnaires place les caméras rapides comme une solution intéressante. Suivre un nombre important de points à un même instant sans contact tout en développant un modèle adapté permet aussi d'identifier les paramètres modaux d'une structure en la décomposant en temps réel sur ses déformées modales.

Ces problématiques nous offrent l'occasion de proposer plusieurs apports. Le cadre d'étude est d'abord décrit dans le chapitre 1 afin de fournir un maximum d'informations au lecteur sur le contexte entourant ces travaux et éclairer certains choix. Le chapitre 2 se place dans un cadre exclusivement vidéo dans lequel nous avons cherché le lien pouvant exister entre les différentes descriptions d'une scène. En s'appuyant sur des représentations utilisées notamment en mécanique des fluides, nous avons ainsi introduit une double description d'une scène à la fois eulérienne et lagrangienne. Afin de

pouvoir étendre l'interpolation existant pour des petits mouvements à des grands mouvements, une méthode tirant profit d'une composition de mouvements est proposée : une description lagrangienne est utilisée pour décrire le mouvement dans le repère de l'image tandis qu'une description eulérienne est utilisée dans le repère local de l'objet.

Si elle permet de faire de l'interpolation et de la compression tout en synthétisant des séquences vidéos, nous constatons que cette méthode, si elle est générique, ne recèle pas assez d'informations physiques pour atteindre tous nos objectifs. Une orientation différente est prise dans le chapitre 3 qui se place dans le cadre de l'assimilation de données dans le but d'introduire des modèles physiques plus complexes et plus spécifiques. Un apport bibliographique sur les méthodes séquentielles y est présenté avec une double approche à la fois stochastique et d'assimilation. Cette section regroupe et compare de nombreuses extensions des filtres de Kalman orientées pour un utilisateur. Un schéma d'assimilation de données est alors proposé afin de relier un modèle mécanique sous forme différentielle avec des observations vidéos du système. Ce premier schéma d'assimilation est appliqué à un cas test réel et permet d'exploiter le modèle pour faire de la prédiction, de l'identification de paramètres, de l'interpolation respectant la physique du modèle tout en permettant de compresser les séquences vidéos sans pertes de manière sémantique à la scène.

Le chapitre 4 traite de l'analyse modale et présente une solution aux problématiques exposées plus haut dans ce domaine en utilisant une approche d'assimilation de données. Deux schémas d'assimilation différents sont proposés : le Kinematic-SAMI pour *Kinematic multiSensors Assimilation Modal Identification* (identification modale par assimilation avec un modèle spatial) et le Blind-SAMI pour *Blind multiSensors Assimilation Modal Identification* (identification modale par assimilation aveugle). Ils permettent de faire de l'identification de paramètres modaux par des mesures vidéos dans un schéma temps réel. Les SAMI présentent des avantages différents. Le Kinematic-SAMI utilise une base modale à priori pour filtrer les observations et permet de faire de l'identification, de l'interpolation et de la prédiction en utilisant des observations partielles pour décrire une structure complète. En plus des paramètres modaux de fréquence et d'amortissement, le Blind-SAMI estime les déformées modales expérimentales. La méthode Kinematic-SAMI est ensuite utilisée pour une autre application : un schéma d'identification modale multi-caméra sur un modèle 3D sans passer par les étapes de reconstruction 3D classiques. Enfin, une méthode de magnification de mouvement différente de celles présentées dans l'état de l'art est introduite. Utilisable avec les deux schémas, elle permet de synthétiser des vidéos en amplifiant le comportement modal de la structure observée, agissant comme une « loupe modale » sur les images de la structure.



# 1

## Problématiques de l'étude

*Everything should be made as simple as possible, but not simpler.*

– Albert Einstein

### Sommaire

---

I	Contexte de la CIFRE . . . . .	11
II	Cadre technologique de l'étude . . . . .	14
III	Conclusion : apports proposés . . . . .	23
	Bibliographie du chapitre 1 . . . . .	25

---

### Résumé

Ce chapitre introductif détaille le contexte des travaux présentés dans ce manuscrit afin de donner un maximum de clés de lecture. Après avoir introduit le projet FAMOSYS qui englobe ces travaux à travers ses objectifs et ses acteurs, il est proposé de s'immerger dans le contexte de la vidéo rapide. Un état de l'art matériel introduit les technologies des caméras rapides tandis que nous précisons les hypothèses de travail utilisées dans le reste du manuscrit découlant de ce contexte particulier. Enfin nous résumons les apports segmentés par chapitre afin de permettre au lecteur d'avoir une vision d'ensemble sur les travaux.



## I Contexte de la CIFRE

Prendre connaissance d'une étude n'est pas toujours facile pour le lecteur. L'exercice ne l'est pas moins pour un rédacteur soucieux de la bonne compréhension de son message. Cette transmission du message se fait d'autant mieux que l'émetteur connaît le destinataire et les filtres inhérents à leur communication qui peuvent être culturels, professionnels ou encore contextuels. Aussi l'écriture amène un problème complexe pour le rédacteur puisqu'il se trouve en défaut d'information sur l'ensemble des personnes qui prendront connaissance de son document. Notre démarche consiste alors à inverser les rôles. En plus de la pédagogie recherchée dans ce manuscrit, nous avons souhaité exposer le contexte afin d'expliquer les hypothèses et les choix qui ont été faits. Nous avons donné à ce chapitre introductif le rôle de présenter ce contexte, c'est-à-dire le projet, ses acteurs et ses objectifs. Il nous permettra également de pousser la porte du monde de la caméra rapide, duquel nous tirerons des hypothèses de travail.

### I.A. Présentation du projet

L'imagerie est de plus en plus employée dans de nombreux domaines professionnels pour surveiller, contrôler, qualifier, diagnostiquer ou mesurer des phénomènes dynamiques. Son utilisation en remplacement de capteur mono-dimensionnels est de plus en plus fréquente pour des raisons de maturité de la technologie, de coût mais également parce qu'elle possède l'avantage de capturer un grand nombre de points à un même instant, donnant directement accès à un champ de mesure.

Le projet se focalise sur un certain type d'appareils d'imagerie : les caméras hautes vitesses. L'enregistrement de séquences d'images numériques à haute définition et haute cadence permet d'accéder à des mesures optiques de phénomènes rapides avec une problématique majeure : la quantité de données. Ces données sont en majorité non-significatives et les identifier permettrait au niveau caméra de limiter les données à transférer tout en abaissant considérablement le stockage nécessaire. Les performances actuelles des caméras hautes cadences et des logiciels de traitement d'image associés ne permettent pas un traitement en temps réel ou légèrement différé des séquences capturées : la plupart des utilisateurs sont contraints d'utiliser des post-traitements.

Le projet FAMOSYS (pour *FAst MOving SYStems*) a pour objectif de développer des méthodes permettant d'extraire les pixels utiles en temps réel. Il propose d'aller plus loin en réduisant les données aux seules informations nécessaires à l'analyse ou la modélisation : en interprétant ces informations de manière physique, il est possible d'ouvrir un champ d'applications interactives entre l'utilisateur et le phénomène observé. L'utilisation de caméras étant de plus en plus répandue dans l'industrie pour des tâches allant de la surveillance à la métrologie, il est particulièrement intéressant de pouvoir proposer d'utiliser ces observations pour mesurer des grandeurs "invisibles" (au sens non-visuelle) sur l'image. L'utilisation de modèles mécaniques d'une scène observée par une caméra quelconque peut permettre d'estimer des paramètres intervenant dans le modèle. Une caméra ou un ensemble de caméras peuvent ainsi être utilisées comme des capteurs logiciels en plus de leur usage initial. Enfin, si elle est problématique par son coût de calcul, la donnée image est très profitable en sortie de la chaîne pour son potentiel d'interaction avec l'utilisateur. En proposant une liaison entre un modèle et des données vidéos, il est possible de projeter les informations du modèle dans le plan image afin de venir augmenter la sortie vidéo. L'utilisateur peut de cette manière interagir facilement avec des grandeurs qu'il ne perçoit pas par la vision. Outre les applications industrielles en plein essor dans l'industrie 4.0, il existe également un fort potentiel didactique. Provoquer cette

## I. Contexte de la CIFRE

interaction permet de confronter directement l'intuition intrinsèque aux sens avec des grandeurs habituellement intellectualisées. Dans un monde de plus en plus centré sur l'image, l'utilisation de capteurs permettant de capturer plus d'informations que ne le permet une caméra classique est un avantage indéniable, à condition de pouvoir exploiter conjointement ces informations. La caméra haute vitesse fait partie de ces capteurs, puisqu'elle permet de capturer davantage d'informations dynamiques sur la scène qu'elle observe.

Le projet propose ainsi de lier en particulier deux domaines en apparence éloignés : l'utilisation de modèles mécaniques avec des données issues de séquences vidéos rapides après traitement vidéo. Afin de faire ce lien, il est proposé dans ce projet d'utiliser des filtres de Kalman. L'avantage d'utiliser ces filtres est multiple : le coût de calcul est faible, la possibilité de relier un modèle avec des observations d'un système réel et un algorithme séquentiel s'adaptant à nos contraintes temps réel. L'approche choisie dans ces travaux a été celle de l'assimilation de données qui développe des méthodes permettant une liaison entre observations et modèle dont les filtres de Kalman font partie. Nous proposons ainsi dans ces travaux d'assimiler des données images avec des modèles mécaniques à l'aide de méthodes d'assimilation de données.

### I.B. Acteurs

Une thèse CIFRE est un contrat tripartite entre un laboratoire, un industriel et l'ANRT (Agence Nationale de la Recherche et des Technologies). Cette dernière est indispensable pour la mise en place juridique et contractuelle. Elle n'intervient pas dans les travaux autrement qu'en apportant un support financier essentiel sans lequel ces recherches n'auraient pas eu lieu. Les compétences se partagent entre un laboratoire et une entreprise :

**Le laboratoire Quartz - Supméca** et plus spécifiquement l'équipe VAST (pour Vibro-Acoustique et Structures) possède une forte expertise sur le développement de modèles mécaniques et de méthodes stochastiques notamment appliquées à l'analyse vibratoire.

**L'entreprise Vannier-Kinoptik** est un industriel historique de la caméra haute vitesse. Si l'entreprise a un bel âge (depuis 1886), son activité en imagerie rapide a démarré dans les années 70. Elle a fait partie du début de l'aventure de la haute vitesse avec à l'époque des produits de type caméra film. Le virage numérique a été pris en s'adossant à de prestigieux partenaires de la haute vitesse comme Redlake, Weinberger ou AOS Technologies. Avec la plus longue antériorité dans ce domaine sur le marché français, Vannier-Kinoptik continue d'équiper une majorité de grands comptes pour des applications aéronautiques civiles ou militaires, des crash-tests, ou de la métrologie. C'est sur les systèmes de vidéo rapide et le traitement d'image que l'entreprise apportera son expertise.

### I.C. Objectifs de la thèse

Entre la problématique énoncée d'isoler les pixels utiles et la solution proposée d'utilisation de modèles traduisant la dynamique de la scène filmée pour assimiler les données images entrantes, plusieurs directions ont été dessinées en fonction des axes du laboratoire et de l'entreprise. Il s'agit de manière générale de définir un cadre d'application à des modèles mécaniques ou plus généralement à des systèmes dynamiques de manière à pouvoir extraire des informations de la scène et les utiliser en temps réel. Nous verrons plus tard (sous-section II.B. de ce chapitre) que le cadre de la vidéo haute cadence nous permet de travailler sur des scènes dont le scénario est connu à l'avance. Sur ce type de scène, différents objectifs sont attendus.

**L'utilisation de schémas temps réel** tout au long des travaux. Le choix exclusif de schémas temps réel comme les schémas séquentiels pourra permettre une transposition pour une utilisation temps réel des résultats de ces travaux. La contrainte de produire des algorithmes temps réel n'est pas exigée mais il est nécessaire d'avoir la réflexion sur la possibilité d'utiliser les algorithmes développés dans un contexte temps réel. C'est de cette exigence que nous définissons le concept de schéma temps réel : un algorithme dont l'architecture et la charge de calcul permettent une fois optimisé d'être implémentables dans un contexte temps réel.

**L'identification** de paramètres du système. L'utilisation d'un modèle fait nécessairement entrer en jeu des paramètres. Un choix judicieux au niveau de la modélisation permettra de faire intervenir les paramètres que l'on souhaite évaluer.

**La prédiction** de la scène à partir des observations passées du système réel. La capacité de l'algorithme à prédire la scène pourra constituer une sortie pour permettre d'agir sur le système en prévention.

**L'interpolation** d'images entre deux images successives pourra permettre d'augmenter numériquement la vitesse à laquelle on observe le phénomène physique et augmenter ainsi le ralenti de la scène.

**Étudier les applications de ces méthodes pour la compression de données.** Le stockage de séquences vidéos rapides étant problématique pour l'ensemble des utilisateurs, le développement de méthodes permettant de décrire l'évolution de pixels utiles dans une séquence pourrait permettre de reconstruire la vidéo à partir de données fortement réduites et du modèle utilisé.

**La synthèse de vidéos** à partir de données issues du modèle permet de reconstruire une réalité différente de celle qui a été filmée. À la manière de systèmes optiques comme le microscope, sa plus-value réside dans la déformation de l'image. La synthèse de vidéos permet de manière similaire d'offrir à l'utilisateur un résultat visuel.

Ces objectifs seront traités dans chaque chapitre sous des angles différents. Il nous a semblé important de montrer d'autres applications de ces travaux et ainsi de fixer d'autres objectifs

**La prédiction** évoquée ci-dessus est temporelle et s'appuie directement sur le schéma de fonctionnement des filtres de Kalman. Cependant, en fonction du modèle développé, il est possible de faire intervenir des informations géométriques sur la structure. Les données spatiales pourront être utilisées à la fois pour une **interpolation spatiale**, c'est-à-dire l'ajout de points de mesure ou d'observation entre des points réels, mais également en extrapolation que nous appellerons **prédiction spatiale** lorsque le modèle n'est observé que partiellement. Un exemple dans le chapitre 4 illustre comment il permet de reconstruire des zones masquées.

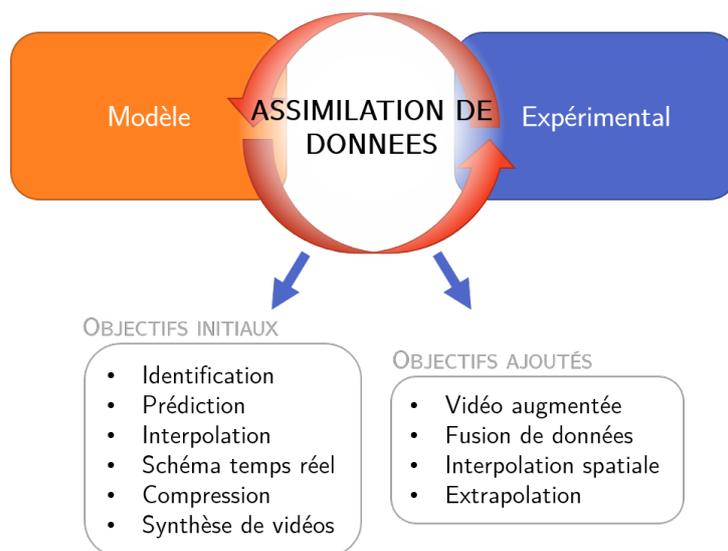
**La fusion de données** est sous-jacente dans ce type de méthodes. Nous verrons dans le chapitre 3 de quelle manière la fonction d'observation peut permettre d'utiliser des sources d'informations issues de différents appareils de mesures. Une application sera présentée également dans le chapitre 4. Bien que ces travaux soient centrés sur l'utilisation de données vidéos, il est important de souligner que ces méthodes permettent une incorporation facile de capteurs de différents types.

**L'utilisation d'un cadre de réalité augmentée** pour l'affichage des résultats. Bien qu'il n'y ait aucun développement scientifique sous-jacent, les algorithmes développés devront comporter une sortie proposant une donnée image augmentée des résultats obtenus.

La figure 1.1 permet de représenter les apports de notre approche. Les applications proposées de ces concepts se déclineront ainsi dans les différents domaines.

— Pour la mécanique de grands mouvements les méthodes développées ici permettront de confronter modèle et expérience. L'utilisation conjointe de données réelles avec des modèles

## II. Cadre technologique de l'étude



**Figure 1.1** – Illustration des objectifs attendus de l'utilisation de méthodes d'assimilation de données.

également connue sous la dénomination *DDDAS* (pour *Dynamic Data Driven Assimilation System*) peut permettre de valider le modèle et le recalibrer par rapport aux observations.

- Au niveau de l'utilisation pour la vidéo rapide, les objectifs de compression mais également d'extrapolation ont des applications directes. L'interpolation temporelle et la synthèse vidéo permettront de générer des séquences ralenties des vidéos en entrée et la vidéo pourra être compressée de manière sémantique à la scène observée, c'est-à-dire en sélectionnant les données inutiles à compresser avec perte. L'identification permet de proposer d'utiliser une caméra comme un capteur logiciel tandis que les applications en prédiction sont utiles pour des systèmes de contrôle comme sur des chaînes de production.
- Les concepts appliqués à l'analyse vibratoire expérimentale nous amènent à continuer l'introduction de l'utilisation de traitement vidéo pour proposer des moyens de mesure sans contact. L'utilisation de modèles spécifiques et les particularités du contexte modal permettront de développer des méthodes en schéma temps réel en reliant l'espace réel à l'espace modal. Il s'agira alors soit d'utiliser un modèle de structure pour augmenter les observations initiales, soit d'utiliser uniquement les observations pour faire apparaître les défauts de la structure.

## II Cadre technologique de l'étude

Le projet FAMOSYS étant centré sur l'utilisation de séquences vidéos haute vitesse, il nous a semblé primordial d'exposer les technologies utilisées dans ce domaine. Le but est de faire ressortir les points bloquants pour améliorer ces caméras : augmenter à la fois la vitesse d'acquisition et la résolution. Après un rapide tour d'horizon de l'état de l'art, nous exposerons les hypothèses qui peuvent être posées à partir de ce cadre particulier. Nous utiliserons ces hypothèses dans le reste du manuscrit.

## II.A. Introduction à la vidéo rapide

Accélérer ou ralentir une vidéo est apparu dès l'invention des premières caméras films. En effet, une caméra film consiste à impressionner tour à tour des images qui défilent avec une certaine vitesse. La projection du film consiste à repasser à la même vitesse ces images, qui sont interprétés par l'œil comme continues donnant ainsi l'illusion du mouvement grâce à la persistance rétinienne.

Accélérer ou ralentir le défilement des images impressionnées sur la pellicule lors de la capture permet alors de modifier la fréquence d'échantillonnage et ainsi capturer plus ou moins d'images pour une même durée d'enregistrement. La séquence sera en revanche repassée à la même vitesse que pour une séquence vidéo filmée à vitesse normale. Une caméra vidéo haute vitesse consiste à **enregistrer à haute vitesse** une scène afin de pouvoir ensuite la rejouer **en ralenti**<sup>1</sup>. Le développement de caméras hautes cadences s'est attaché au fil des années à accélérer les vitesses de captures. Vannier-Kinoptik atteindra jusqu'à 1 000 000 d'images par seconde ou *frame per second (fps)* sur des installations militaires en caméra film à l'aide de miroirs rotatifs. La poursuite de cette aventure en numérique est totalement différente au niveau de la technologie que nous détaillerons dans la suite mais se situe dans la continuité de la course « folle » à l'augmentation du débit d'images capturées. Cette introduction se base sur des documents internes Vannier-Kinoptik, mais également sur l'introduction de Slangen *et al.* (2016) et les cours de Arques (2011).

### II.A.1. Chaîne de mesure

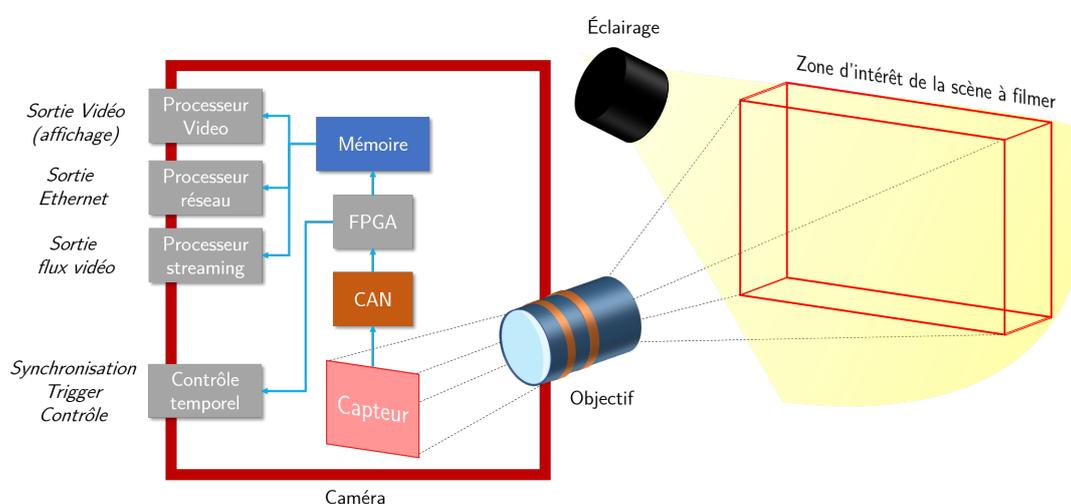


Figure 1.2 – Schéma d'une caméra rapide

Comme pour tout appareil de mesure, l'organe principal de la caméra est son capteur. Ce capteur possède certaines performances et zones de fonctionnement optimales par rapport aux données entrantes. La figure 1.2 présente la décomposition de la caméra comme une chaîne de mesure classique pour un capteur numérique.

Nous remarquerons la présence de l'éclairage qui ne fait pas figure d'option en vidéo rapide. En effet, si l'on considère que la lumière d'un environnement est globalement constante, c'est que le débit de

1. L'annexe A. développe cet aspect sur la perception du temps dans une séquence vidéo

## II. Cadre technologique de l'étude

photons entrants dans la caméra sera globalement constant au cours du temps. Le capteur proposant de faire une intégrale du nombre de photon reçus par le temps d'ouverture de l'appareil, diminuer le pas de temps revient à diminuer du même facteur le débit de photon, et donc la luminosité. La solution en imagerie rapide consiste à utiliser des éclairages puissants afin d'augmenter le débit de photons entrant dans la caméra.

### II.A.2. Notions optiques

Le signal mesuré étant de la lumière, on utilise une optique pour mettre en forme ce signal. Plusieurs paramètres peuvent être réglés par l'optique :

- La netteté de l'image qui consiste à faire coïncider tous les rayons lumineux issus d'un même point sur un point du capteur. Minimiser le flou consiste à faire le point sur l'objectif en modifiant l'écart inter lentille.
- L'objectif est caractérisé par une longueur de focale qui détermine à la fois la distance minimale possible à la scène, mais également la distance pour laquelle la scène se situera à l'infini. L'objectif introduit également un grandissement qui permet de dimensionner la taille du capteur en fonction de sa diagonale.
- L'objectif comporte également un diaphragme qui permet de limiter les rayons entrants dans la caméra en partant des bords. Ainsi, un diaphragme fermé ne laissera passer que les rayons colinéaires à l'axe optique, avec très peu de rayons entrants : le réglage du diaphragme permettra de jouer sur la profondeur de champ, c'est-à-dire la profondeur de la scène sur laquelle on a de la netteté et sur la luminosité entrante.

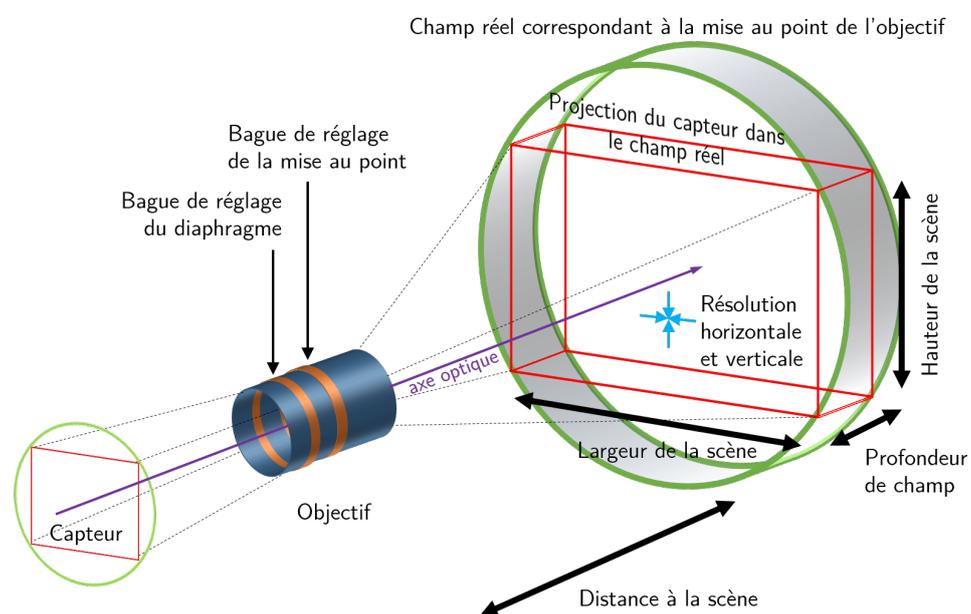


Figure 1.3 – Schéma d'une caméra rapide

Lors du choix de l'optique, il faut prendre en compte le grandissement qui permet d'obtenir la taille de l'image au niveau du capteur, mais également la précision de l'optique en terme de résolution. De la même manière que pour un système électronique, une optique peut être caractérisée en fréquence spatiale selon différentes directions. Cette grandeur est caractérisée par la Fonction de Transfert de Modulation (FTM) de l'objectif et permet de qualifier la résolution de l'image sur le capteur. Nous

avons résumé sur la figure 1.3 l'ensemble des paramètres réglés par l'optique de manière à obtenir une image bien conditionnée dans le plan du capteur :

- positionnement de la caméra ou choix de l'optique pour que la scène ait les dimensions désirées ;
- maximisation de la taille du capteur impressionné par les rayons lumineux par un choix judicieux d'objectif ;
- contrôle du flou souhaité ;
- réglage de la profondeur de champ ;
- réglage de la luminosité entrante.

On notera la notion importante de projection effectuée par l'optique qui amène à considérer la perpendicularité de la scène par rapport à l'axe optique pour caractériser la relation entre la scène et l'image du capteur.

### II.A.3. Le capteur

Derrière l'objectif se situe le capteur<sup>2</sup>. Le capteur est caractérisé par des paramètres repris dans le tableau 1.1.

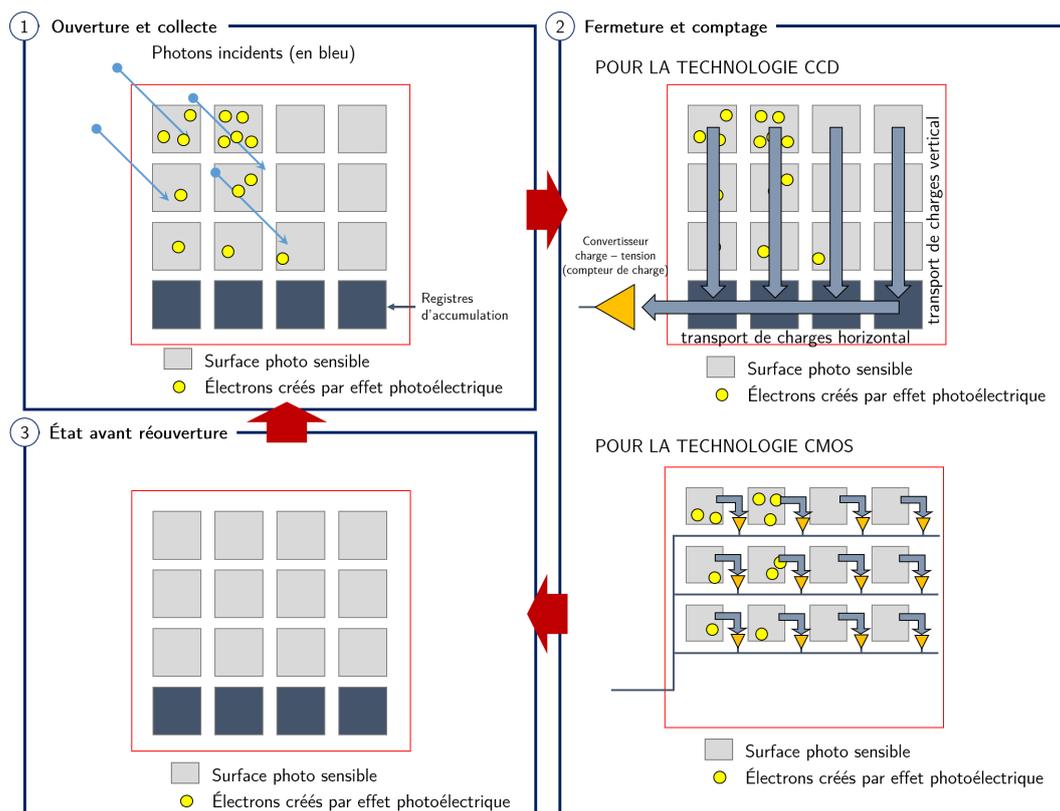


Figure 1.4 – Fonctionnement d'un capteur pour les technologies CMOS et CCD

Le capteur est constitué d'une matrice de pixels. Chaque pixel est structuré comme indiqué sur la figure 1.4, c'est-à-dire avec une surface de collecte de photons et une partie réservée à l'électronique.

2. Nous ne parlerons ici que des technologies de capteurs numériques

## II. Cadre technologique de l'étude

Un capteur fonctionne globalement en deux étapes :

- 1 - **Ouverture** pendant le temps d'exposition réglé au préalable. Les photons sont collectés au niveau de chaque pixel et transformés en électrons grâce à l'effet photoélectrique.
- 2 - **Fermeture et comptage** pendant lequel les électrons sont convertis et transmis à un convertisseur analogique numérique en dehors de la matrice de pixels.

La figure 1.4 ne détaille pas tous les types de capteurs mais fait ressortir les avantages et inconvénients des deux principales technologies sur le marché. Le capteur CCD possède une surface photosensible beaucoup plus élevée que le capteur CMOS ce qui lui permet d'être plus sensible car il collectera plus de photons. En revanche, un mécanisme de transport de charges avant le comptage possède beaucoup de désavantages en terme de pertes et d'indépendance des pixels. Le CMOS propose de résoudre ces désavantages en implantant un compteur de charges directement au niveau de chaque pixel. Nous ne rentrerons pas plus dans les différences entre les capteurs CCD et CMOS, où le second est en train de supplanter le premier sur le marché.

Globalement, si l'on suit l'évolution de la technologie, la tendance est à sacrifier de la surface de collecte pour implanter de plus en plus de la chaîne de traitement électronique au niveau du pixel. Nous verrons dans la sous-sous-section II.A.5. les suites de cette évolution.

Comme on l'aura compris, la luminosité réglée par l'objectif est transformée selon la sensibilité en signal électrique. Un gain peut être appliqué mais il est préférable de jouer si possible sur le temps d'exposition<sup>3</sup>. Ce temps d'exposition doit être par ailleurs limité par la vitesse du phénomène observé : un temps d'exposition trop long par rapport à un objet en mouvement ne figera pas assez l'objet et engendrera un flou de bougé. Soit  $B$  le flou de bougé,  $\mathcal{V}$  la vitesse d'un objet observé,  $T_{exp}$  le temps d'exposition,  $N_{pix}$  le nombre de pixels utilisés dans la région d'intérêt de la vidéo ( $ROI$ ) et  $S$  la surface de la  $ROI$ , on a la relation  $B = \mathcal{V} \times T_{exp} \frac{N_{pix}}{S}$ . De la même manière que le diaphragme est un compromis entre la profondeur de champ et la luminosité, le choix du temps d'exposition est un compromis entre le flou de bougé et la luminosité. Nous préférons noter qu'en pratique une utilisation adaptée de l'éclairage peut s'avérer judicieuse. En effet, pour figer un mouvement, il est également possible d'utiliser des éclairages stroboscopiques synchronisés avec l'acquisition de la caméra.

**Table 1.1** – Paramètres réglables et intrinsèques au capteur utilisé

Paramètre	Définition	réglable
Vitesse d'acquisition	Nombre d'images prises pendant une seconde.	oui
Temps d'exposition	Temps pendant lequel le capteur reçoit des photons pour une image.	oui
Sensibilité	Rapport entre le signal de sortie (en électrons) et le nombre de photons nécessaires pour avoir un signal : plus les pixels sont gros, plus ils sont sensibles.	non
Taille du capteur	Dimensions du capteur en hauteur et largeur ou de la diagonale.	non
Résolution	Nombre de pixels par ligne et par colonne.	non
Zone d'intérêt ( $ROI$ )	Zone réduite de la matrice sur laquelle les données sont collectées et transmises.	oui
Débit	Quantité de données transmise par secondes.	non

La collecte de charges est perturbée par certains phénomènes qui occasionnent des bruits de mesure. Au-delà des inconvénients des technologies CMOS et CCD déjà évoqués, d'autres bruits existent

3. Le gain possède l'inconvénient d'amplifier le bruit du capteur.

dans le capteur. L'apparition de charges parasites dues à des phénomènes d'agitation thermique crée un courant appelé **courant d'obscurité**. Ce courant est faible parce que le temps d'exposition est très court. Le fonctionnement du détecteur introduit un bruit dit **bruit de photon**. Ce bruit poissonnien se modélise à l'aide de la formule de Schottky : il reste faible avec une dépendance au temps d'exposition.

Un convertisseur analogique numérique permet ensuite de coder les informations en niveaux de gris. Ce codage se fait classiquement sur 8 à 12 bits. Les bonnes pratiques lors de la préparation de la scène consistent comme pour toute acquisition à utiliser un maximum de la dynamique du capteur sans saturation. Dans le contexte vidéo, il s'agit de ne pas avoir de pixel saturé en noir ni en blanc, mais d'avoir tous les pixels étendus sur des valeurs de gris au maximum de la plage de mesure. Ce convertisseur introduit également un bruit de quantification qui reste cependant faible étant donné que les capteurs sont souvent suréchantillonnés à 10 ou 12 bits.

Les limitations de performances en fréquence d'acquisition d'une caméra numérique ne se mesurent plus en vitesse de défilement d'un film, mais en débit de données sortant du capteur : en numérique ce débit représente le flux de données à transporter à la seconde. Comme indiqué sur le tableau 1.2, il est égal à la multiplication de la vitesse, la résolution et le nombre de niveau de gris. On s'aperçoit que le numérique a introduit une relation qui n'existait pas pour les caméras films entre la vitesse de capture, la précision du niveau de gris mesuré et la résolution spatiale.

Ce débit est actuellement le point limitant pour une augmentation de la vitesse. Les technologies CCD et CMOS ont été ainsi munies d'interrupteurs permettant de désactiver des lignes entières de pixels, ce qui permet de gagner du temps lors du comptage et surtout de diminuer la masse de données à transférer. Ce mécanisme de sélection du nombre de ligne est commun à la plupart des systèmes de vision rapide. Il permet de diminuer la résolution pour augmenter la vitesse, d'où la précision de la résolution dans les spécifications de vitesse d'une caméra : par exemple 1000 fps pleine définition ou 10 000 fps à 640 × 480.

**Table 1.2** – Débits de données suivants différents paramètres

Vitesse	Résolution	Codage	Flux	Stockage (pour 1 s)
30 fps	640 × 480 (SD)	8 bits	$\sim 0,75 \times 10^{-3}$ Gbit/s	0,1 Mo
30 fps	1920 × 1080 (HD)	8 bits	$\sim 0,5$ Gbit/s	62 Mo
30 fps	4096 × 3072 (4K)	8 bits	$\sim 3$ Gbit/s	375 Mo
30 fps	1920 × 1080 (HD)	24 bits	$\sim 1,5$ Gbit/s	188 Mo
1000 fps	1920 × 1080 (HD)	8 bits	$\sim 16,5$ Gbit/s	2 Go
1000 fps	4096 × 3072 (4K)	8 bits	$\sim 100$ Gbit/s	12,5 Go

#### II.A.4. Stockage

Cet important débit doit être traité, afin d'être transformé selon la sortie utilisée puis transmis. Chaque modèle de caméra est limité par un flux maximum. On distingue plusieurs catégories de haute vitesse :

**la vidéo grand public** dont les débits restent inférieurs à 1 Gpixel/s, les vitesses allant jusqu'à 100 fps ;

**les machines de vision ou vidéo haute vitesse** dont les débits vont jusqu'à 5 Gpixel/s, les vitesses allant jusqu'à 1000 fps. Ces machines permettent de faire de la décision en ligne et

## II. Cadre technologique de l'étude

sans flou ;

**la très haute vitesse** permet de travailler avec des débits allant jusqu'à 30 Gpixel/s, les vitesses allant jusqu'à 50 000 fps. Elle est utilisée pour de la vidéo rapide. Les traitements sont tous effectués a posteriori ;

**l'ultra haute vitesse** permet de travailler avec des débits supérieurs à 30 Gpixel/s, les vitesses allant jusqu'à 200 Mfps. Les événements filmés sont extrêmement courts et les systèmes de déclenchements sont adaptés.

Pour toutes les caméras allant jusqu'à la haute vitesse, il est possible de traiter les données pour les transporter sur une unité de traitement déportée. Afin de transporter ces données au plus vite, quatre technologies sont disponibles et décrites dans le tableau 1.3.

**Table 1.3** – Différentes connectiques d'interface

	Giga Ethernet	USB3	Camera Link	CoaXPress
Bande passante maximum (Mbytes/s)	115	400	655	3600
Longueur de câble maximum (m)	100	5	10	105
Carte d'acquisition spécifique ( <i>Frame grabber</i> )	Non	Non	Oui	Oui

Cette importante quantité de données reste malgré tout un problème puisqu'il s'agit ensuite de la stocker lors de l'enregistrement. Si les disques durs SSD peuvent convenir pour les plus bas débits, l'utilisation d'un enregistrement sur mémoire vive devient obligatoire au-delà de certains débits. Lorsqu'il s'agit d'un enregistrement à conserver, il faut ensuite stocker les séquences après l'enregistrement. Pour minimiser les données, il peut être proposé de sélectionner à nouveau une zone d'intérêt et de sélectionner les images à sauvegarder dans la séquence capturée. Un coup d'œil sur les valeurs de stockage données dans le tableau 1.2 suffit pour s'apercevoir de l'énorme capacité de stockage nécessaire pour faire de la haute vitesse ou de la très haute vitesse. De plus, les utilisations étant à des fins métrologiques ou de diagnostic, ces séquences sont rarement compressées, laissant une problématique de stockage de données ouverte.

Pour les caméras de très haute vitesse les débits sont trop important pour les transmettre pendant l'acquisition : il faut travailler sur mémoire embarquée. Les schémas utilisés pour la mémoire sont de plus en plus élaborés, complexes et spécifiques au constructeur avec l'augmentation de la vitesse de capture. Nous détaillerons uniquement deux schémas.

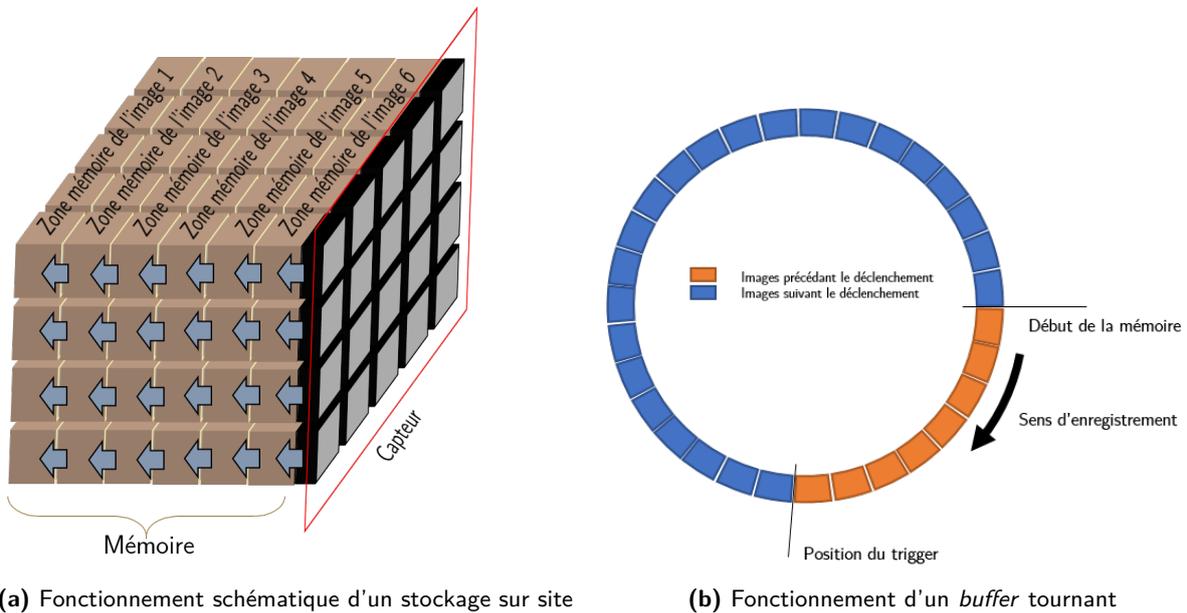
### **Le *buffer* tournant**

Il s'agit d'une organisation de la mémoire qui permet de stocker un certain nombre de pixels ou d'images en fonction de la résolution choisie et de la vitesse. Lorsqu'on arrive à la fin de la mémoire disponible, les données suivantes sont enregistrées en repartant du début en écrasant les anciennes données. Ce mode de fonctionnement permet d'enregistrer à la fois des images passées et futures par rapport au déclenchement de l'enregistrement. La caméra enregistre en direct la scène, lors du déclenchement l'enregistrement va jusqu'à la fin du *buffer* et s'arrête (figure 1.5b). Les caméras *Phantom* utilisent ce type de technologies qui est par ailleurs très répandu.

### **Le stockage sur site ou *In Situ Image Storage (ISIS)***

Lorsque la vitesse d'acquisition est extrêmement importante, les stratégies développées consistent à minimiser le parcours des charges. L'idée de l'*ISIS* est représentée schématiquement sur la figure 1.5a. Il s'agit de stocker les charges dans des registres mémoires situés derrière le capteur. Ainsi à chaque pas de temps les images progressent de proche en proche dans la profondeur de la mémoire. Cette technologie notamment utilisée par Shimadzu permet de fil-

mer un petit nombre d'images (une centaine) pour des vitesses de l'ordre du million d'images secondes (jusqu'à 200 Mfps pour le constructeur Invisible Vision).



**Figure 1.5** – Deux exemples d'architecture de stockage de vidéos rapides pendant l'acquisition.

### II.A.5. Caméras bio-inspirées : une vision d'avenir

Comme nous l'avons souligné précédemment, l'innovation apportée par les capteurs CMOS par rapport aux CCD consistait à venir implanter une électronique pour chaque pixel afin de permettre un réglage du gain individuel.

Par ailleurs, les limitations de flux évoquées et les solutions développées dans ce manuscrit seront basées sur les technologies présentées jusqu'ici : elles mettent en jeu un flux de pixels constant à échantillonnage constant. Quelles que soient les traitements appliqués après l'acquisition, la donnée image à analyser aura les propriétés données lors de l'enregistrement de la séquence c'est-à-dire une résolution imposée et constante, et une vitesse de capture imposée et constante. L'œil ne fonctionne pas de cette manière. Différents récepteurs sont sensibles à la luminosité et sont plus concentrés au centre du champ de vision tandis que d'autres sont sensibles au mouvement : ils n'envoient de l'information que s'ils détectent une variation.

Les caméras décrites précédemment ne permettent en aucun cas de traiter ce type de problématiques au niveau du pixel. L'idée de développer au début des années 1990 des caméras événementielles, asynchrones ou encore bio-inspirées (également appelées *Silicon Retinas*) est de continuer l'intégration de l'électronique au niveau du pixel afin de transformer celui-ci d'un détecteur d'intensité en un détecteur de variation d'intensité. On pourra se référer à la thèse de Ni (2013) pour une vision d'ensemble de la technologie. Ces capteurs permettent ainsi de détecter et transmettre uniquement les pixels utiles, directement au niveau du capteur. Ils permettent ainsi des réductions du flux de données famines : la société Chronocam annonce des vitesses équivalentes de capture de 100 000 fps, une dynamique deux fois plus élevée qu'une caméra standard (120 dB) de la compression intrinsèque au niveau du capteur et une consommation extrêmement faible idéale pour les applications embarquées.

## II. Cadre technologique de l'étude

Ces capteurs représentent pour l'instant une part de marché marginale à cause de défauts de jeunesse de la technologie. Les définitions proposées ne sont pas encore aux normes du marché (au maximum du QVGA c'est-à-dire  $320 \times 240$ ) et surtout l'ensemble des structures de traitements vidéos et de stockage ne sont pas conçues pour les données issues de ces caméras : matrices creuses, données asynchrones... La recherche est cependant active sur ce sujet (Mueggler, 2017) et des applications convaincantes émergent (Dorn *et al.*, 2017).

### II.A.6. Conclusion

Cette courte introduction aux caméras hautes cadences nous a permis de faire ressortir les différentes tendances et le sens d'évolution. On a pu remarquer les problématiques concernant le stockage et le débit d'informations sous-jacentes à l'augmentation de la fréquence d'acquisition. Nous proposons au lecteur souhaitant approfondir de se référer au récent panorama du marché effectué par Reu (2014).

Nous travaillerons dans la suite de ce manuscrit dans un contexte haute vitesse ou très haute vitesse, sur des débits ne permettant pas de traitements en ligne élaborés. Les caméras que nous utiliserons seront souvent à haute définition et fonctionneront en contrepartie à des vitesses plus faibles. Elles fonctionneront à l'aide de la technologie de *buffer* tournant.

### II.B. Contexte particulier d'une vidéo rapide



**Figure 1.6** – Vidéo rapide d'une formule 1 des mers : l'objet rapide de la scène est le bateau, la mer qui n'a pas la même dynamique peut être considérée comme immobile.

La vidéo haute vitesse est généralement utilisée pour observer un phénomène invisible pour des caméras classiques. La capture de ce phénomène se prépare ainsi avec soin : c'est-à-dire avec un temps de préparation non-négligeable, un réglage de la caméra optimisé et une scène dont le scénario est bien défini à l'avance. On est loin d'une vidéo quelconque. Quand cela est possible, les conditions d'éclairage sont contrôlées ou du moins renforcées afin de bien conditionner les informations d'entrée dans le capteur.

La figure 1.6 présente une scène filmée en dynamique rapide. On peut voir sur cette scène plusieurs autres spécificités d'une vidéo rapide : un ou plusieurs objets d'une dynamique élevée sur une zone souvent minoritaire de la vidéo. Les autres objets ont une dynamique lente. Le caractère rapide de la vidéo va augmenter cet effet de contraste de dynamique : les objets de dynamique lente peuvent

être considérés comme statiques dans une vidéo rapide. La figure 1.6 donne l'exemple de la mer qui semble figée sur les images de la séquence.

Nous avons souhaité définir pour le reste de ce manuscrit un cadre restreint pour les vidéos issues d'enregistrements avec des caméras rapides, que nous appellerons **vidéos rapides** dans la suite.

Nous prendrons les hypothèses suivantes :

- bonnes conditions d'éclairage ;
- bonne qualité de la séquence vidéo ;
- les objets de dynamique lente sont supposés statiques et appartenir au fond ;
- les objets ayant des dynamiques rapides représentent une petite zone d'intérêt par rapport à la taille de l'image ;
- le scénario de la vidéo rapide est connu mais pas quantifié.

Ces hypothèses posées nous permettent de travailler dans un cadre particulier. Au niveau du traitement vidéo, les hypothèses nous permettront d'utiliser certaines méthodes. L'hypothèse d'un scénario connu nous conforte dans la pertinence du projet de développer des modèles spécifiques au contenu mécanique de la scène.

### III Conclusion : apports proposés

Les travaux présentés dans ce manuscrit proposent des réponses aux problématiques du projet FAMOSYS. Nous avons choisi de structurer ces travaux en trois grandes parties dans lesquelles nous présenterons différents apports. Cette approche nous permettra d'introduire progressivement l'état de l'art et ainsi d'immerger progressivement le lecteur dans les trois domaines principaux de ces travaux : le traitement vidéo, l'assimilation de données et l'analyse vibratoire.

**Dans le chapitre 2** nous reprendrons les hypothèses définies précédemment en enlevant l'hypothèse de scénario connu. Pour répondre à nos problématiques notamment d'interpolation et de compression nous présenterons un état de l'art concernant le traitement vidéo et en particulier la compression de données. Partant du constat qu'il est difficile de faire de l'interpolation d'image pour de grands déplacements, nous proposerons un cadre théorique à la vidéo reprenant les concepts eulériens et lagrangiens permettant une double approche par le traitement vidéo et de manière mécanique. Après avoir fait un état de l'art concernant les méthodes permettant de passer d'une représentation à l'autre, nous proposerons un schéma basé sur une composition de mouvements multi-échelle permettant de faire de l'interpolation à grande échelle et de la compression de manière sémantique aux objets réels de la scène. Deux exemples d'application seront présentés afin de permettre un recul critique sur le cadre d'application de la méthode.

**Dans le chapitre 3** nous nous placerons cette fois dans les hypothèses du projet FAMOSYS. La connaissance de la scène nous permet d'envisager le développement d'un modèle issu des équations de la mécanique. Afin de relier les données images issues de la caméra avec le modèle de la scène, nous introduirons un double état de l'art sur l'assimilation de données et sur le contexte bayésien offrant au lecteur plusieurs angles de vue sur les filtres de Kalman. Ce chapitre présente également un apport bibliographique en proposant une revue transverse aux domaines scientifiques des différentes versions des filtres de Kalman. Nous proposerons un cadre général d'utilisation d'un schéma d'assimilation avec des modèles mécaniques en utilisant les résultats du chapitre 2 illustré par une application sur un modèle de règle glissant

### III. Conclusion : apports proposés

sur une table à deux points de contact. À travers l'assimilation de ce modèle nous proposons d'utiliser des modèles sous forme de systèmes différentiels avec des observations synchrones ou aléatoires. Les résultats sur des cas tests réels remplissent les objectifs énoncés précédemment, à savoir identification, prédiction et extrapolation tout en respectant un schéma temps réel.

**Dans le chapitre 4** nous appliquerons les concepts proposés dans le chapitre 3 à l'analyse vibratoire. L'état de l'art sera segmenté en deux parties. Une partie matérielle met en avant les avantages de l'utilisation de caméras rapides puisque ces dernières offrent l'avantage d'une acquisition sans contact multipoints sur des phénomènes transitoires ou aléatoires. Une seconde partie introduisant l'état de l'art des techniques de traitement de données utilisées pour faire de l'identification modale montre qu'il serait judicieux d'utiliser notre cadre d'assimilation de données pour proposer une nouvelle méthode permettant de faire de la séparation de contributions modales, de l'identification de paramètres modaux non linéaires avec des schémas temps réel. En utilisant des modèles dynamiques d'évolution du signal comme proposé par *Dion et al. (2013)*, nous proposons deux schémas d'assimilation aux visées différentes basés sur des observations par vidéo rapides. Le premier utilise un modèle de la structure et permet de faire de l'identification des paramètres modaux de la structure. L'utilisation d'un modèle de la structure autorise des extrapolations intéressantes à la fois spatiales et temporelles mais également un lissage des données acquises. Le second propose d'estimer en plus des paramètres modaux les différentes déformées modales de la structure. Ces développements sont appliqués sur un cas test réel de poutre encastree libre et montrent des applications pour de l'extrapolation spatiale (ajout de points, observation partielle de la structure), de l'identification et également une application en fusion de données issues de capteurs de nature différente. Un algorithme basé sur ces schémas est proposé pour permettre une application sur des structures 3D multi-caméras sans utiliser d'étapes de reconstruction stéréo, apportant ainsi un gain de temps significatif. Enfin une nouvelle méthode de magnification de mouvements est proposée en combinant l'utilisation de ces schémas et les méthodes de synthèses vidéos développées dans le chapitre 2.

À travers ces différents apports nous proposons une réflexion sur l'utilisation de la donnée vidéo avec des modèles mécaniques. Les travaux sont centrés sur les avantages d'une utilisation de caméras rapides mais également sur ceux d'algorithmes d'assimilation de données qui permettent de reconstituer une réalité entre observations d'un système réel et son modèle. La synthèse d'image utilisée offre une manière de matérialiser cette réalité en produisant des matrices de pixels auxquelles notre cerveau donne un sens en appliquant également des modèles...

## Bibliographie du chapitre 1

Marc Arques : Images sensors : Cours sur les capteurs d'image, décembre 2011.

Jean-Luc Dion, Cyrille Stephan, Gaël Chevallier et Hugo Festjens : Tracking and removing modulated sinusoidal components : A solution based on the kurtosis and the Extended Kalman Filter. *Mechanical Systems and Signal Processing*, volume 38, pages 428–439, juillet 2013. ISSN 08883270.

Charles Dorn, Sudeep Dasari, Yongchao Yang, Garrett Kenyon, Paul Welch et David Mascareñas : Efficient Full-Field Operational Modal Analysis Using Neuromorphic Event-Based Imaging. In Julie M. Harvie et Javad Baqersad, éditeurs : *Shock & Vibration, Aircraft/Aerospace, Energy Harvesting, Acoustics & Optics, Volume 9 : Proceedings of the 35th IMAC, A Conference and Exposition on Structural Dynamics 2017*, pages 97–103. Springer International Publishing, Cham, 2017. ISBN 978-3-319-54735-0.

Elias Mueggler : *Event-Based Vision for High-Speed Robotics*. Thèse de doctorat, University of Zurich, Zurich, juillet 2017.

Zhenjiang Ni : *Asynchronous Event Based Vision : Algorithms and Applications to Microrobotics*. Thèse de doctorat, Université Pierre et Marie Curie-Paris VI, 2013.

Phillip L. Reu : The evolution of high and ultra-high speed imaging from qualitative to quantitative, 2014.

Pierre Slangen, Nicolas Long et Pascal Picart : Imagerie numérique ultrarapide. *Techniques de l'ingénieur Métrologie optique et photonique*, volume base documentaire : TIB143DUO., 2016.



# 2

## Analyse de vidéos par approche eulérienne basée sur un modèle cinématique élémentaire

*Quand l'eau courbe un bâton, ma raison le redresse.*

– Jean de la Fontaine

### Sommaire

---

I	Introduction du contexte vidéo et présentation de deux descriptions d'une même scène vidéo . . . . .	30
II	Estimer le mouvement des pixels et interpoler leur trajectoire . . . . .	36
III	Caractériser un objet dans une séquence d'images : un problème mal posé en 2D . . . . .	47
IV	Une nouvelle méthode pour l'interpolation et la compression à l'aide de la composition cinématique du mouvement . . . . .	62
V	Formalisation de la chaîne complète . . . . .	67
VI	Conclusion . . . . .	77
	Bibliographie du chapitre 2 . . . . .	79

---

### Résumé

Dans un cadre élargi de vidéos rapides, à fond fixe, dont le scénario est inconnu, il n'est pas possible de développer un modèle spécifique décrivant suffisamment bien le phénomène observé pour y ajouter de l'information. Ce chapitre propose d'utiliser des modèles cinématiques élémentaires pour décrire des objets quelconques. Il développe une double approche liant la mécanique du solide et l'estimation de mouvements en vidéo à travers la définition d'une double description de la scène à la fois eulérienne et lagrangienne ; il introduit une nouvelle méthode basée sur une composition de mouvements permettant de faire de l'estimation de mouvements sur de grands déplacements, impossible dans un cadre eulérien. Cette méthode offre alors la possibilité de faire de la compression sémantiquement liée à l'objet observé ainsi que de l'interpolation et de la synthèse de vidéos au rendu amélioré. Des applications sont proposées sur des scénarios testant les hypothèses de la méthode.



Les problématiques exposées dans le premier chapitre nous amènent dans un premier temps à répondre à la question suivante : comment identifier les pixels utiles ? Nous allons proposer une première réponse dans un cadre d'hypothèses élargi à une vidéo quelconque en cherchant comment modéliser les trajectoires de pixels à l'aide d'un modèle générique. Comme résumé sur la figure 2.1, notre démarche part de la proposition d'une double description eulérienne et lagrangienne de scènes vidéos composées d'objets évoluant sur des fonds stationnaires. En décrivant comment observer le mouvement à l'échelle du pixel avec une description eulérienne de la scène nous concluons qu'il est difficile d'estimer le mouvement d'un objet lorsque celui-ci subit un « grand déplacement » dans un temps de calcul raisonnable pour nos applications. Nous mettrons ensuite en évidence la difficulté d'extraire un objet et les différentes méthodes permettant de les extraire. Ces méthodes nous permettront de passer d'une description eulérienne à lagrangienne. Nous proposerons une nouvelle approche modulaire utilisant une description lagrangienne appliquée aux objets de la scène dans le repère global de l'image alliée à une description eulérienne dans le repère local de l'objet pour permettre de décrire les mouvements de pixels pour des « grands déplacements ». Nous décrirons nos choix de développement puis nous appliquerons notre algorithme à deux cas tests.

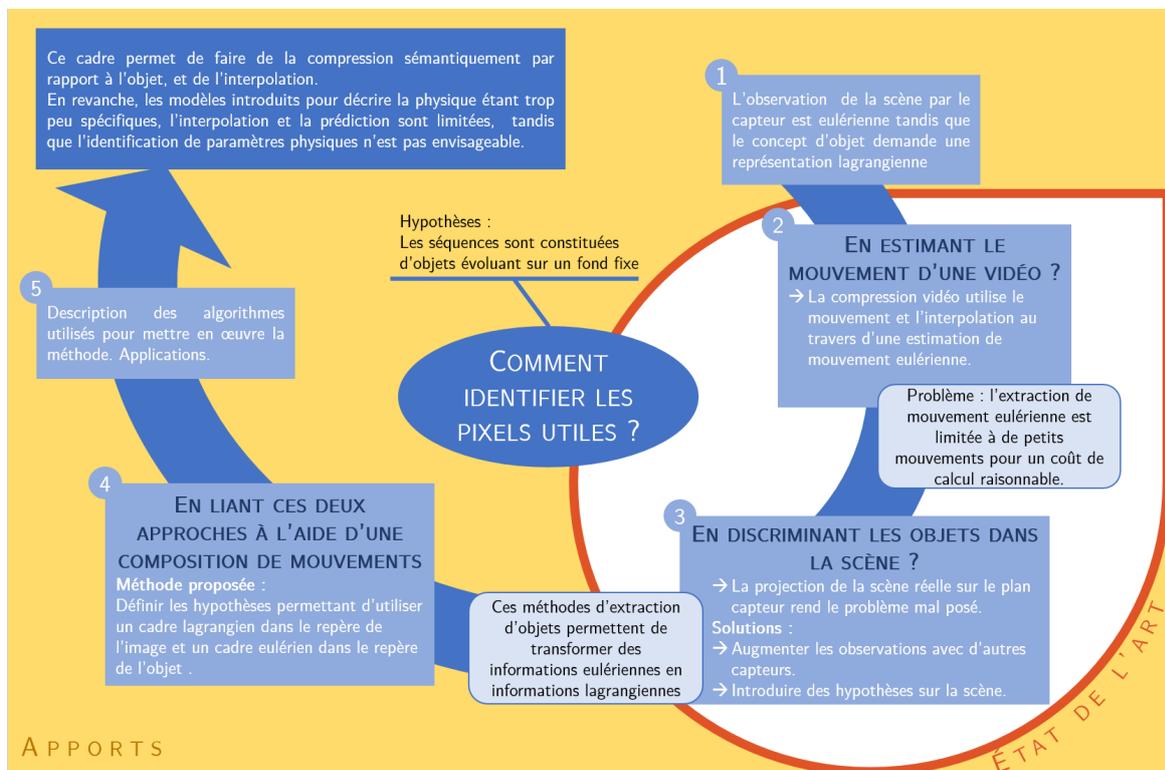


Figure 2.1 – Résumé graphique présentant la démarche et les résultats du chapitre 2

Cette première approche nous permet ainsi :

- de définir un cadre liant la mécanique et le traitement vidéo ;
- de proposer une nouvelle approche d'interpolation pour des grands mouvements ;
- d'effectuer de la compression sans perte sémantiquement à la scène ;
- de conclure sur l'impossibilité de faire de l'extrapolation, de l'identification ou de la prédiction avec notre modèle trop générique à la scène.

## I Introduction du contexte vidéo et présentation de deux descriptions d'une même scène vidéo

Les problématiques exposées dans ce chapitre sont souvent liées à la nature numérique des signaux traités ; elles ne sont apparues que par l'explosion de la numérisation des données. Aussi cette première partie propose de revenir au début de la numérisation afin de bien comprendre les problématiques qui vont émailler ce chapitre. Dans un deuxième temps, nous introduirons dans le cadre particulier de nos vidéos comportant des objets deux descriptions différentes : l'une eulérienne et l'autre lagrangienne. On notera que le terme de *vidéo numérique* ne sera utilisé que dans cette section pour différencier ces séquences des *vidéos analogiques* issues de caméras films et autres appareils de capture d'images analogiques : toutes les vidéos et caméras traitées dans ce manuscrit sont numériques. Ainsi tous les termes *vidéos* rencontrés par la suite mentionneront implicitement leur caractère numérique.

### I.A. Introduction à la vidéo numérique

#### I.A.1. Le signal vidéo : concept et discrétisation

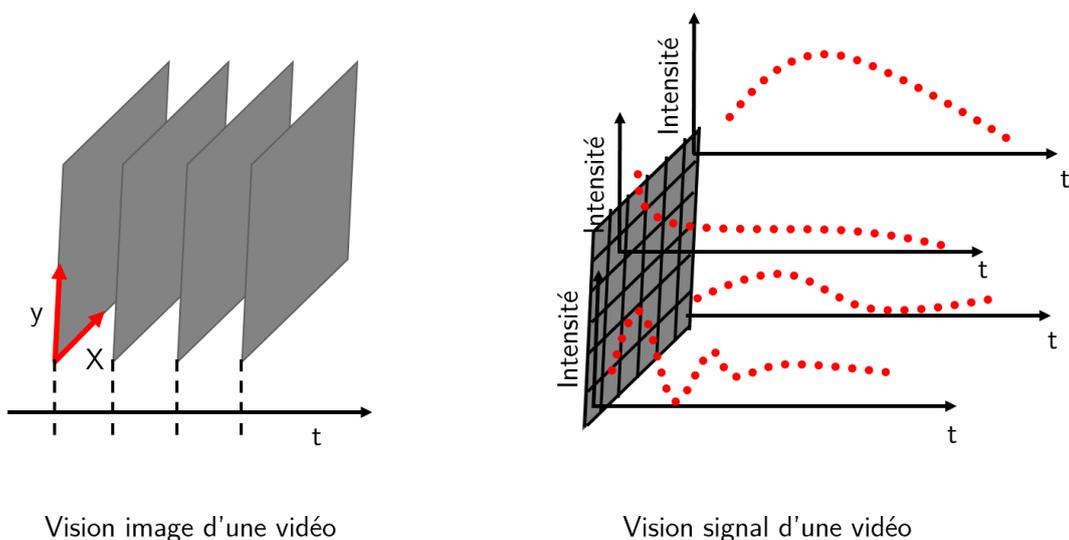


Figure 2.2 – Deux manières différentes de voir le signal vidéo

La nature spatiotemporelle du signal vidéo peut porter à confusion, souvent à cause de la culture signal monodimensionnelle issue du traitement de signal classique. Aussi, afin de bien expliquer le concept de la vidéo numérique, on propose d'opérer par analogie avec deux autres concepts plus communs que sont le signal 1D et l'image. Le signal vidéo peut alors être vu de deux manières différentes illustrées sur la figure 2.2 :

**Un signal 1D** classique qui évolue en fonction du temps. Le signal numérique de chaque pixel d'une vidéo est défini de la même manière que pour une autre grandeur numérique, c'est à dire en définissant une fréquence d'échantillonnage ou un pas de temps entre chaque mesure. La généralisation de ce concept pour une matrice  $M \times N$  de signaux 1D nous permet de

lui donner une consistance spatiale et de définir de cette manière la vidéo numérique. La figure 2.2 schématise cette vision sur sa partie droite.

**Une image 2D** standard est une matrice de taille  $M \times N$ . Cette matrice constitue une discrétisation spatiale d'une image réelle projetée sur le capteur d'image. On peut ainsi définir une fréquence d'échantillonnage spatiale selon chaque direction de la matrice. Cette fois, c'est en juxtaposant différentes images prises régulièrement suivant un pas de temps donné que l'on définit la vidéo numérique. La figure 2.2 illustre ces concepts sur la partie gauche.

Ces deux approches soulignent les confusions possibles entre les fréquences d'échantillonnage spatiales et temporelles, faisant de ces dernières deux obstacles différents mais similaires. Ainsi les phénomènes de repliement de spectre interviennent en vidéo sur deux aspects. Sur l'exemple connu de la roue à rayon dont le mouvement s'inverse au-delà d'une certaine vitesse, il s'agit d'un *aliasing* temporel. Sur le port de marinières à l'écran dont les lignes semblent se courber : il s'agit cette fois d'*aliasing* spatial.

D'où vient ce phénomène d'*aliasing* ? Le signal numérique est créé à partir d'un signal analogique en l'échantillonnant. On obtient ainsi un ensemble discret de points. Le signal continu peut alors être reconstruit à partir de cet ensemble de points. Le phénomène d'*aliasing* se mesure lorsqu'on compare le signal ré-obtenu avec le signal original. Il est une cause directe du théorème de Shannon.

L'apparition des problématiques de discrétisation temporelle date de l'invention des premières caméras argentiques en 1888 par John Carbutt. Les premières caméras « analogiques » opèrent en réalité déjà cette discrétisation effectuée par l'obturateur. En utilisant le principe du stroboscope, les premiers Kinéoscopes permettaient de recréer l'impression de mouvement par le phénomène physiologique de persistance rétinienne. Aussi, lors de l'arrivée des premières caméras numériques, la fréquence d'échantillonnage admissible par l'œil de 25 images par secondes (ou *frame per second - fps*) était déjà connue et éprouvée. L'apparition des problématiques de discrétisation spatiale remonte également avant l'apparition de la vidéo numérique : lors de la projection de séquences vidéos sur des téléviseurs. En reprenant les standards de la télévision analogique terrestre, on définissait en Europe 625 lignes pour une transmission télévisuelle hertzienne. Cette valeur est également issue de paramètres physiologiques : du pouvoir de résolution de l'œil et de la distance d'observation d'un objet en fonction de ses dimensions.

Lorsque caméras et vidéos numériques ont été mises au point, les problématiques de discrétisation spatiales et temporelles ainsi que les valeurs acceptables à atteindre étaient déjà en place. La transition numérique s'est déroulée petit à petit pour la vidéo. Le numérique facilitant les retouches colorimétriques mais également les incrustations et de nombreux effets spéciaux, les composants numériques ont d'abord été inclus dans la chaîne analogique à grand renfort de convertisseurs analogiques numériques (CAN) avec certains inconvénients comme la double conversion analogique numérique puis numérique analogique ou la maniabilité des données. La numérisation totale du flux permettait ainsi d'éviter ces conversions, et d'offrir un support de stockage plus pérenne que celui des bandes magnétiques. Cependant d'autres problématiques sont apparues, notamment au niveau du stockage de ces vidéos numériques. La réponse à ces problématiques a nécessité le développement de techniques de compression et de restitution du signal.

### I.A.2. Codage et stockage

La vidéo numérique couleur peut se définir comme un signal multidimensionnel :

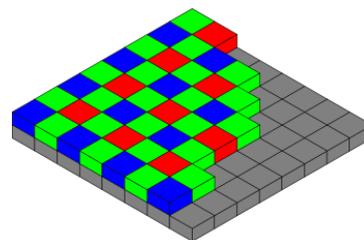
- une composante temporelle ;

## I. Introduction du contexte vidéo et présentation de deux descriptions d'une même scène vidéo

- deux composantes spatiales ;
- 3 canaux de couleur.

Introduisons le pixel comme l'entité spatiotemporelle à la base de tous les algorithmes évoqués dans la suite de ce manuscrit. Chaque pixel correspond à une position fixe sur la matrice du capteur et contient la valeur numérisée de l'intensité lumineuse correspondant au nombre de photons collectés pendant le temps d'ouverture. Cette valeur varie beaucoup en fonction de la qualité des caméras : généralement codée sur 8 bits. Il est courant de rencontrer des caméras avec des codages sur 12 voire 16 bits dans le monde industriel pour des applications spécialisées.

Compte-tenu de l'évolution pas toujours linéaire des standards et des technologies, plusieurs formats de stockage ont cohabité et nous offrent aujourd'hui différents formalismes pour représenter les couleurs. La couleur provient de l'utilisation de filtres colorimétriques sur le flux lumineux incident<sup>1</sup>. En exploitant la nature additive de la lumière pour la décomposer sur trois composantes, la lumière est décomposée sur trois couleurs : rouge, bleu et vert. Étant donné que l'on doit extraire trois informations à partir de la même quantité de photons incidente sur le capteur, il est nécessaire de « sacrifier » de la précision sur une grandeur au profit de l'obtention de l'information colorimétrique.



**Figure 2.3** – Schéma de Bayer utilisé sur les capteurs noir et blanc afin de permettre de capturer les couleurs. On réduit de moitié la résolution du capteur

Plusieurs technologies existent et agissent sur différents paramètres. En utilisant une roue à filtre, on divise par trois la fréquence temporelle et on peut recomposer toutes les trois images une image couleur : c'est évidemment une méthode jamais employée en haute vitesse. Une autre méthode consiste à utiliser une lentille dichroïde qui permet de séparer le rayon incident en trois et récupérer ainsi sur trois capteurs différents ; elle divise le flux de photon incident, nécessite plus de luminosité et plus de capteurs. Le capteur *Foveon* propose de superposer sur chaque pixel différents photosites de couleur différentes. Cette fois les inconvénients se mesurent sur l'indépendance entre les couleurs ainsi que sur la précision de la couleur captée par la couche la plus profonde. Bien que la lentille dichroïde et le capteur *Foveon* soient utilisés pour des applications spécifiques, la très grande majorité des caméras utilisent un schéma de Bayer (voir figure 2.3). Les capteurs sont recouverts d'un damier permettant de transformer chaque pixel en pixel d'une bande de couleur spécifique : rouge, bleu, vert. Utiliser un schéma de Bayer a pour inconvénient de perdre en résolution spatiale.

### **Le format RVB (Rouge Vert Bleu )**

Moyennant une interpolation bilinéaire entre les différents pixels de même couleur, il est possible d'obtenir en sortie de ces capteurs trois canaux de couleur rouge, vert et bleu (RVB ou *RGB* en anglais) sur une image. Ce format de stockage est donc le format natif d'acquisition des caméras mais également le format natif d'affichage des écrans numériques, ce qui en fait le format numérique par excellence. La figure 2.4 propose à gauche une représentation de l'espace colorimétrique RVB.

### **Le format YCbCr**

Le format YCbCr ou YUV est hérité de la diffusion des flux vidéos analogiques. Il est composé d'un canal de luminance, la *luma* Y, et de deux canaux colorimétriques dits de chrominance,

1. voir le chapitre 1 pour une vision plus large de la chaîne de mesure d'une caméra

l'un dans le rouge (Cr) l'autre dans le bleu (Cb). L'apparition de ce format a permis la compatibilité avec les téléviseurs noir et blanc en introduisant un canal de signal comportant la somme des trois couleurs.

### Les formats Hue : *HSL* et *HSV*

Les formats issus du formalisme de Hue fonctionnent selon le même principe que YCbCr qui sépare la luminance de la chrominance. La couleur est codée sur un disque dont l'angle définit la valeur de couleur et la distance au centre définit la saturation (figure 2.4). L'espace engendré par le modèle Teinte-Saturation-Valeur (TSV ou *Hue-Saturation-Value - HSV*) ou Teinte-Saturation-Luminosité (TSL ou *Hue-Saturation-Luminance - HSL*) peut être représenté par un cylindre dont la hauteur représente la luminosité. La différence entre « luminosité » et « valeur » reposant sur le calcul de la luminosité en fonction des composantes trichromatiques. La figure 2.4 propose au milieu une représentation de l'espace colorimétrique HSV et à droite l'espace colorimétrique HSL.

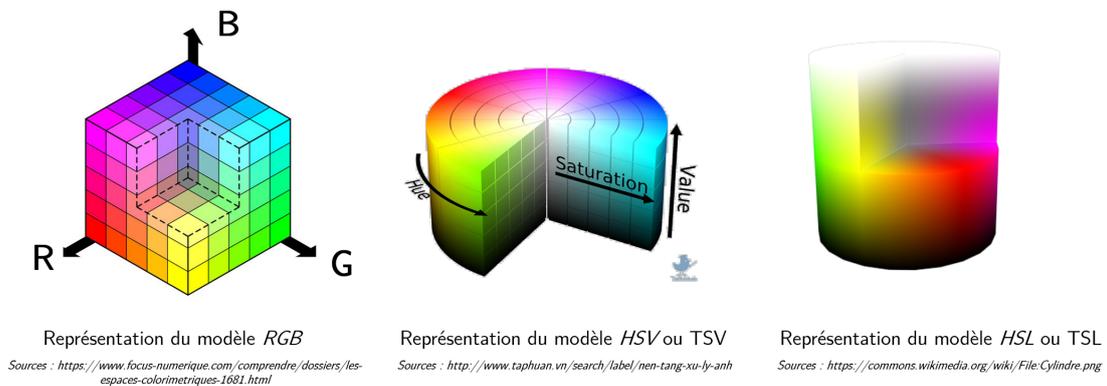


Figure 2.4 – Quelques représentations de modèles colorimétriques

Une multitude d'autres formats existent parmi lesquels on citera le *RG Chromaticity* très utilisé en vision industrielle proposant une version normalisée de l'espace *RGB*. Le sous-paragraphe a. nous donnera l'occasion d'insister sur l'importance du choix de l'espace colorimétrique pour distinguer des objets dans l'analyse d'images et de vidéos. En pratique les espaces de type *HSV* offrent une segmentation colorimétrique plus adéquate pour identifier à partir des informations colorimétriques.

### I.B. Représentation eulérienne et lagrangienne : proposition d'un formalisme pour décrire une scène vidéo

Comme l'a justement remarqué Rubinstein (2014), il existe deux manières différentes de caractériser le rapport entre les pixels et les objets réels de la scène. Ces différentes approches que nous proposons de développer sont utilisées notamment en mécanique des milieux continus : la description eulérienne et la description lagrangienne.

## I. Introduction du contexte vidéo et présentation de deux descriptions d'une même scène vidéo

### I.B.1. La vidéo numérique : une vision eulérienne sous-jacente

Comme on a pu le voir pour la discrétisation, la caméra est un capteur complexe à cause de sa structure matérielle mais surtout à cause de la nature des informations qu'il permet d'acquérir. Comme indiqué dans le chapitre 1, le capteur d'image se situe derrière un système optique qui permet de mettre en forme le flux lumineux. Il sert notamment à maximiser le champ visuel entrant sur la surface du capteur, mais également à régler les propriétés de flou, de quantité de lumière, ou de sélectionner la direction des rayons lumineux entrant dans l'objectif à l'aide du diaphragme.

Le monde réel capturé par la caméra possède quatre dimensions : nos trois dimensions spatiales et la dimension temporelle. La caméra procède à une simplification puisqu'elle effectue une projection pour passer du monde réel aux deux dimensions du capteur. Les images capturées par la caméra correspondent donc à la projection des informations lumineuses du monde réel sur une grille fixe de pixels. L'analyse de vidéos va reposer sur l'analyse spatiotemporelle des valeurs de ces pixels. Ce cadre correspond à celui de la **description eulérienne** définie en mécanique des fluides. En se plaçant dans le repère du capteur, la vidéo acquise revient à considérer l'évolution cinématique d'une scène à travers la grille fixe du capteur. Considérons l'exemple d'un objet évoluant sur un fond statique pour fixer les idées (figure 2.5). Lorsque l'on regarde un objet évoluer en fonction du temps, un pixel va « voir » l'objet pendant son temps de passage dans son champ de vision.

Considérer les objets de la scène à travers les pixels revient à étudier la scène avec un point de vue eulérien. On propose alors de continuer la similitude plus loin en l'étendant à la description lagrangienne.

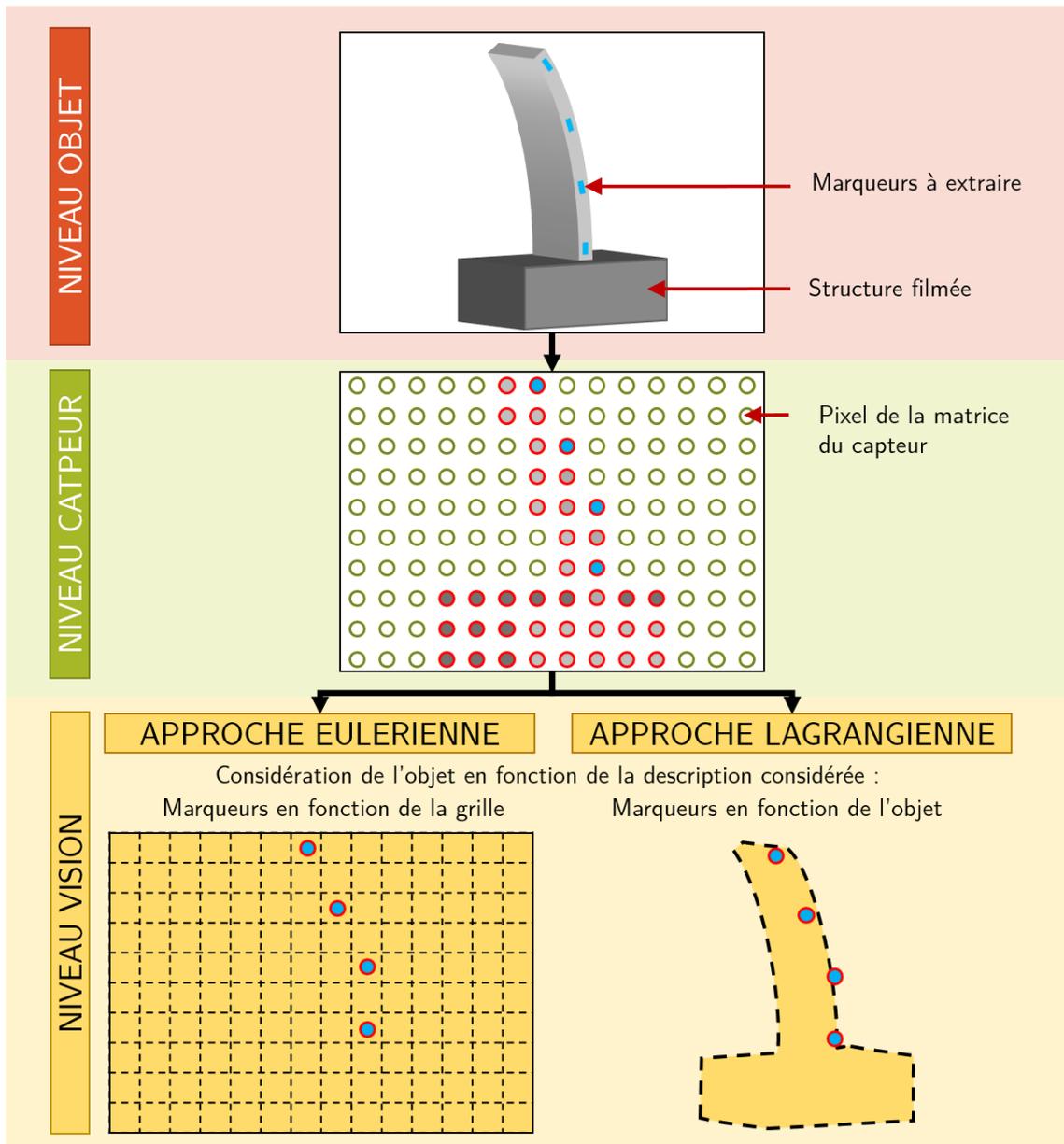
### I.B.2. Introduction d'une description lagrangienne en vision

La représentation lagrangienne des objets évoluant dans une scène revient à considérer l'objet comme entité. Ce point de vue sur la scène qui nous semble familier est pourtant contre-intuitif pour une caméra. En effet, alors que la représentation eulérienne possède un véritable sens physique au niveau du capteur, la représentation lagrangienne n'en a aucun. Elle donne de l'importance à des groupes de pixels associés selon des règles spatiales ou temporelles et fait intervenir des *a priori* dans l'analyse. La multiplicité des situations rencontrées et la projection faite au niveau du capteur complexifient toute généralisation. De plus, la représentation lagrangienne implique non plus de considérer une grille de pixels fixe observant un objet mobile mais cette fois des pixels mobiles appartenant à un objet dans le repère du capteur.

En reprenant l'exemple d'un modèle évoluant sur un fond fixe, on peut considérer l'objet à chaque pas de temps par les pixels qui le composent. Dans cet exemple, plusieurs règles de segmentations peuvent être appliquées pour regrouper les pixels : une segmentation selon le mouvement de l'objet par rapport au reste de l'image, une segmentation selon l'intensité du niveau de gris des pixels, et une segmentation spatiale de proximité de l'objet.

### I.B.3. Intérêt du formalisme dans notre cadre d'étude

À travers cet exemple simple, on s'aperçoit de la complexité du passage entre les deux descriptions et surtout l'absence de bijectivité. Les algorithmes de vision s'intègrent bien dans ce formalisme, certains restant dans le cadre eulérien, d'autres permettant le passage d'une description eulérienne à lagrangienne (lors de la détection par exemple) et enfin des algorithmes uniquement lagrangiens



**Figure 2.5** – Représentation Eulérienne et Lagrangienne pour une règle encastrée libre avec des marqueurs extraits par des méthodes de vision

(c'est le cas du traitement de masques binaires). Ces aspects seront approfondis dans la section III. Dans le domaine du traitement d'image et vidéo, on attache peu d'importance à cette formalisation car les scènes présentées par les vidéos sont considérées quelconques. Prenons l'exemple suivant : la compression de données doit permettre de compresser n'importe quelle vidéo, et non se focaliser sur un type de vidéo en particulier comme dans notre cas avec l'hypothèse de fond fixe.

Dans ces travaux, nous proposons de travailler à l'interface entre plusieurs domaines que sont la mécanique et la vision. Dans ce chapitre spécifiquement, on souhaite utiliser le cadre des hypothèses des vidéos rapides définies dans le chapitre 1 et y introduire des concepts issus de la mécanique. Dans ce domaine, l'étude du mouvement se fait de manière lagrangienne par rapport aux objets considérés.

## II. Estimer le mouvement des pixels et interpoler leur trajectoire

Il faudra définir les relations entre les objets au sens vidéo et les solides au sens mécanique, mais également trouver de quelle manière il sera possible de passer d'une description eulérienne à une description lagrangienne de la scène afin de pouvoir introduire un modèle mécanique.

La définition de ce formalisme que résume la figure 2.5 nous permet d'illustrer les deux différentes approches pour caractériser une même scène ; c'est en combinant leurs avantages que nous allons proposer dans ce chapitre une méthode d'interpolation dans le cadre particulier des vidéos à fonds fixes.

## II Estimer le mouvement des pixels et interpoler leur trajectoire

Est-il possible dans notre cas d'étude avec des techniques de traitement vidéo de pouvoir modifier après l'acquisition la fréquence d'une vidéo ? Appliqué à nos problématiques, quelles sont les solutions numériques permettant de ralentir une vidéo ? Ce chapitre propose une première réponse essentiellement basée sur l'interpolation spatiotemporelle de séquences vidéos et cette partie propose de décrire l'état de l'art des méthodes appliquées au niveau du pixel. L'évolution des techniques de compression sera introduite avant de se focaliser sur deux aspects : l'interpolation de trajectoire et les méthodes permettant de décomposer puis reconstruire un mouvement.

### II.A. Cadre de l'interpolation vidéo temps réel : quels modèles pour quelles informations ?

L'interpolation vidéo est intervenue rapidement après l'arrivée de la vidéo numérique, et il est devenu difficile de s'en abstraire de nos jours tellement les technologies actuelles s'appuient sur ce concept. Cependant, dans le domaine de la vidéo, l'interpolation est loin d'être un problème simple à cause de la nature projective des informations sur chaque image : ce problème ne peut pas être résolu sans approximations ou hypothèses particulières d'un cadre d'étude.

On propose dans cette section de définir les modèles spatiaux et temporels classiquement utilisés (*Motion Models*) puis de décrire comment observer ce mouvement dans les vidéos (*Observation Models*).

#### II.A.1. Modéliser le mouvement dans une vidéo

La modélisation du mouvement dans une vidéo est un problème complexe en raison de la diversité des scènes filmées à laquelle s'ajoute la projection 2D faite par la caméra. Pourtant des modèles de mouvement pour la vision existent. Nous proposons de rappeler les principaux modèles utilisés d'après l'état de l'art de Bovik (2009). Cette approche d'estimation de mouvement s'appuie sur des modèles élémentaires : linéaires ou quadratiques.

##### Modèles spatiaux

Deux modèles spatiaux sont principalement utilisés dans la littérature. En considérant que l'objet observé est un objet complexe tridimensionnel et que la projection sur le capteur est une projection orthographique, on suppose que la projection de la vitesse 3D entre deux obturations peut s'exprimer dans l'image comme une translation 2D entre deux images comportant deux composantes. Deux modèles existent, l'un reliant position et vitesse de manière

linéaire, l'autre proposant un modèle affine reliant une position  $x$ , et une vitesse  $v$  :

$$\forall x, \quad v(x) = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + \begin{pmatrix} b_3 & b_4 \\ b_5 & b_6 \end{pmatrix} \cdot x \quad (2.1)$$

La complexité du modèle a un coût d'estimation non négligeable puisqu'on passe de deux variables à estimer à six variables. Cependant ce modèle permet de décrire des mouvements complexes de l'entité considérée en proposant un paramétrage de la translation  $(b_1, b_2)$ , de la rotation  $(b_3, b_4, b_5, b_6)$  et permettant de décrire également le zoom et la déformation longitudinale.

### Modèles temporels

Pour les mêmes raisons que pour les modèles spatiaux, le suivi temporel des points dans une séquence d'image peut s'avérer impossible à cause de phénomènes pourtant connus : occlusions, nature tridimensionnelle des objets observés, variation de luminosité... De la même manière, il est impossible de définir un modèle complexe d'évolution temporelle des entités suivies. Deux modèles sont utilisés dans le cas général. Le premier proposant une linéarisation du déplacement  $x(t)$  entre deux images prises aux temps  $t$  et  $\tau$  avec l'introduction d'une vitesse instantanée  $v(x)$  :

$$x(\tau) = x(t) + v(x) \cdot (\tau - t) \quad (2.2)$$

Le second propose une linéarisation de la vitesse à travers un modèle quadratique introduisant une accélération  $a(x)$  :

$$x(\tau) = x(t) + v(x) \cdot (\tau - t) + \frac{1}{2}a(x)(\tau - t)^2 \quad (2.3)$$

Ces modèles ont été volontairement définis sur des entités quelconques car ils peuvent être appliqués à différents niveaux de l'image :

- sur un pixel ;
- sur une grille d'analyse ou un groupe de pixels définis selon leur position dans l'image ;
- sur une zone segmentée irrégulière ou un groupe de pixels définis selon des caractéristiques communes ;
- sur une image complète.

Après avoir défini les principaux modèles d'évolution utilisés dans les vidéos, on propose de présenter les principales méthodes pour observer un mouvement dans des vidéos.

#### II.A.2. Observer le mouvement dans une vidéo

Dans un cadre de vidéo général, l'observation du mouvement peut se faire sur différentes entités de l'image décrites précédemment. Ces approches ne traduisent pas une réalité physique de la scène, mais plutôt une réalité numérique au niveau de l'image. Les algorithmes développés pour extraire des informations de séquences ou de flux vidéos ne prennent pas en compte le contenu de la scène pour fonctionner. Nous proposons ainsi de décrire dans cette partie les principales méthodes utilisées pour extraire des champs de mouvements au niveau de pixels ou de groupes de pixels.

## II. Estimer le mouvement des pixels et interpoler leur trajectoire

### a. Notions de flot optique

Le flot optique a été introduit dès les années 40 par le psychologue James Jerome Gibson dans une étude sur la vision humaine. L'introduction en robotique du flot optique s'est faite bien plus tard par Horn et Schunck (1981), suivis la même année par Lucas et Kanade (1981) qui proposent pour la première fois de déterminer le flot optique sur des images numériques.

Le but de ces méthodes est d'estimer le mouvement entre deux images d'une séquence  $I(x, y, t)$  et prises à un intervalle  $\Delta t$  pour chaque pixel de coordonnées  $(x, y)$ , sous une hypothèse d'invariance locale :

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t) \quad (2.4)$$

En effectuant un développement de Taylor à l'ordre 1 du second membre, on obtient :

$$\frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t = 0 \quad (2.5)$$

$$\iff \frac{\partial I}{\partial x} v_x + \frac{\partial I}{\partial y} v_y + \frac{\partial I}{\partial t} = 0 \quad (2.6)$$

$$(2.7)$$

$$\iff v(x, y) \cdot \nabla \cdot I + \frac{\partial I}{\partial t} = 0 \quad (2.8)$$

Cette dernière équation donne un modèle d'évolution pixélique valable au voisinage de chaque pixel. Cependant cette équation de conservation locale seule ne peut permettre d'identifier les deux composantes du vecteur vitesse mais uniquement la composante normale du vecteur vitesse. On définit la **densité** du flot comme un indicateur de résolution du flot calculé. On parle de flot optique dense ou de densité de 100 % lorsqu'on a un vecteur de mouvement par pixel.

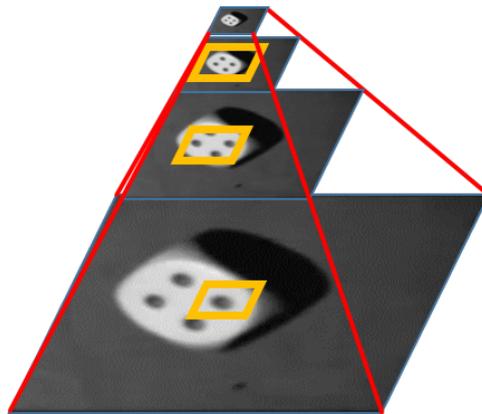
Il est important de noter que la détermination du flot optique s'avère être en pratique une opération loin d'être triviale. Parmi les difficultés, la nécessité de relier l'échelle locale des mouvements et l'échelle globale. Cette liaison est faite classiquement en traitement d'image et vidéo par l'utilisation très large de la pyramide gaussienne. Le concept de la pyramide est décrit sur la figure 2.6. Une pyramide gaussienne nécessite une image et un coefficient positif inférieur à un. La pyramide est construite de différents étages : à chaque étage les dimensions de l'image sont multipliées par le coefficient.

Cette pyramide offre plusieurs avantages :

- accélérer les calculs en travaillant sur les niveaux dont la résolution est la moins élevée dans un premier temps puis raffiner en revenant sur les résolutions originales ;
- permettre de prendre en compte des mouvements à différentes échelles avec des paramètres fixés pour des mouvements compris dans une certaine fenêtre d'amplitudes.

Soulignons que le flot optique peut être utilisé pour décrire plusieurs notions similaires :

- Le flot optique peut désigner le résultat d'une méthode d'estimation de mouvements comme nous allons le voir dans cette section.
- Le flot optique désigne également les méthodes différentielles s'attachant à résoudre le problème de l'estimation de mouvements en résolvant l'équation de flot optique décrite ci-dessus.



**Figure 2.6** – Pyramide gaussienne en rouge avec une fenêtre d'analyse spatiale en jaune : la pyramide permet d'utiliser la même fenêtre pour traiter des échelles différentes.

Nous retrouverons le flot optique également dans la section V.A.. Nous l'appliquerons cette fois sur de la détection d'objet. La suite de cette section propose de détailler les différentes méthodes permettant de trouver le flot optique appliqué à des pixels ou des blocs de pixels. Nous finirons la section en décrivant les méthodes différentielles de flot optique.

### b. Estimation globale du mouvement

L'estimation globale propose de se placer dans le cadre du second modèle spatial présenté dans l'équation (2.1). L'entité considérée ici est l'ensemble des pixels de l'image complète.

Ces méthodes permettent alors d'extraire à partir de l'image les mouvements d'ensemble en utilisant les intensités des pixels. On définit la fonction de coût pour le pixel  $n$  sur l'image  $I_k$  prise au temps  $t_k$  :

$$\epsilon[n] = I_k[n] - I_{k-1}[n - v(n) \cdot (t_k - t_{k-1})] \quad (2.9)$$

Le mouvement défini par la vitesse  $v(n)$  permet alors notamment d'estimer les mouvements globaux de la caméra par rapport à la scène. Ce paramètre peut être estimé en procédant à une minimisation sur un critère quadratique :

$$\min_{b_i} (\mathcal{E}(v)), \quad \mathcal{E}(v) = \sum_n \epsilon^2[n] \text{ avec } i \in \llbracket 1, 6 \rrbracket \quad (2.10)$$

Étant donné le caractère introductif de cette partie, on ne développera pas les méthodes permettant ce type de minimisations.

### c. Correspondance de blocs (*Block matching*)

Cette famille de méthodes d'estimation de mouvements est une des plus célèbres de l'essor de la compression vidéo. Si elle s'exécute sur la totalité de l'image, elle propose une estimation locale du mouvement sur une grille de pixels le plus souvent fixée comme un multiple de la grille physiquement

## II. Estimer le mouvement des pixels et interpoler leur trajectoire

définie par le capteur. Les entités ayant donc la forme de blocs, ces méthodes recherchent alors le bloc le plus semblable dans l'image suivante. Cette famille de méthodes obtient de très bons résultats à condition de limiter la recherche à un rayon défini autour de chaque bloc.

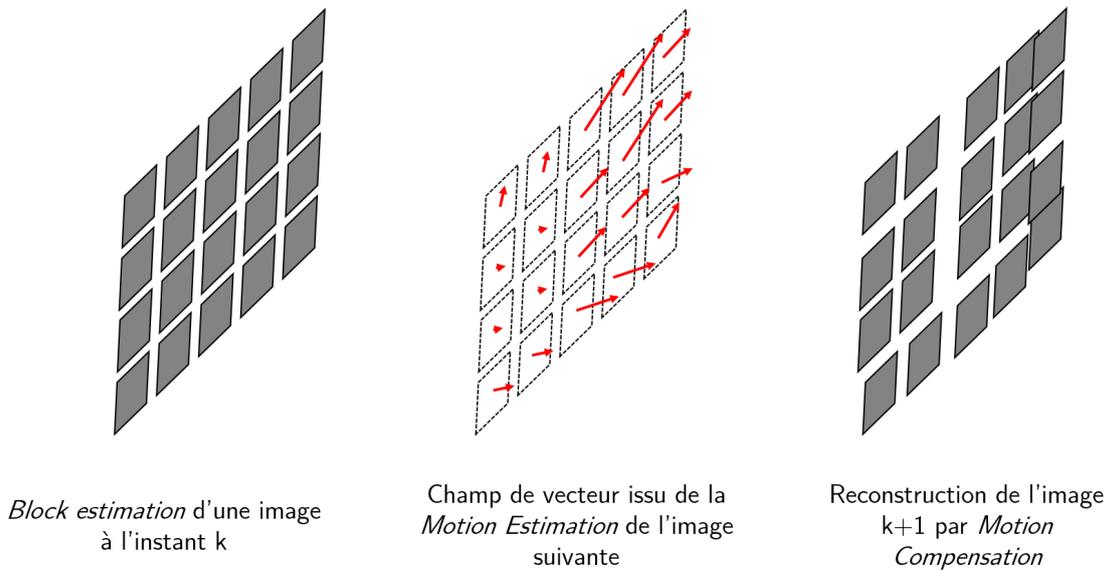
Les hypothèses du modèle temporel de l'équation (2.2) sont sous-jacentes à ces méthodes, ainsi que l'hypothèse d'une faible variation de luminosité locale pour une utilisation optimale.

D'une manière générale, les méthodes proposent la minimisation suivante :

$$\min_{d[m] \in \mathcal{D}} (\mathcal{E}(d[m])), \quad \mathcal{E} = \sum_{n \in \mathcal{B}_m} \Phi(I_k[n] - I_{k-1}[n - d[m]]) \quad (2.11)$$

Dans cette minimisation, on définit un domaine  $\mathcal{D}$  autour du bloc  $\mathcal{B}_m$ , le but étant de minimiser au sens du critère  $\Phi$  le déplacement  $d[m]$  du bloc  $m$ .

Le critère classiquement utilisé a longtemps été quadratique avant d'évoluer vers un critère valeur absolue pour des raisons de rapidité. La plupart des méthodes classiques de recherche de correspondance utilisées dans la littérature sont détaillées par Barjatya (2004) et Jakubowski et Pastuszak (2013).



**Figure 2.7** – Schéma montrant la reconstruction d'image à l'aide du champ de mouvement des différents blocs de l'image précédente

### d. Corrélation de phases

Contrairement à l'ensemble des méthodes développées dans cette section, la famille de méthodes de corrélation de phases propose une résolution du problème dans le domaine des fréquences spatiales.

En posant la transformée de Fourier spatiale  $\mathcal{F}$  d'une image  $I[n]$  telle que  $\mathcal{F}\{I[n]\} = \hat{I}[f_u]$ , avec  $f_u$  la fréquence spatiale 2D, on a si on considère le vecteur  $z$  :

$$\mathcal{F}(I_k[n - z]) = \hat{I}_{k-1}(f_u) \cdot e^{-2\pi j f_u \cdot z} \quad (2.12)$$

Autrement dit, la phase de la transformée de Fourier spatiale de l'image contient l'information spatiale de son déplacement. Cette propriété peut être applicable directement pour récupérer des informations sur les déplacements globaux de l'image. Les informations sur les déplacements globaux peuvent être extraites en cherchant les maximums d'un opérateur de corrélation croisée normalisée entre les deux images, ce qui donne :

$$\Psi_{k-1,k}(n) = \mathcal{F}^{-1} \left\{ \frac{\hat{I}_k(f_u) \hat{I}_{k-1}^*(f_u)}{|\hat{I}_k(f_u) \cdot \hat{I}_{k-1}^*(f_u)|} \right\} \quad (2.13)$$

$$\Psi_{k-1,k}(n) = \mathcal{F}^{-1} \left\{ e^{-2\pi j f_u \cdot z} \right\} \quad (2.14)$$

$$\Psi_{k-1,k}(n) = \delta(n - z) \quad (2.15)$$

avec  $\delta$  le symbole de Kronecker. Pour utiliser la phase dans un cadre plus local, il suffit de changer d'entité et de prendre une grille de blocs puis de regarder leur déplacements respectifs de la même manière que le *Block matching* étudié dans le paragraphe précédent. Dans ce cadre, on ne peut pas prendre en compte les rotations ou les zooms des blocs considérés. La méthode s'avère peu robuste lors de changements brusques de luminosité et de précision pixélique. Pour obtenir une meilleure approche multiéchelle, Simoncelli et Freeman (1995) proposent d'utiliser une banque de filtres basée sur le concept des pyramides gaussiennes. Cette approche demande une charge de calcul plus importante mais permet d'obtenir une phase locale de l'image.

#### e. Méthodes différentielles

Comme il était précisé dans la partie introductive de cette section, la détermination du flot optique de manière différentielle est le plus souvent associée au terme « flot optique », parce qu'elle s'attache à la résolution de l'équation du flot optique évoquée plus haut (équation (2.8)).

Dans les hypothèses de cette approche, de nombreuses méthodes ont été utilisées pour résoudre ce problème. Horn et Schunck (1981) proposent une méthode de résolution basée sur une régularisation globale donnant un résultat malheureusement bruité et avec un comportement temporel moyenné. La même année, Lucas et Kanade (Lucas et Kanade, 1981) proposent une approche différente basée sur une régularisation locale du champ de vitesses. Cette approche plus robuste au bruit sacrifie la résolution spatiale pour conserver la résolution temporelle. L'état de l'art de référence d'avant les années 2000 est bien synthétisé à travers une comparaison des différentes techniques existantes par Barron *et al.* (1994).

Le problème de résolution différentielle a été exhaustivement traité depuis, les applications étant extrêmement nombreuses et variées : détection de caractéristiques de l'environnement (Ohnishi et Imiya, 2006), vision pour véhicules autonomes (Jia *et al.*, 2008) sur des applications temps réel (Jia *et al.*, 2015), photogrammétrie, compression, suivi d'objet, structure 3D à partir du mouvement (Lee *et al.*, 2009). Pour permettre ces nombreuses utilisations, nous proposons de rappeler les grandes lignes de l'état de l'art récent du flot optique. L'équation du flot optique étant mal posée, il s'agit d'ajouter une condition pour pouvoir l'utiliser. Farnebäck (2003) propose d'utiliser une approximation polynomiale à une échelle locale pour permettre de s'abstraire du caractère discret de l'image. Cette approche est très robuste au bruit et permet de donner de bons résultats à condition de paramétrer correctement la pyramide gaussienne d'images traitées. Bernard (1999) étudie une approche différente : il répond à la contrainte multiéchelle par une décomposition sur une base d'ondelettes. Brox *et al.* (2004) reprennent les approches de régularisation spatiale et temporelle

## II. Estimer le mouvement des pixels et interpoler leur trajectoire

et forment une version spatio-temporelle. Le point faible de toutes les méthodes de flot optique définies précédemment reste que leur résolution de l'équation de mouvement n'est exacte que dans un voisinage local. L'unique solution pour s'abstraire de ce voisinage reste la pyramide, ce qui rend le mouvement estimé imprécis.

Alvarez *et al.* (2000) soumet le premier une méthode ne linéarisant pas les contraintes et pouvant traiter de grands mouvements. Brox *et al.* (2009) combine des méthodes de segmentation automatique avec les méthodes classiques de flot optique et obtient des résultats concluants pour de grands déplacements. La philosophie est la même pour Denman *et al.* (2009) qui propose en plus d'une estimation dense, de l'appliquer au cadre du *tracking* (suivi) d'objets et à l'extraction d'informations comme les occlusions. Enfin Brox et Malik (2011) vont introduire l'utilisation de descripteurs dans l'image permettant de caractériser les zones de l'image selon les amplitudes de mouvements. Sans être exhaustifs, on finira ce tour d'horizon des méthodes différentielles en citant les plus récents travaux de Lepoittevin (2015) qui résout l'équation de flot optique en utilisant des méthodes d'assimilation de données. Il montre une application sur les images satellites de météorologie mais aussi sur du suivi d'objets<sup>2</sup>.

### f. Conclusion

La littérature fait état d'une large palette de solutions pour résoudre le problème de l'estimation de mouvements. Ce problème étant mal posé, beaucoup de stratégies différentes ont été proposées. Parmi ces stratégies les hypothèses de résolution sur un voisinage local donnent de bons résultats mais limitent les résultats à de « petits mouvements » d'objets entre deux images consécutives. Plus récemment, plusieurs approches ont été développées permettant de travailler sur les différentes échelles de l'image : la pyramide gaussienne qui permet d'appliquer un même filtre sur des images de tailles différentes, l'utilisation de fonctions de régularisation du flot optique, la formulation du problème sous une forme énergétique, ou encore l'utilisation d'une segmentation automatique pour travailler sur des sous-ensembles de l'image. Si ces méthodes permettent de travailler sur de grands mouvements, elles possèdent deux inconvénients : un coût de calcul élevé et l'utilisation de l'entité pixel pour extraire les mouvements, classique pour une approche de traitement vidéo.

Dans notre but de construire un algorithme utilisant une approche de composition de mouvement, nous utiliserons ces algorithmes en restreignant leur application sur des petits mouvements. Il est à noter que la même littérature existe au niveau de l'implémentation des méthodes et de la vitesse de calcul. Nous n'avons pas retenu de méthodes utilisant de la corrélation croisée ou des approches par ondelettes pour la suite par souci de garantir une solution temps réel.

Afin de pouvoir continuer cet état de l'art, la partie suivante introduit un domaine utilisant ces méthodes d'extraction de mouvement : la compression de données.

## II.B. Application à travers la philosophie des normes de compression MPEG

L'histoire de la compression de donnée et des standards populairement connus *MPEG* (*Movie Pictures Experts Group*) ont été initiés par le comité normatif *ISO* (*Organisation internationale de normalisation*) à la fin des années 80. En partant des normes *JPEG* (*Joint Photographic Experts Group*) et de la norme H.261<sup>3</sup>, le comité a pour objectif de définir les paramètres des formats

2. De nombreuses revues existent dont on conseillera la lecture de Fortun *et al.* (2015).

3. Toutes les informations concernant les normes *MPEG* sont issues du très bon site éponyme *Moving Picture Experts Group* (2017).

standards (débit, résolution, fréquence, format) et les décliner à travers une norme structurée en plusieurs parties ; parmi celles-ci nous nous intéresserons plus spécifiquement à la partie vidéo (partie 2 des différentes normes) et la partie compression (partie 10). Bien que ces normes ne soient que des standards qui ne sont éventuellement pas implémentés faute de solution technique, il nous a semblé pertinent d'exposer les différents mécanismes utilisés dans la compression en rapport avec l'estimation de mouvement. De la même manière les différents profils pour définir du contenu multimédia définis par la norme MPEG-4 seront détaillés, notamment le codage vidéo prenant en compte des objets. Ce dernier point nous permettra de définir les contours d'une philosophie de développement dans le domaine vidéo.

### II.B.1. Les méthodes d'estimation de mouvements et de compensation de mouvements

Les méthodes décrites dans la section précédente ont permis ainsi le développement d'une chaîne qui a entraîné l'essor de la compression. L'estimation de mouvement (ou *Motion Estimation - ME*) se base ainsi sur une détection de mouvement pour obtenir le mouvement entre deux images. Il est alors possible de construire une image intermédiaire en appliquant les modèles spatiotemporels choisis (sous-section II.B.). Cette méthode s'appelle la compensation de mouvements (*Motion Compensation* ou *MC*). Un exemple d'implémentation est illustré sur la figure 2.7.

Les méthodes de compression proposent de travailler sur une grille de pixels, et après une première phase d'estimation de mouvements, de venir comparer à l'aide de critères d'erreur comme le *RMSE*<sup>4</sup> le résidu pour chaque bloc prédit avec le bloc réel. Lorsque l'erreur est en deçà d'un certain seuil, le bloc peut être supprimé et les valeurs du vecteur de mouvement retenues : la vidéo vient d'être compressée.

Cette technique va être très largement utilisée dans les différentes normes et codecs<sup>5</sup> de compression que nous décrivons dans la suite de cette section.

### II.B.2. MPEG-1 à MPEG-2

Le MPEG-1 fait donc suite en 1993 à la norme H.261 et au standard JPEG. Le MPEG-1 est ainsi peu novateur mais propose des premières améliorations. Tout d'abord, il reprend la compensation de mouvements prédictive qui existait pour établir une compensation de mouvement bidirectionnelle. Ainsi, on peut voir sur la figure 2.8 comment sont organisées les images successives dans une vidéo. Cette succession introduit le concept de « groupe d'images » (*Group Of Pictures - GOP*) qui définit une série d'images calculées à partir de la première et de la connaissance du champ de mouvement. On distingue ainsi les images initiales (I), celles prédites à partir de ces images initiales (P) et celles prédites de manière bidirectionnelles à partir d'images I ou P.

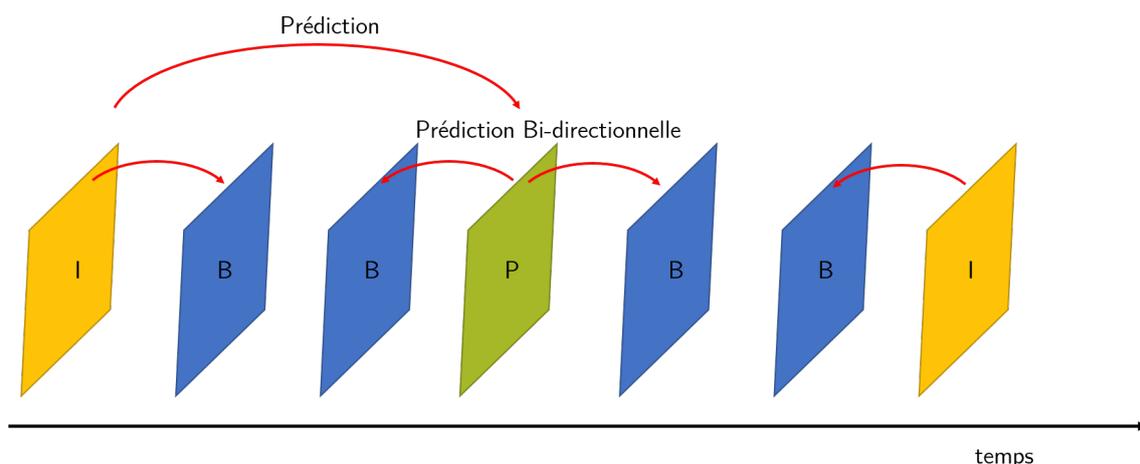
Les exigences sur cette compensation de mouvements sont également plus élevées : on passe d'une précision au pixel à 0,5 pixel et l'étendue d'un vecteur de mouvement passe de  $\pm 15$  pixels à  $\pm 512$  pixels.

Bien entendu, la transformée *DCT* (pour *Discrete Cosinus Transform*) introduite par le *JPEG* est utilisée sur une grille de blocs, des améliorations sur les tailles de blocs sont apportées ainsi que l'introduction de codage à taille variable (*Variable Length Coding*) que l'on ne développera pas.

4. *Root Mean Square Error* ou racine carrée de l'erreur quadratique moyenne.

5. Un codec est une implémentation matérielle d'un standard.

## II. Estimer le mouvement des pixels et interpoler leur trajectoire



**Figure 2.8** – Succession des images dans une vidéo : seules les images jaunes I existent dans le flux compressé, puis sont reconstruites les images P en vert et enfin les images B en bleu

La norme MPEG-2 a été introduite en 1994. Elle n'a pas apporté beaucoup de changements au niveau de l'estimation de mouvement si ce n'est l'introduction du codage d'images entrelacées. Ce codage introduit la notion de trame : une image est composée de deux trames entrelacées<sup>6</sup>. Le MPEG-2 permet de définir ces trames en tant qu'images I, P ou B, avec l'avantage en cas d'image I de pouvoir définir une trame de manière prédictive par rapport à l'autre.

Le MPEG-2 a surtout introduit des profils de décodeurs permettant de décoder le flux avec plus ou moins de détails, offrant la possibilité à des appareils moins puissants de lire les vidéos (par exemple certains décodeurs ne décodent pas les image-B). La norme MPEG-2 a ainsi permis le passage définitif de la vidéo dans le numérique en la démocratisant pour de nombreux supports.

### II.B.3. Les normes MPEG-4 et MPEG-7

#### a. La compression selon MPEG-4

Il a fallu attendre près de 10 ans pour établir la norme suivante : le MPEG-4. Cette norme comporte dans sa partie 10 la suite des standards de compression initiés plus haut. Elle améliore la précision en diminuant la taille des blocs utilisés dans le processus d'estimation/compensation et propose d'abaisser la précision au quart de pixel. La concordance temporelle du *GOP* est mise de côté pour préférer le concept de tampon (*buffer* appelé aussi *Dynamic GOP*). 16 images variables constituent un tampon à partir duquel les images prédites de manière simple ou bidirectionnelle seront calculées. Ainsi une image peut être composée de blocs de différentes images et une image prédite peut elle-même faire partie du tampon. Elle permet de définir un bloc prédit par une moyenne pondérée de deux blocs issus d'images différentes. Enfin, de manière duale à la prédiction temporelle, elle introduit l'*Intra-prediction* ou prédiction spatiale. Cette prédiction spatiale intervient en complément de la DCT qui était la seule méthode à venir limiter les redondances spatiales. Cela revient concrètement à prédire le mouvement d'un bloc à partir de celui de ses voisins et donc de supprimer le vecteur de mouvement correspondant. Ce type de prédiction est très efficace pour des mouvements globaux de l'image au niveau du fond ou des fonds fixes.

6. chaque image est décomposée en deux trames contenant chacune une ligne sur deux de l'image. La création de ces trames est uniquement due aux limitations du balayage des premiers téléviseurs

Contrairement aux idées reçues, ce n'est pas l'amélioration des normes de compression qui a atténué la discrétisation spatiale en petits carrés si caractéristiques des versions précédentes, mais l'ajout d'un filtre antiblocs que l'on applique après la compensation : *In-Loop Deblocking Filter*.

La norme de compression définie par MPEG-4 appelée également H.264 ou MPEG-4/AVC (pour *Advanced Video Coding*) continue finalement dans la même direction que les normes précédentes en introduisant de nombreuses nouveautés. Cependant, la dernière évolution de ces trois normes donne l'impression d'aboutir à la déstructuration complète à la fois spatiale et temporelle. La vidéo compressée devient un flux de blocs et de vecteurs de mouvements : cette compression algorithmique perd le sens sémantique de la vidéo. Les applications spécifiques de vidéos rapides étant souvent utilisées en photogrammétrie ou vidéogrammétrie, il est indispensable qu'il n'y ait aucune perte dans la compression. Néanmoins, une compression utilisant la sémantique de l'image pourrait être plus adaptée et ainsi préserver la qualité de la vidéo sur les zones d'intérêt tout en compressant le reste de l'image.

### b. La vidéo selon MPEG-4 et MPEG-7

Si le but du standard de compression H.264 était d'être implémenté rapidement sous forme de codec, ce n'est pas le cas d'autres parties de la norme qui définissent un cadre d'application très large comme la vidéo 3D ou la vidéo contextuelle. Si certains aspects ont depuis été développés, d'autres standards définis dans la partie 2 de la norme restent encore sujets à innovation.

Il est intéressant de s'attarder sur les standards énoncés par la norme MPEG-4 en terme de flux vidéo qui introduit le concept de codage « orienté objet ». Sans mauvais jeu de mot avec le langage de programmation éponymes, il s'agit d'identifier, de suivre et de considérer les objets dans les vidéos comme séparés entre eux et du fond de la vidéo. Cette vision beaucoup plus sémantique est par ailleurs appuyée par la norme MPEG-7 (2002) qui propose non seulement de mettre en place des descripteurs visuels permettant d'extraire les informations comme la couleur, la texture, la forme ou encore le mouvement, mais de les organiser spatialement et temporellement à travers une structure de type XML. On propose de détailler la partie 2 du MPEG-4 afin de voir quelles orientations sont proposées pour de futures implémentations.

Le standard MPEG-4/Part 2 définit le concept d'Objet Audio Visuel (AVO) caractérisé par sa texture et sa forme. Le standard ne définit pas d'algorithme de segmentation permettant d'extraire ces objets, laissant ainsi la liberté d'utiliser toute la diversité de ces méthodes. Les objets sont décrits par des imajettes carrées en bitmap avec une composante alpha codant la transparence de l'image (essentiel pour respecter la forme de l'objet). La vidéo revient alors à l'évolution d'objets définis par une série d'imajettes, sur un fond. La définition du fond est d'ailleurs généralisée avec la notion de *Sprite* qui propose de fusionner le fond de plusieurs images pour pouvoir définir un panorama sur lequel la caméra évolue. Enfin, des évolutions significatives sont proposées pour l'animation de l'intérieur des objets avec l'utilisation d'un maillage (*mesh*). Afin de réduire les erreurs, un maillage triangulaire est utilisé et permet d'animer et de coder l'évolution interne de l'objet. Cette introduction de modèles dédiés est liée à quelques objets spécifiques : animation du visage et animation du corps.

Cette philosophie de compression vidéo est plus orientée dans la sémantique de l'objet et pourrait davantage convenir pour les problématiques énoncées dans ce projet. Étant donné que l'état de l'art ne propose pas de méthode complète selon cette approche, la méthode que nous détaillerons dans la section IV sera développée en suivant cette philosophie : extraire les objets de la scène pour les

traiter séparément du reste de la scène.

### II.C. L'interpolation pour modifier la fréquence temporelle

Étant donné notre cahier des charges, il s'agit de proposer une méthode qui permet de compresser des séquences, mais également de faire de l'extrapolation en synthétisant des images supplémentaires entre les images acquises. Ces techniques de reconstruction sont souvent basées sur des approches de type *ME-MC* où l'on cherche dans un premier temps les correspondances entre les pixels de deux images pour ensuite reconstruire le trajet de ces pixels et obtenir des positions intermédiaires de pixels entre deux instants. Les objectifs sont différents des méthodes de compression qui recomposent le mouvement entre deux images pour le comparer aux images acquises et obtenir un résidu. Ici le mouvement est recomposé pour créer de nouvelles images et faire ainsi une interpolation. Ces interpolations sont couramment utilisées pour des adaptations de fréquence d'affichage ou dans les outils vidéos qui permettent d'effectuer des ralentis en venant modifier la fréquence de restitution, ce qui nécessite d'ajouter des images<sup>7</sup>. Du point de vue du traitement du signal, il s'avère que ces applications esthétiques sont pourtant en contradiction avec le théorème d'échantillonnage qui nous empêche de voir des mouvements au-delà de deux fois la fréquence d'acquisition. Cependant, s'agissant plus d'application de **rendu** que de photogrammétrie, la question ne se pose pas dans ce domaine. Les premières applications proposent d'utiliser le modèle linéaire (cf. section II.A.1.) pour des adaptations de fréquence d'affichage (Castagno *et al.*, 1996). Les applications liées aux objets démarrent avec l'arrivée du standard MPEG-4, comme le souligne Benois-Pineau et Nicolas (2002) avec une approche « 2,5D » alliant segmentation par régions et identification des différents plans de la scène. Ces méthodes coexistent pour des utilisations en interpolation et sont regroupées sous le terme de *FRUC* (pour *Frame Rate-Up Conversion*). Dans un domaine où toute implémentation est extrêmement concurrentielle l'état de l'art est assez vaste à ce sujet (Ling *et al.*, 2008; Wang *et al.*, 2010a,b) et l'interpolation de trajectoires basée sur des objets commence à apparaître sans toutefois avoir de correspondance entre un objet physique et les objets segmentés dans la vidéo Brox *et al.* (2009). D'autres applications utilisent des descripteurs (sous-sous-section III.B.2.) pour caractériser les zones de l'image en terme de mouvement, forme, contenu (Brox et Malik, 2011; Cho *et al.*, 2013; Kim *et al.*, 2015)<sup>8</sup>. Certaines méthodes proposent d'utiliser des manières différentes pour estimer le mouvement. Mahajan *et al.* (2009) et Yan et Chen (2013) proposent par exemple une approche basée sur les chemins de pixels avec un coût de calcul plus élevé. Meyer *et al.* (2015) propose des méthodes sous-pixéliques spectrales pour faire de l'interpolation au niveau pixel pour des coûts de calculs significativement plus faibles que les méthodes classiques.

Nos hypothèses de fond fixe offrent aux méthodes utilisant l'opposition objet/fond un cadre d'application cohérent. Dans ce cadre, nous proposons une méthode basée sur la trajectographie des objets évoluant sur la scène. Lee *et al.* (2013) suggère dans un contexte différent d'annotation vidéo d'utiliser des splines cubiques pour suivre les trajectoires des objets. L'état de l'art ne fait pas état d'un rapprochement entre les concepts d'objets en mécanique et d'objets en vidéo qui pourraient être appliqués dans une méthode *ME-MC*.

Aussi, nous resterons dans les orientations définies par les normes MPEG-4 Part 2 et MPEG-7 pour proposer une nouvelle méthode orientée objet combinant avantageusement une approche mécanique

7. L'annexe A. détaille ces effets de restitution.

8. Plus d'informations sur les descripteurs pourront être trouvés dans l'agréable lecture francophone de Garcia (2009) et dans l'ouvrage complet de Krig (2014).

grands mouvements et une approche vidéo de l'évolution de l'objet dans le plan image. Cette méthode permettra de faire de l'estimation de grands mouvements avec une nouvelle approche et des applications en compression et extrapolation.

Le développement d'une correspondance entre objet vidéo et objet mécanique demande de développer l'état de l'art des méthodes de traitement vidéo concernant la détection, le suivi et l'extraction d'objets. Nous introduirons cette nouvelle méthode dans la section IV après avoir détaillé cet état de l'art.

### III Caractériser un objet dans une séquence d'images : un problème mal posé en 2D

Pour caractériser un objet dans une séquence d'images, il faut s'intéresser à son contenu sémantiquement lié au monde réel. Le problème n'est pas simple et consiste à identifier des groupes de pixels appartenant à un objet de la scène. Cette partie expose un état de l'art sur les techniques d'extraction d'objets utilisées en vision, selon nos hypothèses de travail : fond fixe, caractérisation d'objets et schéma temps réel. Ces méthodes peuvent provenir soit du traitement vidéo, soit du traitement d'images. Cet état de l'art nous permet également de définir des méthodes permettant de passer d'informations eulériennes, c'est-à-dire liées à la matrice de pixels du capteur, à des informations lagrangiennes liées cette fois aux objets de la scène. Ces méthodes nous serviront par la suite comme des applications permettant de relier les deux descriptions (cf. section IV).

Deux parties seront distinguées :

#### La détection

Il s'agit du niveau le plus bas de traitement d'images, c'est-à-dire le premier niveau de transformation des pixels. Ces méthodes ont en entrée des pixels eulériens et fournissent en sortie des informations eulériennes ou des informations lagrangiennes. Nous traiterons essentiellement ce dernier cas, étant donné nos objectifs d'extraire des informations liées aux objets vidéos.

#### L'extraction

Ce niveau permet d'extraire des informations sur les objets préalablement détectés. Durant cette phase, la sortie sera alimentée par des informations lagrangiennes alors qu'à l'entrée elles seront variables en terme d'objets considérés (pixels, groupe de pixels).

Avant tout, nous avons souhaité donner un aperçu plus orienté « applications industrielles » des techniques utilisées actuellement pour pallier ce cadre de manière matérielle.

#### III.A. Cas particulier de dispositifs matériels ajoutant des informations à l'image 2D

Les méthodes employées tout le long de la chaîne de traitement de l'image travaillent sur un problème mal posé : caractériser une scène 3D à partir d'informations 2D de cette scène. Le problème est impossible à résoudre sans informations complémentaires sur la scène. Une première manière de le résoudre est matérielle : l'ajout d'un autre capteur permet de compléter les mesures faites par le premier et ainsi retrouver la position en trois dimensions des points d'intérêt dans la scène.

- Ces informations peuvent être apportées par d'autres capteurs. La stéréovision propose en effet d'utiliser deux points de vue différents pour permettre de retrouver une information 3D. Le projet DICCIT (Mathieu *et al.*, 2015) propose de développer et de normaliser l'utilisation de

### III. Caractériser un objet dans une séquence d'images : un problème mal posé en 2D

mesures par stéréovision pour des caméras rapides avec des stratégies longues de traitements après l'acquisition. Ces processus définissent l'espace de travail de manière relativement libre concernant les hypothèses de la scène tout en exigeant une phase de calibration spécifique, ainsi qu'un traitement individuel de chaque mesure. D'autres caméras moins rapides à ce jour proposent directement de fournir des données 3D en utilisant des *GPU*<sup>9</sup> embarqués sur un capteur comportant déjà deux caméras. Les capteurs proposés par *StéréoLabs* sont actuellement les plus polyvalents : faible flux de données, précision dans l'espace de mesure, taille de l'espace de mesure et facilité de mise en place. La scène peut alors être reconstruite en 3D dans un espace déterminé par la position des caméras et leur profondeur de champ.

- D'autres approches utilisent un nombre quelconque de caméras de positions connues pour déterminer complètement la position des objets dans la scène. C'est le cas par exemple des caméras embarquées qui peuvent prendre plusieurs points de vue d'une scène et ainsi effectuer une reconstruction 3D. C'est le cas par exemple des produits sur drone que propose *Pix4D*.
- Des approches utilisent la fusion de données et peuvent fournir les informations manquantes. C'est le cas des capteurs *RGB-D* (pour *Red Green Blue - Depth* ou *RVB - profondeur* en français) associant une matrice classique *RGB* avec un capteur pouvant mesurer la profondeur sur une grille de point de l'espace. Il s'agit la plupart du temps d'un laser infrarouge. Des systèmes grands public comme la *Kinect™* de Microsoft® utilisent ce type de capteur. Ils sont cependant contraints par une faible distance de travail et minorés par le capteur le moins performant.
- Enfin, des approches par reconstruction existent sans effectuer de fusion de données. Les approches utilisant la lumière structurée permettent ainsi de reconstruire la profondeur à l'aide de la déformation d'un pattern projeté sur l'objet par l'éclairage. Les caméras *ToF* (pour *Time of Flight* ou temps de vol) permettent également d'obtenir la profondeur, cette fois en mesurant les différences de temps d'arrivée des photons sur le capteur. Cette technologie est prometteuse, mais n'atteint pas pour l'instant des niveaux de précision suffisants pour nos applications. Pour finir, une élégante méthode développée ces dernières années consiste à extraire une carte de profondeur à partir du flou des différents objets a été appliquée sous forme de caméra par *Trouvé (2012)*. À travers le concept du *Depth from Defocus*, il est proposé d'estimer la profondeur à partir du flou issu de la défocalisation. Ces recherches sont maintenant commercialisées par la start-up *Cyclopus*.

Lorsque le matériel ne permet pas de reconstruire les données manquantes, il faut malheureusement traiter des données incomplètes. Il est cependant possible d'ajouter une hypothèse pour permettre de traiter les images et pallier au manque d'information. La section suivante propose de traiter ce point en passant en revue d'abord les hypothèses globales puis des hypothèses locales.

#### III.B. Détecter et suivre

La détection de pixels dans une image 2D peut se faire sans à priori sur la scène (cf. section II.B.). Cependant le cadre de notre étude est plus restreint et nous autorise à utiliser des méthodes de détection plus robustes grâce à la connaissance du scénario de la vidéo.

Détecter un pixel consiste à trouver des manières de caractériser l'image pour en déduire un classement, une appartenance, une corrélation en fonction des données spatiales ou temporelles. Ces caractéristiques sont multiples : positions, mouvement et photométrie (terme regroupant couleur,

---

9. Processeurs Graphiques utilisés notamment pour leur capacité à traiter les calculs de manière parallèle.

forme et fréquence spatiale). Suivant les hypothèses que l'on peut formuler sur la scène, c'est-à-dire en fonction du scénario attendu, l'état de l'art se sépare en deux approches : une approche globale appliquée à l'ensemble de l'image, opposée à une approche locale calculée sur des points caractéristiques. C'est en suivant cette segmentation que nous allons aborder cette section.

Nous avons préféré exposer les différentes méthodes de suivi avec les méthodes de détection car le résultat est identique : une détection de pixels. Nous garderons néanmoins à l'esprit que les méthodes de suivi nécessitent une étape d'initialisation et utilisent récursivement l'estimation faite à l'instant précédent.<sup>10</sup>

### III.B.1. Approches globales

#### a. Hypothèse d'à priori spatial

La manière la plus naturelle de segmenter un objet est de le définir à travers une caractérisation spatiale. Au niveau global, de nombreuses manières de caractériser l'objet ont été développées.

#### Couleur

La manière la plus intuitive de caractériser spatialement est d'utiliser le niveau d'intensité ou la couleur. Sélectionner les pixels de cette manière s'appelle une opération de seuillage. Lorsqu'on travaille en nuance de gris, on projette les couleurs issues d'une représentation colorimétrique en trois dimensions sur un espace comportant une seule dimension. Les informations colorimétriques en nuances de gris ou intensité ne sont pas assez significatives. Dans le cas particulier d'un éclairage contrôlé, il est possible de faire de la segmentation à condition que celui-ci soit homogène sur toute la scène. De telles conditions expérimentales demandent une préparation très minutieuse et sont plus ou moins difficiles à mettre en œuvre selon les objets.

L'utilisation de la couleur comme caractéristique de détection s'avère très efficace pour des images couleurs. Le choix de l'espace colorimétrique est déterminant. La plupart des machines de vision industrielles<sup>11</sup> utilisent le *RG Chromacity*. Ces machines sont souvent placées dans des environnements contrôlés en éclairage de manière à maîtriser la luminosité, principal paramètre dégradant la robustesse de l'identification colorimétrique. Le désavantage de ce type d'espace réside principalement dans la difficulté de réglage puisqu'il ne permet pas de découpler la couleur de la luminosité. Aussi, un seuillage sur un domaine de couleur pour un intervalle de luminosité dans un espace de type *RGB* reviendra à une série de domaines disjoints.

D'autres modèles colorimétriques comme le *YCbCr* ou le *HSL* permettent justement d'effectuer cette séparation. Il est ainsi plus aisé de paramétrer un intervalle de couleur. Cependant comme on peut le voir sur la figure 2.9 ces espaces comportent des zones où les couleurs sont difficilement déterminables : l'influence de l'éclairage reste encore prépondérante. Malgré des conditions initiales contrôlées, ces plages peuvent être atteintes pendant le déroulement de la scène par l'objet en mouvement principalement à cause de ses caractéristiques spéculaires.

Ces techniques ne permettent pas de modéliser tous les cas. En effet, l'image réelle possède souvent des "textures" et non des plages de couleur unies. La caractérisation de ces textures se fait classi-

10. Nous conseillons au lecteur souhaitant approfondir le suivi de se reporter aux revues de Yilmaz *et al.* (2006) et Smeulders *et al.* (2014).

11. Les machines de vision industrielles désignent des systèmes utilisant l'ensemble des technologies optiques, électroniques et logicielles permettant d'acquérir des images et d'y détecter ou mesurer des caractéristiques prédéfinies.

### III. Caractériser un objet dans une séquence d'images : un problème mal posé en 2D

quement à l'aide de la notion d'entropie. Cette notion utilise les probabilités d'apparition de chacun des  $N$  niveaux d'intensité possibles  $p_i$  pour calculer l'entropie de Shannon :  $E = -\sum_{i=1}^N p_i \log(p_i)$ . Pour de faibles valeurs, on peut caractériser une texture par des seuillages classiques puis en utilisant des opérations morphologiques afin d'utiliser la cohérence spatiale pour assembler les différents pixels détectés. Cependant, des approches bien plus robustes existent en caractérisant la couleur à l'aide de distributions. Différentes méthodes existent utilisant des opérateurs de distance pour comparer ces distributions (euclidienne, Mahalannobis, Bhattacharyya) (Dubuisson, 2011) ou des mixtures de gaussiennes pour les modéliser. D'autres méthodes sont basées sur la logique floue ou utilisent des moments statistiques pour les caractériser<sup>12</sup>. La littérature propose également des algorithmes permettant de segmenter la totalité de l'image en objets vidéo, sans avoir toujours beaucoup de sens avec les objets réels (*K-means* ou *Watershed*). Ces algorithmes plus coûteux mais robustes servent souvent à initialiser une segmentation de l'image pour raffiner ensuite le résultat par d'autres méthodes nécessitant une initialisation (contour actif, méthodes de suivi...).

D'autres méthodes de segmentation bas niveau basées sur les coupes de graphes s'avèrent efficaces pour segmenter une image en plusieurs sous-parties. Ces approches sont des méthodes d'optimisation qui minimisent une fonctionnelle d'énergie composée d'un terme de couleur ou de texture et d'un terme de cohérence spatiale. Elles peuvent être utilisées pour faire notamment du suivi (Fu *et al.*, 2015; Malcolm *et al.*, 2007; Wang *et al.*, 2009). Cependant, la construction d'un graphe s'avère extrêmement coûteuse et ces algorithmes souffrent également de défauts d'erreur de segmentation similaires aux méthodes précédemment citées<sup>13</sup>.

Enfin certaines méthodes caractérisent la colorimétrie par son histogramme de couleur, c'est-à-dire le nombre de pixels en fonction des différents niveaux colorimétriques. C'est notamment de cette manière que fonctionne l'algorithme de *MeanShift* (Bugeau, 2007). Ces algorithmes partent d'une zone de départ et convergent vers la zone de l'image la plus proche ayant un histogramme similaire sur l'image suivante. L'algorithme CAMshift est une adaptation pour les images et permet de faire varier la taille de la fenêtre d'analyse en fonction des moments du second ordre (Artner, 2008).

#### Contour

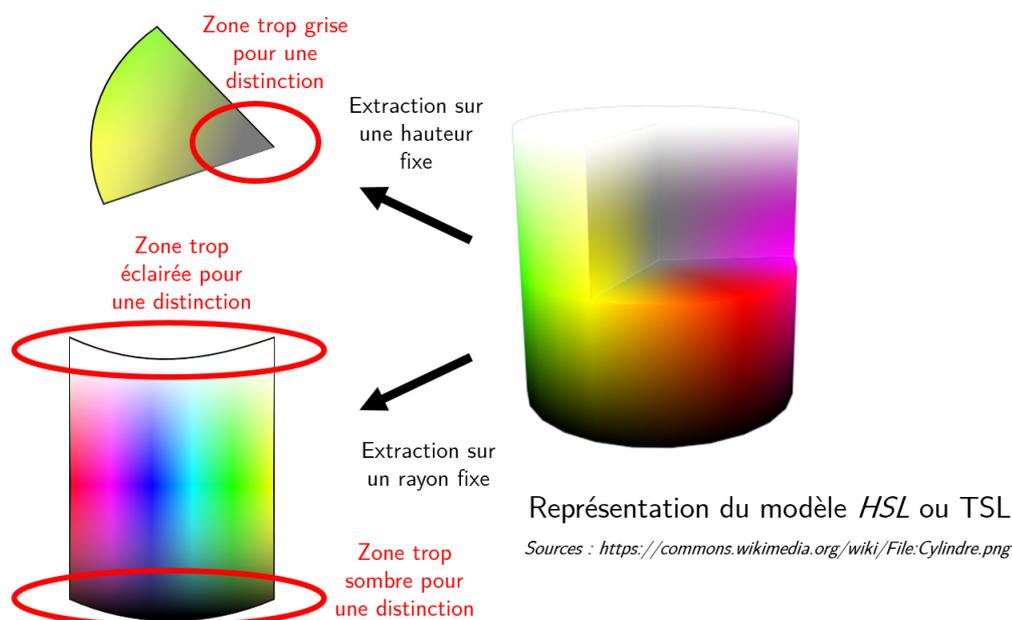
Identifier le contour d'un objet est également une manière efficace de le caractériser. La manière la plus simple d'extraire le contour est d'utiliser l'approche issue du traitement d'image : utiliser un filtrage haute fréquence. Ce filtrage peut être fait par convolution à l'aide de noyaux basés sur des opérateurs de gradients directionnels (filtrage de Sobel) ou des noyaux laplaciens. Une approche plus complexe existe avec le filtre de Canny qui utilise dans un premier temps un filtre passe-bas pour éliminer les contours parasites en floutant l'image, applique dans un second temps des filtres de Sobel et vient raffiner dans un troisième temps les résultats en fonction de la connectivité des pixels détectés. Cette approche est plus coûteuse mais plus robuste. D'autres approches orientées vidéo se servent de la notion d'objet vidéo pour isoler le contour externe et la texture interne de l'objet en fonction de la persistance de zones du contour (Wang *et al.*, 2013)<sup>14</sup>.

Il n'est pas possible de traiter le suivi de contours sans parler de contours actifs. Cette approche itérative introduite en 1987 permet à chaque pas de faire évoluer un contour initial défini par

12. On se reportera à la lecture de Sezgin et Sankur (2004) pour plus de références.

13. On pourra se référer à Boykov et Veksler (2006) pour plus de détails sur la théorie des graphes.

14. Un état de l'art assez récent davantage orienté traitement d'image que traitement vidéo est proposé par Papari et Petkov (2011).



**Figure 2.9** – Schéma mettant en lumière les zones de l'espace colorimétrique HSL où la distinction colorimétrique est plus difficile

l'utilisateur pour qu'il converge sur les contours d'un objet de l'image. Plusieurs algorithmes existent et sont souvent basés sur une minimisation de fonctionnelle énergétique entre l'intérieur et l'extérieur de l'objet. La revue de Fares (2013) montre que le coût algorithmique de ces méthodes est souvent exhaustif. L'avantage de ce concept reste dans la correspondance du contour identifié avec celle de l'objet réel. On pourra citer également d'autres approches moins coûteuses faisant intervenir des filtres de Kalman. Ces approches permettent de traiter les problèmes d'occultation d'objets en temps réel (Peterfreund, 1999) à l'aide de contours actifs, et peuvent avantageusement être basées sur des splines (Orderud *et al.*, 2007).

Certaines applications se servent des contours pour identifier des objets connus à priori et les suivre dans des séquences d'images. En 2D, on peut citer les approches du type *Chamfer Tracking* qui mettent en correspondance les contours d'une image binaire avec l'image considérée (Liu *et al.*, 2010). Une approche orientée vidéo et suivi d'objets est proposée par Choi et Christensen (2012) en utilisant les projections d'un modèle 3D de l'objet dans l'espace 2D de l'image en temps réel. Ces applications utilisent cependant une forte charge de calcul puisque c'est un filtre particulier qui est utilisé pour estimer l'orientation du modèle 3D en fonction de l'image considérée.

Nous venons de détailler différentes manières de segmenter une image en fonction de ses caractéristiques spatiales. Cette séparation couleur/contour pour expliquer les différents concepts n'est pas établie comme telle et de nombreuses méthodes proposent d'allier ces deux paramètres pour détecter ou suivre des objets. La caractérisation spatiale est fortement teintée image et n'utilise pas la notion de cohérence temporelle qui relie les images consécutives d'une vidéo. Le paragraphe suivant détaille les différentes méthodes utilisant cette cohérence temporelle pour identifier des objets.

#### b. Hypothèse d'à priori de mouvement

Un objet dans une séquence d'images peut être différencié du reste de la scène par son comportement cinématique. Cette proposition a permis le développement d'un grand nombre de méthodes basées sur l'hypothèse que la dynamique du fond est significativement plus lente que celle des objets.

Cette hypothèse a permis de définir plusieurs cadres de travail qui diffèrent dans l'approche mais dont les résultats sont similaires : la soustraction de fond (*background subtraction*) et la détection de premier plan (*Foreground detection*). Ces approches ont un fonctionnement eulérien et permettent de transformer des informations eulériennes en entrée en informations lagrangiennes.

#### Détection de premier plan

La manière la plus intuitive de détecter le mouvement consiste à utiliser la différence d'image (ou *Interframe*). Une simple différence d'image pose le problème d'un objet qui est présent deux fois : une pour chaque instant. L'*interframe*  $\Delta I$  est définie de manière centrée en utilisant deux images autour de l'image  $I$  à l'instant  $k$  :

$$\Delta I = |I_k - I_{k-1}| \cdot |I_k - I_{k+1}| \quad (2.16)$$

On peut aisément utiliser l'*interframe* à l'aide d'un seuillage simple ou intégrée dans un contexte probabiliste. Cette méthode est très peu coûteuse en mémoire et en calcul, mais elle ne détectera pas une partie de l'objet si cette dernière reste commune à deux des trois images considérées. Le manque de robustesse de cette méthode la rend difficile à utiliser seule. On notera l'approche originale de Sappa et Dornaika (2006) qui propose une étape d'extraction de contours avant d'utiliser l'*interframe*, minimisant ainsi les pixels communs. Cette étape d'extraction de contours alourdit cependant le calcul. Bugeau et Pérez (2007) proposent d'utiliser l'*interframe* comme détecteur pour restreindre les zones d'intérêt. Elle utilise ensuite des points d'intérêt et des descripteurs<sup>15</sup> pour achever la détection de l'objet. Ces méthodes garantissent d'avoir l'objet en mouvement sans garantir d'avoir uniquement l'objet.

L'approche de la différence d'image effectuée en réalité est une dérivée temporelle pour chaque pixel, c'est-à-dire à un niveau eulérien. L'approche lagrangienne consiste à se servir du flot optique. Le champ de vecteurs obtenu est alors utilisé de différentes manières pour détecter l'objet. Le problème de la détection d'objet est résolu en cherchant les directions majoritaires dans le champ de vecteur et ainsi de segmenter l'image en un fond et plusieurs objets (Tian et Hampapur, 2005)<sup>16</sup>. La robustesse du flot optique peut rendre cependant cette détection moins juste pour de grands mouvements avec des problèmes de détection.

Ces deux premières approches montrent qu'une approche de détection de premier plan seule peut suffire pour identifier les zones de mouvement dans l'image, sans garantir pour autant que l'objet entier sera détecté. On s'aperçoit qu'il serait judicieux de coupler plusieurs types d'approches afin de gagner en robustesse, ce qui a été fait par exemple pour la couleur et le mouvement (Mittal et Paragios, 2004).

#### Soustraction de fond

Détecter le mouvement dans une vidéo peut être interprété comme un filtrage temporel visant à

15. cf. section III.B.2. sur les approches locales.

16. On pourra lire Karlsson et Bigun (2012) à propos d'interaction entre le flot optique et les objets.

séparer les dynamiques rapides des dynamiques plus lentes. Cette idée a conduit à des approches basées sur une modélisation du fond au niveau du pixel. Stauffer et Grimson (1999) propose une modélisation robuste du fond par une mixture de gaussienne (*Gaussian Mixture Model - GMM*). Le fond étant considéré comme variant peu, on peut construire une distribution des valeurs de chaque pixel. Le passage d'un objet de dynamique plus élevée fait varier de manière évanescence un pixel sur des valeurs n'appartenant pas à la distribution. À chaque instant, les nouveaux pixels sont comparés aux distributions spatialement correspondantes : la probabilité d'appartenir à la distribution revient à tester si un pixel est issu du fond ou non. Cette approche est largement répandue et le caractère adaptatif de la méthode lui fait gagner en robustesse. L'algorithme est cependant plus coûteux en mémoire et en calcul.<sup>17</sup>

Une méthode plus récente s'appuie sur une approche orientée filtrage. Manzanera et Richefeu (2007) considèrent la vidéo en utilisant la description "signal" présentée sur la figure 2.2. Identifier un pixel appartenant au fond revient à caractériser l'écart de sa valeur courante avec sa valeur moyenne. Ce calcul est fait lors de la numérisation d'un signal dans un convertisseur analogique numérique (CAN), avec la supposition que le signal varie beaucoup plus lentement que l'horloge du CAN. C'est en faisant l'analogie avec les propriétés stationnaires des pixels du fond que Manzanera et Richefeu reprennent l'algorithme le plus répandu et le plus robuste pour les CAN : le sigma-delta. La méthode est simple, efficace, et peu coûteuse comme le montre son implémentation en temps réel (Lacassagne et Manzanera, 2009) : elle est idéale pour une détection d'objets bas niveau.

La dernière méthode que nous développons est introduite par Barnich et Van Droogenbroeck (2011) : il s'agit de la méthode *ViBe*. Cette approche utilise un contexte de type sac de mots pour décrire les valeurs de fond. L'algorithme utilise pour chaque pixel une collection de pixel représentatifs du fond dans une mémoire tampon. Le test d'appartenance au fond d'un pixel est défini par une fonction de distance dans l'espace des pixels, non pas en considérant la distance du nouveau pixel par rapport à la collection, mais en considérant le nombre de pixels de la collection inférieurs à un rayon fixé. La manière de mettre à jour la collection est très innovante également puisqu'au lieu de considérer une fonction de comparaison entre la collection et les pixels courants, il est choisi un renouvellement aléatoire, chaque pixel de la collection étant testé suivant une loi exponentielle paramétrée par l'instant d'arrivée dans la collection. Cette approche, très facile à implémenter et peu coûteuse, connaît une forte popularité et d'autres auteurs ont permis d'améliorer la méthode. Jenifa *et al.* (2012) renomme la méthode en ajoutant simplement un module de détection d'ombres ou Jiang et Lu (2017) formule des règles de renouvellement de la collection de pixels plus adaptatives pour obtenir des résultats dans des environnements pouvant changer le fond soudainement.

Les hypothèses de dynamique d'évolution des objets de la scène aboutissent à des méthodes très efficaces pour obtenir une détection des objets évoluant sur la scène. Ces méthodes permettent de convertir des informations eulériennes en informations lagrangiennes. Cependant, elles traitent à chaque pas l'ensemble des pixels de l'image : il est légitime de s'interroger sur la pertinence d'inclure dans les calculs la totalité des pixels, étant donné qu'une majorité appartient au fond. Cette réflexion menée notamment dans les travaux de Bugeau (2007) aboutit à considérer une approche locale des caractéristiques de l'image (position, mouvement et photométrie), que nous développons dans le paragraphe suivant.

---

17. On encouragera la lecture de la revue de Jodoin (2010) pour plus de détails et de variétés sur les méthodes de soustraction de fond.

#### III.B.2. Approches locales

L'utilisation de méthodes globales appliquées à la totalité de l'image aboutissent à l'extraction d'informations discriminantes pour identifier des objets. On passe d'informations eulériennes sur des pixels de coordonnées fixes et de couleur variable à des informations lagrangiennes, ce qui revient à grouper des pixels et à considérer cette fois que les pixels ont une couleur fixe et des coordonnées variables. Comme indiqué dans la section précédente, certains auteurs ont obtenu de meilleurs résultats en utilisant une mesure de plusieurs critères. Considérer la globalité de l'image devient alors coûteux en calcul, et pose des problèmes d'artefacts : par exemple lors d'un changement brutal de luminosité dans une large région de l'image, un détecteur basé sur l'ensemble sera biaisé. L'utilisation d'une approche locale permettrait de contourner ces inconvénients. Il serait possible de faire des mesures locales plus variées demandant plus de calculs mais sur des zones réduites, tout en restant à l'abri de biais présents sur des zones sans intérêt pour l'application.

En réalité, l'utilisation d'approches locales est beaucoup plus variée. Nous allons introduire ces concepts dans le paragraphe suivant en séparant deux parties : où chercher ces informations à la fois discriminantes et invariantes et comment les extraire.<sup>18</sup>

##### a. Détecteur de points d'intérêt

Avoir une méthode d'extraction locale n'a d'intérêt que si l'on définit comment trouver des points de l'image où l'on peut trouver des informations pertinentes pour effectuer cette détection de manière **répétable**, **efficace** et **représentative** (au sens statistique). Deux grandes approches existent : une grille fixe sur l'image permettant une approche plus eulérienne où une recherche de points d'intérêt. La problématique de cette recherche de points d'intérêts revient donc à se demander comment définir des zones de l'image qui vont rester invariantes pendant l'évolution de la scène. Ces points peuvent être appelés selon le contexte *keypoints*, *landmark interest points* ou *anchor points*. Moravec (1980) est le premier à considérer la détection de contenu haute fréquence d'une image comme répétable lors du déroulement de la scène. Les contours définissent alors des lignes et ne permettent pas d'assurer une localisation invariante de points caractéristiques. Il s'intéresse ainsi à l'intersection de contours comme points d'intérêts de l'image : des coins. La force de cette approche a été de trouver des points dont la détection est répétable à la fois spatialement et temporellement. Pour isoler ces points, prenons une image de nuances de gris  $I$ . On considère une fonction de fenêtrage binaire définie par  $w(u, v)$  pour  $u$  et  $v$  les coordonnées des pixels d'une imagerie. On calcule la somme des différences du carré de l'imagerie  $I_m$  avec sa translatée de  $(x, y)$  :

$$S(x, y) = \sum_{(u,v) \in I_m} w(u, v) (I(u+x, v+y) - I(u, v))^2 \quad (2.17)$$

L'approche initiale avait quelques inconvénients :

- Le détecteur restait très bruité et sensible aux pixels isolés.
- Le déplacement de la fenêtre se faisait pixel par pixel en diagonale comme en horizontal et vertical.
- La fonction  $S$  ne considérant qu'un seul pixel, le détecteur restait sensible aux contours.

---

18. Cette partie sur les différents suiveurs (ou *trackers*) utilisés en traitement d'images et vidéo s'est appuyée sur plusieurs ouvrages : Krig (2014, 2016) dont la lecture est recommandée à un lecteur voulant approfondir ces notions.

C'est en perfectionnant cette méthode que Harris et Stephens (1988) donnent naissance au célèbre détecteur de coins éponyme. Ils utilisent un développement de Taylor au premier ordre en introduisant le gradient selon la direction horizontale  $I_x$  et verticale  $I_y$ . Ils utilisent alors une mise sous forme matricielle :

$$I(u+x, v+y) = I(u, v) + x \cdot I_x(u, v) + y \cdot I_y(u, v) + o(x^2, y^2) \quad (2.18)$$

$$\text{Soit } S(x, y) \simeq \sum_{(u,v) \in I_m} w(u, v) [x \cdot I_x(u, v) + y \cdot I_y(u, v)]^2 \quad (2.19)$$

$$\iff S(x, y) \simeq \begin{pmatrix} x & y \end{pmatrix} \cdot M \cdot \begin{pmatrix} x \\ y \end{pmatrix} \quad (2.20)$$

$$\text{avec } M = \sum_{(u,v) \in I_m} w(u, v) \begin{pmatrix} I_x^2 & I_x \cdot I_y \\ I_x \cdot I_y & I_y^2 \end{pmatrix} \quad (2.21)$$

La généralisation des calculs de gradient sur la totalité de l'imagette améliore la précision. De plus, l'utilisation d'un fenêtrage gaussien limite l'influence des pixels isolés<sup>19</sup>.

Le problème revient alors à un problème de recherche de valeurs propres  $(\lambda_1, \lambda_2)$ . Harris et Stephens préfèrent alors limiter la charge de calcul en approximant cette recherche par une minimisation de la fonction  $\det M - \mu \cdot \text{Tr } M^2 = \lambda_1 \cdot \lambda_2 - \mu \cdot (\lambda_1 + \lambda_2)$  avec un paramètre  $\mu$  de réglage. Shi et Tomasi (1994) développent plus tard une version exacte du calcul de minimum ci-dessus, dans une version du détecteur de coin plus précise. Le tracker Kanade-Lucas-Tomasi (KLT) que nous utiliserons dans le chapitre 4 est issu notamment de cet article.

Nous remarquerons sur cet exemple que les points d'intérêt détectés sont invariants face à des changements d'illumination mais pas invariants pour des changements d'échelles. Cette remarque nous amène à la raison du développement de la grande famille des détecteurs de points d'intérêts. Le détecteur Harris-Laplacien propose une approche multi-échelle pour pallier ce problème (Mikolajczyk et Schmid, 2001). Les détecteurs FAST (?) testent l'écart d'intensité entre un point et ses voisins sur un cercle de périmètre donné. La décision est alors prise non pas sur l'écart mais sur le nombre de points du cercle dépassant un seuil. Les détecteurs ORB pour *Oriented FAST and Rotated BRIEF* (Rublee *et al.*, 2011) utilisent un détecteur FAST et trient les points en utilisant le score de Harris. Les meilleurs coins sont alors sélectionnés et l'orientation du point d'intérêt est calculée en utilisant les moments statistiques 2D (cf. section III.C.). De même des améliorations sont proposées en changeant la fonction de dissimilarité utilisée par Moravec (une somme de différences au carré) ou en inversant l'approche : en utilisant des fonctions de similarité (comme un opérateur de corrélation croisée) (Arnaud, 2015).

De manière plus générale, les points d'intérêt sont donc basés sur l'extraction de plusieurs types de caractéristiques :

- des coins de type rampe ;
- des coins de type toit ;
- des coins de type échelon ;
- des contours ;
- des maxima de squelettes (basés sur la transformée en distance).

19. appelés également *salt and pepper noise* ou bruit poivre et sel

### III. Caractériser un objet dans une séquence d'images : un problème mal posé en 2D

- des maxima de région.

La plupart des points d'intérêt s'appuient sur les concepts de :

- norme du gradient ;
- direction du gradient ;
- orientation du laplacien ;
- matrice hessienne.

Sont ensuite utilisés des outils comme le déterminant, la trace, l'analyse en composantes principales ou la décomposition en valeurs propres pour extraire les informations pertinentes.

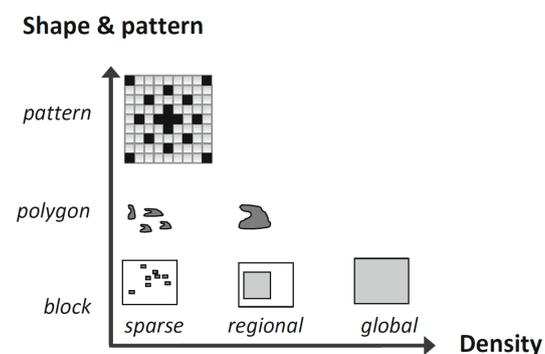
Ces détecteurs sont conçus pour trouver des points **répétables** et **représentatifs** dans une séquence d'images. Les scores de sélection de ces points peuvent permettre de les discriminer mais selon les scènes les résultats ne sont pas toujours suffisant. Le calcul de traits via des descripteurs numériques apporte alors une information supplémentaire.

#### b. Descripteurs de traits

Les descripteurs font intervenir d'autres moyens d'analyser l'image. Avec une fenêtre d'analyse plus grande et une collecte d'informations de différents types, ils permettent de caractériser l'environnement proche du point d'intérêt. Les descripteurs se différencient sur trois critères :

##### Forme et schéma

Les descripteurs agissent sur une région proche d'un point d'intérêt. Chaque descripteur définit alors un schéma et une forme particulière continue (polygone, bloc) ou discrète (schéma). Quelques exemples de schémas sont donnés figure 2.10. Ils peuvent être très complexes pour permettre de capturer au mieux les informations comme le descripteur *FREAK* (pour *Fast REtinA Keypoint*) qui s'inspire du système de vision humain pour définir son schéma.



**Figure 2.10** – Différentes manières d'analyser l'image, d'un point de vue de la densité et d'un point de vue de localisation et concentration des informations. Figure extraite de Krig (2014) page 192.

##### Densité

La densité des descripteurs varie en fonction des schémas et des méthodes. L'analyse du descripteur peut être faite sur des entités disjointes : on parle d'analyse creuse. Elle peut également se faire sur une région de l'entité ou encore sur la globalité de celle-ci. La densité caractérise la quantité d'informations traitées par le descripteur.

##### Traits

C'est la catégorie la plus vaste et nous ne serons pas exhaustifs. Les traits utilisent ainsi différents types de données qui peuvent être des scalaires, des vecteurs, des histogrammes,

ou des espaces multi-variables. Ils peuvent caractériser le contraste, la couleur à l'aide des niveaux d'intensité, la direction ou la courbure en quantifiant le gradient.

À la différence des détecteurs dont les objectifs sont d'être **répétables**, **efficaces** et **représentatifs**, les descripteurs doivent être plutôt **invariants** et **discriminants**. Le caractère discriminant des descripteurs est essentiel pour garantir le caractère fonctionnel du descripteur, ce qui peut avoir un coût de calcul non négligeable. Il est également intéressant de mentionner une approche par filtrage plus connue sous le nom de **cascade** (Viola et Jones, 2001). Ces schémas sont constitués d'une série de filtres, d'une complexité croissante : chaque filtre est à la fois plus discriminant et plus coûteux que le précédent. Cette approche permet de réduire la quantité de données à traiter au fur et à mesure que la complexité du filtre augmente et donc de gagner en efficacité. Les filtres en cascade sont paramétrés pour détecter un objet spécifique. Ils demandent une phase d'apprentissage qui peut être longue servant au réglage. Cette structure en série permet de traiter directement des images sans extraire de points d'intérêt.

Nous remarquerons que certains descripteurs de points comprennent également leur propre détecteur de points d'intérêt. C'est le cas des méthodes *SIFT*, *ORB*, *SURF*, ou *BRISK*. Ces traits ne peuvent être associés immédiatement pour faire de l'apprentissage puisqu'ils ne sont pas exprimés dans des espaces de même nature. On utilise alors des *Bags of visual Words* (Sacs de mots visuels ou *BoW*).

Les exemples traitant des fonctionnements et performances des descripteurs classiques comme *SIFT* et *SURF* ont été abondamment traités dans la littérature (Oyallon et Rabin, 2015). Aussi nous

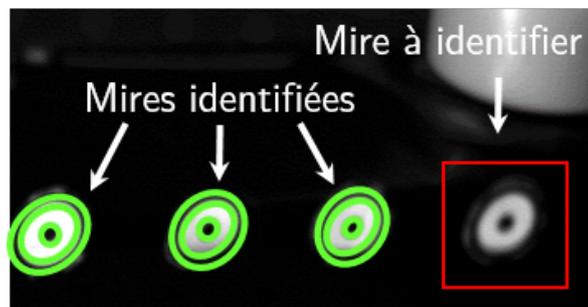


Figure 2.11 – Exemple de détection MSER sur les mires qui seront utilisées dans le chapitre 4

préférons donner comme exemple un descripteur ayant une approche par polygone : le descripteur *MSER* (pour *Maximally Stable Extremal Regions*). Ce descripteur est proposé par Matas *et al.* (2004) pour identifier des *blobs* (taches). Il est invariant aux déformations, aux changements d'illumination et au niveau du flou. Le fonctionnement de MSER est basé sur une double approche :

- repérer les régions restant stables pour différentes valeurs de seuillages ;
- calculer les moments de Hu (cf. section III.C.) afin de caractériser géométriquement ces régions.

Le résultat donnera une série d'ellipses dont les caractéristiques pourront être interprétées. Nous utiliserons ce détecteur dans la sous-section IV.A. du chapitre 4 pour identifier des mires dans une scène (figure 2.11).

### c. Comparaison de traits

La détection et le calcul de traits d'une image a un sens statique dans le cas de la reconnaissance

### III. Caractériser un objet dans une séquence d'images : un problème mal posé en 2D

d'objet. Par contre, ils n'offrent pas un appareillage des points entre les images consécutives, élément de notre cahier des charges. Il faut alors procéder à une troisième étape de mise en correspondance des traits entre deux images. Cette étape n'est pas simple puisqu'elle doit appareiller les points deux à deux alors que les ensembles de départ et d'arrivée ont pu varier. Ce sont des fonctions de distances qui permettent selon le type de données entrant de comparer les éléments d'une collection. Ces éléments peuvent ainsi être des points, des vecteurs, des vecteurs binaires ou des histogrammes multi-canaux. Classiquement les fonctions de distance suivantes sont utilisées Krig (2014) :

- distance euclidienne ou carré de la distance euclidienne permettant de comparer deux couples de points ;
- la distance cosinus permettant d'obtenir l'angle entre deux vecteurs ;
- la norme L1 qui s'exprime par la somme de la norme des différences ;
- la norme L2 déjà vue pour le détecteur de Harris-Stephens ;
- la distance de Manhattan qui mesure la distance entre deux points en suivant les mailles d'une grille ;
- la distance de Chebyshev permettant de mesurer la distance maximum entre deux vecteurs sur les cases d'une grille, rendant ainsi un triangle isocèle en rectangle équilatéral ;
- la distance de Mahalanobis qui possède une propriété d'invariance d'échelle intéressante et s'applique sur des variables aléatoires supposées gaussiennes. Ainsi pour calculer la distance entre un vecteur aléatoire  $x = (x_1, \dots, x_n)^T$  avec un vecteur de valeur moyenne  $m = (m_1, \dots, m_n)^T$  et de covariance  $\Sigma$ , on a :

$$D_M(x) = \sqrt{(x - m)^T \cdot \Sigma^{-1} \cdot (x - m)} \quad (2.22)$$

Des fonctions de distances sont également utilisées pour les histogrammes. On définit ici la distance de Bhattacharyya utilisée par la suite pour comparer deux histogrammes de couleur. On a pour deux distributions de probabilité  $p$  et  $q$  :

$$D_B(p, q) = -\ln(BC(p, q)) \quad (2.23)$$

$$\text{avec le coefficient de Bhattacharyya } BC(p, q) = \sum_x \sqrt{p(x) \cdot q(x)} \quad (2.24)$$

Ces fonctions distances permettent d'appareiller les points. Il faut ensuite les raffiner en un sous-groupe (ou *cluster*) ayant des caractéristiques similaires. Le domaine de la vision utilise alors des techniques de filtrage robuste. Cette famille de méthodes fait appel à des critères de distance différents pour modéliser les données. Les plus utilisées pour les descripteurs sont le RANSAC, la méthode des K plus proches voisins (*KNN Clustering*), les M-estimateurs, mais aussi les approches par apprentissage. Ces méthodes sont décrites plus en détail dans (Krig, 2014) mais on pourra également se référer à Malis et Marchand (2005) pour une approche non exhaustive et pédagogique.

#### d. Remarque sur la dimension des détecteurs et descripteurs

Une interrogation pourrait se poser dans la description faite des approches locales. En effet, ces approches sont directement issues de techniques de traitement d'images, par exemple initialement pour faire de la reconstruction 3D ou de la structure à partir de mouvement ou encore de la création d'images panoramiques. Nous travaillons donc dans un espace en deux dimensions et le lien avec le temps se fait à l'aide de mise en correspondance.

Pourquoi ne pas utiliser directement des descripteurs en trois dimensions ? Cette approche existe et utilise les mêmes concepts précédemment évoqués étendus sur les deux dimensions de l'image et sur le temps. Ainsi, une recherche de coins de Harris peut être faite en 3D. Cette approche très populaire en reconnaissance de gestes ou d'actions reste cependant éloignée de nos problématiques. En effet, bien que nous recherchions des invariants sur un objet d'intérêt, nous cherchons justement à caractériser ses mouvements. Trouver ses invariants de mouvement pourrait être une autre approche qui ne sera pas développée ici.<sup>20</sup>

### III.B.3. Conclusion

La détection ou le suivi d'objets dans les vidéos 2D est un problème mal posé. Si de nombreuses méthodes matérielles ou logicielles proposent d'augmenter les données vidéos afin de lever l'indétermination sur la profondeur, nous avons détaillé ici le cas de vidéos 2D. Deux différentes approches existent.

- Les **approches locales** font appel à des concepts **différentiels** et **géométriques** de l'image. Elles sont structurées en deux temps : la recherche de **points d'intérêt** à la fois **répétables**, **efficaces** et **représentatifs**, puis leur caractérisation en calculant des traits **invariants** et **discriminants** autour de ces points. Ces approches permettent de caractériser un objet et de le suivre dans une scène de manière robuste. Deux inconvénients existent pour cette approche : elle n'assure pas l'extraction exacte de l'objet réel dans la vidéo et ces points d'intérêts peuvent être perdus lors d'occlusions.
- Les approches de détection **globales** sont plus orientées sur une approche **statistique**. Différentes manières de caractériser l'objet ont été présentées : statiques comme pour la forme et le contenu colorimétrique et dynamiques avec les familles de soustraction de fond. Ces méthodes sont quant à elles capables d'extraire un objet vidéo correspondant à l'objet réel mais requièrent des hypothèses supplémentaires sur le scénario de la vidéo.

Ces deux approches permettent de passer d'informations pixéliques eulériennes à des informations lagrangiennes : sous forme de points discrets ou sous la forme de groupes de pixels. L'utilisation de ces techniques doit alors être couplée à des méthodes permettant de mesurer des informations sur les groupes de pixels ou des collections de descripteurs pour caractériser les objets de la séquence d'images.

### III.C. Extraire et mesurer

La détection nous fournit des collections de groupes de pixels ou des collections de points appareillés. Un masque de  $n$  groupes de pixels est labellisé pour pouvoir extraire  $I_{b_n}$  masques binaires comportant chacun un objet. L'image peut ainsi être interprétée comme une somme de couches indépendantes et de leur résidu (la partie restante de l'image segmentée). Dans cette partie quelques méthodes permettant d'extraire des mesures de l'image sont décrites.

#### La méthode des moments statistiques de Hu

Hu (1962) propose une transposition des moments statistiques dans les espaces 2D. Il en déduit des propriétés lorsque ces moments statistiques sont appliqués à un masque binaire  $I_b$ . On rappelle l'expression générale d'un moment  $M_{ij}$  d'ordre  $i+j$  et de son moment central

20. On pourra commencer par lire Shabani *et al.* (2012) et approfondir avec Laptev (2004).

### III. Caractériser un objet dans une séquence d'images : un problème mal posé en 2D

normalisé  $m_{ij}$  :

$$M_{ij} = \sum_{x \in I_b} \sum_{y \in I_b} x^i \cdot y^j \cdot I_b(x, y) \quad (2.25)$$

$$m_{ij} = \sum_{x \in I_b} \sum_{y \in I_b} (x - \bar{x})^i \cdot (y - \bar{y})^j \cdot I_b(x, y) \quad (2.26)$$

$$(2.27)$$

$$\text{avec} \begin{cases} \bar{x} = \frac{M_{10}}{M_{00}} \\ \bar{y} = \frac{M_{01}}{M_{00}} \end{cases} \quad (2.28)$$

Le moment d'ordre 0 correspond à l'aire de l'image tandis que  $(\bar{x}, \bar{y})$  correspond au barycentre de l'image, que l'on considère comme le centre de l'objet vidéo. Les moments d'ordre plus élevés nous permettent d'obtenir d'autres informations sur l'objet. Le moment central du second ordre est défini tel que :

$$m'_{ij} = \frac{m_{ij}}{m_{00}} \quad (2.29)$$

et la matrice de covariance de l'image  $I(x, y)$  est définie par :

$$\text{cov} [I(x, y)] = \begin{pmatrix} m'_{20} & m'_{11} \\ m'_{11} & m'_{02} \end{pmatrix} \quad (2.30)$$

Cette matrice nous donne alors ses axes principaux en calculant ses valeurs propres, mais également son orientation que l'on calcule par :

$$\theta = \frac{1}{2} \arctan \left( \frac{2m'_{11}}{m'_{20} - m'_{02}} \right) \quad (2.31)$$

Avec ces moments statistiques 2D, Hu définit le cadre d'un outil très puissant et rapide puisqu'il permet de mesurer l'objet en position, en aire, de calculer son centre apparent et de l'inclure dans une ellipse dont on extrait les demi-axes, l'excentricité et l'orientation.

#### La transformée de Hough

En 1962 également, Hough introduit une transformée utilisée pour reconnaître des droites. Cette transformée ne va être utilisée qu'après une généralisation à d'autres formes développée par Duda et Hart (1972). Elle permet par une méthode coûteuse d'énumération et de test non seulement de détecter des lignes, mais également des cercles ellipsoïdaux.

L'idée générale est de travailler dans l'espace paramétrique  $(\rho, \theta)$ . Or, une droite passant par un point dans un plan image de coordonnées  $(x, y)$  peut s'exprimer par  $\rho = x \cdot \cos(\theta) + y \cdot \sin(\theta)$  : une droite passant par un point s'exprime sous la forme d'un point dans l'espace paramétrique  $(\rho, \theta)$ . L'algorithme de Hough propose ainsi d'extraire tous les contours de l'image, puis, pour chaque point du contour, de venir prendre toutes les droites possible, ce qui va donner une courbe dans l'espace paramétrique. En remarquant simplement qu'une droite est composée de plusieurs points, il suffit ainsi de repérer les intersection de courbes dans l'espace de Hough pour en déduire le paramétrage d'une droite de l'image.

La transformée de Hough est également très populaire pour détecter des cercles et des ellipsoïdes. Le paramétrage d'un cercle souffre d'un problème puisqu'il comporte trois composantes : deux pour son centre et une pour le rayon. Les calculs précédents doivent se faire dans

un espace 3D ce qui augmente encore le temps de calcul. D'autres approches existent, comme par exemple limiter l'espace des paramètres ou encore travailler sur le gradient du contour. En effet le gradient de chaque point du périmètre du cercle va pointer vers son centre. Le problème peut donc être séparé en deux étapes : une étape de recherche du centre, puis une étape de recherche du rayon, qui pourra être faite en maximisant l'histogramme des distances.

### Les constellations

Les méthodes précédentes sont très utilisées au niveau global de l'image : ce sont en effet des méthodes statistiques. Pour terminer cette section sur la mesure d'information, nous proposons d'introduire une méthode utilisée cette fois plutôt pour des approches locales : les méthodes de constellation. Ces méthodes sont utilisées en reconnaissance d'images et on pourra se reporter à la thèse de Erus (2008) qui décrit les différents modèles utilisés. Étant donné que nous n'utiliserons pas ces méthodes, nous n'irons pas plus loin dans la présentation.

### Les méthodes de mesure de champ

Il est difficile de ne pas parler des méthodes de mesures de champs dans un contexte de mesure, particulièrement les méthodes de corrélation croisée (ou *Digital Image Correlation - DIC*). Cependant ces méthodes sont très coûteuses pour de larges mouvements et ne correspondent pas à la diversité des scènes étudiées. Il faut cependant souligner que les mesures effectuées par corrélation croisée peuvent atteindre le centième de pixel pour des scènes suffisamment bien préparées.

Deux approches existent dans la littérature : une approche locale et une approche globale. Les deux approches utilisent l'opérateur de corrélation croisée déjà défini dans l'équation (2.13).

— L'approche locale est la plus largement employée. Elle consiste à simplifier le problème en travaillant sur des imagerie de taille réduite. Le problème revient alors à tester pour quel déplacement l'imagerie maximise la fonction de corrélation croisée par rapport à une imagerie de référence. Les codes de corrélation actuels cherchent également la déformation et la rotation pour chaque imagerie.

— Une approche globale existe également. Elle consiste à ajouter des hypothèses cinématiques définies par l'utilisateur selon la scène ou l'objet à étudier. Le résultat des comparaisons d'images (déplacement, orientation, déformation) sont alors projetées sur un maillage type élément fini qui permet d'homogénéiser et de lisser les déplacements obtenus sur les images locales. Ces approches sont plus précises encore que des approches locales.

Ces approches nous éloignant de notre objectif temps réel, nous n'approfondiront pas plus<sup>21</sup>. Des études récentes montrent cependant que l'utilisation du flot optique pourrait alléger le calcul de corrélation croisée (Liu *et al.*, 2015) tout en conservant des niveaux de performance similaires.

## III.D. Conclusion et remarque sur une simplification par le passage de la 2D à la 3D

Cette section résume les différentes méthodes permettant de détecter des objets lagrangiens puis d'en extraire des caractéristiques à partir d'une séquence d'images eulériennes. La détection ou le

21. Le lecteur pourra approfondir en consultant l'article de Hild et Roux (2012) qui traite des différentes approches utilisées dans ce domaine.

#### IV. Une nouvelle méthode pour l'interpolation et la compression à l'aide de la composition cinématique du mouvement

suivi se sépare sous deux approches : une **approche locale** plus orientée sur des méthodes **différentielles** et **géométriques** et une **approche globale** faisant appel à des **méthodes statistiques** et exigeant des hypothèses sur le scénario de la vidéo. Cette première couche permet de transformer l'information pixélique eulérienne en information lagrangienne qui est utilisée dans une deuxième couche d'extraction d'informations.

Cette partie propose un développement non exhaustif des différentes méthodes permettant d'extraire et de caractériser un ou plusieurs objets à partir d'une image. Les caractéristiques extraites portent sur la photométrie ou sur le mouvement et permettent d'en déduire les positions, orientations, vitesses, contour et parfois la déformation de l'objet que l'on appellera par la suite **objet vidéo**.

Avant de passer à l'introduction de la méthode que nous avons développée, nous proposons un instant de nous abstraire du cadre de travail 2D auquel nous nous étions attaché. Il s'avère que ce cadre est pratique pour tester des concepts et également pour comprendre l'articulation de certains algorithmes en réduisant l'espace de travail à deux dimensions. Cependant, comme nous l'indiquons dans la partie III.A., le problème de suivi d'objet est un problème mal posé en deux dimensions. Aussi, l'ensemble des méthodes développées précédemment devient rapidement complexe.

Travailler directement sur trois dimensions en utilisant les méthodes logicielles ou matérielles nous permet de définir l'objet d'une manière plus simple. L'ensemble des approches de mesures sur des images ou des séquences d'images que sont la **photogrammétrie** ou la **vidéogrammétrie** utilisent une représentation 3D de la scène pour s'abstraire des erreurs de parallaxe et permettre une mesure la plus précise possible. Ce type d'approche comporte néanmoins des étapes de mise en correspondance ou de fusion qui offrent d'autres problématiques. Elle occasionne une charge de calcul plus importante due à une plus grande quantité de données à traiter. L'évolution technologique et les besoins du marché poussent à considérer la 3D comme l'outil de demain. C'est pourquoi nous avons tenu à concevoir une chaîne de mesure isolant la partie extraction de l'objet des autres, de manière à pouvoir facilement faire évoluer l'algorithme dans un cadre tridimensionnel.

#### IV Une nouvelle méthode pour l'interpolation et la compression à l'aide de la composition cinématique du mouvement

Les approches proposées par la philosophie MPEG et le traitement vidéo se sont fortement orientées sur l'importance de caractériser l'objet. La projection du monde réel sur le capteur pose problème pour caractériser l'évolution de l'objet dans la scène. L'impossibilité de décrire le comportement d'un objet dans une scène inconnue à partir des données vidéos rend l'ensemble des objectifs du projet FAMOSYS inatteignables<sup>22</sup>. Cependant, il nous a semblé pertinent de se focaliser sur une résolution partielle de ces problématiques dans un cadre de traitement vidéo avec pour avantage non-négligeable de pouvoir appliquer les concepts sur un large spectre de vidéos rapides. Nous proposons dans cette partie de développer une méthode basée sur une équivalence entre objet mécanique et objet vidéo afin d'étendre l'interpolation petit mouvement si efficace à des mouvements plus larges. Cette méthode nous permettra ainsi de faire de l'interpolation, une prédiction limitée et une compression sans perte.

---

22. Objectifs qui sont interpolation, extrapolation, identification, compression et réalité augmentée.

#### IV.A. Lien entre l'approche eulérienne et lagrangienne

Bien que la notion ait été évoquée dans les sections précédentes, il est préférable de rappeler les différences sous-jacentes entre objet mécanique et objet vidéo.

##### Objet vidéo

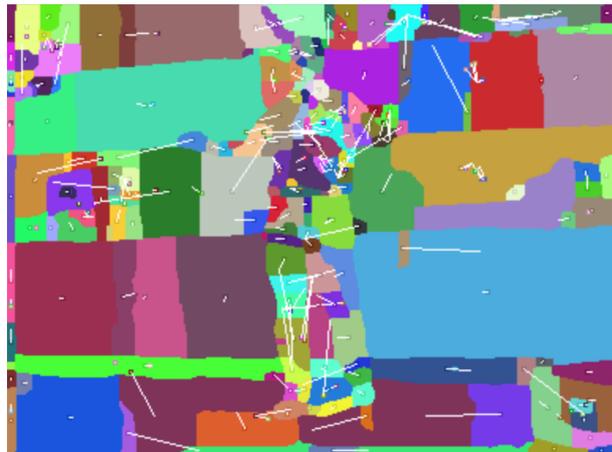
Comme on peut le constater dans la littérature, un objet vidéo est défini comme un ensemble de pixels regroupés selon des caractéristiques spatio-temporelles communes. Les objets vidéos sont à la frontière entre une représentation eulérienne et lagrangienne : ils peuvent être paramétrés selon les deux approches. La notion de groupe de pixels fait intervenir la notion eulérienne définie en début de chapitre puisque les pixels vont tour à tour appartenir ou non à l'objet selon son évolution dans la scène. Il est important de souligner qu'il n'y a aucune raison pour qu'un objet vidéo ne corresponde à un objet réel (la figure 2.12 montre la segmentation obtenue par Brox *et al.* (2009)).

##### Objet mécanique

L'objet mécanique auquel nous faisons allusion est celui considéré en mécanique du solide : un objet dont l'étude du mouvement peut être réduite à un point matériel et un repère. Cet objet est intrinsèquement 3D sauf si le problème est ramené à un cas particulier 2D. Par contre le mouvement de l'objet 3D peut être projeté en 2D. Le paramétrage de ces objets fait intervenir a minima un vecteur de position et de rotation sinon le torseur des vitesses associé dans une vision plus cinématique.

Faire correspondre ces deux concepts peut paraître délicat. On peut mesurer la différence entre ces deux aspects à travers l'exemple d'une segmentation d'une scène donnée figure 2.12 : les objets vidéos (taches de couleur) ne correspondent pas à l'objet réel de la scène (un homme sur un fond fixe). On s'aperçoit que cette segmentation suffit pour séparer la scène en plusieurs objets et pour les traiter. Pourtant ces objets n'ont pas de sens mécanique : un objet mécanique peut correspondre à plusieurs objets vidéos et un objet vidéo pourrait englober plusieurs objets mécaniques, et ce à différents instants temporels d'une même scène.

La projection de la scène complexifie la trajectoire des différents objets ce qui nous contraint à la mise en place de modèles élémentaires. Cependant, notre cadre d'étude nous plaçant dans des vidéos à fond quasi-fixe, il est possible d'isoler un objet vidéo, et de décrire son évolution temporelle d'une manière mécanique à travers l'extraction de sa trajectoire, sa vitesse et sa rotation en le considérant comme 2D : la projection de l'objet est alors négligée. L'idée de cette méthode repose sur la possibilité de faire correspondre un objet vidéo avec un objet mécanique ; il serait alors possible d'effectuer une composition entre les grands mouvements qui peuvent être définis dans le cadre de l'approche lagrangienne de la mécanique et les petits mouvements décrits par une approche eulérienne d'estimation de mouvement. Cette approche eulérienne peut également décrire les mou-



**Figure 2.12** – Figure présentant la segmentation automatique d'une scène d'un homme sur fond fixe. Chaque zone colorée représente un objet qui ne correspond pas à un objet réel : c'est ce type d'objet que nous appellerons objet vidéo.

#### IV. Une nouvelle méthode pour l'interpolation et la compression à l'aide de la composition cinématique du mouvement

vements de pixels causés par la projection de l'objet réel qui engendre des déformations du contour et des variations internes. Cette approximation est envisageable dans le cadre d'objets aux contours lentement déformables.

De manière plus formelle, il s'agit de considérer un solide dans une scène décrit par la position de son centre de gravité  $G$  qui a pour coordonnées  $X(t)$  et un repère qui définit une rotation  $\Omega(t)$ . Ce solide est constitué d'un ensemble de points  $M_i$  tel que :

$$S = \{M_i(t)\}$$

On définit alors un opérateur de projection sur le plan du capteur  $P_{\text{réel} \rightarrow \text{capteur}}$  tel que :

$$\left\{ \begin{pmatrix} u(t) & 0 \\ v(t) & 0 \\ 0 & \alpha(t) \end{pmatrix} \right\} = P_{\text{réel} \rightarrow \text{capteur}} \left( \{ \vec{X}(t), \vec{\Omega}(t) \} \right)$$

Le solide projeté sur le plan capteur sera ainsi défini par  $G^p$  le centre de gravité géométrique de sa projection de coordonnées dans le plan capteur  $(u(t), v(t))$ . Le solide projeté aura une orientation  $\alpha(t)$  et on définit son contour  $C^p(t)$ ; il est donc défini par un ensemble de points projetés  $M_i^p$  tel que :

$$S^p = \{M_i^p(t)\}$$

Nous considérons l'objet comme plein, ce qui nous permet de paramétrer le groupe de points  $M_i^p(t)$  comme l'ensemble des points intérieurs du contour  $C^p(t)$ .

On définit maintenant un objet vidéo  $S'$  comme un groupe de  $n$  pixels  $M_i'(t)$  évoluant dans le temps correspondant à l'endroit du capteur où est projeté  $S^p$ . L'objet est défini à un instant donné par :

$$S' = \{M_i'(t)\}$$

On propose une correspondance entre  $S^p$  et  $S'$  sous certaines hypothèses :

- On définit un centre de gravité apparent  $G'(t) = \frac{1}{n} \sum_i^n M_i'(t)$ .
- On va supposer que ce solide est indéformable localement, ce qui peut se traduire par :

$$\forall \epsilon \ll 1, \quad \frac{d\overrightarrow{G^p M^p}}{dt} \ll \epsilon \quad (2.32)$$

Le cadre de la vidéo rapide nous permet de justifier cette supposition et nous assure un déplacement du centre de gravité apparent négligeable.

- On définit la correspondance uniquement valable pour les points appartenant au contour de  $S_p$  qui correspond à celui de  $S'$ . Cette restriction nous permet de séparer le grand mouvement du contour tout en envisageant des variations interne à celui-ci.
- On suppose les conditions d'éclairages et la scène filmée de manière « réfléchie » pour éviter les effets classiques de non concordance entre champs de mouvement vidéos et mouvements réels<sup>23</sup>.

23. Comme le problème d'ouverture en flot optique.

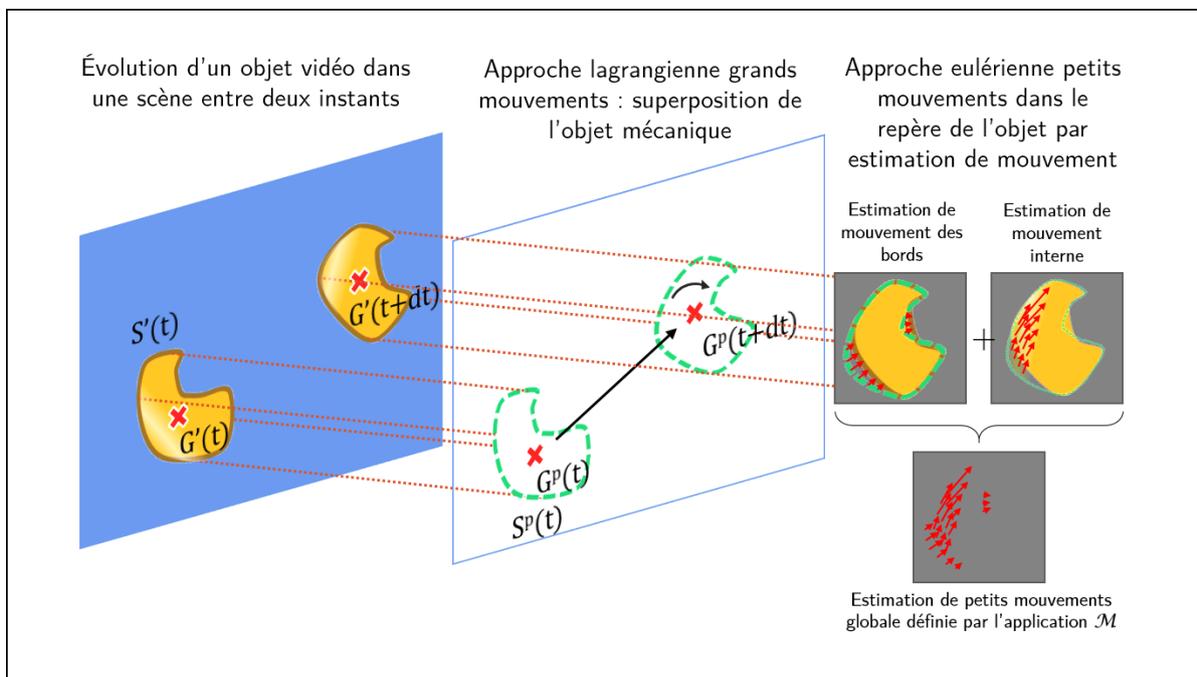
Sous ces hypothèses, on peut alors poser l'évolution temporelle de l'objet vidéo par l'évolution disjointe de son contour  $C'(t)$  et des pixels internes contenus dans son complémentaire :

$$S' = C'(t) \cup C'^C(t) \quad (2.33)$$

$$S'(t + dt) = C'(t + dt) \cup C'^C(t + dt) \quad (2.34)$$

$$S'(t + dt) = [C_p(t + dt) + E_\delta(t)] \cup C'^C(t + dt) \quad (2.35)$$

L'écriture ci-dessus développe simplement que l'objet se sépare en un contour et son complémentaire par rapport à l'objet et que le contour de l'objet vidéo correspond à celui de l'objet mécanique projeté. Nous avons modélisé l'approximation par l'apparition d'un résidu représenté par l'ensemble des résidus des différents points  $E_\delta(t)$  pour le contour.  $E_\delta(t)$  sera un ensemble de vecteurs d'erreurs. Il est possible d'appliquer les équations de la cinématique définies pour un solide uniquement pour le contour de cet objet, ce que nous allons voir dans la partie qui va suivre. Afin de préciser les notations, la figure 2.13 résume la méthode décrite dans ce paragraphe.



**Figure 2.13** – Schéma résumant notre méthode de décomposition de mouvement permettant l'extension de l'estimation de mouvement pour de grands déplacements

#### IV.B. Utilisation de l'interpolation petite échelle pour des grands mouvements

Lorsque l'on travaille au niveau du pixel, il est possible de se placer dans le cadre offert par l'estimation de mouvements dense. Pour une image donnée, l'estimation de mouvements propose de trouver un champ de mouvements optimal associé à chaque pixel de l'image afin de pouvoir reconstituer l'image suivante.

En reprenant les notations du précédent paragraphe, il est possible ainsi de définir l'évolution temporelle d'un groupe de pixel entre deux instants. En notant  $\mathcal{R}_O$  le repère de l'image de centre  $O$ , on

#### IV. Une nouvelle méthode pour l'interpolation et la compression à l'aide de la composition cinématique du mouvement

a alors pour tout  $M'_i \in S'$  :

$$\overrightarrow{OM'_i}(t + dt) = \mathcal{M}_i \left( \overrightarrow{OM'_i}(t) \right) + E_{\mathcal{M}}(t) \quad (2.36)$$

En définissant le champ de vecteur par l'application  $\mathcal{M}$  d'estimation de mouvements avec  $E_{\mathcal{M}}(t)$  le résidu associé :

$$\mathcal{M} : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} \mathcal{M}_1(x_1, \dots, x_n) \\ \vdots \\ \mathcal{M}_n(x_1, \dots, x_n) \end{pmatrix} \quad (2.37)$$

Afin de faire apparaître cette approximation petit mouvement dans le mouvement général de l'objet, on propose de reprendre l'équation (2.33) appliquée à un point  $M_i^p \in C^p$  et  $\overrightarrow{\delta}_i(t) \in E_{\delta}(t)$  en considérant le repère local de l'objet  $\mathcal{R}_G$  de centre  $G$ . Le centre  $G^p$  étant supposé fixe entre deux instants, on applique la formule de transport des vitesses :

$$\overrightarrow{OM_i^p}(t + dt) = \overrightarrow{OM_i^p}(t) + \overrightarrow{V_{M_i^p | \mathcal{R}_O}}(t) \cdot dt \quad (2.38)$$

avec :

$$\overrightarrow{V_{M_i^p | \mathcal{R}_O}}(t) = \overrightarrow{V_{G | \mathcal{R}_O}}(t) + \overrightarrow{\alpha_{\mathcal{R}_G / \mathcal{R}_O}}(t) \wedge \overrightarrow{GM_i^p}(t) \quad (2.39)$$

Les mouvements étant trop larges, on ne peut pas appliquer l'estimation de mouvement sur  $\overrightarrow{OM'_i}$  qui pourrait causer de sévères distorsions. Par contre, le repère local  $\mathcal{R}_G$  s'y prête bien. En travaillant dans ce repère, on propose de faire deux approximations sur l'objet vidéo en différenciant les points appartenant au contour de l'objet des points internes à l'objet :

- Le mouvement des points internes peut être approximé avec l'application d'estimation de mouvement définie plus haut. On considère  $M_i^p$  le point translaté selon l'équation précédente,  $M'_i$  le point de l'objet vidéo, et le résidu  $\overrightarrow{\delta}_i^{int}(t) \in E_{\mathcal{M}}(t)$  :

$$\forall (M_i^p, M'_i) \in C^{pC} \times C'^C, \quad \overrightarrow{GM'_i}(t + dt) = \overrightarrow{GM_i^p}(t) + \mathcal{M} \left( \overrightarrow{M_i^p M'_i}(t) \right) + \overrightarrow{\delta}_i^{int}(t) \quad (2.40)$$

- Les variations lentes du contour sont contenues dans le résidu  $E_{\delta}(t)$ . L'estimation de mouvement nous permet de raffiner ce résidu :

$$\forall (M_i^p, M'_i) \in C^p \times C', \quad \overrightarrow{GM'_i}(t + dt) = \overrightarrow{GM_i^p}(t) + \mathcal{M} \left( \overrightarrow{M_i^p M'_i}(t) \right) + \overrightarrow{\delta}_i^{bords}(t) \quad (2.41)$$

avec  $\overrightarrow{\delta}_i^{bords}(t) \in E_{\mathcal{M}}(t)$

Cette étape de raffinement nous permet ainsi d'opérer une décomposition complète du mouvement. Cette décomposition complète n'est valable que pour une correspondance des centres de gravité

géométriques  $G^p$  et  $G'$ , ce qui est le cas avec nos hypothèses :

$$\begin{aligned} \forall (M'_i(t), M'_i(t + dt)) \in S'(t) \times S'(t + dt), \\ \overrightarrow{OM'_i}(t + dt) = \overrightarrow{OG_i^p}(t) + \left[ \overrightarrow{V_{G_i^p/\mathcal{R}_O}}(t) + \overrightarrow{\Omega_{\mathcal{R}_G/\mathcal{R}_O}}(t) \wedge \overrightarrow{G_i^p M'_i}(t) \right] \cdot dt \\ + \overrightarrow{G_i^p M'_i}(t) + \mathcal{M} \left( \overrightarrow{M_i^p M'_i}(t) \right) \\ + \overrightarrow{\delta_i^{int}}(t) + \overrightarrow{\delta_i^{bord}}(t) \end{aligned} \quad (2.42)$$

En faisant correspondre un objet vidéo avec un objet mécanique, il est possible de combiner avantageusement les techniques vidéos pour les petits mouvements avec une représentation cinématique pour les grands mouvements.

## V Formalisation de la chaîne complète

Les concepts théoriques introduits dans le paragraphe précédent proposent d'utiliser une composition de mouvements pour des objets vidéos dans le but de permettre une interpolation que nous appellerons **rendu**. Cette approximation est soumise à plusieurs hypothèses. Certaines sont afférentes à la scène et sont vérifiées dans le cadre du projet FAMOSYS, mais d'autres doivent être vérifiées par les algorithmes mis en place.

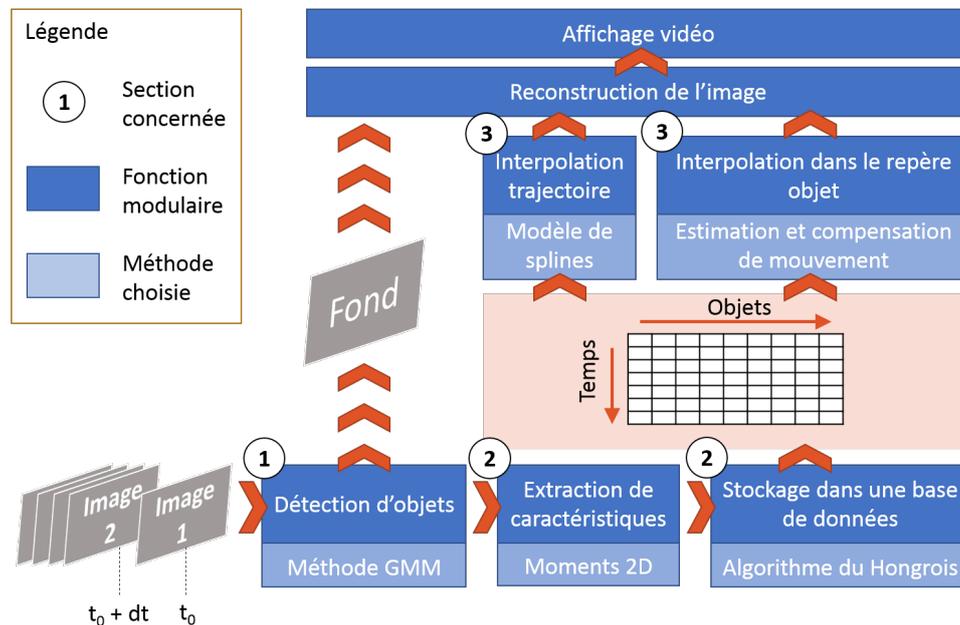


Figure 2.14 – Schéma de la structure algorithmique choisie pour l'implémentation de notre méthode.

L'implémentation que nous détaillons ici apporte une solution au problème de détermination de flot optique pour les grands mouvements. L'utilisation de cet algorithme nous permettra de faire de l'interpolation entre deux images consécutives qui offrira un meilleur rendu en synthétisant de nouvelles images dans la séquence vidéo originale. Son fonctionnement en deux étapes amène à la

## V. Formalisation de la chaîne complète

création d'une base de donnée sur l'évolution des différents objets, ce qui nous permet de faire de la compression sémantique par rapport à la scène.

Nous présentons les choix d'implémentation effectués : en détection, en extraction, sur la structure choisie pour répertorier les objets et sur la reconstruction effectuée à partir des données interpolées. Pour introduire la section, un résumé de la structure algorithmique en fonction des paragraphes est présenté figure 2.14. Nous appliquerons ensuite la méthode sur deux vidéos réelles : une balle rebondissante et un dé.

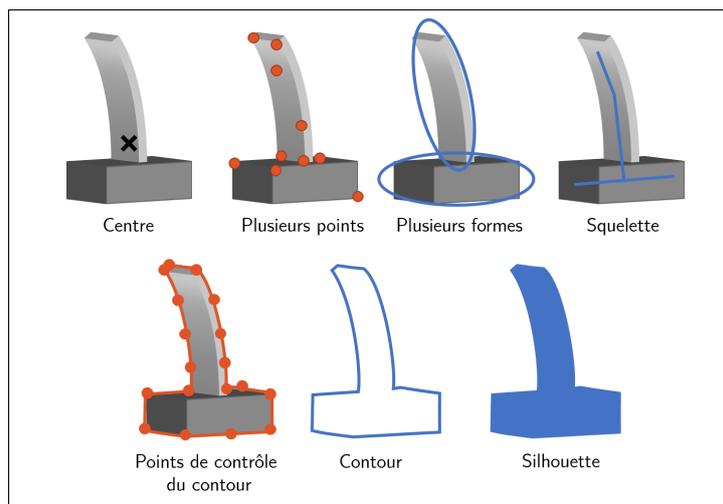
### V.A. Détection et extraction

L'hypothèse majeure de la méthode présentée dans la section précédente est la correspondance entre l'objet vidéo et l'objet mécanique. Cette correspondance est restrictive, puisqu'elle nous oblige :

- à extraire les objets de la scène en respectant leur contour ;
- à éviter les fausses détections qui pourraient induire des déplacements du centre de gravité ;
- à éviter d'utiliser des moyennes temporelles qui pourraient fausser les vitesses ;
- à conserver un schéma temps réel.

Ces trois points ci-dessus restreignent considérablement les méthodes décrites dans la section III pouvant être appliquées. L'ensemble des méthodes peuvent être segmentées en plusieurs catégories, résumées par Yilmaz *et al.* (2006) et que nous illustrons avec la figure 2.15 :

- des objets décrits par leur centre ;
- des objets décrits par plusieurs points ;
- des objets décrits dans une ou plusieurs formes (rectangles, ellipses, englobant ou non) ;
- des objets décrits par leur squelette ;
- des objets décrits par des points de contrôle du contour ;
- des objets décrits par leur contour ;
- enfin, des objets décrits par leur silhouette, c'est-à-dire un masque de l'objet.



**Figure 2.15** – Représentation des différentes manières de détecter un même objet. Dans notre cas, seules les trois dernières descriptions conviennent pour faire correspondre objet vidéo et objet mécanique.

De ces différentes classifications, seules les trois dernières approches permettent d'extraire un objet vidéo correspondant exactement à la projection de l'objet 3D. Compte tenu de nos hypothèses de fond fixe, nous avons choisi d'utiliser des méthodes d'extraction de fond ou de détection de mouvement. Parmi les méthodes exposées précédemment, nous avons testé plusieurs méthodes.

### Les différences d'images

Les différences d'images sont très séduisante étant donné le faible coût de calcul. Cependant la méthode est très inefficace pour des mouvements comportant des positions successives d'objets superposés. La méthode de Sappa et Dornaika (2006) propose de limiter cet effet en ne travaillant que sur le contour. Elle est efficace et ne requiert pas d'initialisation cependant, elle repose principalement sur les caractéristiques spatiales de l'image : elle sera inefficace pour des contours faiblement marqués ou des variations de luminosité. Nous avons développé une couche supérieure à ce détecteur en utilisant la cohérence des informations spatiales de couleur afin de rendre plus robuste le détecteur, ainsi qu'en classant les pixels détectés selon trois catégories : pixel appartenant au fond, pixel appartenant à un objet actif, et pixel appartenant à un objet passif. Ce classement permet d'effectuer un suivi de l'objet en mouvement mais également à l'arrêt. La méthode de différence d'image travaille directement sur une dérivée discrète globale de l'image, et la robustesse de l'image se fait sur le caractère spatial : la détection est statique, ce qui en fait son principal défaut de robustesse.

### Le détecteur sigma delta

Dans notre recherche de rapidité, nous avons testé la méthode de détecteur sigma-delta proposé par Manzanera et Richefeu (2007). Ce détecteur s'avère être très peu coûteux et plutôt robuste par rapport à la vitesse d'exécution. Le fonctionnement simple de la méthode en fait un soustracteur de fond adaptable d'une efficacité redoutable. Comme toute détection, elle reste cependant bruitée, et demande une couche de robustesse pour supprimer le bruit. La qualité de détection nous permet de limiter les traitements à une opération morphologique d'ouverture pour avoir des résultats exploitables. L'utilisation d'un filtre de cohérence spatiale dans le cas d'images couleurs permet de fusionner les informations pour ajouter de la robustesse au détecteur.

### Les contours actifs

La méthode de contour actif semblerait la mieux placée pour répondre à notre besoin, puisqu'elle garantit un suivi de l'objet par ses contours, de manière robuste. La composante temporelle sert à initialiser la méthode. Cependant cette dernière méthode est extrêmement onéreuse en temps de calcul puisqu'elle déplace itérativement les contours de l'objet de l'image précédente pour les faire correspondre à l'objet de l'image suivante. Cette résolution nécessite plusieurs itérations par image. De plus pour des positions d'objets successifs éloignés, la méthode pourra être plus longue voire diverger. Étant donné son aspect hautement itératif très prohibitif pour nos applications : celle-ci n'a pas été retenue.

### Le flot optique

Que le flot optique soit bien déterminé ou pas, il permet de détecter les pixels en mouvement. Cette affirmation se vérifie sur les contenus haute fréquence de l'image où l'équation du flot optique est mal définie, détectant à tort des déplacements de perpendiculaires. Ainsi un détecteur basé sur le flot optique sera efficace pour les contours, moins efficace pour les basses fréquences, nous obligeant à compenser par des opérations morphologiques visant à remplir l'intérieur des contours détectés. Ces méthodes génèrent nativement des erreurs de

## V. Formalisation de la chaîne complète

détection au niveau de l'emplacement précédent de l'objet. Cependant ces fausses détections peuvent souvent être régularisées et la détection des pixels en mouvement s'avère malgré tout très efficace Karlsson et Bigun (2012). Les approches d'estimation de flots denses font malheureusement intervenir des calculs itératifs de minimisation d'un résidu ce qui s'éloigne de notre horizon temps réel (Brox *et al.*, 2009).

### La méthode *ViBe*

La méthode développée par Barnich et Van Droogenbroeck (2011) s'avère être un détecteur plus robuste que le sigma delta et moins lourd que la mixture de gaussienne tout en restant très adaptatif. Nous avons ajouté une couche supérieure de cohérence et de filtrage morphologique pour permettre une extraction robuste et complète de l'objet. La méthode s'avère comme attendu plus efficace sur les hautes fréquences que sur les basses fréquences de l'image. Le bruit de mesure généré par le détecteur se ramène souvent à des pixels isolés, ce qui rend le filtrage plus aisé. Malheureusement, l'application de cette méthode *ViBe* à notre méthode se trouve encore dans les perspectives.

### La mixture de gaussienne *GMM*

La méthode de mixture de gaussienne introduite précédemment permet de segmenter à chaque instant les pixels du fond des pixels des objets (partie III.B.). Cette méthode propose une modélisation complète du comportement d'un pixel permettant ainsi de discriminer les pixels. La méthode modélise ainsi chaque pixel de l'image par une somme de gaussienne. La charge en mémoire n'est pas négligeable, mais le cadre probabiliste confère à la méthode à la fois des propriétés adaptatives et une inertie nécessaire pour détecter les mouvements lents. De la même manière que pour les autres méthodes, nous avons ajouté une couche de traitements morphologiques pour compenser les problèmes de détection (fragmentation d'objets, zones basses fréquences) ainsi qu'un filtre de cohérence spatiale dans le cas d'images couleurs.

Nous utiliserons le détecteur *GMM* pour la suite de la méthode. Ce choix a été fait à la fois pour garantir un schéma temps réel mais également pour assurer plus de fidélité dans la détection. Néanmoins les méthodes *ViBe*, sigma delta ou celles basées sur le flot optique pourraient également être utilisées. L'architecture de l'algorithme a volontairement été conçue pour permettre de changer facilement de méthode de détection. Chaque méthode précédemment citée a été associée à un bloc de correction comportant des étapes spécifiques de filtrage morphologique ou de cohérence spatiale. Ce bloc extrait alors une série de masques à l'aide d'une étape de connexité, chacun comportant un seul objet. Cette série de masque constitue l'entrée de la partie caractérisation que nous allons détailler maintenant.

## V.B. Caractérisation des objets

La caractérisation des objets a été pensée dans la philosophie de la norme MPEG-7 et les recommandations de la vision objet de la norme MPEG-4 qui ont été exposées plus haut. Le but pour effectuer notre reconstruction est d'acquérir la position, l'orientation principale et les vitesses associées de chaque objet. Avec ces informations une base de donnée est construite à laquelle nous ajoutons les pixels concernés ainsi que l'histogramme de chaque objet, l'aire, et le plus petit rectangle englobant l'objet. Une version plus compacte a également été proposée en ne retenant que le périmètre de chaque objet au lieu de la totalité de ses pixels.

À partir des masques issus de la partie précédente, chaque objet est caractérisé à partir de son masque binaire, en utilisant les moments 2D de Hu introduits précédemment. La vitesse est calculée à partir des positions en utilisant une dérivée numérique basée sur une approximation centrée du second ordre de la dérivée :

$$f(t + \Delta t) = f(t) + \Delta t \cdot f'(t) + \frac{\Delta t^2}{2} f''(t) + o(\Delta t^2, t) \quad (2.43)$$

$$f(t - \Delta t) = f(t) - \Delta t \cdot f'(t) + \frac{\Delta t^2}{2} f''(t) + o(\Delta t^2, t) \quad (2.44)$$

$$\text{ce qui donne : } \frac{f(t + \Delta t) - f(t - \Delta t)}{2\Delta t} = f'(t) + \frac{\Delta t^2}{6} f^{(3)}(t) + o(\Delta t^2, t) \quad (2.45)$$

avec  $f$  la fonction position à dériver,  $\Delta t$  le pas de dérivation et  $t$  le temps. Ainsi, à chaque pas de temps, une collection d'objets  $O(t)$  est détectée. Les positions du centre image sont exprimées dans le repère de l'image, tandis que les autres informations sont exprimées dans le repère objet, de manière à travailler dans un repère relatif.

L'association des objets de la collection  $O(t)$  avec ceux du temps  $O(t + \Delta t)$  se fait alors à l'aide d'un algorithme du hongrois (Kuhn, 1955) basé non pas sur la distance, mais sur une fonction de coût mélangeant à la fois la distance entre deux positions et la distance entre les histogrammes calculée à l'aide de la formule de Bhattacharyya. La distance pourrait également être pénalisée en fonction de l'orientation du vecteur vitesse, cependant le bruit de la vitesse ajouté au caractère imprévisible des objets sur la scène rendent cette pénalisation très aléatoire. Elle génère plus d'erreurs d'associations qu'elle ne permet de lever des incertitudes.

Cette base de données permet à elle seule de visualiser l'évolution de la scène avec les formes d'objets uniquement. Elle permet ainsi de limiter le contenu de la vidéo aux seuls objets en mouvements.

### V.C. Reconstruction et rendu amélioré

La reconstruction de la vidéo se fait à l'aide de la base de donnée. Il s'agit d'utiliser une image de fond fixe, et d'utiliser les positions des pixels caractérisant l'objet pour faire évoluer les masques des objets sur la scène en fonction de la trajectoire du centre. Une version plus lourde de l'extraction peut également permettre de stocker dans la base de donnée la valeur des pixels de l'objet en plus de la position, ce qui permet de reconstituer une scène exacte. Afin de permettre de restituer les variations du fond, une compression classique peut être effectuée étant donné qu'il ne présente pas d'intérêt particulier pour la scène, ainsi qu'un rafraîchissement sous-échantillonné.

Cette reconstruction offre une compression de la vidéo au sens de l'objet sous les hypothèses de travail énoncées dans la partie IV.A. : nous l'appelons **compression sémantique**. L'autre objectif de cette méthode était de permettre d'obtenir un rendu amélioré de l'image en effectuant une interpolation entre chaque image. La méthode décrite précédemment a été mise en place : interpoler la trajectoire des centres de chaque objet dans le repère image à l'aide d'un modèle de splines d'une part et d'autre part effectuer dans le repère objet une estimation de mouvements petite échelle suivie d'une compensation de mouvements. Plusieurs algorithmes ont été utilisés, dont les résultats seront présentés avec les algorithmes de flot optique de Farnebäck (2003) et de Brox *et al.* (2004) proposé par Liu (2009).

Encore ici, la structure a volontairement été choisie modulaire afin de pouvoir incorporer d'autres méthodes d'estimation de mouvements. La carte de champ de mouvements obtenue, la reconstruc-

tion se fait ensuite par compensation de mouvement à l'aide d'une méthode présentée par Chan *et al.* (2010).

### V.D. Applications

L'algorithme a été testé sur plusieurs séquences à des fins académiques ou industrielles. Nous présenterons ici quelques applications afin de permettre au lecteur de juger de la qualité du rendu amélioré selon quelques méthodes d'estimation de mouvement. Les difficultés présentées dans la partie A. sont les mêmes au niveau de l'interpolation de la trajectoire. Au niveau de l'hypothèse des petits mouvements, les algorithmes de flot optique diffèrent sur le rayon du « petit mouvement » pour une estimation du mouvement réussie aussi bien que sur leur capacité à effectuer des calculs en temps réels.

#### V.D.1. Scénario du dé

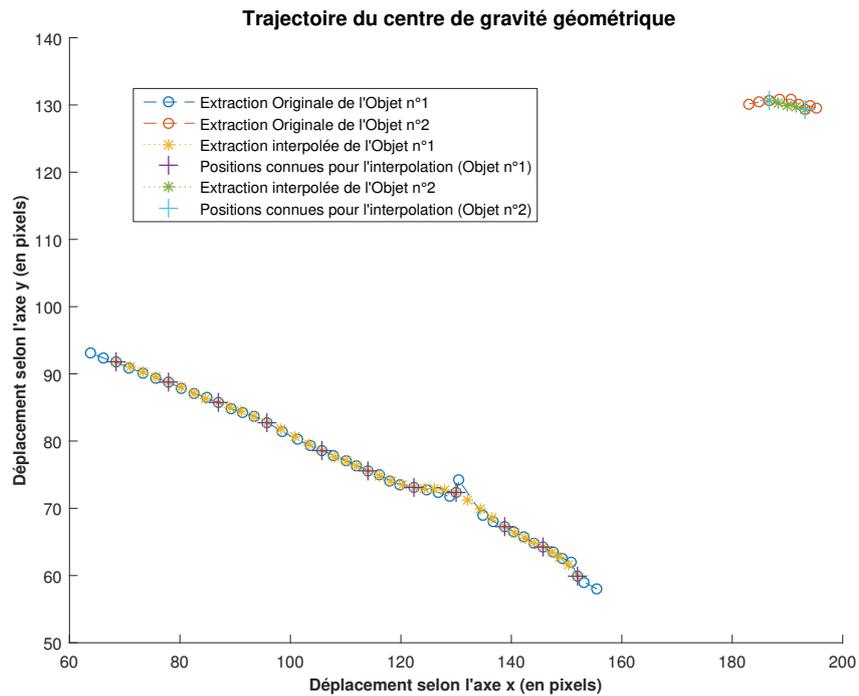
Le premier cas test est un dé sur une table que l'on peut voir figure 2.16. La séquence a été choisie particulièrement courte et centrée sur l'objet afin de bien se focaliser sur la vérification de nos hypothèses. Le dé nous permet de pouvoir tester le cas du contour qui varie faiblement par rapport à la vitesse de capture, mais également tester le cas de mauvaise détection, puisqu'ici l'extraction d'objet est moins robuste, compte tenu de l'absence de couleurs. Ainsi pour ce cas test, on testera une interpolation comparée sur quatre positions. Les méthodes de *block matching* ayant des résultats nettement moins bons, ils ne seront pas présentés ici. Cette deuxième méthode est beaucoup plus lourde que la première (un facteur 10 sur le temps d'exécution), mais il nous semblait important de pouvoir utiliser plusieurs méthodes, afin de montrer la modularité de l'algorithme.

Afin de valider l'approche composition de mouvement, nous commençons par les résultats de la trajectoire. La figure 2.17 propose la trajectoire des différents objets dans la scène (en l'occurrence, le dé et son ombre).

Cette figure superpose une séquence initiale sous la forme de cercles avec la séquence interpolée qui n'a eu que des informations partielles de la séquence originale (ici une image pour quatre). L'interpolation est volontairement poussée aux limites de l'information capturée par l'acquisition afin d'observer des aberrations et de vérifier l'analyse menée dans l'annexe A.. Si l'interpolation de l'ombre n'est pas significative avec deux points, l'interpolation du dé est très bonne dans la majeure partie de la trajectoire, reflétant son comportement continu en vitesse. Un décrochage est visible autour de la position [130,75] : il correspond à une erreur de détection. Cette erreur apportant une perturbation haute fréquence locale, notre modèle d'interpolation répondra de manière inappropriée

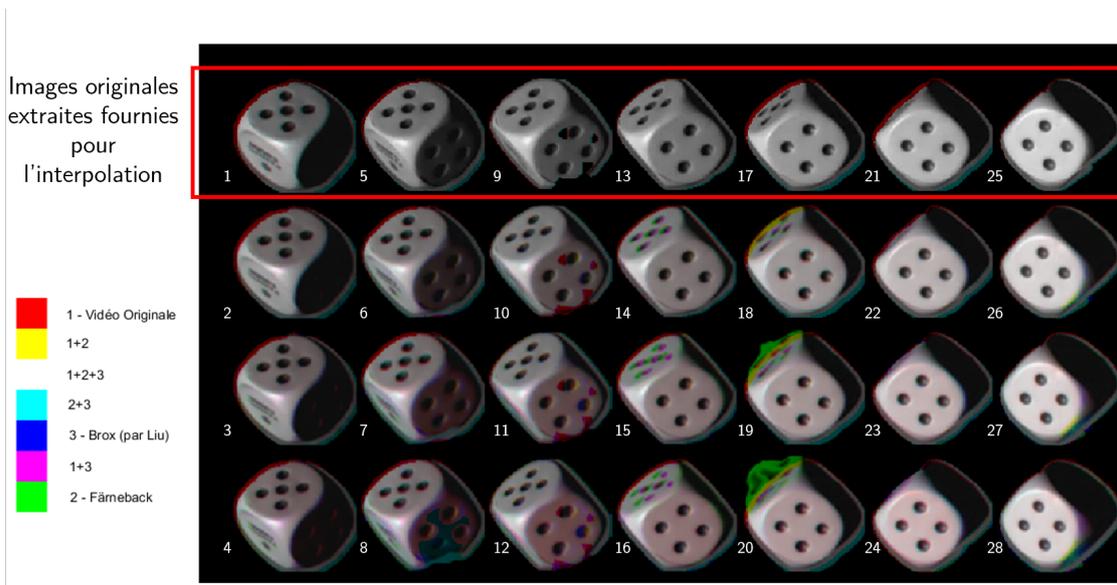


**Figure 2.16** – Vidéo originale du dé. Les vidéos suivantes se serviront de certaines images de cette vidéo initiale et les interpolations seront comparées à cette référence.



**Figure 2.17** – Représentation de la trajectoire du centre de gravité du dé (en bas à gauche) et de son ombre (en haut à droite)

dans le cas général (voir annexe A.). Dans ce cas, le comportement du dé entrant dans le cadre d'un objet de vitesse continue et peu soumise à accélération, notre modèle de spline se trouve approprié et son comportement permet de lisser la trajectoire.



**Figure 2.18** – Représentation des différences entre les images originales (canal rouge), les images reconstruites à partir du flot optique de Färneback (canal vert) et celles reconstruites à partir du flot optique de Brox (canal bleu)

## V. Formalisation de la chaîne complète



**Figure 2.19** – Vidéo présentant les masques interpolés des objets superposés à la vidéo originale. Cette interpolation est faite à partir d'une image sur quatre de la vidéo originale.



**Figure 2.20** – Vidéo présentant la vidéo interpolée avec notre méthode.

Nous proposons ensuite de nous intéresser à la figure 2.18 qui superpose les estimations de mouvement dans le repère local de l'objet. Les canaux de couleur nous permettent de superposer sur une même image les différents résultats. Ainsi, sur le canal rouge on aura les images originales, sur le canal vert, les images reconstruites à partir du flot optique de Färneback, et sur le canal bleu, celles reconstruites à partir du flot optique de Brox. Ainsi, lorsque les images sont superposées, l'images a les couleurs originales du dé. Lorsque les images ne sont plus superposées, la vidéo originale apparaîtra rouge, Färneback en vert et Brox un bleu. Nous avons précisé dans la légende la figure les teintes observées pour une superposition de deux canaux sur les trois. Nous aurons du cyan lorsque les flots coïncideront sans la vidéo originale, du jaune lorsqu'il s'agira de Färneback et la vidéo originale et du magenta pour Brox et l'original. Un dégradé de luminosité des couleurs sera observé en fonction de l'intensité originale des pixels.

La première ligne de cette figure présente l'ensemble des images fournies pour l'interpolation. Il faut donc lire chaque colonne de haut en bas en partant de la gauche. Pour faciliter cette lecture, les numéros de chaque image sont ajoutés.

Il est intéressant d'observer en détail cette figure. Tout d'abord, on peut constater des erreurs de centrage d'images sur les images observées. Ce problème de centrage est causé par l'utilisation d'une fenêtre d'analyse dont les dimensions sont définies comme les maximums des fenêtres encadrant les objets détectés. En plus de cette erreur sur la référence, les contours n'étant pas fixés, il est possible que le centre de l'image réelle et celle interpolée ne correspondent plus exactement, produisant un décalage de l'image par rapport à sa référence. Les résultats sont convaincants, même si l'absence de contrôle de résidu peut produire des résultats biaisés comme on peut le constater sur la cinquième colonne de la figure 2.18. Ces erreurs d'estimation de mouvements proviennent cette fois d'une absence d'informations. Entre la première image de la cinquième colonne et la première image de la sixième, on remarque qu'on a une disparition d'une face. Les deux méthodes d'estimation traitent cette disparition de manière différente. Les variations lentes de contour sont par ailleurs bien modélisées par l'estimation de mouvement.

Deux vidéos présentent les résultats dynamiques pour permettre au lecteur d'apprécier le résultat

de cette interpolation aveugle. L'estimation de mouvement dans ces vidéos est celle de Brox. La figure 2.19 présente le résultat en affichant les masques des images interpolées par dessus la vidéo originale. Malgré que trois images sur quatre ont été masquées, les évolutions des masques correspondent à celles des objets de la scène. La figure 2.20 propose l'interpolation complète de la vidéo.

L'apport au niveau de la compression est immédiat puisque nous proposons de conserver les zones d'intérêt que sont les objets en pleine définition et de compresser les données non-significatives. Le gain dépendra du rapport entre les pixels appartenant à des objets par rapport à la taille des images. Mais la compression se fera ici dans la sémantique de l'objet.

Ce cas test permet de valider notre méthode sur les faibles variations de contours et le bon fonctionnement de la composition de mouvements. Cependant, il propose un scénario ne mettant pas à mal les hypothèses de trajectoires des splines. Ce que nous allons aborder dans un second exemple.

### V.D.2. Scénario balle de tennis

Un second exemple est présenté avec un scénario de balle de tennis qui rebondit. Ce scénario permet de tester deux autres aspects : une trajectoire discontinue et une mauvaise acquisition (un flou de mouvement).

La vidéo originale est présentée sur la figure 2.21. Nous présenterons les résultats de la même manière que pour l'exemple précédent. La figure 2.22 montre la trajectoire réelle de la balle avec une interpolation basée sur une valeur sur quatre de l'acquisition. Ce cas test présente une trajectoire pseudo périodique, avec des variations d'accélération de l'objet. Deux types de changements de direction des vitesses selon l'axe y sont observés : le sommet du rebond où ce changement est continu et le rebond où il est discontinu. Il est intéressant de vérifier que l'on observe bien les conséquences du théorème d'échantillonnage qui n'apparaît qu'à partir d'une certaine vitesse. Les trois premiers rebonds sont correctement reconstruits sur la partie continue, la vitesse étant suffisamment importante pour garantir un échantillonnage suffisant. Deux remarques peuvent être faites sur ce cas test. Comme présenté dans l'annexe A, le modèle se comporte mal au niveau des rebonds : il ne peut se fier qu'aux informations mesurées et on observe de surcroît un phénomène en dehors de ses hypothèses d'application. On remarquera également que les points interpolés appartiennent bien à la trajectoire mais ne correspondent pas aux points réels : cela traduit l'incapacité de ce modèle à capturer la dynamique de la scène ce qui en fait une limitation.

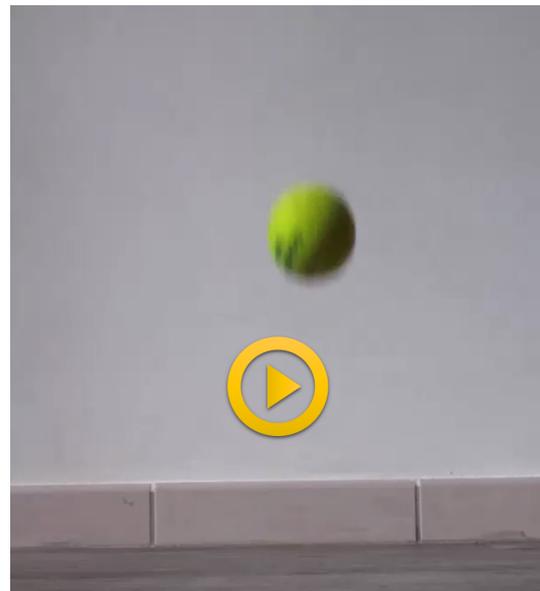


Figure 2.21 – Vidéo originale « Balle de Tennis ».

L'utilisation de la méthode montre des limites sur l'interpolation aveugle de la trajectoire. Néanmoins dans une utilisation en compression de données les positions seront toutes connues. Il n'y aura donc pas d'erreurs dues à la dynamique de la scène. Le gain de compression ici sera plus important que

## V. Formalisation de la chaîne complète

dans l'exemple précédent puisque les pixels relatifs à l'objet n'occupent en moyenne que 3% de l'image.

De la même manière que pour l'exemple précédent, nous proposons dans la figure 2.23 d'analyser la superposition des différentes méthodes avec l'original. Malgré le même problème que précédemment, les résultats correspondent. On notera également que le flou engendré par une vitesse d'obturation mal adaptée n'a pas gêné l'estimation de mouvement : le phénomène devrait d'ailleurs améliorer les résultats puisqu'il atténue les hautes fréquences de l'objet.

Pour conclure sur cet exemple, nous présentons deux vidéos montrant une amélioration du **rendu** de la vidéo originale. Nous avons choisi arbitrairement de rajouter 9 images entre chaque image de la vidéo précédente. La figure 2.24 montre l'évolution du masque en fonction de la vidéo initiale tandis que la figure 2.25 montre les résultats de l'amélioration de la séquence. Dans cette séquence, la balle est plus flottante : la dynamique du mouvement est très nettement lissée. Par contre, on souligne la netteté de détails comme l'écriture sur la balle ou sa texture, ce qui montre que notre algorithme est bien capable de faire de l'estimation et compensation de mouvements dans un cadre de grands déplacements.

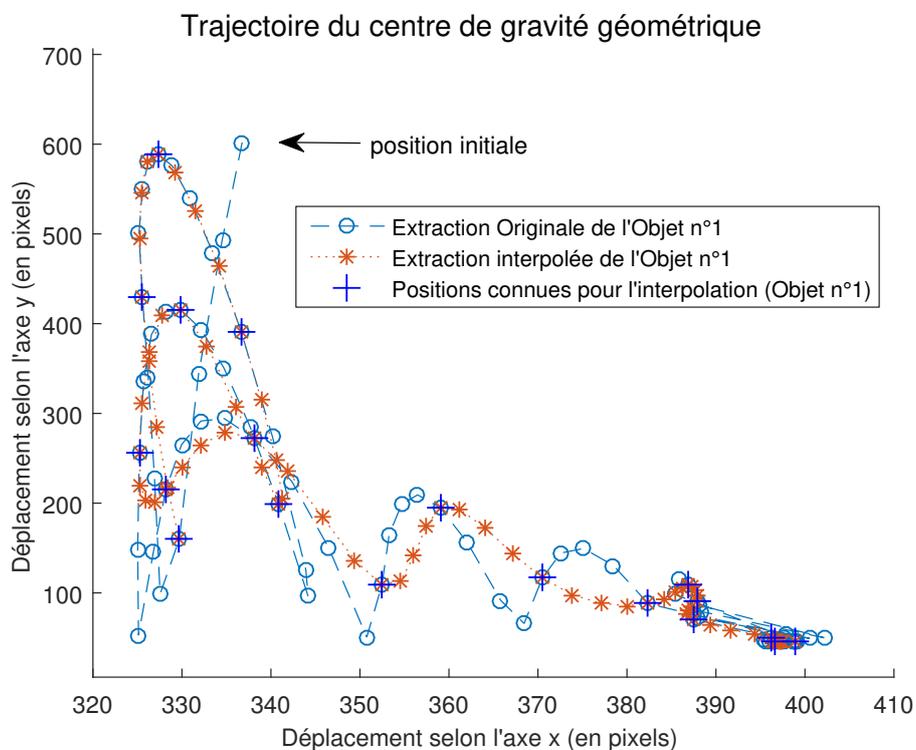
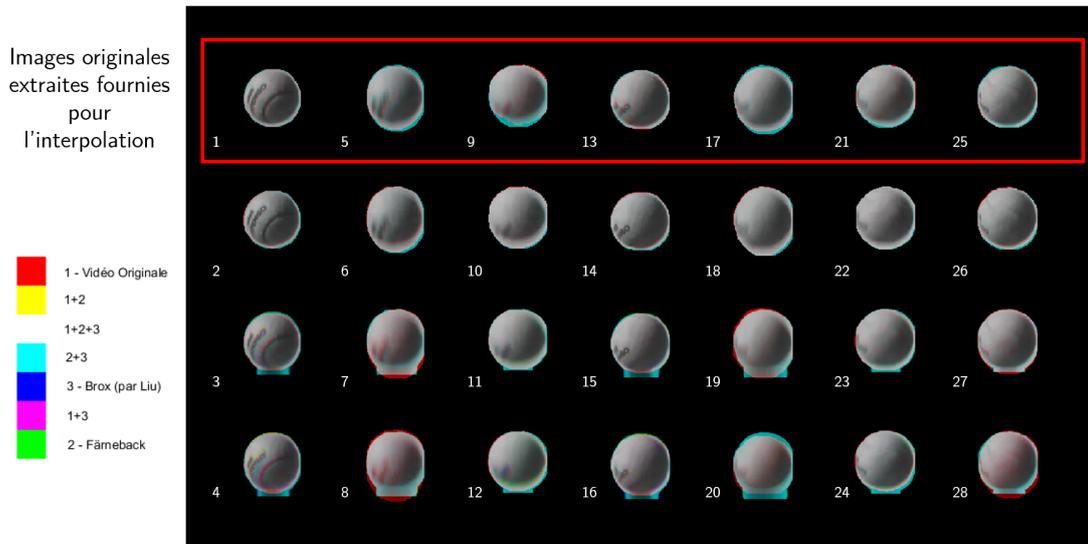
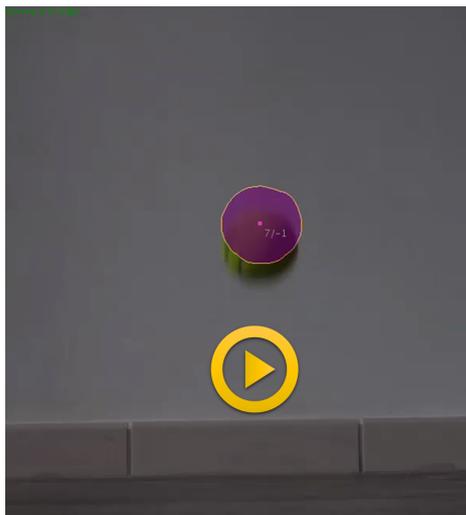


Figure 2.22 – Représentation de la trajectoire du centre de gravité de la balle de Tennis



**Figure 2.23** – Représentation des différences entre les images originales (canal rouge), les images reconstruites à partir du flot optique de Färneback (canal vert) et celles reconstruites à partir du flot optique de Brox (canal bleu)



**Figure 2.24** – Vidéo superposant le masque d'une amélioration de rendu de facteur 10 avec la vidéo originale : on multiplie par 10 les images dans la nouvelle séquence.



**Figure 2.25** – Vidéo synthétisée proposant une amélioration de rendu d'un facteur 10.

## VI Conclusion

Pour pouvoir mieux positionner les problématiques énoncées dans le chapitre introductif, un état de l'art général sur le traitement vidéo actuel a été réalisé. La problématique traitée dans ce chapitre a été différente par rapport à celle formulée dans le chapitre 1 puisqu'on a considéré le cas de vidéos rapides sans modèle spécifiquement développé pour la scène. Les travaux exposés dans ce chapitre

## VI. Conclusion

répondent ainsi partiellement aux objectifs dans ce cadre plus élargi.

Un état de l'art spécifique basé sur les normes vidéos a été développé afin de bien détailler les processus sous-jacents à la compression, les principales techniques de traitement vidéo et la philosophie de développement dans ce domaine. Étant focalisé sur les objets de la scène nous avons remarqué qu'en identifiant les zones d'intérêt de la vidéo, il était possible de travailler spécifiquement sur ces objets afin de faire de l'interpolation ou de la compression.

En reprenant l'approche classique d'estimation de mouvements applicable pour de petits déplacements, une méthode d'estimation de mouvement applicable à de grands déplacements a été proposée. Cette méthode s'articule autour de la correspondance entre deux approches de l'évolution d'un objet dans une scène 2D : l'une mécanique et l'autre vidéo. Pour effectuer ce lien nous avons introduit un double cadre de description des objets dans une vidéos : l'un eulérien et l'autre lagrangien. Nous avons présenté et caractérisé des méthodes permettant de passer entre ces deux représentations. Forts de ce double contexte, une nouvelle méthode a été présentée, basée sur une composition de mouvements afin d'utiliser avantageusement les deux descriptions. Un modèle élémentaire et cinématique de splines a été utilisé pour reconstruire la vidéo. Si les applications en compression sont immédiates, elles sont moins évidentes en interpolation : la **généricité** de notre modèle ne nous permet pas d'ajouter d'information et nous **limite** à la **dynamique capturée** lors de l'acquisition. Néanmoins, cette méthode peut simplement améliorer le rendu visuel d'une séquence et offre une nouvelle solution au problème d'estimation de mouvement de grands mouvements. Ces concepts ont été testés sur deux séquences types choisies pour illustrer les limites de nos hypothèses.

Ce chapitre décrit un algorithme modulaire d'extraction et reconstruction d'une vidéo rapide sans informations préalables. Ce modèle a été posé selon des hypothèses : nous avons exposé ses intérêts et ses limites. Parmi celles-ci, le cahier des charges initialement prévu n'a pas été respecté : notre méthode ne permet de faire ni d'identification ni de prédiction, l'unique raison étant que notre modèle n'ajoute pas d'informations suffisamment précises à l'acquisition. Cette approche permettant l'enrichissement d'observation grâce à un modèle plus complexe paraît séduisante. C'est la direction que nous proposons de suivre dans le chapitre suivant.

## Bibliographie du chapitre 2

- Luis Alvarez, Joachim Weickert et Javier Sánchez : Reliable Estimation of Dense Optical Flow Fields with Large Displacements. *International Journal of Computer Vision*, volume 39, pages 41–56, août 2000. ISSN 1573-1405.
- Elise Arnaud : Extractions de caractéristiques - les points d'intérêt, mars 2015.
- Nicole M. Artner : A comparison of mean shift tracking methods. *In 12th Central European Seminar on Computer Graphics*, pages 197–204, 2008.
- Aroh Barjatya : Block matching algorithms for motion estimation. *IEEE Transactions Evolution Computation*, volume 8, pages 225–239, 2004.
- O Barnich et M Van Droogenbroeck : ViBe : A Universal Background Subtraction Algorithm for Video Sequences. *IEEE Transactions on Image Processing*, volume 20, pages 1709–1724, juin 2011. ISSN 1057-7149, 1941-0042.
- J. L. Barron, D. J. Fleet et S. S. Beauchemin : Performance of optical flow techniques. *International Journal of Computer Vision*, volume 12, pages 43–77, février 1994. ISSN 1573-1405.
- Jenny Benois-Pineau et Henri Nicolas : A New Method for Region-Based Depth Ordering in a Video Sequence : Application to Frame Interpolation. *Journal of Visual Communication and Image Representation*, volume 13, pages 363–385, septembre 2002. ISSN 10473203.
- Christophe Bernard : *Ondelettes et Problèmes Mal Posés : La Mesure Du Flot Optique et l'interpolation Irrégulière*. Thèse de doctorat, Ecole Polytechnique X, Centre de Mathématiques Appliquées, 1999.
- Alan C. Bovik, éditeur. *The Essential Guide to Video Processing*. Academic Press/Elsevier, Amsterdam ; Boston, 2nd ed édition, 2009. ISBN 978-0-12-374456-2.
- Yuri Boykov et Olga Veksler : Graph cuts in vision and graphics : Theories and applications. *In Handbook of Mathematical Models in Computer Vision*, pages 79–96. Springer, 2006.
- Thomas Brox, Christoph Bregler et Jitendra Malik : Large displacement optical flow. *In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference On*, pages 41–48. IEEE, 2009.
- Thomas Brox, Andrés Bruhn, Nils Papenberg et Joachim Weickert : High accuracy optical flow estimation based on a theory for warping. *In Computer Vision-ECCV 2004*, pages 25–36. Springer, 2004.
- Thomas Brox et Jitendra Malik : Large displacement optical flow : Descriptor matching in variational motion estimation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, volume 33, pages 500–513, 2011.
- Aurélié Bugeau : *Détection et Suivi d'objets En Mouvement Dans Des Scenes Complexes, Application a La Surveillance Des Conducteurs*. Thèse de doctorat, Université Rennes 1, Rennes, 2007.
- Aurélié Bugeau et Patrick Pérez : Detection and segmentation of moving objects in highly dynamic scenes. *In Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference On*, pages 1–8. IEEE, 2007.

## Bibliographie du chapitre 2

- Roberto Castagno, Petri Haavisto et Giovanni Ramponi : A method for motion adaptive frame rate up-conversion. *Circuits and Systems for Video Technology, IEEE Transactions on*, volume 6, pages 436–446, 1996.
- Stanley H. Chan, Dung T. Vo et Truong Q. Nguyen : Subpixel motion estimation without interpolation. *In Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference On*, pages 722–725. IEEE, 2010.
- Yang-Ho Cho, Ho-Young Lee et Du-Sik Park : Temporal Frame Interpolation Based on Multiframe Feature Trajectory. *IEEE Transactions on Circuits and Systems for Video Technology*, volume 23, pages 2105–2115, décembre 2013. ISSN 1051-8215, 1558-2205.
- Changhyun Choi et Henrik I. Christensen : 3D textureless object detection and tracking : An edge-based approach. *In Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference On*, pages 3877–3884. IEEE, 2012.
- Simon Denman, Clinton Fookes et Sridha Sridharan : Improved Simultaneous Computation of Motion Detection and Optical Flow for Object Tracking. *In 2009 Digital Image Computing : Techniques and Applications*, pages 175–182. IEEE, décembre 2009. ISBN 978-1-4244-5297-2.
- Séverine Dubuisson : Tree-structured image difference for fast histogram and distance between histograms computation. *Pattern Recognition Letters*, volume 32, pages 411–422, février 2011. ISSN 01678655.
- Richard O. Duda et Peter E. Hart : Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, volume 15, pages 11–15, 1972.
- Güray Erus : *Reconnaissance d'objets Cartographiques Dans Les Images Satellitaires à Haute Résolution*. Thèse de doctorat, Paris 5, 2008.
- Wassima Ait Fares : *Détection et Suivi d'objets Par Vision Fondés Sur Segmentation Par Contour Actif Basé Région*. Thèse de doctorat, Université Paul Sabatier-Toulouse III, septembre 2013.
- Gunnar Farneback : Two-Frame Motion Estimation Based on Polynomial Expansion. *In Josef Bigun et Tomas Gustavsson, éditeurs : Image Analysis : 13th Scandinavian Conference, SCIA 2003 Halmstad, Sweden, June 29 – July 2, 2003 Proceedings*, pages 363–370. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003. ISBN 978-3-540-45103-7.
- Denis Fortun, Patrick Bouthemy et Charles Kervrann : Optical flow modeling and computation : A survey. *Computer Vision and Image Understanding*, volume 134, pages 1–21, 2015.
- Huazhu Fu, Dong Xu, Bao Zhang, Stephen Lin et Rabab Kreidieh Ward : Object-Based Multiple Foreground Video Co-Segmentation via Multi-State Selection Graph. *IEEE Transactions on Image Processing*, volume 24, pages 3415–3424, novembre 2015. ISSN 1057-7149, 1941-0042.
- Vincent Garcia : *Suivi d'objets d'intérêt Dans Une Séquence d'images : Des Points Saillants Aux Mesures Statistiques*. Thèse de doctorat, Université Nice Sophia Antipolis, Nice, avril 2009.
- C. Harris et M. Stephens : A Combined Corner and Edge Detector. pages 23.1–23.6. Alvey Vision Club, 1988.

- F. Hild et S. Roux : Comparison of Local and Global Approaches to Digital Image Correlation. *Experimental Mechanics*, volume 52, pages 1503–1519, novembre 2012. ISSN 0014-4851, 1741-2765.
- Berthold K. Horn et Brian G. Schunck : Determining optical flow. *In 1981 Technical Symposium East*, pages 319–331. International Society for Optics and Photonics, 1981.
- Ming-Kuei Hu : Visual pattern recognition by moment invariants. *information Theory, IRE Transactions on*, volume 8, pages 179–187, 1962.
- M. Jakubowski et G. Pastuszak : Block-based motion estimation algorithms — a survey. *Opto-Electronics Review*, volume 21, pages 86–102, mars 2013. ISSN 1896-3757.
- R. Amali Therese Jenifa, C. Akila et V. Kavitha : Rapid background subtraction from video sequences. *In 2012 International Conference on Computing, Electronics and Electrical Technologies (ICCEET)*, pages 1077–1086, mars 2012.
- Baozhi Jia, Rui Liu et Ming Zhu : Real-time obstacle detection with motion features using monocular vision. *The Visual Computer*, volume 31, pages 281–293, mars 2015. ISSN 0178-2789, 1432-2315.
- Zhen Jia, Arjuna Balasuriya et Subhash Challa : Vision based data fusion for autonomous vehicles target tracking using interacting multiple dynamic models. *Computer Vision and Image Understanding*, volume 109, pages 1–21, janvier 2008. ISSN 10773142.
- Shengqin Jiang et Xiaobo Lu : WeSamBE : A Weight-Sample-Based Method for Background Subtraction. *IEEE Transactions on Circuits and Systems for Video Technology*, pages 1–1, 2017. ISSN 1051-8215, 1558-2205.
- Pierre-Marc Jodoin : Comparative study of background subtraction algorithms. *Journal of Electronic Imaging*, volume 19, page 033003, juillet 2010. ISSN 1017-9909.
- Stefan M. Karlsson et Josef Bigun : Lip-motion events analysis and lip segmentation using optical flow. pages 138–145. IEEE, juin 2012. ISBN 978-1-4673-1612-5 978-1-4673-1611-8 978-1-4673-1610-1.
- Kyungah Kim, Minwoo Kim, Deokho Kim et Won Woo Ro : True motion compensation with feature detection for frame rate up-conversion. *In Image Processing (ICIP), 2015 IEEE International Conference On*, pages 2260–2264. IEEE, 2015.
- Scott Krig : *Computer Vision Metrics : Survey, Taxonomy, and Analysis*. Apress, Berkely, CA, USA, 1st édition, 2014. ISBN 1-4302-5929-9 978-1-4302-5929-9.
- Scott Krig : Feature Learning and Deep Learning Architecture Survey. *In Computer Vision Metrics : Textbook Edition*, pages 375–514. Springer International Publishing, Cham, 2016. ISBN 978-3-319-33762-3.
- Harold W Kuhn : The Hungarian method for the assignment problem. *Naval Research Logistics (NRL)*, volume 2, pages 83–97, 1955.
- Lionel Lacassagne et Antoine Manzanera : Motion detection : Fast and robust algorithms for embedded systems. *In ICIP*, pages 3265–3268, 2009.

## Bibliographie du chapitre 2

- Ivan Laptev : *Local Spatio-Temporal Image Features for Motion Interpretation*. Theses, KTH, Stockholm, 2004.
- Dah-Jye Lee, Paul C. Merrell, Brent E. Nelson et Zhaoyi Wei : Multi-frame structure from motion using optical flow probability distributions. *Neurocomputing*, volume 72, pages 1032–1041, janvier 2009. ISSN 09252312.
- Jin-Hwal Lee, Kee-Sung Lee et Geun-Sik Jo : Representation Method of the Moving Object Trajectories by Interpolation with Dynamic Sampling. In *Information Science and Applications (ICISA), 2013 International Conference On*, pages 1–4. IEEE, 2013.
- Yann Lepoittevin : *Estimation de La Dynamique à Partir Des Structures Observées Dans Une Séquence d'images*. Thèse de doctorat, Université Pierre et Marie Curie, 2015.
- Yang Ling, Jin Wang, Yunqiang Liu et Wenjun Zhang : A novel spatial and temporal correlation integrated based motion-compensated interpolation for frame rate up-conversion. *Consumer Electronics, IEEE Transactions on*, volume 54, pages 863–869, 2008.
- Ce Liu : *Beyond Pixels : Exploring New Representations and Applications for Motion Analysis*. Thèse de doctorat, Citeseer, 2009.
- Ming-Yu Liu, Oncel Tuzel, Ashok Veeraraghavan et Rama Chellappa : Fast directional chamfer matching. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference On*, pages 1696–1703. IEEE, 2010.
- Tianshu Liu, Ali Merat, M. H. M. Makhmalbaf, Claudia Fajardo et Parviz Merati : Comparison between optical flow and cross-correlation methods for extraction of velocity fields from particle images. *Experiments in Fluids*, volume 56, août 2015. ISSN 0723-4864, 1432-1114.
- Bruce D. Lucas et Takeo Kanade : An iterative image registration technique with an application to stereo vision. In *IJCAI*, volume 81, pages 674–679, 1981.
- Dhruv Mahajan, Fu-Chung Huang, Wojciech Matusik, Ravi Ramamoorthi et Peter Belhumeur : Moving gradients : A path-based method for plausible image interpolation. *ACM Transactions on Graphics (TOG)*, volume 28, page 42, 2009.
- James Malcolm, Yogesh Rathi et Allen Tannenbaum : Multi-object tracking through clutter using graph cuts. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference On*, pages 1–5. IEEE, 2007.
- E. Malis et E. Marchand : Méthodes robustes d'estimation pour la vision robotique. In *Journées Nationales de La Recherche En Robotique, JNRR'05*, 2005.
- Antoine Manzanera et Julien C Richefeu : A new motion detection algorithm based on Sigma-Delta background estimation. *Pattern Recognition Letters*, volume 28, pages 320–328, 2007.
- Jiri Matas, Ondrej Chum, Martin Urban et Tomás Pajdla : Robust wide-baseline stereo from maximally stable extremal regions. *Image and vision computing*, volume 22, pages 761–767, 2004.
- Florent Mathieu, François Hild et Jérémie Bellec : Le projet DICCIT : Des outils pour la comparaison essai/calcul et l'identification. Lyon, France, août 2015. AFM, Association Française de Mécanique.

- Simone Meyer, Oliver Wang, Henning Zimmer, Max Grosse et Alexander Sorkine-Hornung : Phase-based frame interpolation for video. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1410–1418, 2015.
- Krystian Mikolajczyk et Cordelia Schmid : Indexing based on scale invariant interest points. *In Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference On*, volume 1, pages 525–531. IEEE, 2001.
- Anurag Mittal et Nikos Paragios : Motion-based background subtraction using adaptive kernel density estimation. *In Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference On*, volume 2, pages II–II. IEEE, 2004.
- Hans Peter Moravec : *Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover*. Thèse de doctorat, Stanford University, Stanford, CA, USA, 1980.
- Moving Picture Experts Group : MPEG | The Moving Picture Experts Group website. <http://mpeg.chiariglione.org/>, 2017.
- Naoya Ohnishi et Atsushi Imiya : Dominant plane detection from optical flow for robot navigation. *Pattern Recognition Letters*, volume 27, pages 1009–1021, juillet 2006. ISSN 01678655.
- Fredrik Orderud, Jøger Hansgård et Stein I. Rabben : Real-time tracking of the left ventricle in 3D echocardiography using a state estimation approach. *In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, pages 858–865. Springer, 2007.
- Edouard Oyallon et Julien Rabin : An Analysis of the SURF Method. *Image Processing On Line*, volume 5, pages 176–218, juillet 2015. ISSN 2105-1232.
- Giuseppe Papari et Nicolai Petkov : Edge and line oriented contour detection : State of the art. *Image and Vision Computing*, volume 29, pages 79–103, février 2011. ISSN 02628856.
- Natan Peterfreund : Robust tracking of position and velocity with Kalman snakes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, volume 21, pages 564–569, 1999.
- Michael Rubinstein : *Analysis and Visualization of Temporal Variations in Video*. Thèse de doctorat, Massachusetts Institute of Technology, 2014.
- Ethan Rublee, Vincent Rabaud, Kurt Konolige et Gary Bradski : ORB : An efficient alternative to SIFT or SURF. *In Computer Vision (ICCV), 2011 IEEE International Conference On*, pages 2564–2571. IEEE, 2011.
- Angel D. Sappa et Fadi Dornaika : An edge-based approach to motion detection. *In Computational Science–ICCS 2006*, pages 563–570. Springer, 2006.
- Mehmet Sezgin et Bülent Sankur : Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, volume 13, pages 146–168, 2004.
- Amir H. Shabani, David A. Clausi et John S. Zelek : Evaluation of local spatio-temporal salient feature detectors for human action recognition. *In Computer and Robot Vision (CRV), 2012 Ninth Conference On*, pages 468–475. IEEE, 2012.
- Jianbo Shi et Carlo Tomasi : Good features to track. *In Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference On*, pages 593–600. IEEE, 1994.

- Eero P. Simoncelli et William T. Freeman : The steerable pyramid : A flexible architecture for multi-scale derivative computation. *In Image Processing, 1995. Proceedings., International Conference On*, volume 3, pages 444–447. IEEE, 1995.
- Arnold W. M. Smeulders, Dung M. Chu, Rita Cucchiara, Simone Calderara, Afshin Dehghan et Mubarak Shah : Visual Tracking : An Experimental Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 36, pages 1442–1468, juillet 2014. ISSN 0162-8828, 2160-9292.
- Chris Stauffer et W. Eric L. Grimson : Adaptive background mixture models for real-time tracking. *In Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference On.*, volume 2. IEEE, 1999.
- Ying-Li Tian et Arun Hampapur : Robust salient motion detection with complex background for real-time video surveillance. *In Application of Computer Vision, 2005. WACV/MOTIONS'05 Volume 1. Seventh IEEE Workshops On*, volume 2, pages 30–35. IEEE, 2005.
- Pauline Trouvé : *Digital-Optical Codesign of an Imaging System with 3D Ability*. Theses, Ecole Centrale de Nantes (ECN), décembre 2012.
- Paul Viola et Michael Jones : Rapid object detection using a boosted cascade of simple features. *In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference On*, volume 1, pages I–I. IEEE, 2001.
- Chaohui Wang, Martin de La Gorce et Nikos Paragios : Segmentation, ordering and multi-object tracking using graphical models. *In Computer Vision, 2009 IEEE 12th International Conference On*, pages 747–754. IEEE, 2009.
- Demin Wang, André Vincent, Philip Blanchfield et Robert Klepko : Motion-Compensated Frame Rate Up-Conversion - Part II : New Algorithms for Frame Interpolation. *IEEE Transactions on Broadcasting*, volume 56, pages 142–149, juin 2010a. ISSN 0018-9316, 1557-9611.
- Demin Wang, Liang Zhang et André Vincent : Motion-Compensated Frame Rate Up-Conversion - Part I : Fast Multi-Frame Motion Estimation. *IEEE Transactions on Broadcasting*, volume 56, pages 133–141, juin 2010b. ISSN 0018-9316, 1557-9611.
- Oliver Wang, Martina Dumcke, Aljoscha Smolic et Markus Gross : Distinguishing Texture Edges From Object Boundaries in Video. *IEEE Transactions on Image Processing*, volume 22, pages 5063–5070, décembre 2013. ISSN 1057-7149, 1941-0042.
- Bo Yan et Yue Chen : Low complexity image interpolation method based on path selection. *Journal of Visual Communication and Image Representation*, volume 24, pages 661–668, août 2013. ISSN 10473203.
- Alper Yilmaz, Omar Javed et Mubarak Shah : Object tracking : A survey. *ACM Computing Surveys*, volume 38, pages 13–es, décembre 2006. ISSN 03600300.

# 3

## Analyse de vidéos par approche lagrangienne basée sur un modèle dynamique physique

*Notre esprit ne se soucie pas de la vérité. C'est une chaîne d'information en continu qui déforme constamment la réalité.*

– Shane Kuhn

### Sommaire

---

I	Apports d'un modèle dynamique . . . . .	88
II	L'assimilation de données : une famille de méthodes pour combiner modèle et expérience . . . . .	90
III	Courte Kalmanographie pour les expérimentateurs à vocations non linéaires	102
IV	Proposition d'un schéma d'assimilation de vidéos rapides sur des modèles dynamiques complexes et réduits . . . . .	132
V	Conclusion . . . . .	149
	Bibliographie du chapitre 3 . . . . .	150

---

### Résumé

En remarquant qu'il est possible de formaliser la connaissance de la scène par un modèle physique, nous proposons de lier ce modèle aux observations faites sur le système réel à l'aide de caméras à l'aide de méthodes d'assimilation de données. L'utilisation de ce modèle permet ainsi de venir augmenter les acquisitions et faire à la fois de l'interpolation temporelle, de la prédiction tout en identifiant des paramètres physiques du modèle. Ce chapitre présente les méthodes d'assimilation de données sous un double éclairage. Ces méthodes étant utilisées dans un nombre important de domaines, une synthèse de l'état de l'art multidomaine est proposée ici avec une focalisation particulière sur la sous-famille d'assimilation séquentielle. Connues également sous le nom de filtres de Kalman ces schémas sont utilisables en temps réel. Nous proposerons ensuite un cadre d'association entre ces méthodes, des observations par caméra et des modèles mécaniques sous forme différentielle. Une validation sur un cas test réel de règle glissante nous permettra de conclure.



### Chap. 3 : Analyse de vidéos par approche lagrangienne basée sur un modèle dynamique physique

Les modèles utilisés dans le chapitre 2 se sont avérés trop génériques pour permettre de caractériser suffisamment bien les scènes. Ce chapitre réintroduit une hypothèse formulée au début de ces travaux sur la connaissance à priori du scénario de la scène. Cette hypothèse nous permet de développer un modèle mécanique souvent sous forme différentielle spécifique à une scène. Dans le but d'utiliser le modèle avec les observations, nous avons choisi de développer les méthodes d'assimilation de données (section II). Ces méthodes permettent l'utilisation conjointe de modèles d'état de complexité et de taille importante dans les domaines de la météorologie ou l'océanographie. Nous nous intéresserons plus spécifiquement aux schémas d'assimilation dits « séquentiels » connus également comme filtres de Kalman, autorisant l'estimation de l'état courant en fonction des observations passées. Nous proposons dans une troisième partie (section III) un éclairage stochastique afin de présenter sous forme d'outils un certain nombre d'extensions de ces filtres, leurs évolutions à travers les différents domaines et leurs potentiels d'utilisation. La section IV présente un schéma d'assimilation permettant de lier un modèle mécanique sous forme différentielle avec des observations vidéos comportant la position et la vitesse de points spécifiques de l'objet observé. L'application sur une règle glissante illustre l'identification sur des paramètres invisibles comme le coefficient de frottement, la prédiction sur la position d'arrêt de la règle, et l'interpolation selon la physique de la scène. Ce cadre autorise alors une compression sémantique à la scène et l'utilisation de réalité augmentée. La figure 3.1 présente graphiquement les étapes suivies par ce chapitre.

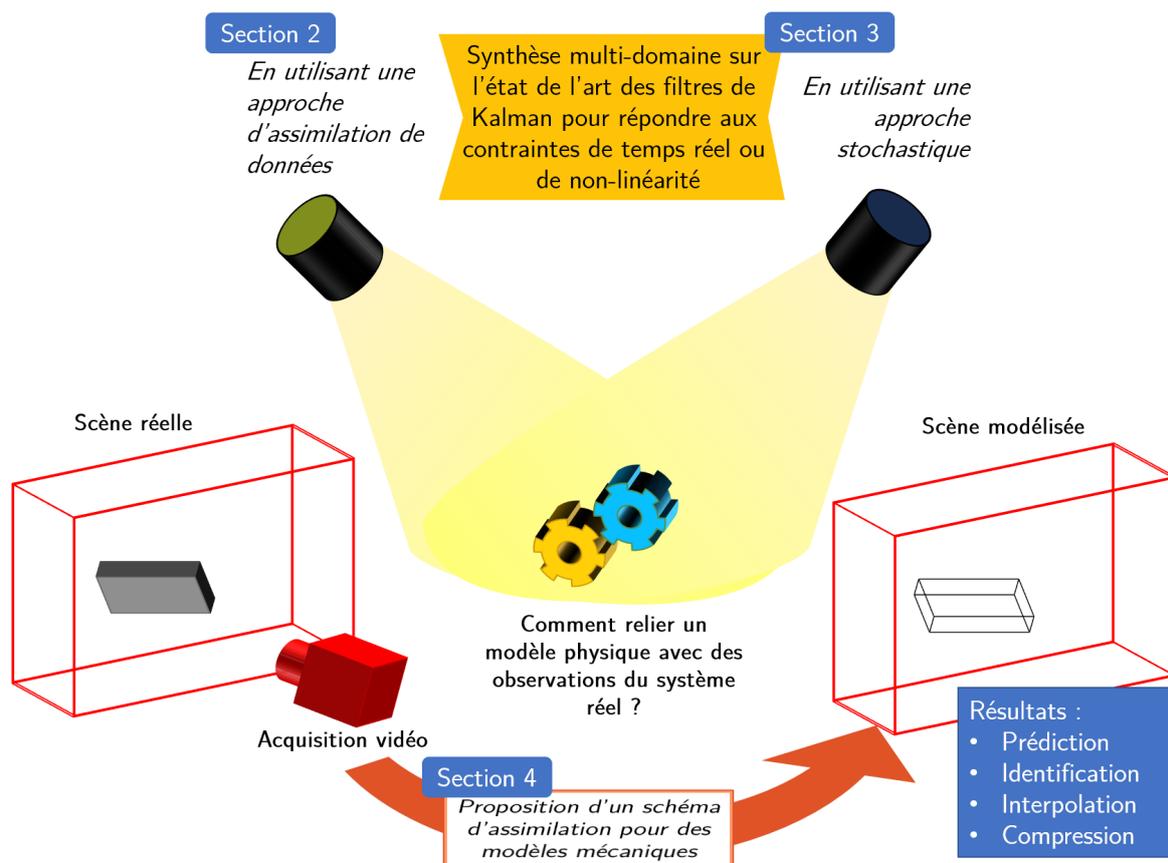


Figure 3.1 – Résumé graphique présentant la démarche et les résultats du chapitre 3

### I Apports d'un modèle dynamique

Pour répondre aux exigences du projet FAMOSYS, le chapitre 2 développe une méthode permettant de faire de l'interpolation et de la compression en se basant sur des hypothèses particulières aux vidéos rapides :

- un fond fixe ;
- un ou plusieurs objets déformables dont le contour et l'intérieur évoluent lentement par rapport à sa cinématique ;
- un scénario quelconque non connu.

Cette première approche développe une méthode cinématique basée sur la composition de mouvement entre les grands déplacements et une approximation des petits déplacements dans le repère objet par de l'estimation de mouvements. Le modèle utilisé ne permet d'apporter que peu d'information, ce qui rend les objectifs de prédiction et d'identification inatteignables par cette méthode. Cependant les hypothèses initiales du projet FAMOSYS étaient plus restrictives puisque **les scènes vidéos à analyser sont supposées être connues mais non quantifiées**. Nous allons ainsi pouvoir définir précisément un modèle d'évolution temporelle du ou des objets considérés dans la scène. L'exploitation de ce modèle nous permettrait ainsi d'ajouter de l'information afin d'identifier, de prédire, et d'interpoler de manière plus fine au sens du modèle.

#### I.A. Définition des modèles physiques

Cette nouvelle hypothèse de scène connue à l'avance, classique en vidéo rapide, nous permet ici d'introduire ce que nous appellerons un modèle physique par opposition au modèle cinématique élémentaire introduit précédemment. L'imprévisibilité du cadre vidéo du chapitre précédent nous a obligé à l'instar de l'état de l'art, à travailler sur des modèles d'évolution élémentaires compte tenu des hypothèses. Ces modèles permettent notamment de grouper les pixels ayant un mouvement similaire dans une image. La relation entre l'évolution de la trajectoire d'un pixel et celle de sa projection sur l'objet réel qu'il représente étant trop complexe pour être reconstruite en fonction des observations, des modèles de suivi de trajectoires des pixels sont utilisés, apportant une interprétation cinématique au niveau des pixels. Notre approche de composition de mouvement n'est pas satisfaisante de ce point de vue. Même si elle se place à un plus haut niveau en considérant le déplacement d'un objet vidéo plutôt que les déplacements de pixels, le modèle cinématique utilisé est adaptable et élémentaire au lieu d'être spécifique en traduisant sa physique.

Les hypothèses de ce chapitre nous offrent une connaissance sur l'objet en terme de géométrie, de comportement dynamique et sur le scénario qu'il va rencontrer. Nous pourrions utiliser les modèles élémentaires de la partie précédente, mais ils ne permettent pas de rendre compte de cette nouvelle hypothèse de connaissance. C'est pourquoi nous avons choisi de formaliser cette connaissance sur l'objet et le scénario à l'aide d'un modèle mécanique. Ces modèles sont 2D ou 3D, ils font intervenir des variables et des paramètres internes, c'est-à-dire non mesurables par des mesures vidéos. Ils pourront être exprimés sous forme de modèle d'état issus des équations différentielles de mécanique. Ces modèles sont classiquement non-linéaires et ce sont ces différentes caractéristiques qui amènent à définir les modèles utilisés dans ce paragraphe comme complexes.

Nous proposons ainsi de formaliser ces modèles dynamiques sous une forme générale de **modèles d'états**. Les modèles obtenus utilisent un vecteur de commande  $u(t)$  et un vecteur d'état  $x(t)$  comprenant des variables et des paramètres. Les paramètres sont définis par le fait que leurs états restent

constants au cours du temps. Le modèle d'état s'exprime alors sous la forme de deux équations : une équation d'évolution et une équation d'observation. Dans un formalisme à la fois non-linéaire et continu, ces équations s'expriment sous la forme<sup>1</sup> :

$$\begin{cases} \dot{x}(t) = \mathcal{F}(x(t), u(t)) \\ z(t) = \mathcal{H}(x(t)) \end{cases} \quad (3.1)$$

Ces équations font ainsi intervenir une fonction d'évolution  $\mathcal{F}$ , une fonction d'observation  $\mathcal{H}$  et les observations du système  $z(t)$ . Sans équivalence avec le système précédent, en posant  $k$  le pas du système, on a le système d'état discret :

$$\begin{cases} x_{k+1} = f(x_k, u_k) \\ z_k = h(x_k) \end{cases} \quad (3.2)$$

Ces modèles font principalement intervenir des variables issues de la dynamique du mouvement. Ces variables peuvent faire partie du vecteur d'état comme pouvant être internes à la fonction  $f$ . Nous pourrions par ailleurs définir des applications d'observations spécifiques permettant d'obtenir pour un instant donné le vecteur des paramètres et le vecteur des variables à partir du vecteur d'état.

L'utilisation d'un modèle dynamique va permettre l'estimation de paramètres et de variables issues du modèle. Ces informations pourront être utilisées pour augmenter la vidéo, comme pour identifier des paramètres spécifiques.

Comme une large partie des méthodes employées faisant intervenir des modèles numériques, il s'agit d'identifier une série de paramètres du modèle tout en faisant correspondre les variables observées. Le recalage de modèle pour l'identification de paramètres est un domaine où la littérature est abondante, et nous citerons la thèse de Pottier (2010) qui détaille les grandes méthodes utilisées.

Cependant, le cadre d'étude défini par le projet FAMOSYS nous impose à la fois de faire de la prédiction, de l'interpolation, et de l'extrapolation dans un contexte temps réel. Certains domaines ont déjà ce type de contraintes. C'est par exemple le cas pour la navigation de robots qui doit permettre d'utiliser sa localisation et les faire correspondre à une cartographie (en anglais *Simultaneous Localization And Mapping* ou *SLAM*). La résolution de ce type de problèmes se fait classiquement en utilisant dans un contexte bayésien des méthodes séquentielles : le filtre de Kalman ou le filtre particulaire (détaillés dans la section III). On pourrait pourtant s'interroger sur la capacité de ces méthodes à utiliser des modèles plus complexes puisque le *SLAM* n'utilise qu'un modèle cinématique.

Si l'on enlève la problématique de temps réel, nos problématiques sont similaires à celles rencontrées dans les domaines des géosciences, de l'océanographie et surtout de la météorologie. Ces domaines manipulent de très gros modèles à la fois locaux et globaux et ont pour objectif de faire à la fois un filtrage sur les mesures de capteurs de nature ou de mesurandes différentes, de la prédiction ou de l'extrapolation. Ces domaines utilisent un cadre exploré en parallèle de celui bayésien développé dans la communauté du suivi ou de la théorie du contrôle en automatique : l'assimilation de données que nous allons définir dans la partie qui va suivre.

---

1. Pour faciliter les développements et les lourdeurs d'écriture, nous avons négligé le cas peu courant où la fonction d'observation dépend également de la commande.

## II L'assimilation de données : une famille de méthodes pour combiner modèle et expérience

Le principe de faire correspondre une série d'acquisitions avec un modèle a été développé dans des domaines très variés. Avant de détailler ces multiples approches dans la section suivante, nous proposons d'introduire le point de vue de l'assimilation de données, une famille de méthodes dont nous présenterons les caractéristiques avant de voir les rapprochements possibles avec notre problématique.

### II.A. Principes de l'assimilation

#### II.A.1. Introduction d'une famille de méthodes liant la modélisation à l'expérimental

Pour introduire cette famille de méthode, nous proposons de commencer par un exemple aussi courant qu'essentiel tout au long de l'histoire : les prévisions météorologiques.

Les conditions météorologiques ont une influence considérable sur nos différentes sociétés à travers l'histoire. De par l'importante complexité des phénomènes entrant en jeu, les conditions météorologiques sont longtemps restées mystérieuses.

Entre grands événements comme la bataille de Waterloo ou encore les mauvaises récoltes précédant la révolution française, la maîtrise des prévisions météorologiques s'est rapidement imposée comme essentielle pour assurer la stabilité des sociétés, mais également un symbole inconscient de puissance. Longtemps gouvernée par les dictons, ce sont les inventions telles que les baromètres, les thermomètres ou les sondes atmosphériques, qui vont permettre aux scientifiques d'observer et de quantifier les phénomènes climatiques à partir du début du XX<sup>ème</sup> siècle. C'est à partir de ces mesures que la science météorologique a pu se développer. La météorologie étudie les lois régissant la dynamique des fluides des différentes couches atmosphériques afin de prédire leur comportement futur. Cette prédiction a conduit au développement de modèles dynamiques pour caractériser les différents phénomènes en jeu.

La prédiction météorologique peut donc se faire de différentes manières. Ainsi, à l'instar des dictons et proverbes, une prédiction possible peut être basée sur une vision statistique. Bien qu'inconsciente, cette prédiction statistique est également associée à une hypothèse de localisation des mesures qui permet de réduire l'espace observable du phénomène et donc sa complexité. Cette première méthode statistique consiste à extrapoler les différentes variables en fonction de leurs tendances.

Une seconde approche peut se faire en utilisant les modèles dynamiques développés. Ces modèles permettent d'avoir une vision plus globale des phénomènes engendrant les mesures locales obtenues. Le modèle d'évolution permet cette fois de rendre compte du caractère temporel de l'évolution des données. Cependant, la complexité de ces modèles est souvent difficile à manipuler et leur utilisation demande une initialisation précise pas toujours facile à obtenir.

Daget (2008) propose de reprendre la démarche scientifique face à un système inconnu pour mieux mettre en valeur l'éventail des solutions offertes par l'assimilation de données. Une première manière de comprendre un système inconnu est de développer un modèle paramétrique simulant les composantes internes du système. L'amélioration du modèle se fait ensuite par comparaison entre les

sorties du modèle et les résultats expérimentaux. Une deuxième manière plus inspirée de l'automatique considère le système comme une boîte noire et consiste à construire une fonction de transfert du système en fonction de l'étude de ses entrées/sorties.

Ces deux approches proposent donc de résoudre un problème inverse<sup>2</sup>. L'assimilation de données propose également de résoudre un problème inverse sous un autre angle. Selon la jolie formule de Blayo et Nodet (2012) l'assimilation de données serait la « science des compromis fructueux ». Autrement, c'est un ensemble de méthodes qui permet de combiner de manière optimale ou non les informations disponibles sur un système :

- des modèles dynamiques ;
- des observations expérimentales ;
- des statistiques d'erreurs.

Ces méthodes permettent ainsi d'estimer l'**état caché** ou l'**état vrai** du système en prenant en compte les multiples erreurs pouvant biaiser les calculs : modèle imparfait, interpolation de données, erreurs d'observations. Ce processus d'estimation permet ainsi de faire du recalage de modèle en estimant les paramètres du modèle, de la modélisation inverse en estimant des séries de paramètres, mais également de l'analyse de données en utilisant le modèle comme un interpolateur ou de la prévision.

## II.A.2. Formalisme et notations

L'assimilation de données offre trois cadres différents :

- un cadre probabiliste élémentaire qui propose d'extraire les informations vraies à partir des statistiques sur les variables du vecteur d'état ;
- un cadre variationnel qui vise à minimiser une fonction de coût construite de manière à minimiser les erreurs sur le vecteur d'état. Ces méthodes supposent que les informations postérieures et antérieures sont connues ;
- un cadre séquentiel qui estime les états postérieurs en fonction d'estimateurs statistiques sur les variables du vecteur d'état. L'estimation de l'état postérieur s'effectue uniquement en fonction de la connaissance des états antérieurs.

Les méthodes peuvent être appliquées de manière intermittente, en utilisant de petits intervalles de temps ou de manière continue (figure 3.2). Les données fournies à ces méthodes sont importantes puisqu'elles permettent d'agréger des données non homogènes sous certaines hypothèses. Le respect de ces hypothèses mais également la réduction en coût ou en complexité des calculs peut également amener à considérer une sélection locale des données en fonction du point d'analyse.

L'analyse est l'étape qui permet d'estimer un état à un instant donné. Voici un inventaire des variables utilisées en assimilation de données :

- Nous noterons  $x^a$  le vecteur d'état analysé.
- L'analyse est effectuée à partir d'une ébauche que nous noterons  $x^b$ , c'est-à-dire une estimation à priori de l'état. Cette ébauche peut par exemple dans un cas séquentiel, être issue d'un état d'analyse antérieur.
- L'état caché du système continu est noté  $x^{tc}$  et les erreurs de discrétisations en font un état caché discrétisé que nous noterons  $x^t$ .
- Dans l'espace des observables, on définit  $z^t$  le vecteur des observations cachées et  $z^o$  le vecteur d'observations mesuré.

---

2. Pour une définition claire et pédagogique de la notion de problème inverse, on pourra se référer à Pottier (2010).

## II. L'assimilation de données : une famille de méthodes pour combiner modèle et expérience

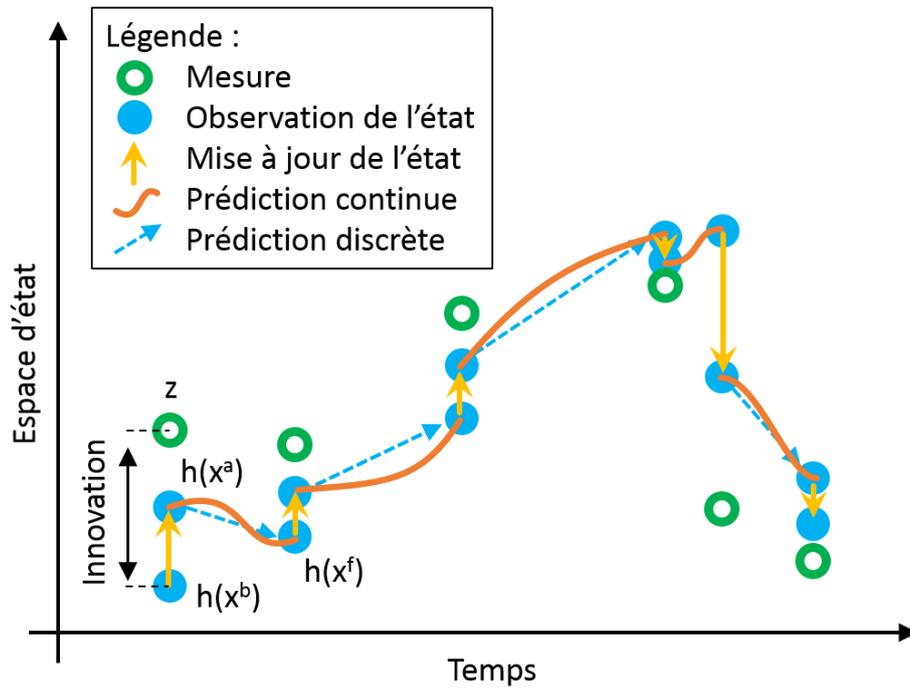


Figure 3.2 – Schéma synthétisant une assimilation de données

- On définit  $\mathcal{H}$  la fonction d'observation continue du système et  $H$  sa version discrétisée. On définit l'observation parfaite du système continue en posant  $z^t = \mathcal{H}(x^{tc})$ .
- On définit  $\mathcal{F}$  la fonction d'évolution du système. Elle est posée continue ici, mais peut tout aussi bien être discrète.

En pratique, le cas plus général d'un modèle d'observation faiblement non linéaire, c'est-à-dire dont les erreurs de linéarisation restent acceptables est pris. On définit la fonction d'observation par  $(x) \mapsto \mathcal{H}(x)$  avec  $\mathcal{H} : \mathbb{R}^n \mapsto \mathbb{R}^p$  de manière à travailler sur l'opérateur d'observation linéaire défini par la jacobienne de  $\mathcal{H}$  calculée à chaque instant en un point  $x = x_0$ . Pour simplifier l'écriture, nous la noterons ici  $H$  comme l'opérateur d'observation linéaire, mais nous l'appellerons dans la partie séquentielle  $H_k$  pour insister sur le fait que cet opérateur n'est qu'une version locale linéarisée de l'opérateur  $\mathcal{H}$ .

$$H \in \mathcal{M}_{n \times p} | [H]_{ij} = \frac{\partial \mathcal{H}}{\partial x_i}$$

avec  $\mathcal{M}_{n \times p}$  l'espace des matrices de taille  $n \times p$ .

Avec ces notations, il est possible de définir les différentes erreurs résidant entre l'estimation de ces différentes grandeurs. Ces erreurs sont résumées dans le tableau 3.1.

La figure 3.3 permet de synthétiser l'ensemble des variables utilisées. Elle détaille notamment l'erreur d'observation en faisant intervenir l'erreur de mesure  $\epsilon^m$ , une erreur induite par la discrétisation du vecteur d'état dans l'espace des observations que l'on appelle erreur de représentativité  $\epsilon^r$  et l'erreur de discrétisation de la fonction d'observation  $\epsilon^i$ .

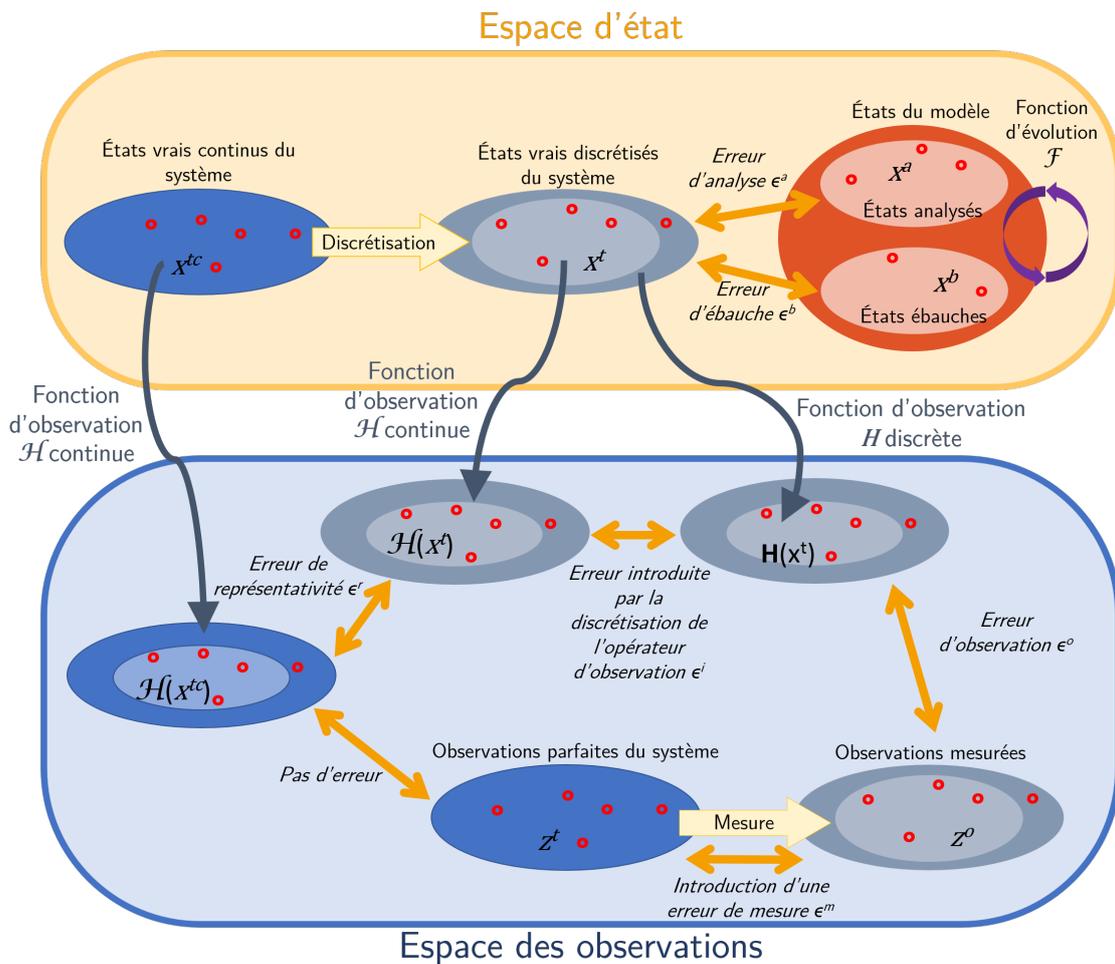


Figure 3.3 – Illustration synthétisant les différents espaces, variables et fonctions ainsi que les erreurs introduites en assimilation de données

### II.A.3. Une première méthode fondamentale : le BLUE

Une première vision statistique d'un estimateur d'une série est donnée par la définition du BLUE ou Best Linear Unbiased Estimator (meilleur estimateur linéaire sans biais). Cet estimateur est issu

Table 3.1 – Synthèse des différentes erreurs entrant en jeu durant le processus de l'assimilation de données

Nom	Définition	Moyenne	Covariance
Erreur de mesure	$\epsilon^m = z^o - z^t$	$\bar{\epsilon}^m$	-
Erreur de représentativité	$\epsilon^r = \mathcal{H}(x^{tc}) - \mathcal{H}(x^t)$	$\bar{\epsilon}^r$	-
Erreur de discrétisation pour H	$\epsilon^i = \mathcal{H}(x^t) - \mathbf{H}(x^t)$	$\bar{\epsilon}^i$	-
Erreur d'observation	$\epsilon^o = \epsilon^m + \epsilon^r + \epsilon^i$	$\bar{\epsilon}^o$	$R = E[(\epsilon^o - \bar{\epsilon}^o)(\epsilon^o - \bar{\epsilon}^o)^T]$
Erreur d'ébauche	$\epsilon^b = x^b - x^t$	$\bar{\epsilon}^b$	$B = E[(\epsilon^b - \bar{\epsilon}^b)(\epsilon^b - \bar{\epsilon}^b)^T]$
Erreur d'analyse	$\epsilon^a = x^a - x^t$	$\bar{\epsilon}^a$	$A = E[(\epsilon^a - \bar{\epsilon}^a)(\epsilon^a - \bar{\epsilon}^a)^T]$

## II. L'assimilation de données : une famille de méthodes pour combiner modèle et expérience

d'un cadre probabiliste dérivé du théorème de Gauss-Markov. Une généralisation a été développée par Aitken en 1935. Il s'applique sur des modèles linéaires uniquement.

Dans un cadre où l'on considère  $H$  comme linéaire, une estimation de l'état du système pourrait être une combinaison linéaire de l'observation et d'une ébauche :

$$x^a = Lx^b + Ky^o \text{ avec } L \in \mathcal{M}(\mathbb{R})_{n \times n}, K \in \mathcal{M}(\mathbb{R})_{n \times p} \text{ et } y^o = Hx^t + \epsilon^o$$

On peut alors définir l'erreur :

$$x^a - x^t = L(x^b - x^t + x^t) + K(Hx^t + \epsilon^o) - x^t \iff \epsilon^a = L\epsilon^b + K\epsilon^o + (L + KH - I)x^t$$

En prenant des hypothèses d'erreurs centrées :  $\overline{\epsilon^a} = \overline{\epsilon^b} = 0$

$$\begin{aligned} \text{Ce qui donne :} \quad L = I - KH &\iff x^a = (I - KH)x^b + Ky^o \\ &= x^b + K(y^o - Hx^b) \end{aligned}$$

On appelle alors l'innovation la quantité  $y^o - Hx^b$  et le gain la matrice  $K$  reliant les sous-espaces d'observation et d'état. Lorsque le gain est linéaire, le problème revient à trouver le gain optimal pour minimiser les erreurs.

Avec un gain optimal, on aurait alors pour les erreurs à postériori :  $\epsilon^a = \epsilon^b + K(\epsilon^o - H\epsilon^b)$  c'est-à-dire :

$$\begin{aligned} P^a &= E[(\epsilon^a)(\epsilon^a)^T] \\ &= E[(L\epsilon^b + K\epsilon^o)(L\epsilon^b + K\epsilon^o)^T] \\ &= LBL^T + KRK^T \end{aligned}$$

avec la supposition que les bruits  $\epsilon^b$  et  $\epsilon^o$  sont décorrélés et que  $K$  est linéaire. La covariance des erreurs à postériori revient alors à :

$$P^a = (I - KH)B(I_KH)^T + KRK^T$$

Améliorer l'estimation du système  $x^a$  par rapport à l'estimation précédente c'est minimiser l'erreur à postériori qui peut être mesurée par  $\text{Tr}(P^a)$ . Chercher une estimation optimale, c'est chercher un gain optimal, ce que l'on va faire en calculant la dérivée de la trace de la covariance de l'analyse et en supposant que  $B$  et  $R$  sont symétriques :

$$\begin{aligned} \frac{\partial(\text{Tr}(P^a))}{\partial K} &= \text{Tr}(-HBL^T - LBH^T + RK^T + KR) \\ &= \text{Tr}(RK^T - HBL^T) + \text{Tr}(KR - LBH^T) \\ &= 2 \text{Tr}(KR - LBH^T) \end{aligned}$$

On obtient alors le gain optimal  $K^*$  pour  $\frac{\partial \text{Tr}(P^a)}{\partial K}$

$$\begin{aligned} K^*R - LBH^T = 0 &\iff K^*(R + HB^TH) = BH^T \\ &\iff K^* = BH^T(R + HBH^T)^{-1} \end{aligned}$$

On appelle cet estimateur BLUE pour *Best Linear Unbiased Estimator* car il est linéaire, sans biais et optimal. Comme le lecteur pourra le voir dans la suite, cette manière de calculer le gain est similaire à celle utilisée pour les filtres de Kalman, mais également pour les méthodes variationnelles présentées dans la section suivante.

#### II.A.4. Assimilation variationnelle

Les méthodes variationnelles ont été développées afin de retrouver les mêmes résultats que le BLUE sous les mêmes hypothèses mais en posant le problème comme la minimisation d'une fonctionnelle. Sasaki (1958) pose les bases de la méthode 3D Var qui consiste à minimiser au sens des moindres carrés les différences entre un état estimé et les sources d'observations disponibles. Cette minimisation se fait à un pas de temps fixe. Lorenc (1986) pose une variante de la méthode en effectuant des analyses à plusieurs pas de temps sur une plage temporelle afin d'obtenir non pas l'état optimal sur un pas de temps donné mais la trajectoire optimale sur une fenêtre de temps donnée. Les méthodes variationnelles sont actuellement très utilisées dans les systèmes de prévision météorologiques.

##### a. Méthode 3D-Var

L'approche variationnelle s'attache à optimiser un gain en minimisant une fonction de coût. On suppose une telle fonction telle que  $g : \mathbb{R}^n \mapsto \mathbb{R}$ . Pour étudier sa variation au point  $x_0$ , on étudie son développement limité :

$$g(x) = g(x_0) + (x - x_0)^T \nabla g|_{x_0} + \frac{1}{2} (x - x_0)^T \nabla^2 g|_{x_0} (x - x_0) + o(\|x - x_0\|^2)$$

avec  $\nabla g|_{x_0}$  le gradient de la fonction et  $\nabla^2 g|_{x_0}$  la hessienne généralisant la dérivée seconde :

$$[\nabla g|_x]_i = \frac{\partial g}{\partial x_i} \quad \text{et} \quad [\nabla^2 g|_x]_{ij} = \frac{\partial^2 g}{\partial x_i \partial x_j}$$

On définit la fonction de coût suivante :

$$J(x) = (x - x^b)^T B^{-1} (x - x^b) + \frac{1}{2} (z^o - Hx)^T R^{-1} (z^o - Hx)$$

En supposant la fonction  $H$  linéaire, et les fonction  $B$  et  $R$  définies positives la fonction coût  $J$  est convexe et possède un point de selle pour  $\nabla J(x) = 0$  :

$$\begin{aligned} \nabla J(x) &= B^{-1} (x - x^b) - H^T R^{-1} (z^o - Hx) \\ \nabla J(x) = 0 &\iff x^* = x^b + (B^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1} (z^o - Hx^b) \end{aligned}$$

La formule de Sherman-Morrison-Woodbury nous dit :

$$BH^T (HBH^T + R)^{-1} = (B^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1}$$

On a bien équivalence avec le BLUE exposé précédemment en posant

$$K^* = (B^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1}$$

Mettre le problème sous un cadre variationnel comporte plusieurs avantages significatifs :

## II. L'assimilation de données : une famille de méthodes pour combiner modèle et expérience

- Cette fonction de coût a comme intérêt majeur de pouvoir être minimisée par une grande variété de méthodes.
- Une autre propriété très intéressante concernant la précision de l'estimation peut être exploitée :

$$\nabla^2 J(x) = B^{-1} + H^T R^{-1} H \text{ soit } P^a = \left( \nabla^2 J(x) \right)^{-1}$$

- Le cadre variationnel permet de s'abstraire de l'hypothèse de linéarité faite pour le BLUE. Avec une approche variationnelle, l'équation du gradient permet d'utiliser une fonction non-linéaire et introduirait la jacobienne  $H_j$  :

$$\nabla J(x) = B^{-1}(x - x^b) - H_j^T R^{-1}(z^o - h(x)) \quad (3.3)$$

- On peut facilement introduire avec ce formalisme des contraintes sur les solutions en ajoutant des multiplicateurs de Lagrange.

### b. 4D-Var

La méthode 4D-Var est une extension de la méthode 3D-Var qui utilise cette fois une fenêtre temporelle afin de trouver une trajectoire optimale à défaut d'optimiser chaque instant. Cette méthode est très coûteuse et se trouve être particulièrement éloignée de toute application temps réel. Cependant, il est intéressant de la mentionner, surtout pour souligner le changement de démarche sur les quantités optimisées.

Succinctement, voici les différences majeures dans l'approche. De la même manière que le 3D-Var, la fonction coût se décompose en un terme lié à l'ébauche,  $J^b(x)$  et un terme lié aux observations  $J^o(x)$  à la différence que  $x^a$  et  $x^b$  dépendent désormais du temps. On a pour la fonction de coût sur la fenêtre  $[t_0; t_N]$  :

$$\begin{cases} J^b(x) = \frac{1}{2}(x - x^b(t_0))^T B^{-1}(x - x^b(t_0)) \\ J^o(x) = \frac{1}{2} \sum_{i=0}^N (z_i - H_i x(t_i))^T R_i^{-1} (z_i - H_i x(t_i)) \end{cases} \quad (3.4)$$

Il n'est pas possible de résoudre en l'état cette deuxième partie de fonctionnelle, du fait que l'on ne connaît pas la valeur de l'état  $x(t_i)$ . L'avantage majeur des méthodes 4D-Var est d'introduire un modèle d'évolution permettant de propager l'état de  $t_0$  à  $t_i$ . Ce modèle d'évolution peut être exprimé comme une fonction non-linéaire  $f$  :

$$x(t_i) = f_{0 \rightarrow i}(x(t_0)) \quad (3.5)$$

Dans le cas où on considère le modèle d'évolution parfait, l'équation (3.4) peut alors s'exprimer :

$$\begin{cases} J^b(x) = \frac{1}{2}(x - x^b(t_0))^T B^{-1}(x - x^b(t_0)) \\ J^o(x) = \frac{1}{2} \sum_{i=0}^N (z_i^o - H_i f_{0 \rightarrow i}(x(t_0)))^T R_i^{-1} (z_i^o - H_i f_{0 \rightarrow i}(x(t_0))) \end{cases} \quad (3.6)$$

La résolution devient alors une minimisation de fonctionnelle de manière similaire à la méthode 3D-Var. Le gradient s'exprime en fonction de la linéarisation de la fonction d'évolution au voisinage de l'ébauche  $F_{0 \rightarrow i}$  :

$$\nabla J(x) = B^{-1}(x - x^b) - F_{0 \rightarrow i}^T H_i^T R_i^{-1} (z^o - h(f_{0 \rightarrow i}(x))) \quad (3.7)$$

Cependant, l'expérience acquise dans les domaines d'applications de l'assimilation de données a montré que les systèmes étudiés peuvent se révéler extrêmement complexes et que des modèles même sophistiqués échouent à produire des erreurs négligeables. On introduit donc ici une dernière erreur du modèle cette fois-ci, définie par :

$$q_i = x^t(t_i) - f_{i-1 \rightarrow i}(x^t(t_{i-1})) \quad (3.8)$$

En introduisant sa covariance associée  $Q_i$ , on peut alors réécrire la fonction de coût avec un terme correctif supplémentaire lié à cette erreur :

$$\begin{aligned} J(x) &= \frac{1}{2}(x - x^b(t_0))^T B^{-1}(x - x^b(t_0)) \\ &+ \frac{1}{2} \sum_{i=0}^N (z_i^o - H_i f_{0 \rightarrow i}(x(t_0)))^T R_i^{-1} (z_i^o - H_i f_{0 \rightarrow i}(x(t_0))) \\ &+ \frac{1}{2} \sum_{i=0}^N q_i^T Q_i^{-1} q_i \end{aligned} \quad (3.9)$$

Cette fonction de coût est dite à contrainte faible et est utilisée dans de nombreux travaux. On remarquera que le 4D-Var à contrainte faible a de nombreuses similitudes avec les méthodes d'assimilation séquentielles que nous allons introduire dans la partie qui va suivre.

### II.A.5. Assimilation séquentielle

L'assimilation de données variationnelles permet d'exprimer sous une forme de minimisation de fonctionnelle le problème de la minimisation des erreurs entre un système d'état et une série d'acquisitions. Cependant, au-delà des coûts de calcul élevés, l'assimilation variationnelle se sert des acquisitions antérieures et postérieures à l'instant d'analyse. Ce schéma n'est donc pas idéal pour une utilisation en temps réel. Le développement de l'assimilation séquentielle a été motivé à la fois par l'introduction d'une fonction d'évolution reprise ensuite dans la méthode 4D-Var et un aspect séquentiel qui permet de limiter les coûts de calculs et de mémoire.

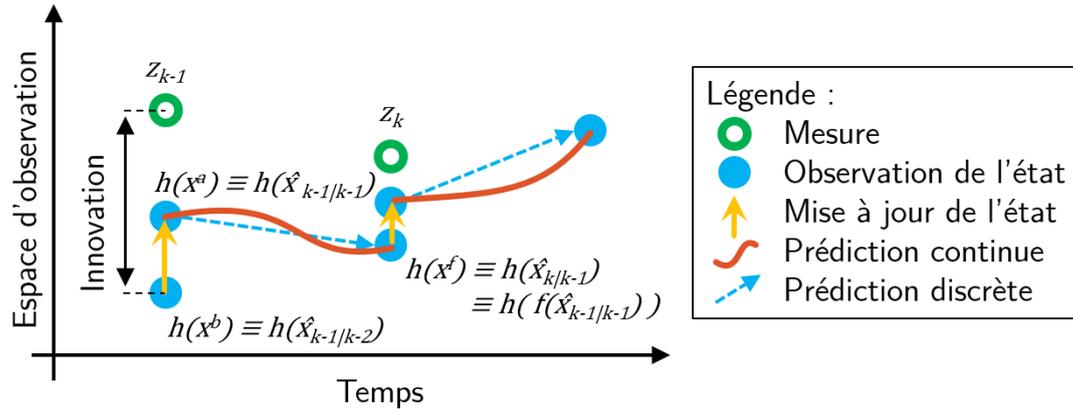
Les méthodes d'assimilation séquentielles sont basées sur une célèbre solution récursive pour les problèmes de filtrage discrets : les filtres de Kalman. Bien que la paternité de ces filtres reste encore sujet à controverse, la version discrète de ces filtres sont communément appelés filtres de Kalman (1960b) tandis que la version continue est généralement nommée filtres de Kalman-Bucy (Kalman et Bucy, 1961). En reprenant le formalisme développé dans le paragraphe précédent, on peut voir le filtre de Kalman comme étant une extension du BLUE pour laquelle l'état analysé d'un état antérieur définit l'ébauche de l'état courant. Ainsi les notations sont très sensiblement différentes puisqu'on utilise la fonction d'évolution pour effectuer une prédiction (d'où l'introduction de l'indice  $f$  pour *forecast*). Ces équivalences sont résumées sur la figure 3.4.

On a les équations d'un système physique discret linéaire pour l'évolution et l'observation :

$$\begin{cases} x_{k+1}^f &= Fx_k + w_k \\ z_k &= Hx_k^f + v_k \end{cases} \quad (3.10)$$

La signification de ces équations est sensiblement différente de celles du système d'état définies équation (3.10). Deux modèles linéaires et discrets sont considérés ici, les fonctions  $F$  et  $H$ . L'écriture

## II. L'assimilation de données : une famille de méthodes pour combiner modèle et expérience



**Figure 3.4** – Schéma comparant les étapes d'assimilation de données avec le formalisme stochastique

ci-dessus implique que ces modèles ne sont pas parfaits, ce qui est représenté par l'ajout de bruits gaussiens centrés non-corrélés pour l'évolution :

$$\exists w_k \quad | \quad E[w_k] = 0 \text{ et } E[w_k \cdot w_l] = Q_k \cdot \delta_{kl} \quad (3.11)$$

et pour l'observation :

$$\exists v_k \quad | \quad E[v_k] = 0 \text{ et } E[v_k \cdot v_l] = R_k \cdot \delta_{kl} \quad (3.12)$$

Ces définitions nous permettent d'introduire la matrice de covariance  $Q_k$  qui définit l'erreur du modèle et celle de l'observation  $R_k$ . On a noté  $\delta_{kl}$  la fonction de Kronecker.

Le filtre de Kalman discret et linéaire entre l'instant  $t_k$  et  $t_{k+1}$  peut se synthétiser en différentes étapes :

— Calcul du gain au temps  $t_k$  :

$$K_k = P_k^f H^T (H P_k^f H^T + R_k)^{-1} \quad (3.13)$$

— Analyse au temps  $t_k$  :

$$x_k^a = x_k^f + K_k (z_k^o - H x_k^f) \quad (3.14)$$

$$P_k^a = (I - K_k H) P_k^f \quad (3.15)$$

— Prévision au temps  $t_{k+1}$  par propagation de l'analyse effectuée en  $t_k$  :

$$x_{k+1}^f = F(x_k^a) \quad (3.16)$$

$$P_{k+1}^f = F P_k^a F^T + Q_k \quad (3.17)$$

$$(3.18)$$

Nous avons introduit dans ce paragraphe une vision linéaire du filtre de Kalman, que nous allons développer sous d'autres formes non-linéaires dans la section III. Ces versions non-linéaires sont

très utilisées en assimilation de données même si nous ne les introduisons que dans un formalisme plus stochastique. Entre autres, les versions Kalman étendu (*Extended Kalman Filter* ou *EKF*) et Kalman d'Ensemble (*EnKF*) sont des versions du filtre largement utilisées en assimilation de données séquentielles. Le filtre EnKF a d'ailleurs été développé spécifiquement dans ce cadre.

Malgré la restriction de l'analyse sur un pas de temps, il s'avère en pratique que le filtre reste coûteux sur des systèmes de grande taille comme en océanographie ou en météorologie (de l'ordre de  $10^7 \times 10^7$  éléments dans la matrice de covariance). Daget (2008) précise que l'étape de loin la plus onéreuse se trouve être l'inversion matricielle effectuée lors du calcul du gain. C'est en partie à cause de ce coût prohibitif que l'algorithme de Kalman présenté ici ne peut être appliqué à des systèmes aussi grands que pour les autres méthodes. Finalement, à travers une approche séquentielle, la philosophie d'assimilation est sensiblement différente puisque l'utilisation du modèle dans les données peut être également orientée comme un filtrage : c'est sous cet autre angle que nous exposerons une série choisie de versions de l'algorithme.

## II.B. Applications du cadre de l'assimilation de données dans d'autres domaines

Les problématiques du projet FAMOSYS définies dans le chapitre 1 peuvent être résolues par de nombreuses méthodes. Entre la théorie du contrôle, l'estimation stochastique ou la minimisation d'une fonctionnelle, le problème peut être résolu par des cadres bien différents. C'est pour cette raison que nous avons choisi l'approche proposée par l'assimilation de données. Au-delà d'être un domaine regroupant des méthodes issues de différents horizons, ces méthodes sont développées pour des buts similaires aux nôtres : prédiction, identification, interpolation ; ceci dans un cadre où les modèles utilisés sont à la fois complexes et les vecteurs d'état importants. Ce cadre nous a semblé être le plus judicieux, d'autant plus qu'il est développé par une communauté expérimentée sur ces problématiques. Cependant la diversité de la littérature nous a mené à proposer une double approche : nous compléterons plus tard en repartant du contexte stochastique.

### II.B.1. Morceaux choisis de l'état de l'art

Il nous a semblé pertinent de ne pas présenter tout de suite un état de l'art exhaustif sur les filtres de Kalman mais plutôt de nous attarder sur les travaux récents de quelques équipes ayant eu une démarche similaire à la nôtre en voulant « exporter » le cadre de l'assimilation de données dans d'autres domaines. Il est important de noter que cette partie donne un avant-goût des notions développées dans la section III. Le lecteur pourra s'y référer.

#### a. L'assimilation de données en dynamique des structures

Dans un contexte de développement des modèles éléments finis en plein essor en 1975, Ladeveze (1977) propose un estimateur permettant de qualifier la confiance que l'utilisateur peut accorder à la solution : l'erreur en relation de comportement (ERC). Basé sur cet estimateur, des méthodes d'identification basées sur le recalage de modèles dynamiques sont proposées par Chouaki et Ladeveze (1998) à posteriori. Plus récemment, Nguyen *et al.* (2008) a proposé un cadre plus robuste pour faire de l'identification sur des données corrompues. Pour ce faire il introduit une ERC modifiée (mERC). Cot (2012) propose alors d'utiliser la fonctionnelle définie par l'ERC dans le cadre de l'assimilation de données. L'utilisation d'une fonctionnelle lui permet ainsi d'apporter des solutions stables aux problèmes de bords, en contraignant la fonctionnelle par des multiplicateurs de Lagrange.

## II. L'assimilation de données : une famille de méthodes pour combiner modèle et expérience

Les applications proposées dans sa thèse sont orientées sur des structures civiles, le partenaire étant EDF ; elles utilisent donc des modèles éléments finis qui pénalisent le temps d'exécution. Il utilise ainsi une méthode séquentielle d'estimation de paramètres basée sur un filtre de Kalman non-parfumé (*Unscented Kalman Filter* ou *UKF*), que nous décrivons dans la section III. Les résultats qu'il propose mettent en exergue la possibilité de travailler conjointement sur un modèle et sur les acquisitions obtenues sur le système réel.

Marchand *et al.* (2015) reprennent le cadre de l'assimilation de données précédemment introduit et posent leurs travaux par rapport au paradigme du dialogue essais-calcul (*Dynamic Data Driven Application System* ou *DDDAS* décrit par Darema (2004)). L'idée de ce dialogue est de créer des systèmes offrant un échange entre un système physique et un modèle numérique pour tirer profit de cette association dans un second temps une fois le modèle recalé. Sans que l'assimilation de données soit particulièrement mise en avant dans ses travaux, le rapprochement entre les objectifs du *DDDAS* et ceux de l'assimilation de données est rapide à faire. L'approche proposée se situe plus dans un contexte temps réel, ce qui incite à utiliser une méthode d'assimilation séquentielle, en l'occurrence un filtre *UKF*. La structure du modèle d'état est choisie sous forme duale et permet de faire intervenir la fonctionnelle proposée par la mERC pour estimer les paramètres. Jusqu'ici le cadre d'étude est similaire à celui proposé par Cot, et s'avère difficile à utiliser dans un cadre temps réel. Marchand *et al.* (2015) proposent ainsi de faire intervenir une étape de réduction de modèle dite PGD (*Proper Generalized Decomposition*) recalculée sur des intervalles de temps. Cette méthode permet en outre d'effectuer un découplage entre les paramètres, le temps et l'espace. Pour finir dans le domaine de l'identification de paramètres sur des structures, nous avons déjà introduit l'article de Hild *et al.* (2016) sur la corrélation d'images dans le chapitre 2 et notamment sur une approche globale orientée structure de la corrélation croisée. Dans cet article récent, il est proposé de coupler des résultats de mesures vidéos 3D avec un modèle mécanique élément fini via l'utilisation d'un algorithme de corrélation : le *Digital Volume Correlation (DVC)* pour donner une vision spatiotemporelle. Malgré la proximité de ces résultats avec les problématiques de l'assimilation, les perspectives ne sont pas les mêmes.

Selon la manière de poser le problème, l'assimilation de données est plus ou moins utilisée dans un domaine ou l'utilisation couplée de modèles avec des données se fait ressentir comme une évidence pour augmenter les informations contenues dans les mesures ou filtrer tout en donnant un sens physique aux modèles numériques.

### b. En biomécanique

L'introduction de méthodes d'assimilation de données dans le domaine de la biomécanique revient à l'équipe MACS de l'INRIA (Chapelle *et al.*, 2006) avec la thèse de Moireau (2009). Ses travaux fournissent de nombreux apports théoriques et pratiques sur l'assimilation de données. La réflexion proposée met en relation les différents aspects de ces méthodes par rapport aux différents domaines où ils peuvent être utilisés. Ainsi, dans un cadre non-linéaire bien déterminé que sont les systèmes hyperboliques, il replace ses travaux par rapport à l'assimilation de données avant de proposer des équivalences entre les différents filtrages envisageables : le schéma de filtrage est proposé tour à tour sous une forme déterministe, probabiliste, mais également à travers le formalisme de la théorie des observateurs. Il propose enfin des applications dans le domaine particulier de la modélisation cardiaque et donc sur des observations principalement 3D de type image, obtenues par des techniques d'imagerie spécifiques. Moireau souligne cependant que les résultats proposés dépassent le cadre de la biomécanique cardiaque. C'est donc dans un cadre plus général de systèmes dynamiques mécaniques qu'une méthode d'estimation jointe état/paramètres robuste est proposée. Inspirée des observateurs

de Luenberger, elle est applicable à n'importe quel observateur à retour d'état (Moireau *et al.*, 2008). Néanmoins l'analyse proposée se limite au cadre de systèmes linéaires ou de faibles variations de paramètres tout en utilisant des mesures en vitesse. Cette méthode est appliquée au cas particulier de mesures de positions en ajoutant un terme correctif dans le système dynamique, se rapprochant ainsi de la donnée brute fournie par les appareils d'imagerie (Moireau *et al.*, 2009). Ces travaux sont basés sur la version étendue des filtres de Kalman ce qui a l'avantage de faciliter le cadre théorique. En proposant de coupler une estimation de paramètres avec l'estimation de l'état d'un système dynamique de grande taille (en l'occurrence le modèle éléments fini du cœur), la correspondance entre les problématiques de ces travaux et l'assimilation de données est claire. L'utilisation de filtres d'assimilation de données dits réduits, c'est-à-dire incluant une réduction dimensionnelle de l'état peut s'avérer particulièrement appropriée à l'intérieur de la méthode d'estimation jointe proposée précédemment. Ainsi, Moireau et Chapelle (2011) proposent une nouvelle version de l'algorithme de filtrage par *UKF* réduite qu'il nomme *Reduced Order Unscented Kalman Filter (ROUKF)*. Ils justifient l'utilisation d'un filtre *UKF* comme étant une meilleure alternative dans un cadre non-linéaire par rapport à d'autres versions de filtres de Kalman. Cet article permet de rapprocher les différents filtres utilisant une réduction que sont les filtres *SEIK* et *SEIK* avec le formalisme de la transformée *Unscented*. Les résultats obtenus par Moireau et Chapelle sont ainsi réutilisés par l'équipe Asclépios de l'INRIA de Nice à travers notamment les travaux de Marchesseau *et al.* (2013a) plutôt orientés sur la modélisation cardiaque.

### c. L'assimilation de données appliquée à l'image

L'équipe projet CLIME de l'INRIA a été fondée par la réunion des chercheurs en traitement de données et image de l'INRIA avec l'équipe spécialisée en assimilation de données issue du Centre d'Enseignement et de Recherche en Environnement Atmosphérique (CEREA-ENPC). Ce n'est que récemment qu'a été proposée une résolution du difficile problème de l'estimation de mouvement par flot optique<sup>3</sup> à l'aide de méthodes d'assimilation de données séquentielles. Les travaux de Lepoittevin (2015) présentent une manière élégante de résoudre le problème en utilisant la variante d'Ensemble des filtres de Kalman. Cette solution permet dans un même algorithme de prendre en compte la dimension spatiale et temporelle du calcul de flot optique en une seule étape. Il développe ainsi deux approches avec son modèle : une séquentielle et l'autre variationnelle et remarque qu'en ajoutant une fonction de régularisation directionnelle qui modifie peu son système d'état il peut identifier les structures de l'image.

Les applications qu'il développe à la fois pour des images satellites et sur des images de vidéo-surveillance montre de très bons résultats notamment en précision puisqu'elle est divisée par deux par rapport aux travaux précédents.

---

3. voir le chapitre 2 pour plus de détails

### III Courte Kalmanographie pour les expérimentateurs à vocations non linéaires

Les filtres séquentiels ont été appliqués et développés dans des domaines variés. Il en résulte différentes approches et interprétations qui apportent chacune un éclairage différent. À travers l'approche de l'assimilation de données, nous avons envisagé ces filtres sous un premier angle. L'approche choisie par les travaux de Moireau s'appuie fortement sur la théorie du contrôle qui permet de définir une classe plus large d'estimateurs englobant les filtres de Kalman. Nous ne travaillerons pas dans ce cadre élargi. Nous proposons dans ce paragraphe de reprendre le cadre initial lors de l'apparition de ces filtres : le cadre stochastique. Dans cette section, nous développerons une brève introduction au contexte bayésien et aux filtres linéaires. Rapidement nous rappellerons plusieurs grandes familles de filtres et nous chercherons à les rapprocher. Le filtre de Kalman se suffit souvent à lui-même, cependant il existe plusieurs structures de filtrages qui peuvent s'avérer plus judicieuses. À travers des extensions sur des modèles non-linéaires, sur des variables non-gaussiennes, sur des fonctions adaptatives de ces filtres, l'idée développée dans ce paragraphe est de fournir au lecteur un précis à vocation pédagogique regroupant différentes approches.

#### III.A. Contexte bayésien

Le cadre bayésien dans lequel nous allons travailler a été développé par le mathématicien Thomas Bayes, publié à titre posthume en 1763. Ce premier chapitre introduit au filtrage bayésien, et plus spécifiquement au filtrage bayésien récursif. Ce filtrage s'appuie sur des modèles statistiques appelés réseaux bayésiens ou modèles de Markov cachés (*Hidden Markov Models* ou *HMM*). Ces modèles reposent sur les chaînes de Markov, qui proposent de modéliser un système comme un processus stochastique possédant une propriété essentielle appelée **propriété de Markov** : l'information utile pour la prédiction du futur est entièrement contenue dans l'état courant. De plus les états du système sont cachés : ils sont uniquement accessibles par des mesures ou observations.

La propriété de Markov nous permet de déduire deux propriétés. Soit  $p$  la probabilité d'un état et  $x_k$  l'état à l'instant  $k$ , et son observation  $y_k$ , on a alors la probabilité de l'état  $k$  :

$$p(x_k | x_{k-1}, x_{k-2}, \dots, x_0) = p(x_k | x_{k-1}) \quad (3.19)$$

À chaque instant, l'état est conditionnellement dépendant de l'état antérieur uniquement. Cela nous donne pour l'observation :

$$p(z_k | x_k, x_{k-1}, \dots, x_0) = p(z_k | x_k) \quad (3.20)$$

Le *HMM* nous permet ainsi de déterminer tout état d'un système à partir de ses mesures et ses états antérieurs :

$$p(x_k, \dots, x_0, z_k, \dots, z_0) = p(x_0) \prod_{i=1}^k p(z_i | x_i) \times p(x_i | x_{i-1}) \quad (3.21)$$

Néanmoins cette détermination de l'état du système peut également s'exprimer de manière récursive. Supposons l'état  $k - 1$  déterminé, on peut alors prédire l'état suivant à l'aide de la formule de

Chapman-Kolmogorov :

$$p(x_k|z_{k-1}) = \int p(x_k|x_{k-1}) \cdot p(x_{k-1}|z_{k-1}) \cdot dx_{k-1} \quad (3.22)$$

La formule de Bayes nous permet alors de mettre à jour cette prédiction avec la nouvelle mesure de l'état  $z_k$  :

$$p(x_k|z_k) = \frac{p(z_k|x_k) \cdot p(x_k|z_{k-1})}{p(z_k|z_{k-1})} \quad (3.23)$$

$$(3.24)$$

Notre cas diffère puisque nous avons une information sur le processus. Cette information est donnée par un modèle d'état dynamique défini par :

$$\begin{cases} x_k = f(x_{k-1}, w_{k-1}) = p(x_k|x_{k-1}) \\ z_k = h(x_k, v_k) = p(z_k|x_k) \end{cases} \quad (3.25)$$

Ce système est défini pour des variables aléatoires  $x_k$  et  $z_k$  avec  $w_k$  une modélisation de l'erreur de modèle et  $v_k$  de l'erreur de mesure.

Le système dynamique nous permet d'effectuer la prédiction en donnant l'état  $k$  à priori connaissant  $k - 1$ . La propagation récursive se fait de la même manière en utilisant la formule de Bayes :

$$p(x_k|z_{1:k}) = \frac{p(z_k|x_k) \times p(x_k|z_{1:k-1})}{\int_{x_k} p(z_k|x_k) \cdot p(x_k|z_{1:k-1}) dx_k} \quad (3.26)$$

L'estimation de la variable d'état consiste à résoudre la densité de probabilité. Différentes manières de procéder existent selon les estimateurs comme minimiser la variance de l'erreur ou maximiser la densité de probabilité.

Si la théorie du filtrage bayésien est séduisante, le calcul de la densité de probabilité à priori s'avère impossible à effectuer dans un cas général. Différentes approches existent pour permettre d'approximer ces densités de probabilité dans le cas général : les méthodes de Monte-Carlo. Cependant un mathématicien Hongrois propose aux États-Unis une méthode de filtrage basée sur une approche en deux étapes dans cadre à la fois linéaire et gaussien qu'on appellera plus tard le filtre de Kalman.

### III.B. Approximer le filtrage bayésien

L'approximation du filtrage bayésien peut être faite selon différentes méthodes<sup>4</sup>. On abordera ici seulement deux méthodes : les filtres de Kalman et le filtrage particulaire.

#### III.B.1. Dans un cas linéaire et discret : filtre de Kalman

Le caractère insolvable des intégrales présentées dans le paragraphe précédent associé aux faibles puissances de calcul disponibles à la fin des années cinquante rendaient les applications de la théorie bayésiennes impossibles. Le contexte de l'époque était à l'explosion des techniques de traitement du

4. Le lecteur pourra se référer à Särkkä (2013).

### III. Courte Kalmanographie pour les expérimentateurs à vocations non linéaires

signal lié aux début de la numérisation et le filtrage était en plein essor. La course à de nouvelles approches était d'autant plus étourdissante que la guerre froide et la conquête spatiale nécessitaient d'obtenir des résultats.

Dans ce contexte l'utilisation d'un modèle pour pouvoir filtrer des données aurait apporté beaucoup plus de pertinence à l'algorithme dans le discernement entre signal et bruit. La publication de Kalman a été contestée puisque des formes semblables de filtres avait été proposées dans l'année précédent la première publication. Pourtant, la communauté scientifique de l'époque a eu du mal à accepter cette nouvelle vision, comme le montre cette publication initiale de Kalman dans un journal mexicain (Kalman, 1960a) qui précède la publication fondatrice (Kalman, 1960b). Kalman et Bucy (1961) publient une autre versions appelée Kalman-Bucy parfois confondue avec les filtres de Kalman. Nous verrons dans la sous-sous-section III.D.2. que cette distinction désigne plutôt la version continue du filtre du Kalman.

Le filtre que nous allons décrire ici est le même que celui déjà présenté dans la section II.A.5. : nous proposons ici une approche sous un angle différent, plus orienté filtrage. Conscient qu'il faut simplifier les hypothèses du problème, Kalman propose de se placer dans un double cas particulier :

- un modèle dynamique **linéaire** ;
- des variables aléatoires **gaussiennes**.

En reprenant le système dynamique 3.10, on définit cette fois l'état  $x_k$ , le bruit de modélisation  $w_k$  et le bruit de mesure  $v_k$  à l'instant  $k$  comme des variables aléatoires gaussiennes telles que :

$$x_k = \mathcal{N}(\hat{x}_k, P_k) \quad (3.27)$$

$$v_k = \mathcal{N}(0, R_k) \quad (3.28)$$

$$w_k = \mathcal{N}(0, Q_k) \quad (3.29)$$

En plus de leur nature gaussienne, les bruits de modélisation et de mesure sont choisis centrés. Les fonctions d'évolution et d'observation, respectivement  $F$  et  $H$ , sont choisies linéaires.

En reprenant les notations précédentes, on peut écrire les densités de probabilité caractérisant l'évolution et l'observation sous leur forme gaussienne, c'est-à-dire :

$$p(x_k|x_{k-1}) = \mathcal{N}(F \cdot \hat{x}_{k-1}, Q_{k-1}) \quad (3.30)$$

$$p(z_k|x_k) = \mathcal{N}(H \cdot \hat{x}_k, R_k) \quad (3.31)$$

On peut ainsi donner l'algorithme complet en détaillant les différentes étapes :

La prédiction au pas  $k$  donne :  $p(x_k|z_{1:k-1}) = \mathcal{N}(\hat{x}_{k|k-1}, P_{k|k-1})$  avec :

$$\begin{cases} \hat{x}_{k|k-1} &= F \cdot \hat{x}_{k-1} \\ P_{k|k-1} &= F P_{k-1} F^T + Q_{k-1} \end{cases} \quad (3.32)$$

L'observation permet une mise à jour de l'état :  $p(x_k|z_{1:k}) = \mathcal{N}(\hat{x}_k, P_k)$ , avec :

$$\begin{cases} \hat{x}_k &= \hat{x}_{k|k-1} + K_k \cdot (z_k - H \hat{x}_{k|k-1}) \\ P_k &= P_{k|k-1} - K_k H P_{k|k-1} \end{cases} \quad (3.33)$$

avec :

$$\begin{cases} K_k &= P_{k|k-1} H^T S_k^{-1} \\ S_k &= [H P_{k|k-1} H^T + R_k] \end{cases}$$

Ces équations nous permettent d'introduire trois quantités :

- le gain  $K_k$  du filtre ;
- l'innovation du filtre donnée par la différence entre les données acquises et l'état projeté dans l'espace des observations. Elle est calculée sur la moyenne  $(z_k - Hx_k)$  et sur la covariance  $S_k$  ;
- la covariance de l'innovation  $S_k$ .

On s'aperçoit à travers cette approche que les filtres de Kalman ne font que donner une solution exacte à l'inférence bayésienne récursive dans le cas particulier de variables aléatoires gaussiennes et de modèles linéaires. Cet estimateur est donc optimal lorsqu'il est utilisé dans ces hypothèses.

Ces filtres présentent beaucoup d'avantages. Ils permettent de faire intervenir un modèle dans une procédure de filtrage tout en modélisant l'erreur de ce modèle avec les données réelles. Ces filtres consistent en une série de multiplications matricielles ce qui en fait un algorithme redoutable d'efficacité : l'étape la plus longue étant l'inversion de la covariance de l'innovation. Enfin, la robustesse de ces filtres vient avant tout de leur fonctionnement. Si l'innovation permet de mesurer l'écart entre l'état et l'acquisition, le gain ne dépend pas de cet écart mais de la covariance de cet écart. Ainsi, le filtre va utiliser la confiance qu'il a dans les variables pour ajuster son gain.

Kalman avait pensé son approche à la manière d'un filtrage mais l'éclairage apporté par d'autres domaines permet de relativiser l'utilisation de ce filtre en fonction des données qu'il traite. Le tableau 3.2 donne l'ensemble des utilisations qui peuvent être faites en fonction du point de vue. On retrouve les applications vues au paragraphe II.B.1. dans un autre contexte : filtrage, lissage, ou prédiction.

**Table 3.2** – Utilisation des filtres de Kalman en fonction de la position des données par rapport à l'état courant du système

Utilisation du filtre	Position de l'estimation	Position de l'observation
Filtrage	Présent	Passé
Lissage	Passé	Passé + Présent
Prédiction	Futur	Passé + Présent

Nous verrons dans les sections suivantes comment les différents auteurs contournent les hypothèses pour continuer à utiliser cet algorithme dans des situations variées, ce qui en fait un algorithme puissant. Avant d'approfondir, nous proposons d'exposer une méthode beaucoup plus récente proposant une autre manière de résoudre, cette fois en approximant la solution de l'inférence bayésienne récursive : le filtre particulaire (*Particle Filter* ou *PF*).

### III.B.2. Filtre particulaire

Ce filtre est très utilisé dans les applications de vision parce qu'il permet non seulement de travailler dans un cadre non-gaussien mais également non-linéaire. Le filtre particulaire est nettement moins rapide que le filtre de Kalman. Bien que fonctionnant en temps réel, le gain de temps est net : pour un même modèle, un filtre de Kalman fonctionne environ aussi vite qu'une particule. Son introduction dans ce paragraphe va nous permettre ensuite de pouvoir mieux comprendre certaines extensions du filtre de Kalman ayant une approche similaire.

Les filtres particuliers utilisent le même schéma que précédemment. Ils ont été proposés de manière

### III. Courte Kalmanographie pour les expérimentateurs à vocations non linéaires

à utiliser les méthodes de Monte-Carlo dans un cadre séquentiel<sup>5</sup>.

L'idée à la base de ces filtres est d'utiliser non pas une distribution, mais un échantillonnage de celle-ci, modélisé par une somme d'impulsions. Le filtre particulaire permet de passer de pas en pas en partant de la connaissance de la densité de probabilité à postériori échantillonnée :

$$p(x) \simeq \frac{1}{N} \sum_{i=1}^N \delta_{x_{k-1}^{(i)}}(x_{k-1}) \quad (3.34)$$

$N$  doit être assez grand pour pouvoir approximer la véritable distribution. L'équation précédente nous permet ainsi d'introduire  $N$  particules  $x^{(i)}$ . Cette équation pose pourtant un problème puisque la distribution  $p(\cdot)$  n'est pas connue. Le filtre particulaire introduit une densité  $q(\cdot)$  **semblable** à  $p(\cdot)$  telle que :

$$p(x) \propto q(x) \iff p(x) \propto \frac{1}{N} \sum_{i=1}^N \delta_{x_{k-1}^{(i)}}(x_{k-1}) \quad (3.35)$$

Cela nous permet d'introduire des coefficients de pondération  $\omega^{(i)}$  que l'on norme en pratique :

$$\omega^{(i)} = \frac{p(x)}{q(x)} \text{ et } \tilde{\omega}^{(i)} = \frac{\omega^{(i)}}{\sum_{j=1}^N \omega^{(j)}} \quad (3.36)$$

$$\text{Ce qui donne : } p(x) \simeq \frac{1}{N} \sum_{j=1}^N \tilde{\omega}_{k-1}^{(j)} \delta_{x_{k-1}^{(j)}}(x_{k-1}) \quad (3.37)$$

Nous pouvons alors écrire l'algorithme de filtrage particulaire. On part de la connaissance de la densité de probabilité à postériori de l'état précédent que l'on connaît sous la forme échantillonnée :

$$p(x_{k-1}|z_{1:k-1}) \simeq \frac{1}{N} \sum_{j=1}^N \tilde{\omega}_{k-1}^{(j)} \delta_{x_{k-1}^{(j)}}(x_{k-1}) \quad (3.38)$$

On effectue l'étape de prédiction à l'aide du modèle d'évolution  $f$  :

$$p(x_k|x_{k-1}) = \frac{1}{N} \sum_{i=1}^N f(\delta_{x_{k-1}^{(i)}}(x_{k-1})) \quad (3.39)$$

On effectue un tirage pour obtenir l'état de chaque particule à priori, ce qui nous permet de reconstruire la densité de probabilité prédite. On utilise ensuite la vraisemblance des particules pour calculer la nouvelle pondération en fonction des observations  $z_k$  :

$$\omega_k^i = \omega_{k-1}^i \frac{p(z_k|x_k^i) \cdot p(x_k^i|x_{k-1}^i)}{q(x_k^i|x_{k-1}^i, z_k)} \quad (3.40)$$

5. La synthèse bibliographique de cette partie est volontaire, nous citerons pour les lecteurs curieux Jouin *et al.* (2016) qui propose une revue récente sur cette famille de méthodes.

En pratique on choisit souvent  $q(x_k^{(i)} | x_{k-1}^{(i)}, z_k) = p(x_k^{(i)} | x_{k-1}^{(i)})$  de manière à pouvoir simplifier :

$$\omega_k^i \propto \omega_{k-1}^i \cdot p(z_k | x_k^i) \quad (3.41)$$

Il reste à effectuer l'étape de mise à jour en calculant  $\tilde{\omega}_k^{(i)}$  les points normés. La densité de probabilité à postériori de l'état  $k$  est obtenue :

$$p(x_k | z_{1:k}) \simeq \frac{1}{N} \sum_{j=1}^N \tilde{\omega}_k^{(j)} \delta_{x_k^{(j)}}(x_k) \quad (3.42)$$

Les filtres particulaires nécessitent un nombre certain de particules afin de pouvoir vérifier l'approximation de Monte-Carlo. Ce nombre de particules important est une des principales causes de la lenteur de ce filtre. Le filtrage particulaire pose cependant un problème : au fur et à mesure des itérations temporelles, une trajectoire de particule devient prépondérante par rapport aux autres au niveau du poids. Cet effet est directement lié à la nature finie de l'échantillonnage. Pour pallier ce problème on ajoute après l'étape de mise à jour une étape de ré-échantillonnage de la distribution obtenue à postériori afin d'itérer l'algorithme avec une distribution bien définie.

Si la distribution de probabilité était gaussienne, nous pourrions extraire de ces particules une moyenne et une covariance approchée, qualifiée d'**empirique** dans la littérature en utilisant la moyenne et covariance de l'échantillon de particules.

Le filtrage particulaire est une autre solution pratique au filtrage bayésien. L'implémentation d'un filtre particulaire utilise les mêmes étapes que le filtre de Kalman, à la différence que les particules sont générées à l'aide de tirages aléatoires à partir des distributions échantillonnées. Le filtre de Kalman possède l'avantage de pouvoir décrire ses variables aléatoires gaussiennes à l'aide de deux paramètres : moyenne et covariance. Nous proposons dans les paragraphes suivants d'introduire d'autres versions du filtre de Kalman qui permettent d'étendre son utilisation au-delà de ses hypothèses initiales, en commençant par l'hypothèse de linéarité.

### III.C. Adaptation du filtre pour traiter des modèles non-linéaires

Devant la force de l'algorithme proposé par Kalman dans sa version initiale, l'outil s'est diffusé dans différentes communautés scientifiques. Beaucoup de chercheurs se sont attelés à modifier l'algorithme afin de pouvoir l'utiliser dans des cas différents. Parmi les obstacles récurrents l'utilisation de fonction non-linéaires, l'impossibilité d'utiliser des systèmes d'état trop grands, la vitesse de convergence ou encore l'utilisation de systèmes continus. Ce paragraphe propose de détailler, comparativement à l'algorithme initial, différentes versions du filtre de Kalman couramment utilisées permettant d'étendre l'hypothèse de linéarité.

#### III.C.1. Linéariser la fonction : le filtre de Kalman étendu (EKF)

Dans un contexte de conquête spatiale, il a fallu quelques années pour que la NASA utilise ces résultats pour concevoir un filtre de Kalman dit « étendu » (*Extended Kalman Filter* ou *EKF*) fonctionnant avec des modèles non-linéaires (Schmidt et McGee, 1985; Smith *et al.*, 1962).

L'idée développée par ce filtre est simple. La propagation d'une variable aléatoire se fait dans le filtre de Kalman en séparant le calcul en deux parties : la moyenne et la covariance. Si le calcul de la

### III. Courte Kalmanographie pour les expérimentateurs à vocations non linéaires

moyenne par une fonction non-linéaire ne pose pas de problème, ce n'est pas le cas de celui de la covariance. La solution proposée par l'*EKF* consiste à linéariser la fonction multivariable en utilisant une matrice jacobienne, issue d'un développement de Taylor au premier ordre. Reprenons le système dynamique donné dans l'équation (3.10) en utilisant les fonctions non-linéaires et différentiables une fois  $f$  pour l'évolution et  $h$  pour l'observation. On définit  $F_k$  et  $H_k$  les jacobienes de ces matrices calculées au point de fonctionnement défini par l'état  $x_k$ . L'algorithme de l'*EKF* peut alors s'écrire :

**Prédiction :**

$$\begin{cases} \hat{x}_{k|k-1} = f(\hat{x}_{k-1}) \\ P_{k|k-1} = F_k P_{k-1} F_k^T + Q_{k-1} \end{cases} \quad (3.43)$$

**Mise à jour :**

**Innovation :**

$$\begin{cases} y_k = (z_k - h(\hat{x}_{k|k-1})) \\ S_k = H_k P_{k|k-1} H_k^T + R_k \end{cases} \quad (3.44)$$

**Gain :**

$$K_k = P_{k|k-1} H_k^T S_k^{-1} \quad (3.45)$$

**Correction :**

$$\begin{cases} \hat{x}_k = \hat{x}_{k|k-1} + K_k \cdot y_k \\ P_k = P_{k|k-1} - K_k H_k P_{k|k-1} \end{cases} \quad (3.46)$$

On peut constater que les modifications apportent des approximations de linéarisation uniquement sur la covariance. Les filtres *EKF* permettent en pratique d'utiliser des fonctions non-linéaires en utilisant un schéma de Kalman. Cependant les erreurs de linéarisations limitent en pratique à l'utilisation de fonction « faiblement non-linéaire ». Les fortes non-linéarités génèrent des erreurs de linéarisations que l'on doit répercuter sur les erreurs du modèle pour éviter une divergence du filtre, ce qui limite sa précision. L'algorithme est également plus coûteux que le filtre linéaire puisqu'il faut effectuer un calcul de Jacobienne pour chaque fonction non-linéaire à chaque itération

Les erreurs de linéarisation étant le point faible de l'*EKF*, d'autres approches ont été développées afin de gagner en précision, comme de proposer un développement en série de Taylor à un ordre plus élevé (Einicke, 2012). Ces approches améliorent la précision du filtre en permettant d'utiliser une meilleure approximation de la fonction non-linéaire. Néanmoins le problème reste le même : si ce gain de précision permet également de diminuer les erreurs, le problème restera le même pour des fonctions fortement non-linéaires. L'utilisation d'un développement de Taylor d'ordre supérieur a également le mauvais goût de dégrader ses performances.

#### III.C.2. Transformer les distributions d'états par la fonction non-linéaire

Le problème de la linéarisation devait être approché différemment pour permettre une amélioration significative. L'idée développée par Julier et Uhlmann (1997) se situe en rupture par rapport au concept développé dans l'*EKF*. Ils proposent avec le filtre de Kalman non parfumé (*Unscented Kalman Filter* ou *UKF*) d'utiliser la fonction originale et de n'apporter des approximations uniquement

sur les distributions de l'état.

Il est plus facile de comprendre la démarche de l'*UKF* en reprenant le fonctionnement d'un filtre particulaire. En effet, Julier et Uhlmann remarquent qu'il est plus juste d'approximer une distribution propagée par une fonction non-linéaire plutôt que de linéariser une fonction non-linéaire. La solution qu'il propose est similaire au filtrage particulaire : en propageant les variables aléatoires échantillonnées par la fonction non-linéaire. Une différence majeure sépare le filtre *UKF* et le filtrage particulaire. L'hypothèse gaussienne des filtres de Kalman est astucieusement utilisée puisqu'elle permet de paramétrer la distribution à l'aide de trois particules par dimension du vecteur d'état : un au niveau de la moyenne et un de chaque côté pour capturer l'information de covariance. En réalité, seul deux points sont nécessaires en reprenant l'hypothèse de symétrie ce qui est appelé « distribution simplex » dans la littérature. Le choix des particules se fait ici de manière déterministe et non aléatoire.

Pour déterminer ces particules les auteurs utilisent la transformée non parfumée (*Unscented Transform* ou *UT*) dont les bases ont été posées dans la thèse de Uhlmann (1995). Cette transformée permet de passer d'un espace multivarié de variables aléatoires d'états à un espace de points d'échantillonnage appelés points de Gauss en français mais que nous appellerons *sigmas points* comme leur dénomination dans l'article original.

#### a. Transformée non-parfumée (*UT*)

La transformée *unscented* propose de générer un échantillon de points à partir d'une distribution gaussienne : la moyenne et la covariance de l'échantillon correspondent à celles de la distribution. Pour générer ces points, on utilise un facteur  $c$  pour régler la distance de la moyenne à laquelle on se place afin d'extraire deux autres points de la distribution.

Pour un état  $\hat{x}$  de dimension  $n$  et de covariance  $P^{xx}$ , on définit  $P^{xx(i)}$  la  $i^{\text{ème}}$  colonne de la covariance. L'algorithme de cette transformée *unscented* peut être résumé comme suit :

$$\chi^{(1)} = \hat{x} \quad (3.47)$$

$$\chi^{(i)} = \hat{x} + \sqrt{c \cdot P^{xx(i)}} \quad (3.48)$$

$$\chi^{(i+n)} = \hat{x} - \sqrt{c \cdot P^{xx(i)}} \quad (3.49)$$

Afin de conserver un formalisme de notations matricielles, nous préférons utiliser cet échantillon de *sigma-points* sous forme de matrice que l'on définit par leur simple concaténation :

$$\chi = [\chi^{(1)} \quad \chi^{(2)} \quad \dots \quad \chi^{(2n+1)}] \quad (3.50)$$

À partir de cet échantillon il s'agit maintenant de poser la transformée inverse. Nous utiliserons les notations matricielles introduites par Särkkä (2007). Similairement à ce que nous avons évoqué dans le filtrage particulaire, on utilise la moyenne et la covariance empirique pour effectuer la transformée inverse. Il serait impossible d'obtenir un résultat significatif pour un échantillon aléatoire de particules sauf que dans notre cas, l'échantillon a été construit pour permettre cette transformation inverse. Cette construction passe par la définition d'une pondération adaptée en fonction de la distance réglée par le paramètre  $c$  défini ci-après. Nous définirons ici la pondération proposée dans l'article

### III. Courte Kalmanographie pour les expérimentateurs à vocations non linéaires

fondateur et nous discuterons des pondérations dans le paragraphe c. . Nous définirons tout de même des poids différents pour la moyenne  $w^m$  et pour la covariance  $w^c$  même s'ils sont égaux pour Julier et Uhlmann :

$$w_0^m = \frac{\lambda}{n + \lambda} \quad (3.51)$$

$$w_i^m = \frac{1}{2(n + \lambda)} \quad (3.52)$$

$$w_0^c = w_0^m \quad (3.53)$$

$$w_i^c = w_i^m \quad (3.54)$$

Cette pondération est reliée au paramètre d'éloignement  $c$  par la relation :

$$c = n + \lambda \quad (3.55)$$

On pose alors les matrices :

$$w^m = [w_0^m \dots w_{2n}^m]^T \quad (3.56)$$

$$W = (I_{2n+1} - [w^m \dots w^m]) \times \begin{bmatrix} w_0^c & & 0 \\ & \ddots & \\ 0 & & w_{2n}^c \end{bmatrix} \times (I_{2n+1} - [w^m \dots w^m])^T \quad (3.57)$$

La moyenne et la covariance empirique se calculent ainsi comme suit et permettent d'effectuer la transformée non-parfumée inverse :

$$E_W(\chi) = \sum_{i=0}^{2n+1} w_i \chi_i \quad (3.58)$$

$$= \chi \cdot w^m \quad (3.59)$$

$$P_W^{xx} = \sum_{i=0}^{2n+1} w_i [\chi_i - E_W(\chi)] [\chi_i - E_W(\chi)]^T \quad (3.60)$$

$$= \chi \cdot W \cdot \chi^T \quad (3.61)$$

Le formalisme de la transformée non parfumée nous permet également de calculer des corrélations entre distributions, soit pour les distributions  $x \sim \mathcal{N}(\tilde{x}, P^x x)$  et  $z \sim \mathcal{N}(\tilde{z}, P^z z)$  représentées respectivement par les échantillons  $\chi$  et  $\zeta$  :

$$P^{xz} = \sum_{i=0}^{2n+1} [\chi_i - \tilde{x}] [\zeta_i - \tilde{z}]^T \quad (3.62)$$

$$= \chi \cdot W \cdot \zeta^T \quad (3.63)$$

Il est intéressant de noter que le paramètre  $\lambda$  est posé comme un degré de liberté supplémentaire permettant d'ajuster les moments d'ordres supérieurs.

**b. Filtre de Kalman non parfumé**

Le filtre de Kalman non parfumé utilise ainsi la transformée du même nom pour pouvoir estimer la densité de probabilité à priori de l'état à partir des échantillons obtenus. On notera la transformée non parfumée par  $UT$ . On a alors l'algorithme suivant :

<b>Prédiction :</b>	$\chi_{k-1} = UT(\hat{x}_{k-1}, P_{k-1})$ (3.64)
	$\begin{cases} \hat{x}_{k k-1} = f_k(\chi_{k-1}) \cdot w^m \\ P_{k k-1} = f_k(\chi_{k-1}) \cdot W \cdot f_k(\chi_{k-1})^T + Q_{k-1} \end{cases}$ (3.65)
<b>Mise à jour :</b>	
<b>Innovation :</b>	$\chi_{k k-1} = UT(\hat{x}_{k k-1}, P_{k k-1})$ (3.66)
	$\zeta_{k k-1} = h_k(\chi_{k k-1})$ (3.67)
	$\hat{z}_{k k-1} = \zeta_{k k-1} \cdot w^m$ (3.68)
	$\begin{cases} y_k = (z_k - \hat{z}_{k k-1}) \\ S_k = \zeta_{k k-1} \cdot W \cdot \zeta_{k k-1}^T + R_k \end{cases}$ (3.69)
<b>Corrélation :</b>	$P_{k k-1}^{xz} = \chi_{k k-1} \cdot W \cdot \zeta_{k k-1}^T$ (3.70)
<b>Gain :</b>	$K_k = P_{k k-1}^{xz} \cdot S_k^{-1}$ (3.71)
<b>Correction :</b>	$\begin{cases} \hat{x}_k = \hat{x}_{k k-1} + K_k \cdot y_k \\ P_k = P_{k k-1} - K_k S_k^{-1} K_k^T \end{cases}$ (3.72)

La complexité du filtre est augmentée puisqu'il s'agit de passer  $2n + 1$  particules pour un pas d'exécution en ajoutant le calcul de la décomposée de Cholesky de la covariance pour le calcul des points. On évite néanmoins les calculs de dérivée et les approximations qui en découlent. Gustafsson et Hendeby (2012) tempèrent les affirmations des auteurs sur la capacité des filtres non parfumés à pouvoir capturer des moments supérieurs à l'ordre 2 tout en soulignant les points communs avec l'EKF. Il a par ailleurs été montré que le filtre UKF a une précision d'un ordre supplémentaire que l'EKF standard<sup>6</sup>. L'utilisation de ces filtres permettent par ailleurs une plus grande flexibilité dans l'ajout de modèle grâce à leur fonctionnement particulière. La figure 3.5 explique de manière animée les différences entre l'EKF et l'UKF.

Le fonctionnement du filtre UKF apporte à l'algorithme une manière déterministe de définir des

6. La démonstration peut être trouvée dans l'article de Moireau et Chapelle (2011)



moments d'ordre supérieur à l'ordre 2, il utilise contrairement à ce qui a été énoncé précédemment des poids différents pour la moyenne et la covariance empirique.

$$w_0^m = \frac{\lambda}{n + \lambda} \quad (3.73)$$

$$w_i^m = \frac{1}{2(n + \lambda)} \quad (3.74)$$

$$w_0^c = \frac{\lambda}{n + \lambda} + (1 - \alpha^2 + \beta) \quad (3.75)$$

$$w_i^c = \frac{1}{2(n + \lambda)} \quad (3.76)$$

$$(3.77)$$

Il définit alors  $\lambda = \alpha^2(n + \kappa) - n$ . Les paramètres  $\alpha$  et  $\kappa$  permettent de déplacer les points plus ou moins loin du centre de la gaussienne, tandis que le paramètre  $\beta$  est introduit pour pouvoir capturer les moments d'ordre supérieur à l'ordre 2.

D'autres améliorations sont apportées par Julier (2002) qui propose une forme différente pour les poids mais également pour le calcul de la covariance empirique pour laquelle il est proposé d'apporter un terme correctif pour capturer le kurtosis<sup>7</sup>. D'autres montrent que les  $2n + 1$  points ne sont en réalité pas l'échantillon minimal et qu'un échantillon de  $n + 1$  points nommé « simplex » est suffisant (Moireau et Chapelle, 2011; Wan-Chun *et al.*, 2007). Plus récemment, des échantillons asymétriques sont proposés afin de pouvoir capturer des distributions non-gaussiennes (Chan *et al.*, 2014). Les paramètres de réglage de la distribution étant nombreux, certains auteurs s'intéressent également aux réglages pour permettre de correspondre au mieux aux distributions observées (Straka *et al.*, 2012a). D'autres approches existent pour capturer des moments statistiques plus élevés mais passent par l'ajout de point dans l'échantillon (Van Der Merwe *et al.*, 2001), ce qui augmente le coût de calcul.

Ainsi la détermination de l'échantillon est un sujet de recherche à part entière et plusieurs auteurs spécialistes du sujet « *Unscented* » proposent de rapprocher aux filtres *UKF* les filtres *QKF* (pour *Quadrature Kalman filter*) et *CKF* (pour *Cubature Kalman Filter*) (Särkkä et Solin, 2012). Ces derniers sont en réalité des cas particuliers de l'*UKF* et correspondent à des réglages particuliers des *sigma-points*. Nous n'avons pas trouvé de classe pour ces filtres de Kalman que nous appellerons à particules déterministes. En revanche Dunik *et al.* (2014) séparent les filtres en deux classes, celles ayant besoin de calculer ou approximer une dérivée tels que les filtres *EKF* et les filtres n'ayant pas besoin de calcul de dérivée qu'il nomme *derivative-free* comprenant les filtres à particules déterministes mais également des filtres basés sur des schémas différentiels comme le *Central Difference Kalman Filter (CD-KF)*, basés sur de l'interpolation polynomiale (*Divided Difference Filter* ou *DDF*) et basés sur des intégrations d'équations stochastiques (Dunik *et al.*, 2013).

Cette première partie nous a permis d'avoir un aperçu de la variété des algorithmes développés pour pouvoir déduire des fonctions non-linéaires et ainsi passer outre l'hypothèse de linéarité. Parmi les filtres évoqués, on remarquera également que certaines solutions traitent des cas non-gaussiens. Le paragraphe qui suit décrit les différentes versions du filtre et propose d'autres améliorations.

7. Moment statistique d'ordre 4.

#### III.D. Quelques versions du filtre

Nous avons évoqué dans la section précédente l'inventivité développée depuis 1960 pour pouvoir utiliser des modèles non-linéaires avec les filtres de Kalman. Ce paragraphe présente quelques morceaux choisis de l'état de l'art concernant d'autres améliorations.

##### III.D.1. Augmenter la vitesse de convergence : les filtres itérés (*IEKF* ou *IUKF*)

Le filtre de Kalman est très efficace à partir du moment où il parvient à recalculer le modèle d'évolution par rapport à l'observation. Ce recalage nécessite cependant un temps minimum certes très court par rapport à d'autres méthodes de recalage mais pouvant tout de même poser problème pour des acquisitions courtes avec des valeurs d'initialisation d'état pouvant être éloignées de la valeur réelle.

Ce problème de durée de convergence est directement lié au schéma en deux étapes des filtres de Kalman. En effet à chaque observation le filtre calcule un gain pour pouvoir corriger l'état. Ainsi, pour un état éloigné de la valeur réelle « cachée », le gain calculé risque d'être important pouvant alors conduire à une divergence du filtre. Sinon plusieurs pas de temps sont nécessaires pour permettre au filtre de converger. Havlík et Straka (2015) soulignent que pour les méthodes utilisant des dérivées, cette approche permet de recalculer la linéarisation de la fonction sur un point de fonctionnement plus proche de la solution, minimisant ainsi les erreurs de linéarisation. Pour pallier ce problème, il expose l'idée d'effectuer durant le même pas de temps une série d'itérations sur les étapes de calcul de la jacobienne, du gain et de la moyenne de la distribution de l'état à postériori. Cette itération permet d'optimiser le gain en effectuant le calcul sur des points de plus en plus proches de l'état caché accessible avec les données observées. Cette méthode permet d'accélérer la convergence du filtre moyennant un surcoût de calcul. Nous avons choisi d'exposer la méthode sur le schéma *EKF* pour des raisons de clarté.

**Prédiction :**

$$\begin{cases} \hat{x}_{k|k-1}^0 = f_k(\hat{x}_{k-1}^0) \\ P_{k|k-1} = F_k P_{k-1} F_k^T + Q_{k-1} \end{cases} \quad (3.78)$$

**Mise à jour :**

**Itération :**

**Calcul de la jacobienne**

$$H_k^j = \frac{\partial h_k(s)}{\partial s} \Big|_{s=x_k^j} \quad (3.79)$$

**Innovation :**

$$\begin{cases} y_k^j = (z_k - h_k(\hat{x}_{k|k-1}^j)) \\ S_k^j = H_k^j P_{k|k-1} H_k^{jT} + R_k \end{cases} \quad (3.80)$$

**Gain :**

$$K_k^j = P_{k|k-1} H_k^{jT} (S_k^j)^{-1} \quad (3.81)$$

**Correction état :**

$$\hat{x}_k^{j+1} = \hat{x}_{k|k-1}^j + K_k^j \cdot y_k^j \quad (3.82)$$

On répète l'itération jusqu'à ce que  $\|\hat{x}_k^{j+1} - \hat{x}_k^j\|$  soit inférieur à un certain seuil ou au-delà d'un certain nombre d'itérations.

**Correction Covariance :**

$$P_k = P_{k|k-1} - K_k^j H_k^j P_{k|k-1} \quad (3.83)$$

Xie et Feng (2012) propose une application sur de l'identification de structure ou les résultats sont probant avec la transposition de cet algorithme sur un filtre non parfumé qu'il nomme *IUKF* (pour *Iterated Unscented Kalman Filter*) (Xie et Feng, 2012).

### III.D.2. Traiter des modèles continus (*KBF*)

La définition du système dynamique exploitée dans le filtre de Kalman est utilisée jusqu'ici comme un système discret. La publication de Kalman et Bucy (1961) a permis de développer toute une branche de filtres applicables sur des systèmes continus. Ces systèmes sont très souvent différenciés des filtres de Kalman puisqu'ils sont appelés filtres de Kalman-Bucy (*Kalman-Bucy Filter* ou *KBF*). Ces filtres admettent quelques différences dans la manière dont ils sont posés. Avec le développement de la numérisation, il est devenu absurde de parler de modèles continus puisque l'observation est désormais disponible de manière discrète. Cependant, nombre de systèmes dynamiques sont régis par des systèmes différentiels continus. C'est cette conjecture qui a permis un renouveau récent de ces filtres sous l'impulsion de Särkkä avec des filtres hybrides à la fois discrets au niveau de

### III. Courte Kalmanographie pour les expérimentateurs à vocations non linéaires

l'observation et continus pour l'évolution. Après avoir posé un cadre formel à l'utilisation de filtres de Kalman-Bucy avec des transformées non-parfumées (Särkkä, 2007), il étend également ce cadre aux *CKF* (Särkkä et Solin, 2012). Afin d'exposer plus clairement les concepts, nous les introduirons avec le filtre linéaire continu puis nous donnerons les algorithmes de l'*EKF* et l'*UKF* appliqués à des systèmes hybrides : l'*EKBF* et l'*UKBF*.

Définissons un système dynamique continu pour le modèle d'évolution et discret pour le modèle d'observation de manière similaire à l'équation (3.10), en introduisant des bruits additifs :

$$\begin{cases} \dot{x}(t) &= \mathcal{F}(x(t), t) + w(t) \\ z(t_k) &= h(x(t_k)) + v(t_k) \end{cases} \quad (3.84)$$

On définit de la même manière que précédemment les covariances associées aux bruits centrés  $w(t)$  et  $v(t)$  respectivement par  $Q(t) = E[w(t)w(t)^T]$  et  $R(t_k) = E[v(t_k)v(t_k)^T]$ . Pour pouvoir faire une analogie avec le modèle discret, on pose deux instants  $t_k$  et  $t_{k-1}$  tels que  $\exists \epsilon \ll 1$ ,  $\Delta t \ll \epsilon$  et  $t_k - t_{k-1} = \Delta t$ . En reprenant l'algorithme initial, on peut ainsi poser le cas continu comme la limite du cas discret énoncé plus haut. On définit le développement limité au premier ordre du modèle d'évolution :

$$\mathcal{F}(\Delta t) = I + \mathcal{F}(\Delta t) \cdot \Delta t + O(\Delta t^2) \quad (3.85)$$

L'intégration de la fonction d'évolution nous permet d'avoir l'état au temps  $t_k$  en fonction de l'état au temps  $t_{k-1}$  :

$$\hat{x}(t_k) = (I + \mathcal{F}(t_{k-1}) \cdot \Delta t) \hat{x}(t_{k-1}) \quad (3.86)$$

La propagation de la moyenne de la distribution de l'état ne pose pas de problème :

$$\frac{\hat{x}(t_k) - \hat{x}(t_{k-1})}{\Delta t} = \mathcal{F}(t_{k-1})\hat{x}(t_{k-1}) \quad (3.87)$$

$$\lim_{\Delta t \rightarrow 0} \frac{\hat{x}(t_k) - \hat{x}(t_{k-1})}{\Delta t} = \frac{\partial \hat{x}(t)}{\partial t} \quad (3.88)$$

$$= \mathcal{F}(t)\hat{x}(t) \quad (3.89)$$

On fait de même avec la propagation de la covariance :

$$P_k = (I + \mathcal{F}(t_{k-1})\Delta t) P_{k-1} (I + \mathcal{F}(t_{k-1})\Delta t)^T + Q_{k-1} \quad (3.90)$$

$$\iff P_k = P_{k-1} + \mathcal{F}(t_{k-1})P_{k-1}\Delta t + (\mathcal{F}(t_{k-1})P_{k-1})^T \Delta t + \mathcal{F}(t_{k-1})P_{k-1}\mathcal{F}(t_{k-1})^T \Delta t^2 + Q_{k-1} \quad (3.91)$$

$$\iff \frac{P_k - P_{k-1}}{\Delta t} = \mathcal{F}(t_{k-1})P_{k-1} + (\mathcal{F}(t_{k-1})P_{k-1})^T + \mathcal{F}(t_{k-1})P_{k-1}\mathcal{F}(t_{k-1})^T \Delta t + \frac{Q_{k-1}}{\Delta t} \quad (3.92)$$

Ce qui nous permet de définir la dérivée de la covariance de l'état  $\dot{P}(t)$  et de poser la covariance du bruit du modèle  $Q(t)$  :

$$\lim_{\Delta t \rightarrow 0} \frac{P_k - P_{k-1}}{\Delta t} = \dot{P}(t) \quad (3.93)$$

$$Q_{k-1} \simeq Q(t)\Delta t \quad (3.94)$$

On a ici les principaux changements dans un filtre continu : l'équation de la propagation de la covariance ainsi que l'apparition d'une dépendance avec le pas de temps  $\Delta t$  pour la matrice de covariance de l'erreur modèle. Appliqué à un modèle non-linéaire, on obtient l'*EKBF*. L'algorithme est donné ci-dessous avec en couleur les modifications par rapport à l'*EKF*. On définit les jacobiniennes de  $\mathcal{F}$  et  $h$  respectivement par  $F_{t_k}$  et  $H_{t_k}$ .

**Prédiction :**

$$\begin{cases} \hat{\hat{x}}(t) = \mathcal{F}(\hat{x}(t), t) \\ \dot{P}(t) = P(t)F_{t_k}(\hat{x})^T + F_{t_k}(\hat{x})P(t) + Q(t) \end{cases} \quad (3.95)$$

Les équations de prédiction sont en pratique intégrées entre  $t_{k-1}$  et  $t_k$ . Elles vont permettre de calculer les moyennes et covariances à priori de l'état :  $\hat{x}^-(t_k)$  et  $P^-(t_k)$ . Le reste des opérations s'effectuent au temps  $t_k$  afin de pouvoir interfacer les observations discrètes.

**Mise à jour :**

**Innovation :**

$$y(t_k) = z(t_k) - h(x(t_k)) \quad (3.96)$$

**Covariance de l'innovation :**

$$S(t_k) = H_{t_k}P^-(t_k)H_{t_k}^T + R(t_k) \quad (3.97)$$

**Gain :**

$$K_k = P^-(t_k)H_{t_k}^T S(t_k)^{-1} \quad (3.98)$$

**Correction :**

$$\begin{cases} \hat{x}(t_k) = \hat{x}^-(t_k) + K(t_k) \cdot y(t_k) \\ P(t_k) = P^-(t_k) - K(t_k)H_{t_k}P^-(t_k) \end{cases} \quad (3.99)$$

De la même manière, on présente l'algorithme *UKBF* proposé par Särkkä (2007). Cet algorithme utilise des poids légèrement différents de ceux exposés précédemment :

$$w_0^m = \frac{\lambda}{n + \lambda} \quad (3.100)$$

$$w_i^m = \frac{1}{2(n + \lambda)} \quad (3.101)$$

$$w_0^c = \frac{\lambda}{n + \lambda + (1 - \alpha^2 + \beta)} \quad (3.102)$$

$$w_i^c = \frac{1}{2(n + \lambda)} \quad (3.103)$$

Comme précédemment, nous avons mis en couleur les différences avec l'algorithme *UKF*.

#### Prédiction :

$$\chi(t_{k-1}) = UT(\hat{x}(t_{k-1}), P(t_{k-1})) \quad (3.104)$$

$$\begin{cases} \dot{\hat{x}}(t) = & f(\chi(t), t) \cdot w^m \\ \dot{P}(t) = & \chi(t) \cdot W \cdot f(\chi(t), t)^T + f(\chi(t), t) \cdot W \cdot \chi(t)^T + Q_{k-1} \end{cases} \quad (3.105)$$

Les équations de prédiction sont en pratique intégrées entre  $t_{k-1}$  et  $t_k$ . Elles vont permettre de calculer les moyennes et covariances à priori de l'état :  $\hat{x}^-(t_k)$  et  $P^-(t_k)$ . La suite des opérations s'effectuent au temps  $t_k$  afin de pouvoir interfacer les observations discrètes. Le reste du filtre reste le même que l'*EKBF* : nous ne rappellerons pas ici pour des raisons de concision.

Les filtres de Kalman continus présentent une manière originale d'intégrer un modèle continu. Cette méthode s'avère de plus particulièrement adaptée pour l'usage contemporain utilisant quasi-systématiquement des observations discrètes. La version non-parfumée présente plusieurs avantages. Elle diminue les erreurs du modèle dans un contexte continu plus sensible à ce type d'erreurs à cause de l'intégration, et elle permet une plus grande flexibilité dans la formalisation du modèle.

#### III.D.3. Utiliser des systèmes d'état de grandes dimensions : filtrage d'ensemble ou filtrage réduits ?

Les systèmes de grandes dimensions sont difficiles à utiliser avec les filtres de Kalman. Les difficultés rencontrées ont plusieurs origines. La propagation de la covariance peut amener pour des systèmes d'état de grandes dimensions à introduire des erreurs numériques. La grande dimension d'un système augmente de façon importante les temps de calcul ce qui peut s'avérer problématique dans certaines applications. C'est l'inversion de la covariance de l'innovation qui est problématique en coût de calcul comme en mémoire (Daget, 2008).

##### a. Approche par réduction de modèle

Pour pallier ce problème, une solution a émergé en réduisant l'ordre du problème. L'hypothèse majeure de ces approches est de supposer qu'un nombre réduit de variables influencent l'évolution du modèle à un instant donné. Plusieurs filtres ont été proposés.

- Dans la famille des filtres utilisant une linéarisation de la fonction, on trouve le filtre *RRSQRT* (pour *reduced rank square-root*) dont le principe est de décomposer à chaque pas la matrice de covariance initiale en mode principaux avant la prédiction et avant la mise à jour puis d'effectuer les étapes suivantes sur la dimension réduite. Cette décomposition de la forme  $P_k = C_k C_k^T$  permet ainsi d'effectuer l'inversion de la covariance de l'innovation sur une matrice de dimension réduite mais également mieux conditionnée améliorant la précision du calcul. Le filtre *SEEK* (pour *Singular Evolutive Extended Kalman filter*) (Brasseur et Verron, 2006) propose de décomposer la covariance de l'état de manière à identifier les directions principales et de négliger l'erreur du modèle sur les autres directions sous une forme classique  $P_k = B_k D_k B_k^T$ . Enfin Simon (2007) propose un filtre d'ordre réduit sans approximation mais

cette fois pour des modèles linéaires : il propose de modifier le gain de Kalman en minimisant la trace de la covariance.

- Dans la famille des filtres de Kalman *derivative-free*, une première approche à été proposée par (Pham *et al.*, 1998). C'est une variante de la méthode *SEEK* nommé *SEIK* (pour *Singular Evolutive Interpolated Kalman* qui propose d'éviter d'utiliser la fonction linéarisée du filtre *EKF* propageant la distribution avec un échantillon de points. Cette approche est très similaire à celle proposée par l'*UKF* bien que le *SEIK* ait été proposé dans le cadre de l'assimilation de données appliquée à l'océanographie. Il faudra attendre une décade pour que Moireau et Chapelle (2011) ne proposent un cadre formel pour une version réduite des filtres *UKF* qu'ils nomment *ROUKF* (pour *Reduced Order Unscented Kalman Filter*). Il est intéressant de noter que l'approche *SEIK* correspond à un échantillon simplex particulier du filtre *ROUKF*.

Étant donné que nous n'utiliserons pas ces filtres, nous ne proposons pas de développer davantage les approches réduites.

## b. Filtre de Kalman d'Ensemble *EnKF*

Une autre solution a été développée par la communauté de l'assimilation de données : le filtre de Kalman d'ensemble. L'impossibilité des filtres de Kalman à être appliqués à des systèmes de grandes tailles a conduit Evensen (2003) à proposer un filtre stochastique en opposition aux filtres *EKF* et *UKF* qui sont déterministes. Comme indiqué plus haut, le problème des systèmes de grande dimension appliqués à l'*EKF* se trouve dans le coût de calcul important et dans les erreurs pouvant intervenir en cas de forte non-linéarité. L'idée du filtre *EnKF* est similaire à celle du *SEIK* introduit plus haut<sup>8</sup> et similaire à la philosophie des filtres à particules déterministes. Il s'agit d'échantillonner par un nombre important de particules l'état initial afin de lui permettre de propager à travers la non-linéarité l'état. Les statistiques d'ordre 1 et 2 sont alors récupérées à l'aide de formules de moyennes et covariances empiriques.

Le filtre d'ensemble est une manière d'utiliser des méthodes de Monte-Carlo dans le formalisme des filtres de Kalman. Il constitue une gradation en plus entre la complexité d'un filtre à particule et un filtre *UKF*. Ces méthodes demandent un coût de calcul important puisque les fonctions d'évolution ou d'observation sont évaluées autant de fois que le nombre de particules. Cependant, le formalisme gaussien donné par les filtres de Kalman permet de simplifier de manière drastique les calculs de vraisemblance en proposant une forme simplifiée aux étapes du théorème de Bayes (équation (3.26)).

Plusieurs algorithmes existent, nous en présentons un qui fait ressortir les différences avec un filtre *EKF*. De manière similaire à l'*UKF*, on définit un ensemble  $X_{k-1|k-1}$  à partir d'un tirage suivant la densité de probabilité à postériori de  $N$  particules  $x_{k-1}^{(i)} \sim \mathcal{N}(\hat{x}_{k-1|k-1}, P_{k-1|k-1})$  à l'état  $k-1$ .

$$X_{k-1|k-1} = [x_{k-1}^{(1)} \dots x_{k-1}^{(N)}] \quad (3.106)$$

Les moyennes et covariances empiriques permettent alors de retrouver les moments d'ordre 1 et 2 de l'état  $x_{k-1|k-1}$ . On les notera respectivement  $\mathbb{E}$  et  $\mathbb{P}$  dans l'algorithme pour alléger l'écriture. Elles

8. En réalité le *SEIK* est postérieur à l'*EnKF* : se reporter à la figure 3.7.

### III. Courte Kalmanographie pour les expérimentateurs à vocations non linéaires

sont alors définies par :

$$\hat{x}_{k-1|k-1} = \frac{1}{N} \sum_{i=1}^N x_{k-1}^{(i)} \quad (3.107)$$

$$P_{k-1|k-1} = \frac{1}{N-1} \sum_{i=1}^N \left[ x_{k-1}^{(i)} - \hat{x}_{k-1|k-1} \right] \cdot \left[ x_{k-1}^{(i)} - \hat{x}_{k-1|k-1} \right]^T \quad (3.108)$$

Ce qui peut s'écrire également sous forme matricielle :

$$\hat{x}_{k-1|k-1} = \frac{1}{N} \cdot X_{k-1|k-1} \cdot \mathbf{1} \quad (3.109)$$

$$P_{k-1|k-1} = \frac{1}{N-1} \left[ X_{k-1|k-1} - \hat{x}_{k-1|k-1} \cdot \mathbf{1}^T \right] \cdot \left[ X_{k-1|k-1} - \hat{x}_{k-1|k-1} \cdot \mathbf{1}^T \right]^T \quad (3.110)$$

avec :  $\mathbf{1} = [1 \dots 1]^T$  vecteur de taille N

La corrélation entre deux ensembles  $X$  et  $Z$  peut également s'écrire comme :

$$P_{k-1|k-1}^{XZ} = \frac{1}{N-1} \left[ X_{k-1|k-1} - \hat{x}_{k-1|k-1} \cdot \mathbf{1}^T \right] \cdot \left[ Z_{k-1|k-1} - \hat{z}_{k-1|k-1} \cdot \mathbf{1}^T \right]^T \quad (3.111)$$

On remarquera qu'il est tout à fait possible de formaliser ces covariances empiriques sous la même forme que pour l'*UKF* en définissant un vecteur de poids pour la moyenne  $w_m$  et une matrice de pondération pour la covariance  $W$ . Un algorithme du filtre de Kalman d'ensemble pourrait être :

<b>Prédiction :</b>		
	$x_{k-1}^{(i)} \sim \mathcal{N}(\hat{x}_{k-1 k-1}, P_{k-1 k-1})$	(3.112)
	$X_{k-1 k-1} = [x_{k-1}^{(1)} \dots x_{k-1}^{(N)}]$	(3.113)
	$\{ X_{k k-1} = f(X_{k-1 k-1}, Q_{k-1})$	(3.114)
<b>Mise à jour :</b>		
<b>Innovation :</b>		
	$Z_{k k-1} = h(X_{k k-1}, R_k)$	(3.115)
	$\{ \begin{aligned} y_k &= (z_k \cdot \mathbf{1}^T - Z_{k k-1}) \\ S_k &= \mathbb{P}[Z_{k k-1}] \end{aligned}$	(3.116)
<b>Corrélation :</b>		
	$P_{k k-1}^{XZ} = \mathbb{P}^{XZ}[X_{k k-1}, Z_{k k-1}]$	(3.117)
<b>Gain :</b>		
	$K_k = P_{k k-1}^{XZ} \cdot S_k^{-1}$	(3.118)
<b>Correction :</b>		
	$\{ X_k = X_{k k-1} + K_k \cdot y_k \cdot \mathbf{1}^T$	(3.119)

Le filtre *EnKF* a l'avantage d'être facile à implémenter, et de pouvoir traiter des systèmes de grande taille sans faire de linéarisation du modèle. L'avantage de ce filtre est que les distributions échantillonnées utilisées pour modéliser les variables aléatoires sont exactes au sens de l'approximation de Monte-Carlo. En revanche les calculs de mise à jour n'utilisent que les moments d'ordre 1 et 2 sans tenir compte de la nature de la distribution. Implicitement le filtre fonctionne en approximant les distributions de l'échantillon de manière gaussienne. Le lien avec l'*UKF* est évident, cependant il est intéressant de lire l'article de Roth *et al.* (2015) pour s'apercevoir que les deux communautés<sup>9</sup> ont développé ces filtres en parallèle en suivant leurs problématiques et sans avoir connaissance l'une de l'autre.

9. Les communautés de l'estimation stochastique et de l'assimilation de données

#### III.D.4. Traiter des bruits non additifs

L'ensemble des développements présentés jusqu'ici se base sur un cas particulier d'erreur du modèle additif. En reprenant l'équation (3.10) le bruit pourrait se modéliser dans un cas plus général :

$$\begin{cases} x_{k+1} &= f(x_k, w_k) \\ z_k &= h(x_k, v_k) \end{cases} \quad (3.120)$$

Wu *et al.* (2005) utilise cette forme dans le cadre de l'*UKF* qu'il appelle *Augmented UKF*. L'idée est d'incorporer à l'état les covariances d'erreurs de manière à les propager également dans le modèle. L'état augmenté  $\tilde{x}_k$  et  $\tilde{P}_k$  devient :

$$\tilde{x}_k = \begin{bmatrix} x_k \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \tilde{P}_k = \begin{bmatrix} P_k & 0 & 0 \\ 0 & Q_k & 0 \\ 0 & 0 & R_k \end{bmatrix} \quad (3.121)$$

Les auteurs de cette forme augmentée précisent qu'elle plus coûteuse que la forme initiale. Le résultat est cependant plus précis et permet de prendre en compte des bruits non-additifs.

#### III.D.5. Ajouter des contraintes

L'utilisation des filtres de Kalman peut poser problème au niveau du modèle utilisé. En effet, les modèles dynamiques peuvent faire intervenir des paramètres particuliers ayant des domaines de variation réduits, ou impliquant des relations d'égalité ou de grandeurs les uns par rapport aux autres. Il est possible de limiter les variations du filtre par un réglage de  $Q$  et une valeur initiale appropriée. Cependant, lorsque ce n'est pas le cas, le modèle dynamique n'est plus sollicité dans la même zone de fonctionnement et peut devenir totalement aberrant. Un tel comportement peut être critique pour certains systèmes. En guise d'illustration, on peut prendre l'exemple d'un coefficient de frottement négatif, de la limitation de la norme d'une accélération (Sircoulomb *et al.*, 2008) ou des liens existant entre les paramètres d'un réacteur aéronautique (Simon et Simon, 2010).

Le problème peut se poser de telle sorte que pour un état  $x_k$  d'un système donné, on ait :

**Pour des contraintes linéaires**, on introduit un couple (matrice, vecteur)  $(D, d)$  qui permet de traduire des contraintes d'égalité ou d'inégalité :

$$Dx_k = d \quad \text{et} \quad Dx_k < d \quad (3.122)$$

**Pour des contraintes non-linéaires** où on introduit cette fois un couple (fonction, vecteur)  $(\tilde{D}, d)$  pour traduire les relations d'égalité ou d'inégalité :

$$\tilde{D}(x_k) = d \quad \text{et} \quad \tilde{D}(x_k) < d \quad (3.123)$$

Cet aspect a été fortement développé il y a quelques années et la revue proposée par Simon (2010) résume de manière très complète les différentes stratégies développées tout en fournissant les codes des différentes approches. Il en conclut qu'il existe différentes solutions pour les filtres *UKF* et *EKF*. Les égalités sont plus faciles à intégrer que les inégalités et la non-linéarité plus délicate à gérer. Parmi

les méthodes développées, on retiendra la méthode des mesures parfaites qui consiste à augmenter l'état avec les contraintes d'égalité. Pour les contraintes d'inégalité non-linéaires elles peuvent être linéarisées pour ensuite appliquer une projection ou un calcul d'optimisation plus complexe. Straka *et al.* (2012b) proposera plus tard pour tout filtre local comme les filtres de Kalman la généralisation d'une autre méthode : celle de la fonction de densité de probabilité tronquée. Cette méthode consiste à calculer la fonction de vraisemblance et de venir annuler la probabilité d'avoir une solution sur les zones où les contraintes ne sont pas vérifiées. La fonction tronquée est ensuite approximée sous forme de gaussienne pour pouvoir être traitée par le filtre. C'est de loin la solution la plus rapide pour traiter un problème de contrainte quelconque. Mais c'est également la moins stable puisque l'espace n'est réduit qu'après le passage du système dynamique.

### III.D.6. Traiter les cas non-gaussiens

La nature gaussienne des variables aléatoires traitées par le filtre de Kalman est souvent le meilleur argument de ses détracteurs. En effet il est difficile d'échapper à des signaux gaussiens à l'intérieur du filtre. Néanmoins des stratégies sont possibles en utilisant les propriétés d'approximation de distribution des modèles de mélange gaussien (Sorenson et Alspach, 1971) les *Interacting Multiple Models (IMM)*. Nous les développerons dans la section III.E.2. puisqu'il s'agit d'un schéma de filtrage utilisant des filtres de Kalman. L'état de l'art fait également état d'autres approches. Mirza (2011) par exemple utilise un estimateur séquentiel robuste de manière à rejeter les mesures non-gaussienne. Si cette approche semble fonctionner, elle ne fait qu'accroître les suspicions sur l'impossibilité d'adapter Kalman avec des mesures non-gaussienne. D'autres approches adoptent le choix déterministe de la transformée non-parfumée pour pouvoir modéliser des distributions différentes (GS-UKF), ou recourent là encore des modèles de mélanges gaussiens (Kottakki *et al.*, 2014). Ces filtres de Kalman modifiés perdent cependant leur efficacité puisque les étapes de mise à jour doivent être à nouveau résolues à partir du théorème de Bayes.

### III.D.7. Quelques outils : initialisation de paramètres et utilisation en lissage.

#### a. Initialisation de paramètres

L'initialisation de l'algorithme est une étape importante pour prévenir d'éventuelles divergences du filtre. Dans le cadre de l'identification de paramètres, il peut être judicieux de calibrer les paramètres. Marchesseau *et al.* (2013b) propose un algorithme simple basé sur la transformée non-parfumée que nous proposons de rappeler ici. On pose un vecteur d'état initial  $x_0$  que l'on transforme dans l'espace non-parfumé sous la forme d'un set de *sigma points*  $\chi_0$  et un vecteur d'observation calculé dans le même espace  $\zeta_0$  correspondant à un vecteur d'observation initial  $\hat{z}_{0|0}$  :

$$\chi_0 = UT(x_0) \quad (3.124)$$

$$\zeta_0 = h(\chi_0) \quad (3.125)$$

On peut alors calculer la corrélation croisée  $P_0^{xz}$  et appliquer l'étape de gain et de correction de Kalman à l'aide des premières observations  $z_{obs}$  pour obtenir le vecteur d'états calibré :

$$x_0^{calib} - x_0 = P_0^{xz} \cdot P_0^{zz-1} (z_{obs} - \hat{z}_{0|0}) \quad (3.126)$$

Cet algorithme est semblable aux dernières étapes du filtre de Kalman. Quelques différences existent puisque si les poids utilisés sont précisés, le paramètre  $c$  définissant l'écart des *sigma points* par rapport à la moyenne ne l'est pas.

#### b. Utilisation en lissage

La nature séquentielle des filtres de Kalman définit un cadre d'estimation de l'état à l'aide d'observations limitées aux états passés. Si les schémas temps réels ont besoin de cette contrainte, il est dommage dans l'estimation de paramètres notamment de se priver à posteriori de la connaissance des états suivants pour effectuer une meilleure estimation. Le problème consiste alors à estimer un état  $k$  en fonction d'observations connues à l'état  $\eta$  avec  $\eta > k$ . Ce qui revient à estimer  $p(x_k | z_1, z_2, \dots, z_k, \dots, z_\eta)$ . Le problème s'exprime de différentes manières :

- le lissage d'un point fixe (*fixed-point smoothing*) en fonction d'un nombre croissant d'observations ;
- le lissage retardé (*fixed-lag smoothing*) fournit l'estimation d'un état antécédent avec un retard fixe ;
- le lissage sur un intervalle fixe (*fixed-interval smoothing*) fournit l'estimation optimale d'une trajectoire sur un intervalle en connaissance des observations de l'intervalle.

Nous proposons ici de décrire une des variantes les plus connues du lissage sur un intervalle fixe : le lisseur de Rauch-Tung-Striebel (RTS)<sup>10</sup>.

<p><b>Inversion temporelle :</b></p> $k = [\eta, \dots, k, \dots, 1, 0] \quad (3.127)$ <p><b>Prédiction :</b></p> $\chi_{k+1} = f(UT(\hat{x}_k, P_k), k) \quad (3.128)$ $x_{k+1}^- = \chi_{k+1} \cdot w^m \quad (3.129)$ $P_{k+1}^- = \chi_{k+1} \cdot W \cdot \chi_{k+1}^T + Q_{k+1} \quad (3.130)$ <p><b>Mise à jour :</b></p> <p style="text-align: center;"><b>Corrélation entre l'état précédent et l'état prédit :</b></p> $P_{k+1}^{xx^-} = \chi_k \cdot W \cdot \chi_{k+1}^T \quad (3.131)$ <p><b>Gain :</b></p> $D_k = P_{k+1}^{xx^-} \cdot [P_{k+1}^-]^{-1} \quad (3.132)$ <p><b>Correction :</b></p> $\begin{cases} x_k^{liiss} = x_k^- + D_k \cdot [x_{k+1}^{liiss} - x_{k+1}^-] \\ P_k^{liiss} = P_k^- + D_k [P_{k+1}^{liiss} - P_{k+1}^-] D_k^T \end{cases} \quad (3.133)$
--

L'algorithme fonctionne en deux étapes : une première avec les étapes classiques du filtre de Kalman sur la fenêtre  $z_1, \dots, z_\eta$  puis une seconde étape en effectuant une inversion temporelle. Cette seconde

10. L'algorithme est tiré de l'article de Särkkä (2010) que nous encourageons à consulter pour approfondir.

étape peut s'écrire comme indiqué sur l'algorithme ci-dessus. Elle est déclenchée à partir du moment où l'état  $\eta$  a été estimé en posant  $x_\eta^{liss} = \hat{x}_{\eta|\eta}$  et  $P_\eta^{liss} = P_{\eta|\eta}$ . Cet algorithme s'effectue donc en inversion temporelle par rapport à la première étape. Il vient lisser de proche en proche tout en évitant d'inverser la fonction d'évolution. Le lissage peut s'avérer judicieux pour estimer des variables ou des paramètres ayant une dynamique plus faible que les variables estimées en temps réel.

### III.E. Schémas de filtrage

Les filtres de Kalman ont été modifiés pour permettre de fonctionner dans d'autres conditions que celles prévues dans les hypothèses de départ. D'autres solutions ont également été développées soit dans la manière de poser le modèle, soit dans la manière d'utiliser les filtres de Kalman. Dans ce paragraphe nous proposons d'introduire quelques schémas utilisés pour traiter l'identification de paramètres, puis des schémas permettant d'utiliser plusieurs filtres de Kalman ensembles.

#### III.E.1. Filtrage simple, augmenté, dual, joint

Dans le paragraphe précédent, une manière de modifier le vecteur d'état et la matrice de covariance associée a été vue pour traiter les bruits non-additifs. Les filtres de Kalman étant souvent utilisés pour faire de l'identification de paramètres, nous détaillons les deux schémas principaux employés : le filtrage joint et le filtrage dual. Dans les deux cas, le problème précédemment posé se reformule en introduisant un vecteur de paramètres  $\theta_k$  à l'état  $k$  qui ne varie pas dans le temps. Dans un cadre discret leur modèle d'état d'évolution s'écrit sous la forme :

$$\theta_k = \theta_{k-1} \quad (3.134)$$

Le système d'état peut alors se réécrire sous la forme :

$$\begin{cases} x_k = f(x_{k-1}, \theta_{k-1}) + w_{k-1} \\ z_k = h(x_k, \theta_k) + v_k \end{cases} \quad (3.135)$$

À partir de ce système d'état, deux stratégies peuvent être mises en place :

**Le filtrage dual** consiste à utiliser deux filtres de Kalman qui évoluent en concurrence (Haykin, 2001). Comme l'illustre la figure 3.6, un premier filtre gère les variables d'état et un second filtre les paramètres. Les équations d'évolutions sont alors coupées en deux parties, une pour l'état et l'autre pour les paramètres (la fonction identité). Le premier filtre a besoin des paramètres pour pouvoir utiliser la fonction d'évolution. Ce n'est pas le cas du second filtre dont l'équation d'évolution des paramètres est constante. En revanche, étant donné que les paramètres ne font pas partie des variables observées, aucune des deux fonctions d'observation ne nécessite d'avoir le vecteur des paramètres. Le second filtre aura tout de même besoin de l'innovation du premier filtre pour pouvoir effectuer l'étape de correction.

**Le filtrage joint** utilise quant à lui un vecteur d'état augmenté en concaténant le vecteur d'état avec les paramètres. Le système dynamique est alors résolu classiquement par un seul filtre de Kalman.

$$\tilde{x}_k = \left[ x_{k-1}^T \dots \theta_{k-1}^T \right]^T \quad (3.136)$$

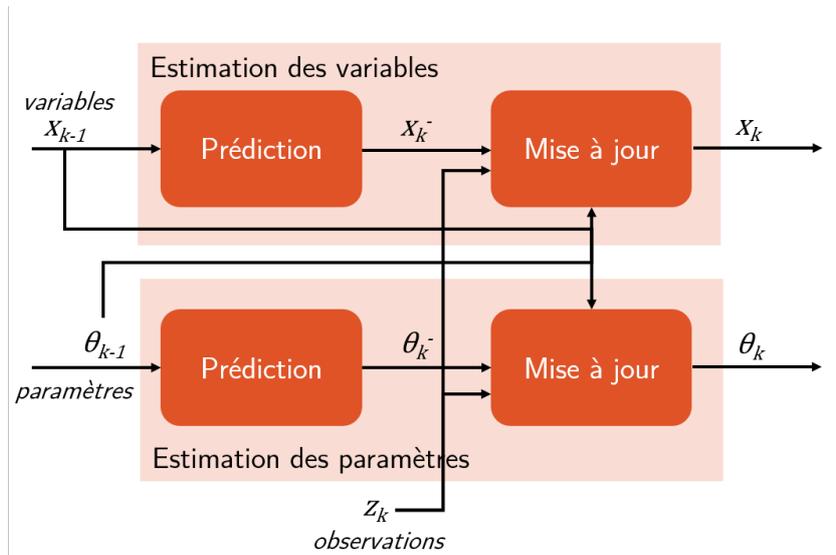


Figure 3.6 – Schéma de la structure d'un filtre dual.

$$\begin{cases} \tilde{x}_k &= \tilde{f}(\tilde{x}_{k-1}) + w_{k-1} \\ z_k &= \tilde{h}(\tilde{x}_k) + v_k \end{cases} \quad (3.137)$$

La littérature n'est pas unanime sur les méthodes les plus performantes. Mariani et Corigliano (2005) affirment que l'approche duale est plus juste que l'approche jointe tout en demandant plus de ressources tandis que d'autres concluent à l'inverse (Hegyí *et al.*, 2006). Nous préférons souligner le fait que la structure duale offre la possibilité d'arrêter l'identification tout en continuant à exécuter le filtre.

### III.E.2. Modèles multiples en interaction (IMM)

Comme nous avons pu le voir précédemment, la simplicité de l'algorithme de Kalman pour résoudre l'inférence bayésienne repose principalement sur deux hypothèses : la linéarité et la nature gaussienne des variables aléatoires. L'hypothèse de linéarité a pu être contournée en sacrifiant l'optimalité de l'estimateur obtenu. Par contre, il est délicat de contourner l'hypothèse sur la nature gaussienne des variables aléatoires. La simplicité de l'algorithme de Kalman repose directement sur cette hypothèse dans le calcul de l'étape de mise à jour. La généralisation à une distribution quelconque nous ferait revenir au cas complexe à traiter de l'inférence bayésienne récursive. L'avantage de rapidité d'exécution du filtre serait perdue.

Nous avons déjà évoqué plus haut quelques auteurs modifiant le filtre pour adopter un modèle de somme de gaussiennes. Le résultat est un filtre qui perd ses propriétés de rapidité sur une partie de l'algorithme. Le filtre multi-modèles interactif (*Interacting Multiple Models* ou IMM) est un algorithme développé et employé dans le domaine du pistage (*tracking*). Dans ce domaine, l'importance d'une réponse temps réel est primordiale. La force de cet algorithme est de pouvoir utiliser plusieurs filtres en même temps. Il suffit alors d'exécuter différents modèles et de sélectionner le comportement le plus vraisemblable par rapport aux données observées.

Rappelons la théorie de cette approche pour illustrer une façon de faire interagir plusieurs filtres de Kalman<sup>11</sup>. Ces modèles sont efficaces lorsque le phénomène observé ne peut être décrit par un modèle unique ou que ce modèle s'avère trop complexe. On parle alors d'un modèle ayant plusieurs modes de fonctionnement. Supposons un cas où le modèle ait  $M$  modes de fonctionnement. On souhaite estimer l'état  $x_k$  à l'aide de données  $z_k$  et de  $M$  filtres de Kalman que l'on dénommera  $M^1, M^2, \dots, M^M$ . La densité de probabilité de l'état  $k-1$  à postériori est posée gaussienne sachant que le filtre  $i$  est sélectionné au pas  $k-1$ , on définit à l'initialisation :

$$P(x_{k-1}|r_{k-1} = i, z_{1:k-1}) = \mathcal{N}(\hat{x}_{k-1}^i, P_{k-1}^i) \quad (3.138)$$

On suppose également les probabilités de mixage connues à postériori :

$$\mu_{k-1}^i = P(r_{k-1} = i|z_{1:k-1}) \quad (3.139)$$

Le filtre *IMM* s'écrit en quatre étapes :

- On définit les probabilités de mixage entre les différents filtres, ce qui revient à calculer la probabilité que le filtre  $M^i$  ait été choisi à l'instant  $k-1$  sachant que le filtre  $M^j$  est choisi à l'instant  $k$  :

$$\mu_{k-1}^{i|j} = P(r_{k-1} = i|r_k = j, z_{1:k-1}) \quad (3.140)$$

$$= \frac{P(r_k = i|r_{k-1} = j, z_{1:k-1}) \cdot P(r_{k-1} = i|z_{1:k-1})}{\sum_{l=1}^M P(r_k = j|r_{k-1} = l, z_{1:k-1})P(r_{k-1} = l|z_{1:k-1})} \quad (3.141)$$

$$= \frac{\pi^{ij} \mu_{k-1}^i}{\sum_{l=1}^M \pi^{lj} \mu_{k-1}^l} \quad (3.142)$$

- On se sert dans un deuxième temps ces probabilités de mixage pour initialiser les  $M$  filtres de Kalman, ce qui donne pour le filtre  $M^j$  la moyenne et la covariance initiale suivante :

$$\hat{x}_{k-1|k-1}^{0j} = \sum_{i=1}^M \hat{x}_{k-1|k-1}^i \cdot \mu_{k-1|k-1}^{i|j} \quad (3.143)$$

$$P_{k-1|k-1}^{0j} = \sum_{i=1}^M \mu_{k-1|k-1}^{i|j} \left[ P_{k-1|k-1}^i + \left( \hat{x}_{k-1|k-1}^i - \hat{x}_{k-1|k-1}^{0j} \right) \left( \hat{x}_{k-1|k-1}^i - \hat{x}_{k-1|k-1}^{0j} \right)^T \right] \quad (3.144)$$

L'exécution des  $M$  filtres de Kalman permet d'obtenir la fonction de vraisemblance pour chaque filtre en incorporant les données nouvelles  $z_k$ , soit pour le filtre  $M_j$  :

$$L_k^j = P(z_k|r_k = j, z_{1:k-1}) \quad (3.145)$$

$$= \mathcal{N}(\hat{x}_k^j, S_k^j) \quad (3.146)$$

Mais également les moyennes  $\hat{x}_{k|k}^j$  et covariances  $P_{k|k}^j$  de l'état.

11. Un état de l'art récent sur ces techniques peut être trouvé dans El Mokhtari (2015).

### III. Courte Kalmanographie pour les expérimentateurs à vocations non linéaires

— On peut alors mettre à jour les probabilités des modèles :

$$\mu_k^j = P(r_k = j | z_{1:k-1}) \quad (3.147)$$

$$= \frac{P(z_k | r_k = j, z_{1:k-1}) \cdot P(r_k = j | z_{1:k-1})}{\sum_{l=1}^M P(z_k | r_k = l, z_{1:k-1}) \cdot P(r_k = l | z_{1:k-1})} \quad (3.148)$$

$$= \frac{L_k^j \cdot \tilde{\mu}_{k-1}^j}{\sum_{l=1}^M L_k^l \cdot \tilde{\mu}_{k-1}^l} \quad (3.149)$$

$$\text{avec } \tilde{\mu}_{k-1}^j = \sum_{i=1}^M \pi_{ij} \cdot \mu_{k-1}^i \quad (3.150)$$

— On recombine alors les estimations des différents filtres :

$$\hat{x}_{k|k} = \sum_{j=1}^M \mu_k^j \cdot \hat{x}_k^j \quad (3.151)$$

$$P_{k|k} = \sum_{j=1}^M \mu_k^j \left[ P_{k|k}^j + (\hat{x}_{k|k}^j - \hat{x}_{k|k}) (\hat{x}_{k|k}^j - \hat{x}_{k|k})^T \right] \quad (3.152)$$

Les filtres multi-modèles interactifs nous ont permis de montrer un exemple d'utilisation des filtres de Kalman dans une architecture de filtrage plus complexe ; elle exploite le contexte bayésien de l'estimateur de Kalman pour pouvoir lier les filtres entre eux et faire de l'estimation multi-hypothèses.

#### III.F. Performances des filtres de Kalman

L'usage de filtres de Kalman amène plusieurs questions concernant le modèle employé. Certains critères peuvent être utilisés pour déterminer certaines caractéristiques propres au modèle permettant de qualifier son domaine de fonctionnement mais également ses performances. Nous rappellerons sans exhaustivité ces notions dans ce paragraphe qui clôturera cette section.

##### III.F.1. Performance : borne de Cramér-Rao

L'utilisation de filtres de Kalman conduit à construire des systèmes permettant d'observer des variables ou des paramètres définis dans un modèle dynamique. Pour pouvoir utiliser ce système il est nécessaire de pouvoir estimer ces performances. Nous avons pu voir dans la section précédente comment assurer la convergence de l'estimateur, nous allons ici rappeler un critère souvent employé pour calculer l'écart minimum obtenu par un estimateur. Le critère utilisé est la borne de Cramér-Rao qui permet d'accéder à la valeur de la borne inférieure de la variance d'un estimateur.

Pour un estimateur d'un paramètre  $\hat{\theta}$  contenu dans un vecteur d'état  $x_k$  on définit la borne de Cramér-Rao comme :

$$\text{var}(\hat{\theta}) \geq J(x_k)^{-1} \quad (3.153)$$

en définissant la matrice  $J(x_k)$  la matrice d'information de Fisher à partir des mesures  $z_k$  :

$$J(x_k) = -E \left[ \nabla_{x_k} (\nabla_{x_k} \ln(p(x_k, z_k)))^T \right] \quad (3.154)$$

Dans le cadre du filtre de Kalman, il a été montré que la borne pouvait être obtenue par une relation de récurrence (Naik *et al.*, 2015). Ainsi pour un modèle dynamique non-linéaire augmenté (équation (3.137)) à l'instant  $k$ , on définit  $F_k$  et  $H_k$  les jacobiniennes de la fonction d'évolution et d'observation, en posant  $Q$  et  $R$  les covariances des erreurs du modèle et de la mesure :

$$J_k = D_k^{22} - D_k^{21} (J_{k-1} + D_k^{11})^{-1} \cdot D_k^{12} \quad (3.155)$$

$$\text{avec : } \begin{cases} D_k^{11} &= E \left[ F_k^T(x_{k-1}) Q^{-1} F_k(x_{k-1}) \right] \\ D_k^{22} &= Q^{-1} + E \left[ H_k^T(x_{k-1}) R^{-1} H_k(x_{k-1}) \right] \\ D_k^{12} = [D_k^{21}]^T &= E \left[ F_k^T(x_{k-1}) \right] Q^{-1} \end{cases} \quad (3.156)$$

Cette borne permet de pouvoir estimer les performances du filtre en donnant une borne inférieure à la covariance de l'état c'est-à-dire la confiance maximale possible que le filtre peut avoir. Cependant rien n'assure que l'estimateur n'atteindra cette borne : si elle est atteinte on parlera d'estimateur efficace.

### III.F.2. Observabilité et détectabilité

Un filtre de Kalman permet de mettre en relation un système dynamique avec des observations. L'ajout de paramètres dans le vecteur d'état permet alors de pouvoir faire de l'identification, à condition que l'estimateur de Kalman converge. Dans notre cas nous étudions exclusivement des systèmes dynamiques sans commande et il est rapide de voir que les intermédiaires entre l'actualisation de l'état et les observations sont les fonctions d'évolution et d'observation : ce sont elles qui vont définir si le système est observable.

L'observabilité d'un système est un critère calculé à partir de ces fonctions pour pouvoir assurer qu'à partir d'un état initial et de mesures l'estimateur converge vers la valeur recherchée. On se restreint à un modèle linéaire, ce qui revient à poser les fonctions d'évolution et d'observation sous les formes matricielles  $F$  et  $H$ . L'état est de dimension  $n$ . Il est classiquement posé la matrice d'observabilité suivante :

$$O = \begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-1} \end{bmatrix} \quad (3.157)$$

L'observabilité de la paire  $(F, H)$  peut s'écrire sous la forme des propositions équivalentes suivantes :

La paire  $(F, H)$  est observable ;

$$\iff \text{rg}(O) = n ;$$

$$\iff \begin{bmatrix} F - \lambda I \\ H \end{bmatrix} \text{ est de rang plein pour } \lambda \in \mathbb{C} \text{ (Test de Popov-Belevitch-Hautus).}$$

L'observabilité permet d'assurer que l'équation de Riccati régissant la covariance possède bien une solution stable : il s'agit d'une observabilité globale. Le cas non-linéaire est plus délicat. L'approche la plus commune est d'utiliser des outils de géométrie différentielle (Boutat, 2009) comme la dérivée de Lie pour pouvoir généraliser l'observabilité au sens du rang de la matrice d'observation. Celle-ci peut être alors calculée globalement mais également localement (Bartosiewicz *et al.*, 1998). Pour un système linéaire décomposé sur ses modes propres, l'observabilité du système exige que tous

les modes du système soient stables. La notion de détectabilité est plus faible et plus appropriée dans le cadre d'une identification puisqu'elle exige uniquement que tous les modes instables du système soient observables<sup>12</sup>. Nous citerons également les travaux de Anguelova (2007) qui propose un cadre à l'observabilité de modèles non-linéaires définis de manière différentiels. Enfin, dans le contexte de l'estimation stochastique, Bar-Shalom *et al.* (2004) pose l'observabilité stochastique d'un système comme l'existence d'une covariance de l'erreur finie, ce qu'il traduit comme une condition d'inversibilité de la matrice de Fisher.

L'observabilité d'un système est un sujet ouvert très connecté à une classe de problèmes englobant les filtres de Kalman étudiés ici : les observateurs non-linéaires. Nous n'approfondirons pas cette courte introduction à la frontière de notre cadre d'étude.

#### III.G. Conclusion

Cette section nous a permis d'introduire une famille de filtres multidomains : les filtres de Kalman. Nous avons essayé de faire ressortir leur diversité issue de multiples domaines et permettant d'utiliser le filtre en-dehors des hypothèses d'origine. Cette introduction n'a pas été exhaustive sur les autres filtres pouvant être utilisés à la place pour des raisons de concision. Nous redirigerons le lecteur curieux vers d'autres références pour en savoir plus sur les filtres  $H^\infty$  (Simon, 2006), les filtres gaussiens (Särkkä et Sarmavuori, 2013) ou encore la classe plus large des observateurs de Luenberger.

Nous avons souhaité poser cette section en perspective avec la précédente sur l'assimilation de données (section II). Ce double éclairage à la fois stochastique et comme solution d'un problème inverse nous a permis de mettre en valeur les différentes utilisations de cette famille de filtres, que ce soit pour une approche de filtrage, de lissage, de prédiction ou de recalage de modèles. Une vision chronologique des évolutions est d'ailleurs proposée sur la figure 3.7. Nous souhaitons insister sur la diversité des utilisateurs et donc de problématiques différentes.

Dans cette partie, nous avons exposé le contexte général de l'inférence bayésienne récursive qui permet théoriquement d'estimer les distributions des états d'un système en fonction de mesures. Nous avons pu voir que ces développements théoriques étaient difficiles à résoudre à la fois dans le cas général et sans approximation. Le filtre de Kalman est alors introduit comme la seule solution exacte avec un cas restreint à la fois linéaire et gaussien : dans ces conditions, l'estimateur de Kalman est optimal. En-dehors de ce cas, les méthodes de résolution sont approximatives et perdent leur optimalité. Elles peuvent être séparées en deux parties (Sorenson, 1974) :

**Les filtres globaux** qui proposent d'approximer la distribution à l'aide d'un échantillonnage fini.

Il s'agit des méthodes de Monte-Carlo, dans notre cadre récursif des filtres à particules, ou encore des filtres de Kalman d'Ensemble.

**Les filtres locaux** qui sont basés sur une approximation de la distribution du système sous la forme de ses deux premiers moments. Les filtres de Kalman utilisant cette approximation font partie de cette catégorie : on ne peut garantir leur convergence que localement. Ce désavantage laisse apparaître néanmoins un cadre gaussien extrêmement favorable à des calculs rapides, ce qui en fait la solution la plus efficace pour la résolution de ce type de problèmes.

À travers la revue des différentes caractéristiques d'un modèle, nous avons pu ainsi distinguer deux grandes familles de filtres non-linéaires :

---

12. Une définition dans un contexte plus stochastique peut être trouvée dans Bar-Shalom *et al.* (2004) qui consiste à vérifier l'inversibilité de la matrice de Fisher du système, matrice définie dans l'équation (3.154).

### Chap. 3 : Analyse de vidéos par approche lagrangienne basée sur un modèle dynamique physique

- les filtres utilisant des approximations des fonctions du modèle dynamiques. C'est le cas des filtres *EKF* ou les variantes utilisant des développements de Taylor ;
- les filtres utilisant directement les fonctions non linéaires ou *derivative-free* dont font partie les filtres à particules déterministes.

Pour clôturer cette partie et inciter une fois de plus le lecteur à utiliser ces filtres, il nous a semblé utile de mentionner quelques-unes des bibliothèques disponibles sur le sujet. Nous citerons celle mise en place par l'équipe de S.Särkkä (Hartikainen *et al.*, 11) mais également une bibliothèque qui a découlé du projet MACS de l'INRIA cité plus haut (sous-sous-section II.B.1.). Cette bibliothèque nommée *Verdandi* implémente notamment deux filtres réduits (Chapelle *et al.*, 2013). L'ensemble des implémentations sur les filtres de Kalman contraints exposées par Simon (2006) sont également disponibles sous MatLab®. Enfin une bibliothèque adaptée au langage Python existe avec une excellente documentation (Labbe, 2015).

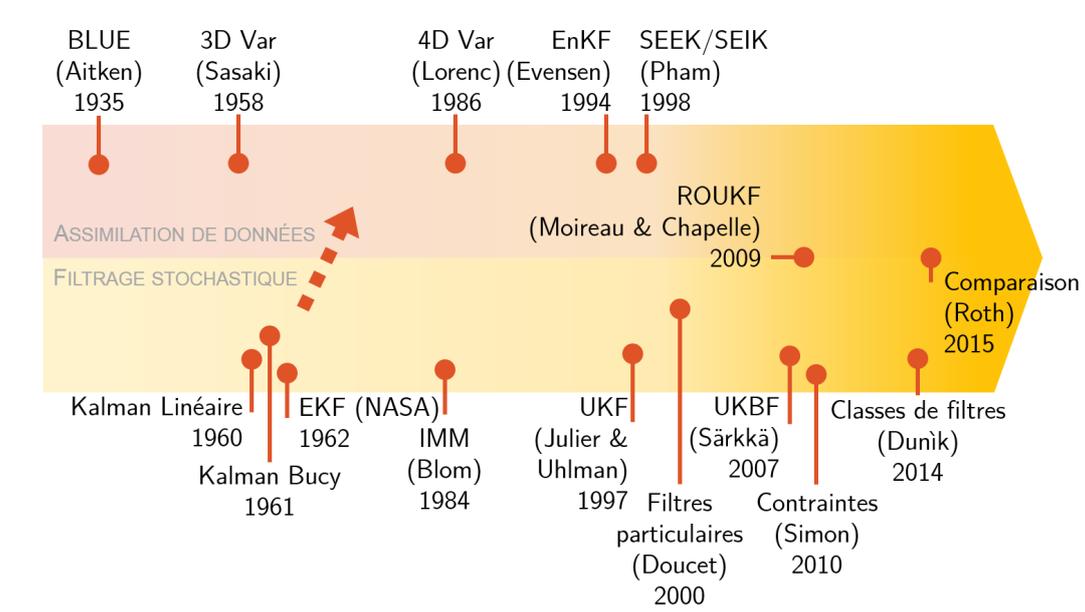


Figure 3.7 – Vision chronologique des articles principaux abordés dans la section.

## IV Proposition d'un schéma d'assimilation de vidéos rapides sur des modèles dynamiques complexes et réduits

Nous avons pu constater dans la section précédente que l'utilisation de modèles élémentaires cinématiques ne permettait pas d'atteindre les objectifs fixés par le projet FAMOSYS : faire de la prédiction, de l'interpolation et de l'identification sur les grandeurs physiques à partir de vidéos rapides avec un schéma temps réel. Nous développerons dans cette section notre approche pour utiliser les données vidéos tout en conservant les propriétés schémas temps réel des filtres de Kalman. Afin d'illustrer ces concepts, une application est ensuite proposée sur le scénario d'une règle glissante : identification du coefficient de frottement, prédiction de la position finale, interpolation d'images.

### IV.A. Identification et extrapolation avec un modèle réduit : une solution pour concilier la donnée image et la rapidité

Les méthodes d'assimilation de données exposées précédemment ouvrent une possibilité de solution en liant un modèle dynamique avec des données. Plusieurs difficultés peuvent être rencontrées pour la mise en place d'une méthode d'assimilation de vidéos rapides. La prédiction du scénario et l'identification de paramètres ne peut se faire que si le modèle utilisé décrit le phénomène observé de manière suffisamment détaillée pour qu'il puisse ajouter de l'information à la scène étudiée. Ainsi il est nécessaire dans le but d'identifier des paramètres de capturer la dynamique d'une scène en faisant intervenir cette fois des modèles physiques. Nous commencerons par étudier quels types de modèles sont susceptibles d'être utilisés. En face du modèle, les données posent également une autre problématique : quelles données doit-on assimiler pour permettre un fonctionnement juste et robuste ? Les méthodes vidéos et schémas de filtrages développés seront décrits après avoir répondu à cette question.

#### IV.A.1. Types de modèles employés

Le chapitre 1 décrit un scénario de vidéo rapide type : il s'agit de l'évolution dynamique d'une scène dont on connaît un scénario qualitatif. La scène est généralement centrée sur l'étude d'un phénomène en particulier. Il est donc possible de décrire chaque scène prise par une modélisation du phénomène étudié dans le cas du scénario. Le modèle introduit devra faire apparaître les paramètres à identifier et les valeurs à estimer.

Une étude dans un cadre mécanique nous permet de trouver une modélisation. Plusieurs origines de modèles sont envisageables : éléments finis, multicorps, analytiques. Nos investigations sur les modèles éléments finis conduisent à des durées de calcul beaucoup trop longues pour pouvoir utiliser ces méthodes dans un cadre temps réel. Cependant l'état de l'art de la sous-section II.B.1. mentionne beaucoup d'auteurs utilisant ce type d'approche et utilisant des stratégies pour aller jusqu'au temps réel (Debunne *et al.*, 2001). Les modèles analytiques réduits utilisent souvent un petit nombre d'équations et de paramètres conduisant à une description de l'objet plus synthétique au niveau des variables utilisées. Ces modèles ont l'avantage de pouvoir être mis en place dans des schémas temps réels tout en restant fidèles aux phénomènes observés. Pour illustrer les principes de la méthode développée, nous choisirons ce type de modèles pour la suite en soulignant deux points négatifs :

- Les modèles utilisés sont souvent de type différentiels, non-linéaires, et avec des variables couplées. Pour un même modèle et un même phénomène, le choix de la forme analytique retenue impacte directement la qualité des résultats.
- La donnée image en entrée fournit des informations concernant la position ou la vitesse des objets de l'image. L'utilisation d'un modèle analytique nous empêche d'utiliser les informations eulériennes de l'image (chapitre 2) puisque ces modèles dynamiques construits sont basés sur une approche lagrangienne.

Posons une version synthétique d'un modèle analytique :

$$\begin{cases} \frac{\partial \mathcal{F}}{\partial t} = \mathcal{F}(x, p, t) \\ z(t) = h(x, t) \end{cases} \quad (3.158)$$

Cette description continue nous permet d'introduire un champ de fonctions d'évolution  $f$  s'appliquant sur un état  $x$  dépendant du temps avec les paramètres  $p$ . Le champ de fonction d'observations  $h$  s'applique alors uniquement à l'état et permet de le projeter dans l'espace des observations.

Le choix de modèles analytiques permet d'avoir une description dynamique de la scène tout en faisant intervenir des paramètres à identifier et des variables à suivre. Ce choix autorise ainsi d'utiliser un modèle rapide et relativement réduit tout en introduisant plus de complexité à travers la description physique du scénario. Deux verrous importants subsistent dans la démarche retenue : l'utilisation en entrée d'informations décrites de manière lagrangienne et l'introduction de systèmes différentiels. Nous proposons de traiter ces deux points en adaptant les traitements vidéos amont et le schéma de filtrage dans les parties qui suivent.

#### IV.A.2. Transformation lagrangienne de la donnée image

Les modèles analytiques réduits introduits dans le paragraphe précédent utilisent une description lagrangienne de l'objet synthétique. En reprenant les développements de la section V du chapitre 2 nous choisirons de la même manière de transformer les données eulériennes des vidéos en données lagrangiennes en utilisant des algorithmes de traitement vidéo. Les hypothèses de nos tests sont identiques à celles utilisés dans le chapitre précédent<sup>13</sup> avec en plus ici l'ajout de la connaissance du scénario de la scène qui n'était pas supposée précédemment. Dans notre cas d'assimilation de données, on retrouvera la problématique de correspondance entre un objet vidéo et un objet mécanique sous une forme plus faible. En effet les méthodes d'extraction de fond que nous utiliserons nous permettront de transformer une série d'image eulériennes en une série d'objets lagrangiens décrits par leurs moments statistiques 2D de Hu (introduit dans la sous-section III.C. du chapitre 2). Ces objets sont des objets vidéos alors que le modèle dynamique que nous utiliserons fera intervenir un objet mécanique. La principale différence avec la partie précédente repose sur l'utilisation d'une méthode d'assimilation de données qui établit un lien entre les deux types d'objets. La flexibilité de ce type de méthodes autorise des erreurs de modèle et nous permet d'envisager d'utiliser un modèle inexact introduisant un biais systématique. L'hypothèse de la connaissance du scénario nous permet de plus de se placer sur des modèles adaptés à la scène et ainsi, de minimiser dès la conception du modèle les erreurs potentielles faites par la caméra dues à la projection 2D.

13. À savoir hypothèse de fond fixe liée à la différence de dynamique entre une faible partie des objets de la scène et le reste de l'image.

## IV. Proposition d'un schéma d'assimilation de vidéos rapides sur des modèles dynamiques complexes et réduits

### a. Choix des données d'assimilation

Parmi les choix de paramétrisation possibles pour décrire la cinématique des objets nous retiendrons les méthodes qui consistent à :

- identifier des points caractéristiques et invariants sur les objets d'intérêt ;
- décrire l'objet à l'aide de la position de son centre de gravité et de son orientation avec ses moments statistiques.

Selon les cas d'étude, notamment lorsque l'objet est déformable, les données d'entrée pourront nécessiter un traitement préalable du support d'étude (positionnement de cibles ou d'un revêtement propice à l'identification des informations cinématiques locales). Dans tous les cas, il s'agira d'extraire des données qui seront des positions, des vitesses ou des angles. Il est important de noter que les données en vitesses peuvent être fortement bruitées à cause de la dérivation numérique induite.

L'utilisation d'une méthode d'assimilation de données est par ailleurs particulièrement adaptée à des instruments hybrides : la structure des méthodes d'assimilation étend intrinsèquement l'intégration de données différentes issues d'autres capteurs, ce que nous verrons dans le chapitre 4 à la section VI.

### b. Chaîne de traitement vidéo

Dans le cas d'un objet indéformable, la même chaîne de traitement vidéo que celle décrites dans le chapitre 2 est employée. Les images sont traitées à l'aide d'un algorithme de soustraction de fond par *GMM*. Le masque binaire résultant contient l'ensemble des objets. Après des opérations de correction morphologique, il est divisé en autant de masques que d'objets afin d'obtenir une collection d'objets. Les moments statistiques 2D sont alors calculés sur chaque masque afin d'obtenir une description réduite de l'objet : position du centre image de l'objet, orientation, longueur des axes de l'ellipse englobant l'objet. Les vitesses sont obtenues par dérivation numérique et la correction du bruit de dérivation est assurée par le processus d'assimilation.

Dans le cas où le modèle développé requiert une série de points décrivant l'objet comme entrée, des méthodes basées sur des recherches de points d'intérêts<sup>14</sup> peuvent être employées pour identifier des parties de l'objet invariantes au cours de la scène. Le suivi de ces points fournit une correspondance entre les détections obtenues à chaque pas de temps. Dans ce cas ces points lagrangiens sont uniquement décrits par leur position tandis que leur vitesse est également issue d'une dérivation numérique.

### IV.A.3. Schéma de filtrage

La forme du modèle analytique proposée au début de cette section possède plusieurs caractéristiques :

- elle est non-linéaire ;
- elle est continue ;
- elle peut se trouver sous une forme différentielle non résolue.

L'emploi d'un modèle analytique de ce type conduit à sortir du cadre linéaire classique avec une mise en équation matricielle. Sorti de ce cadre, l'utilisation de filtres comportant une linéarisation est délicate. Pour conduire le processus d'assimilation stochastique dans ce contexte d'équation différentielles non linéaires, nous avons fait le choix d'employer des filtres de Kalman non parfumés

14. introduites dans la sous-sous-section III.B.2. du chapitre 2.

(UKF). D'une part ce type de filtre permet d'améliorer la précision de l'estimation à l'ordre deux de la fonction. D'autre part un filtre à particules déterministes permet d'utiliser directement le système différentiel non linéaire et continu issu des équations de la physique. Afin de pouvoir utiliser des modèles sous une forme différentielle, nous proposons d'effectuer une intégration à pas variable à l'aide d'une méthode de type Runge-Kutta. Nous avons choisi le schéma proposé par Dormand et Prince (1980). Cependant, sa mise en œuvre nécessite de séparer les variables de la fonction des paramètres, ce qui conduit à utiliser un vecteur d'état augmenté et soit une configuration de type jointe plutôt que duale. On définit la fonction d'intégration de Dormand-Prince de la fonction  $f$  entre les instants  $t_{k-1}$  et  $t_k$  comme étant  $DP(f, x_{k-1}, \theta_{k-1}, t_k - t_{k-1})$  avec  $x$  et  $\theta$  le vecteur d'état et les paramètres. On définit également la fonction  $S$  permettant de séparer un *sigma-point* augmenté  $\tilde{\chi}$  en une partie variable d'état  $\chi$  et une partie paramètres  $\rho$  on a :

**Prédiction :**

**Augmentation du vecteur :**

$$\tilde{x}_{k-1} = \begin{bmatrix} x_{k-1} \\ \theta_{k-1} \end{bmatrix} \quad (3.159)$$

**Transformation non parfumée :**

$$\tilde{\chi}_{k-1} = UT(\tilde{x}_{k-1}, \tilde{P}_{k-1}) \quad (3.160)$$

**Propagation des points par la fonction d'évolution :**

$$(\chi_{k-1}, \theta_{k-1}) = S(\tilde{\chi}_{k-1}) \quad (3.161)$$

$$\chi_{k-1} = DP(f, \chi_{k-1}, \rho_{k-1}, t_k - t_{k-1}) \quad (3.162)$$

$$\tilde{\chi}_k = \begin{bmatrix} \chi_k \\ \rho_k \end{bmatrix} \quad (3.163)$$

$$\text{avec } \rho_k = \rho_{k-1} \quad (3.164)$$

**Transformée non parfumée inverse :**

$$(\tilde{x}_k, \tilde{P}_{k-1}) = UT^{-1}(\tilde{\chi}_k) \quad (3.165)$$

**Mise à jour** La mise à jour est effectuée comme indiqué dans l'algorithme classique.

Nous avons proposé un schéma de filtrage discret utilisant un modèle d'évolution continu. Ce schéma permet de faire de l'identification de paramètres en utilisant un schéma de filtrage joint. À la lecture de l'état de l'art présenté (section III), il semblerait que les modèles continus conduisent à l'emploi d'un filtre de Kalman-Bucy qui permet d'intégrer le système différentiel au niveau du filtre. Un filtre de Kalman discret en intégrant le modèle continu au niveau de la fonction fonctionne en pratique : l'intégration introduira des erreurs qui devront être modélisées comme des erreurs du modèle. Ce choix s'est basé sur une meilleure stabilité du filtre lorsque l'intégration se fait au niveau de la fonction. Notre fonction d'évolution dans ce cas n'est pas un modèle continu mais l'intégration de ce modèle continu. Afin d'illustrer ce schéma nous proposons un cas d'application sur un scénario concret : une règle glissant sur une table.

## IV.B. Application à la dynamique multicorps : la règle glissante

Afin de tester les concepts développés dans ce chapitre, nous proposons une application sur une règle glissant sur une table. Ce scénario répond à plusieurs objectifs :

- être adapté à nos vitesses d'imagerie rapide ;
- illustrer le principe de réalité augmentée ;
- mettre en jeu des paramètres, des variables et des équations différentielles non linéaires issus de modèles physiques ;
- illustrer le concept d'une prédiction ;
- avoir une visée pédagogique afin de pouvoir en faire un démonstrateur des concepts développés ici.

Le scénario ici est simple : on lance une règle glissant sur une table que l'on filme jusqu'à l'arrêt. La caméra est placée en surplomb de la table de manière à considérer celle-ci comme un plan en deux dimensions parallèle au plan image. Ce scénario nous permet moyennant une vitesse d'acquisition moyenne (à partir de 100 images par secondes ou *fps*) de capturer la dynamique du mouvement indépendamment des conditions initiales. Le coefficient de frottement est un paramètre inconnu à identifier. La prédiction de la position finale par extrapolation du modèle permettra en plus d'un affichage en réalité augmentée sur la scène de donner une certaine interactivité avec le lanceur de la règle. La figure 3.16 représente la scène filmée en trois dimensions. Il nous faut développer un modèle ayant pour paramètre le coefficient de frottement et ayant en entrée les positions et vitesses du centre de gravité de la règle avec son orientation.

### IV.B.1. Modèle développé

#### a. Description

Nous proposons une assimilation avec un modèle de glissement comportant deux points de contact. La figure 3.8 illustre la manière dont nous avons posé le problème.

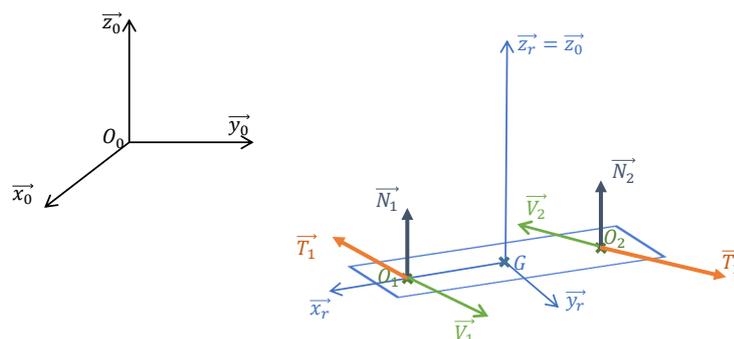


Figure 3.8 – Modèle de la règle glissant sur la table

La règle glisse sur le plan  $(O; \vec{x}_0, \vec{y}_0)$  confondu avec le plan  $(G; \vec{x}_r, \vec{y}_r)$ . Le glissement de la règle sur une table est modélisé par un glissement ponctuel en deux points de contact  $O_1$  et  $O_2$ . La modélisation adoptée et les notations utilisées sont représentées sur la figure 3.8. On note  $\vec{GO}_1 = l_1 \cdot \vec{x}_r$  et  $\vec{GO}_2 = -l_2 \cdot \vec{x}_r$

L'ensemble est soumis à l'accélération de la pesanteur suivant  $-\vec{z}_0$  dans le repère galiléen  $\mathcal{R}_0$  ( $O_0; \vec{x}_0, \vec{y}_0, \vec{z}_0$ ). On note  $\mathcal{R}_x$  le repère lié à la règle ( $G; \vec{x}_r, \vec{y}_r, \vec{z}_r$ ).  $\mu$  désigne le coefficient de frottement de glissement de la règle sur le plan. On note  $\alpha$  l'angle entre  $(\vec{x}_0, \vec{x}_r)$ .

Le développement analytique du modèle est présenté dans l'annexe B.. Nous avons ainsi un modèle défini par le système d'équation (3.166). On pose  $\vec{T}_1, \vec{T}_2$  les forces de frottement et  $\vec{V}_1, \vec{V}_2$  les vitesses aux points  $O_1$  et  $O_2$ . En définissant  $\vec{v}_x, \vec{v}_y$  et  $\dot{\alpha}$  respectivement les vitesses horizontales, verticales et angulaires, on obtient le système suivant :

$$\begin{cases} m \cdot \dot{v}_x & = & \vec{T}_{1x} + \vec{T}_{2x} \\ m \cdot \dot{v}_y & = & \vec{T}_{1y} + \vec{T}_{2y} \\ \ddot{\alpha} \cdot \frac{m(L^2+l^2)}{12} & = & l_1 (\cos \alpha \cdot \vec{T}_{1y} - \sin \alpha \cdot \vec{T}_{1x}) - l_2 (\cos \alpha \cdot \vec{T}_{2y} - \sin \alpha \cdot \vec{T}_{2x}) \end{cases} \quad (3.166)$$

Les forces de frottement sont reliées à la vitesse par un coefficient de frottement telles que :

$$\begin{cases} \vec{T}_1 & = & -\mu \cdot mg \frac{\vec{V}_1}{\|\vec{V}_1\|} \\ \vec{T}_2 & = & -\mu \cdot mg \frac{\vec{V}_2}{\|\vec{V}_2\|} \end{cases} \quad (3.167)$$

Afin de mieux modéliser le mouvement, il nous a semblé judicieux d'introduire une paramétrisation des longueurs  $l_1$  et  $l_2$ . Le mouvement d'une règle glissant avec deux points d'appui comporte déjà une hypothèse forte en supposant que seule la médiane de règle comporte des points de frottement. Il est tout à fait possible que ces points de contact ne soient pas symétriques. Comme illustré sur la figure 3.8 nous introduisons  $d \cdot \vec{x}_r$  la distance  $\overline{O_1 O_2}$  et le paramètre  $\epsilon$  un paramètre de translation par rapport au centre de gravité, ce qui donne :

$$\begin{cases} l_1 & = & \frac{d}{2} + \epsilon \\ l_2 & = & \frac{d}{2} - \epsilon \end{cases} \quad (3.168)$$

Après une étape d'étude de sensibilité, il est apparu que le réglage du paramètre  $d$  influait sur le coefficient de frottement. Aussi nous avons choisi de fixer la distance entre les points de contact à la valeur  $\frac{L}{2}$ . On pose alors le vecteur d'état de nos paramètres comme :

$$\theta_k = \begin{pmatrix} \mu \\ \epsilon \end{pmatrix} \quad (3.169)$$

Le système d'équation (3.166) nous donne des équations dynamiques qu'il faudra intégrer. Il est préférable de relier le modèle aux observations les plus précises dont nous disposons (la position), ce qui nous contraint à avoir un vecteur d'état comportant les vitesses et les positions. De plus l'objet de notre scénario étant particulièrement adapté à une description en utilisant son centre de gravité, nous exprimerons le vecteur d'état au point  $G$  :

$$x_k = \begin{pmatrix} x_G \\ y_G \\ \alpha \\ v_x \\ v_y \\ \dot{\alpha} \end{pmatrix} \quad (3.170)$$

#### IV. Proposition d'un schéma d'assimilation de vidéos rapides sur des modèles dynamiques complexes et réduits

Notre modèle réduit en son centre de gravité développe des variables intermédiaires aux points de contact  $O_1$  et  $O_2$  qu'il nous est possible de récupérer en-dehors du formalisme d'état. L'utilisation de ces variables nous permet de faire de la réalité augmentée sans agrandir la taille du vecteur d'état. Par ailleurs l'implémentation du modèle a mis en lumière un problème d'intégration lorsque l'on se rapproche d'une vitesse nulle, classique avec cette approche de modélisation du frottement. Nous avons fait le choix de mettre les vitesses à zéro pour une vitesse limite. Cela permet de rester dans des temps d'intégration acceptables avec en contrepartie un comportement à l'arrêt moins réaliste.

##### b. Simulation

Nous travaillerons sur deux cas tests différents, l'un avec beaucoup de translation dans les conditions initiales, l'autre avec une contribution équilibrée entre rotation et translation. Le tableau 3.3 rassemble les conditions initiales des cas tests dont les vidéos sont présentées sur les figures ci-dessous.

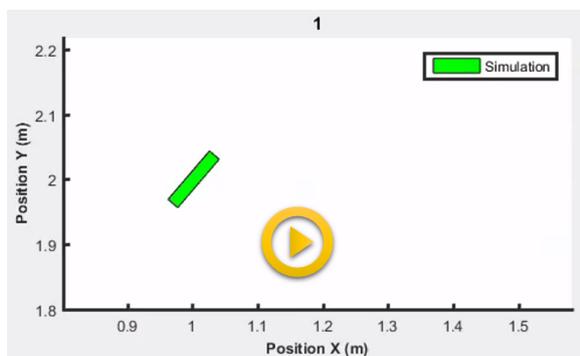


Figure 3.9 – Vidéo du scénario 1

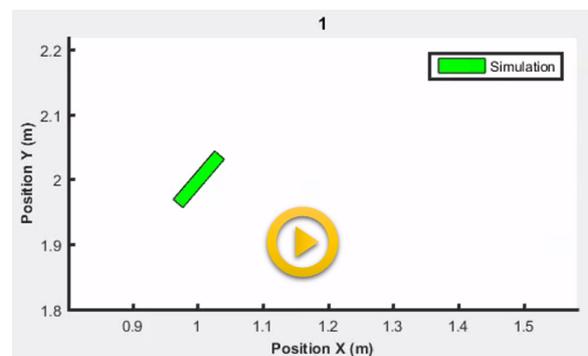


Figure 3.10 – Vidéo du scénario 2

Le premier scénario est plutôt classique; a contrario, nous avons cherché à faire apparaître des comportements plus originaux du modèle dans le second scénario en décalant les points de contacts. Ici notre modèle est assez réaliste pour à la fois générer un cas test et servir de modèle d'assimilation, ce qui n'est pas toujours le cas (par exemple dans les modèles développés dans le chapitre 4). Cette double utilisation du modèle est plutôt pratique puisqu'elle nous permettra de faire une auto-calibration du modèle pour l'identification des paramètres. Les résultats de cette auto-assimilation sont présentés dans la partie suivante.

Table 3.3 – Paramétrage choisi pour les deux scénarios tests présentés

Variable initiale	Paramètre	Valeur scénario 1	Valeur scénario 2
$x_0$		1	1
$y_0$		2	2
$\alpha_0$		$\frac{\pi}{4}$	$\frac{\pi}{4}$
$v_{x0}$		2	2
$v_{y0}$		1	0
$\dot{\alpha}_0$		20	80
	$\mu$	0.4	0.8
	$\epsilon$	0.005	0.09

#### IV.B.2. Assimilation sur simulation

L'assimilation présentée ici correspond aux scénarios générés par le modèle présenté dans la section précédente. Ces simulations sont bruitées avec un bruit blanc de variance  $10^{-2}$  mm pour les positions et  $10$  mm s<sup>-1</sup> pour les vitesses. Ces niveaux sont choisis en fonction des niveaux de bruits obtenus sur des cas réels. Afin de permettre au lecteur d'apprécier les résultats, nous commençons par présenter les résultats sous forme de deux vidéos données figure 3.11 et figure 3.12. Ces vidéos ajoutent à celles présentées dans la section précédente la règle assimilée par le filtre de Kalman en rouge. Nous avons fait le choix de ne pas afficher la covariance de la règle, trop faible pour être visible : nous verrons le résultat sur les courbes données plus loin.

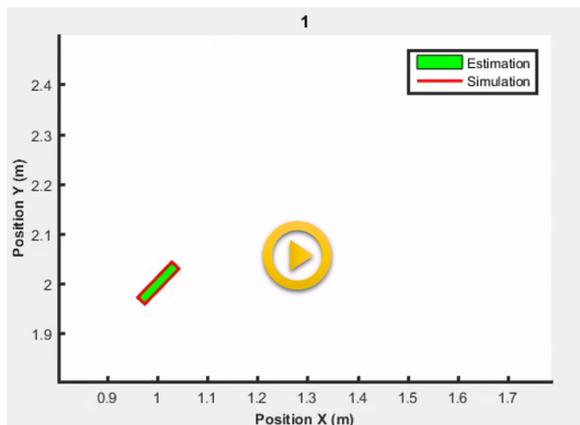


Figure 3.11 – Vidéo du scénario 1

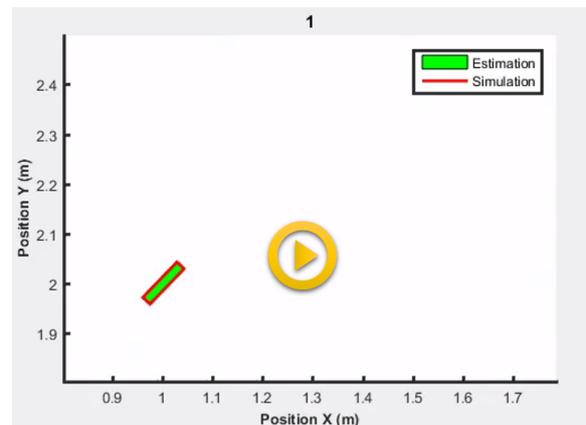


Figure 3.12 – Vidéo du scénario 2

Le paramétrage des valeurs initiales dépend de plusieurs facteurs. La caméra permet facilement d'obtenir une bonne estimation de la position et la vitesse de l'objet à un instant donné. Aussi, compte tenu de la nature locale des filtres de Kalman, il est primordial d'essayer d'initialiser le vecteur d'état proche de la valeur réelle du système. Sur les variables d'état comme pour les paramètres, il nous a semblé judicieux de prendre des valeurs plus éloignées afin de constater la convergence du filtre. On peut constater cet écart sur la figure 3.13 pour le scénario 1 et la figure 3.14 pour le scénario 2. Ces figures présentent en rouge la valeur cachée du système, en bleu les valeurs en entrée du filtre, bruitées avec les niveaux donnés précédemment et la sortie du filtre en jaune. Les covariances du filtre sont cette fois données en vert. On remarquera que celles-ci sont tellement faibles qu'on ne les distingue pas.

Les valeurs des paramètres sont données sur la figure 3.15.

On observe une convergence des différentes variables observées, malgré une initialisation défavorable, ce qui était le comportement attendu. Concernant les paramètres, les deux paramètres convergent vers la valeur indiquée. L'existence d'une valeur absolue sur le paramètre d'écart explique la convergence vers l'opposée de la valeur réelle du système. L'assimilation du modèle prend en moyenne environ 3,8s pour 700 pas de temps. Le modèle est sollicité autant de fois que la dimension de l'espace non parfumé. Néanmoins, le code étant sous Matlab® il est possible de l'optimiser fortement, d'autant plus que la nature particulière et déterministe du filtre ouvre la possibilité d'une parallélisation. Les paramètres convergent jusqu'à atteindre une précision de l'ordre de  $10^{-4}$  en erreur absolue, tandis que le modèle filtre le bruit au niveau des variables d'entrée.

Ces cas tests permettent de valider notre modèle. Nous proposons de l'appliquer sur un cas réel dans

#### IV. Proposition d'un schéma d'assimilation de vidéos rapides sur des modèles dynamiques complexes et réduits

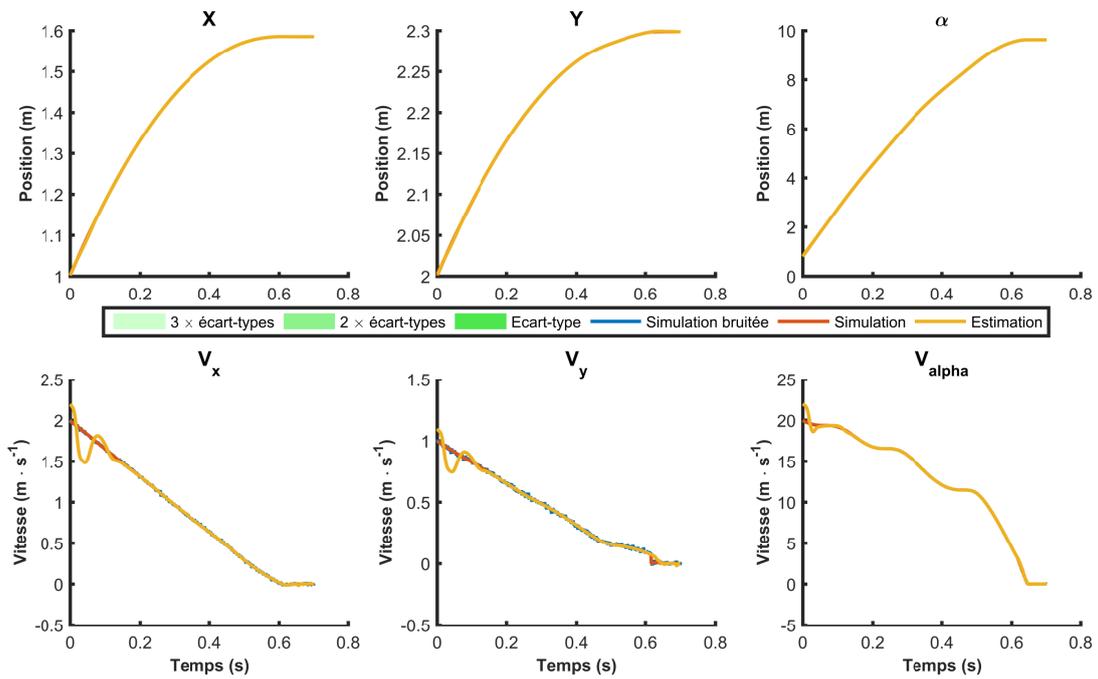


Figure 3.13 – Évolution des paramètres pendant l'assimilation des différentes simulations

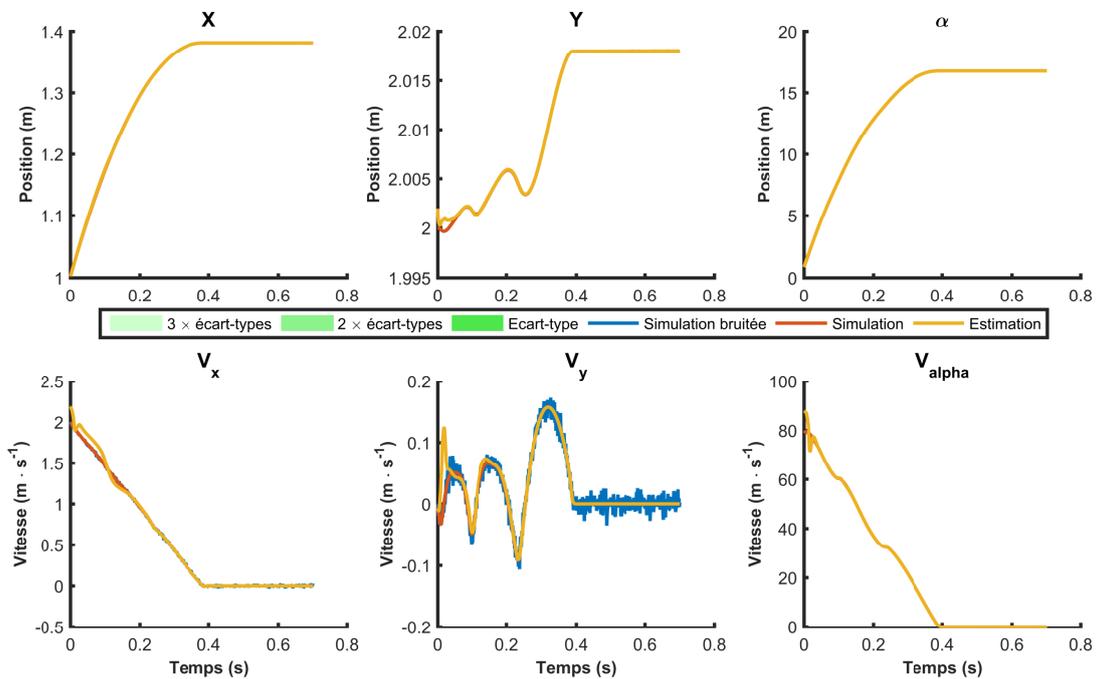


Figure 3.14 – Évolution des paramètres pendant l'assimilation des différentes simulations

le paragraphe suivant.

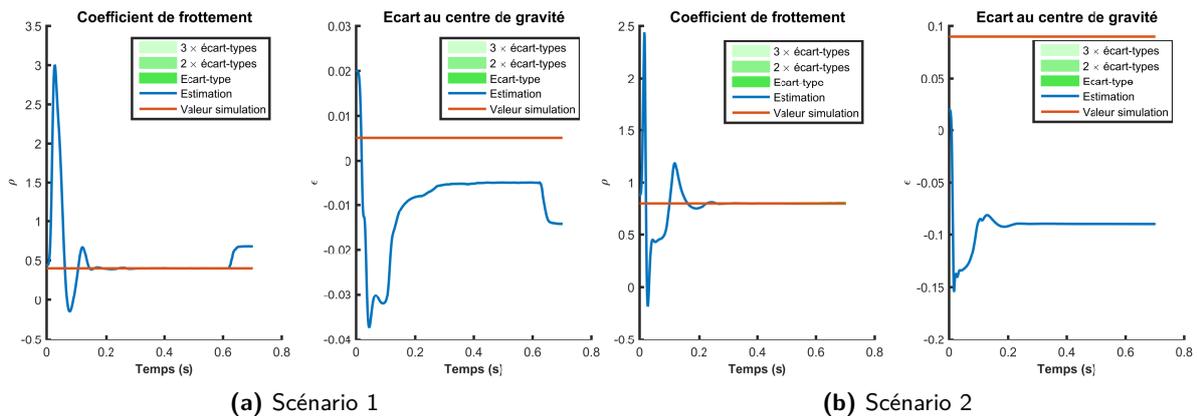


Figure 3.15 – Évolution des paramètres pendant l'assimilation des différentes simulations

#### IV.B.3. Assimilation sur un cas réel

Les cas réels proposés dans cette section ont été effectués avec une caméra AOS technologies dont les caractéristiques sont présentées dans le tableau 3.4. Nous avons employé la caméra avec un pied de manière à se placer au-dessus de la table d'essai dont l'homogénéité en terme de glissement n'est pas assurée : il s'agit d'un établi quelconque avec un revêtement plastique déjà usagé (craquelures, revêtement non-homogène). La caméra a été placée de manière à respecter au mieux les conditions de parallélisme entre le plan image et l'établi. La disposition a été schématisée sur la figure 3.16.

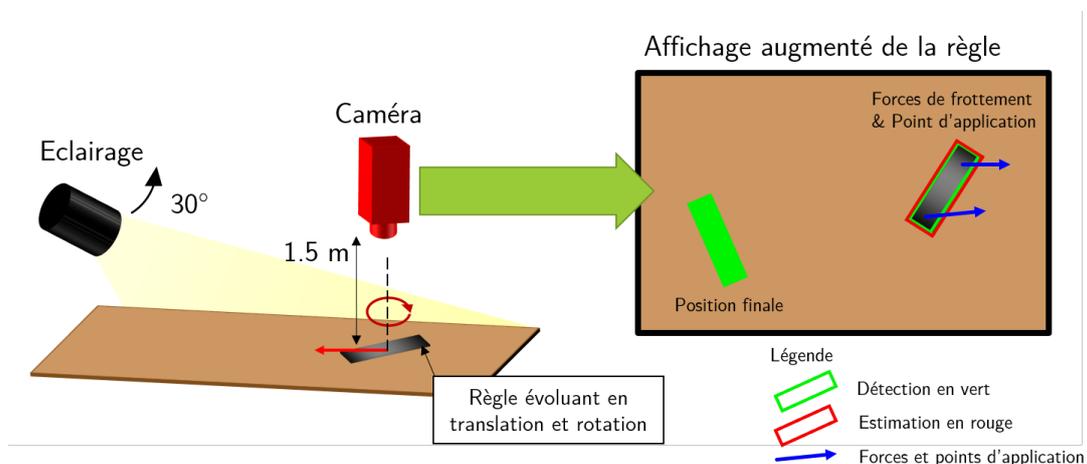


Figure 3.16 – Présentation de la disposition de l'essai d'estimation d'une règle glissante et de l'affichage de l'estimation

Des aberrations existant au niveau des bords de l'image, nous avons préféré éliminer les bords de l'image dans nos cas tests pour garantir une meilleure fidélité.

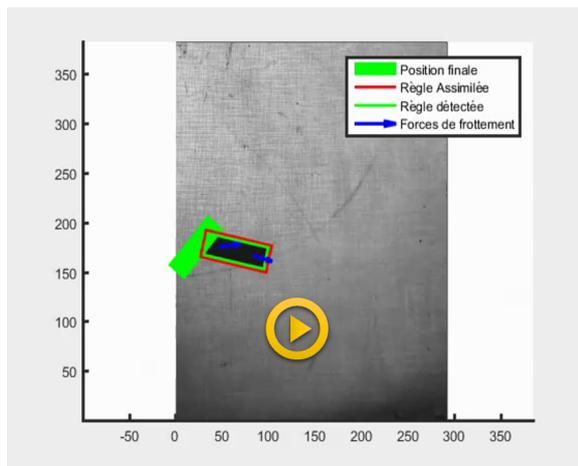
Afin de mieux apprécier les scénarios présentés, nous proposons au lecteur de commencer par regarder les vidéos des deux essais assimilés sur les figures 3.17 et 3.18. Les deux scènes sont filmées en nuances de gris, avec la règle évoluant en noir. Nous avons fait figurer la détection vidéo en vert, afin de rendre visibles les entrées du filtre. Le rectangle rouge représente la règle assimilée, à laquelle ont

#### IV. Proposition d'un schéma d'assimilation de vidéos rapides sur des modèles dynamiques complexes et réduits

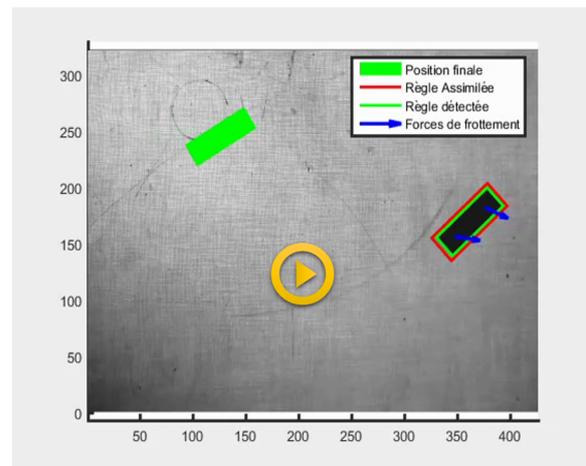
**Table 3.4** – Propriétés de la caméra utilisée

Capteur	CMOS
Vitesse utilisée	1000 Hz
Résolution	900 × 700
Vitesse maximum	16 000 Hz

été ajoutées des informations invisibles : les forces de frottement sur les deux points de contact représentés par des flèches bleues.



**Figure 3.17** – Vidéo du scénario 1 avec affichage des forces de frottement en réalité augmentée et de l'extrapolation de la position finale en vert.



**Figure 3.18** – Vidéo du scénario 2 avec affichage des forces de frottement en réalité augmentée et de l'extrapolation de la position finale en vert.

Le cadre de notre étude nécessitait de pouvoir faire de la prédiction, de l'identification, de l'extrapolation, de l'interpolation et d'illustrer ces résultats à l'aide de réalité augmentée. Les résultats obtenus permettent de valider la méthode proposée dans ce chapitre qui remplit l'ensemble des objectifs attendus.

##### a. Identification et prédiction

Les figures 3.17 et 3.18 affichent la position d'arrêt prédite à partir des informations de chaque état courant. Nous pouvons observer sur les figures 3.19 et 3.20 les courbes des variables observées : le filtre permet de lisser les observations par rapport au modèle tout en effectuant une identification des paramètres de frottement et d'écart au centre de gravité (figures 3.19 et 3.20).

En recalant le modèle manuellement, nous sommes tombés sur une valeur du coefficient de frottement de 0,85 pour un écart nul au centre. Cependant les résultats montraient des erreurs non négligeables au niveau des pentes de vitesse. Certains comportements ne sont pas bien modélisés : c'est le cas par exemple de la légère diminution de vitesse angulaire au début du mouvement pour l'essai 1. Les résultats montrent ici une variation du paramètre uniquement à la fin, ce qui remet en cause nos hypothèses de paramètres fixes. Il est probable que nos deux paramètres ne soient pas constants au cours de l'expérience à cause de la surface de l'établi utilisé ou de la planéité de la règle glissante. L'identification converge bien vers des valeurs de frottement proches de celles obtenues par recalage

Chap. 3 : Analyse de vidéos par approche lagrangienne basée sur un modèle dynamique physique

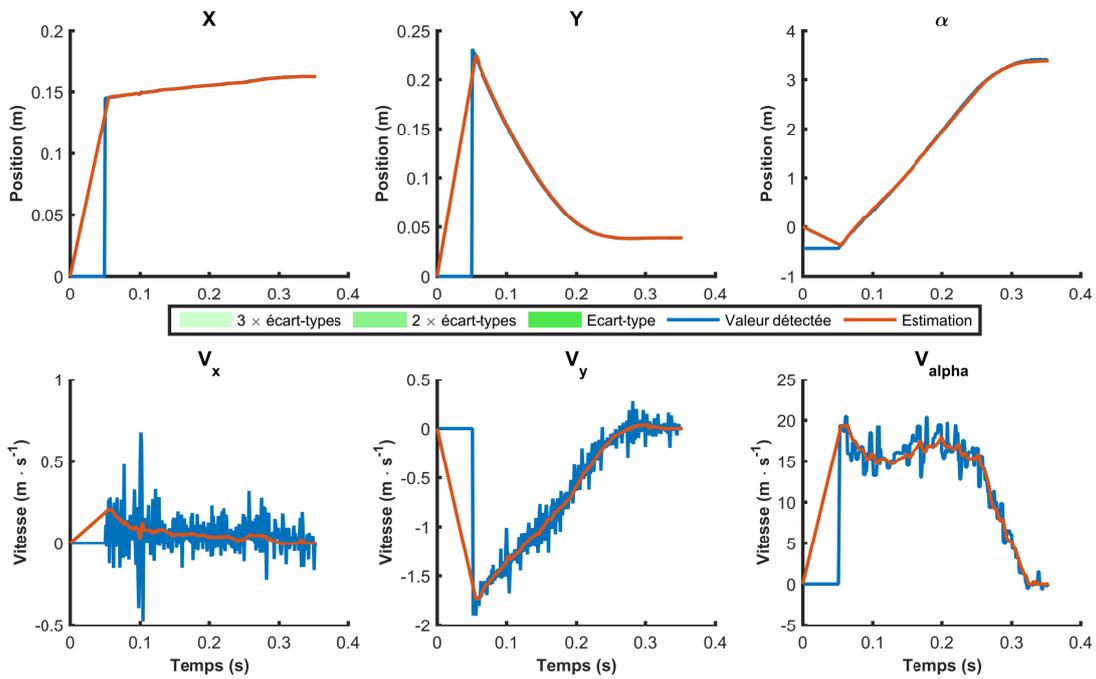


Figure 3.19 – Évolution des variables d'état sur l'essai 1

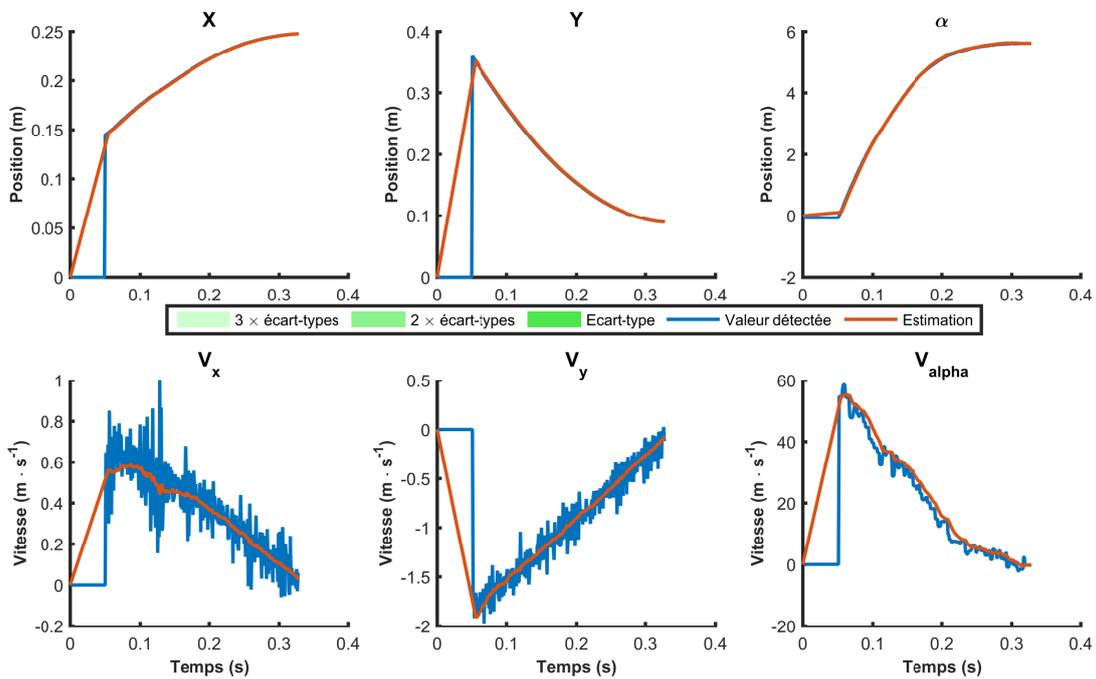


Figure 3.20 – Évolution des variables d'état sur l'essai 2

manuel pour chacun des cas tests. Cependant, nous n'avons pas pu obtenir de convergence rapide similaire à celle obtenue sur les simulations. Il est possible que le cas test ne s'y prête pas entièrement. On remarquera également que l'assimilation utilise dans les deux essais un écart de point dans la phase d'arrêt, ce qui renforce l'hypothèse d'une modélisation imparfaite. L'emploi d'un filtre de

#### IV. Proposition d'un schéma d'assimilation de vidéos rapides sur des modèles dynamiques complexes et réduits

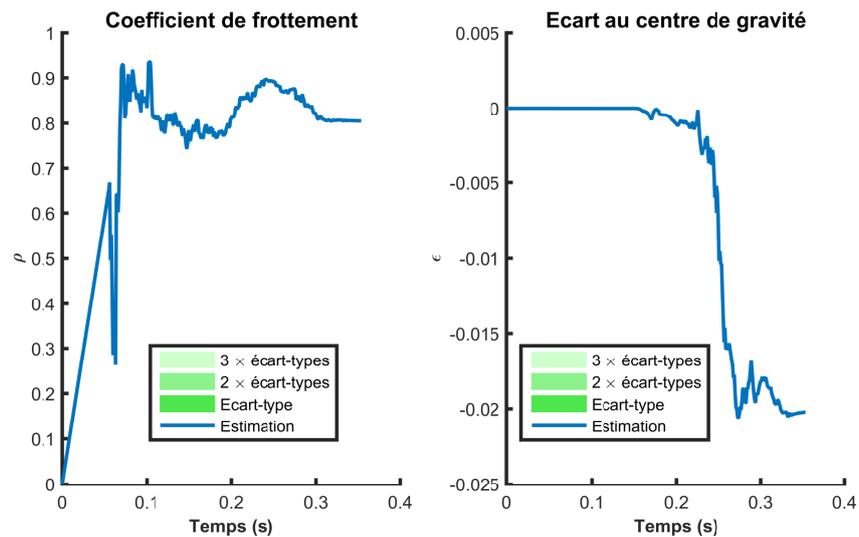


Figure 3.21 – Évolution des variables d'état sur l'essai 1

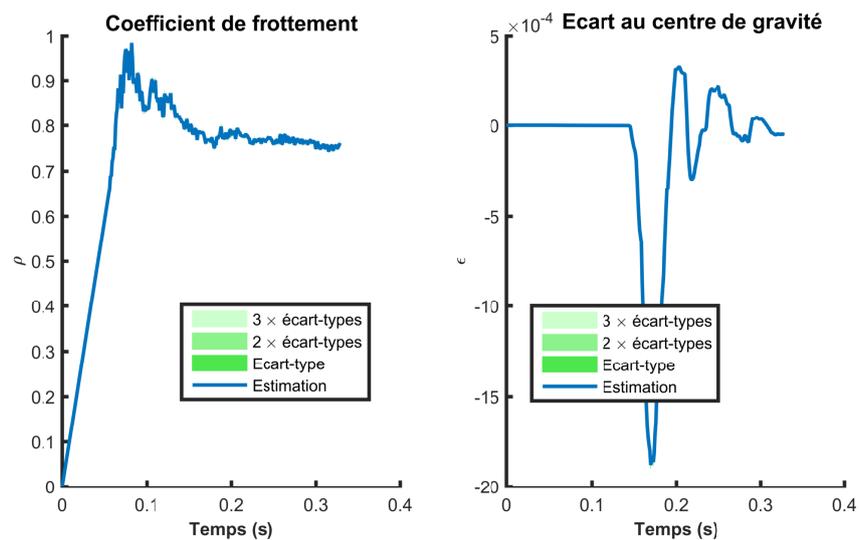
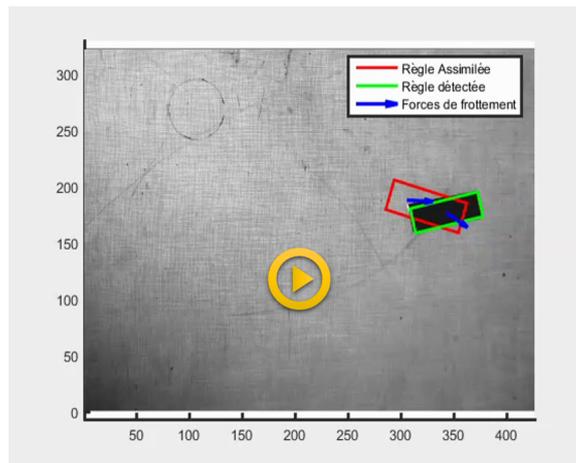


Figure 3.22 – Évolution des variables d'état sur l'essai 2

Kalman nous permet néanmoins de nous servir de ce modèle imparfait avec ces observations, ce qui fait partie des avantages de ces méthodes.

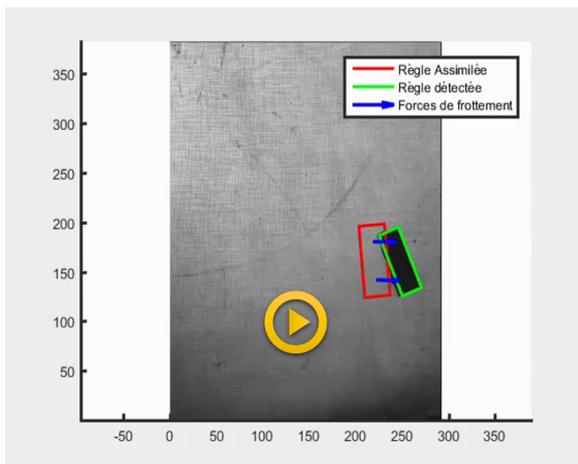
Nous avons également souhaité tester le fonctionnement de cette méthode sur un flux d'observations aléatoires. Les résultats sont proposés sur la figure 3.23. Les mêmes observations que précédemment sont utilisées avec en noir la règle réel sur la table et en vert la visualisation de la règle détectée. Ces observations sont fournies au filtre de manière aléatoire : en rouge la règle prédite par le filtre sur la fréquence initiale de la vidéo. Ces résultats montrent que notre schéma de filtrage utilisant des modèles non-linéaires et sous forme différentielle converge également avec des observations aléatoires, ce qui permet d'élargir le champ des applications possibles.



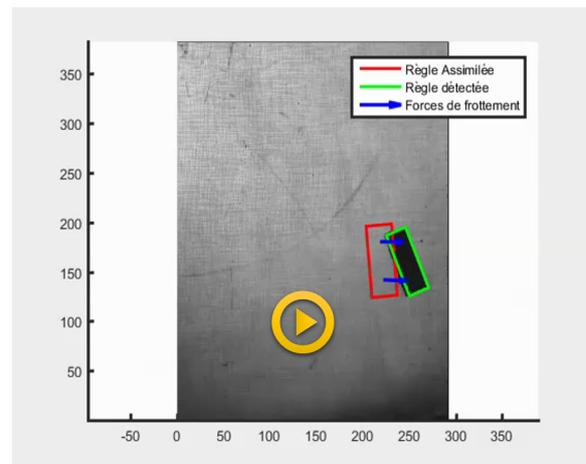
**Figure 3.23** – Vidéo du scénario 1 avec une assimilation sur des observations aléatoires de fréquence moyenne équivalent à un facteur 20

### b. Interpolation

Les résultats du filtre en interpolation sont montrés sur les figures 3.24 et 3.25 pour l'essai 1 et les figures 3.26 et 3.27 pour l'essai 2. La règle assimilée en rouge est interpolée d'un facteur 10 ou 20 par rapport à la détection en vert.



**Figure 3.24** – Vidéo de l'essai 1 assimilé et interpolé d'un facteur 10

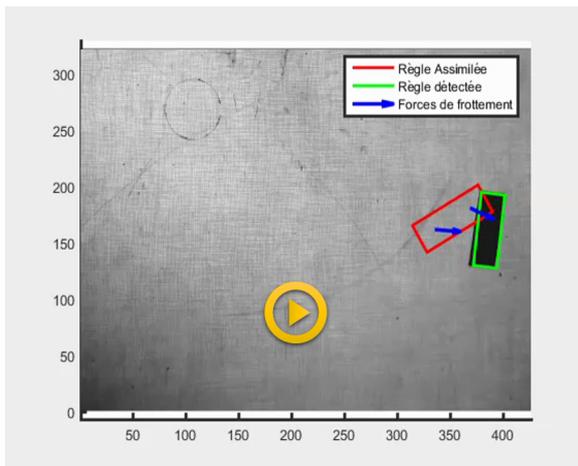


**Figure 3.25** – Vidéo de l'essai 1 assimilé et interpolé d'un facteur 20

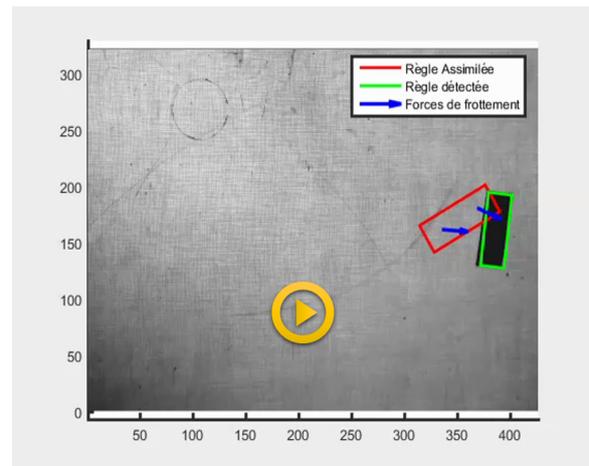
Malheureusement le modèle développé ici n'est pas assez précis pour pouvoir réussir à atteindre de manière correcte un coefficient multiplicateur de 40 qui nous permettrait d'annoncer qu'il nous est possible d'atteindre 1000 fps avec une caméra classique à 25 fps. Nos résultats montrent que cela reste possible de passer d'une acquisition à 50 fps à un rendu de 1000 fps sur un mouvement complexe. Les résultats présentés sur la figure 3.28 présentent des erreurs de l'ordre du pixel en position et de l'ordre du demi degré en angle sur les valeurs interpolées.

L'identification des variables se fait moins efficace dans les cas tests interpolés étant donné qu'il y a moins d'informations disponibles (figures 3.29 et 3.30). C'est également le cas des variables interpolées lorsqu'elles n'ont pas d'observations disponibles et qui rectifient au gré des informa-

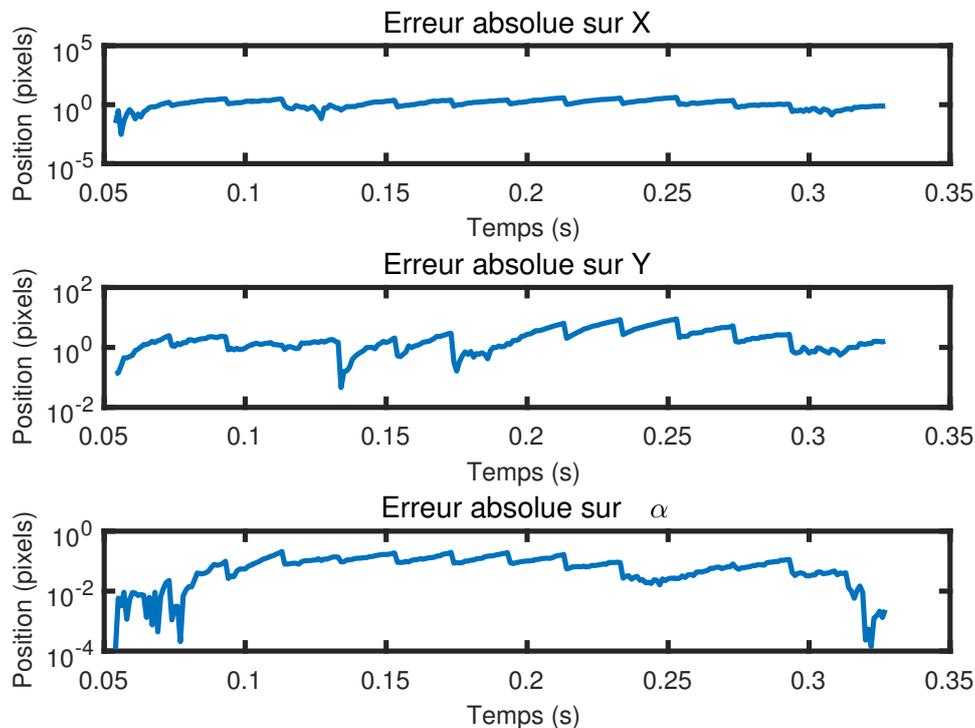
#### IV. Proposition d'un schéma d'assimilation de vidéos rapides sur des modèles dynamiques complexes et réduits



**Figure 3.26** – Vidéo de l'essai 2 assimilé et interpolé d'un facteur 10



**Figure 3.27** – Vidéo de l'essai 2 assimilé et interpolé d'un facteur 20



**Figure 3.28** – Évolution des erreurs de position et d'angle sur l'essai 1

tions entrantes. Dans le but d'un rendu vidéo, il sera possible de synthétiser ces nouvelles images extrapolées en utilisant une méthode de compensation de mouvements<sup>15</sup>.

#### c. Compression

L'utilisation de modèles réduits peut permettre de faire de la compression puisqu'il est possible de décrire la vidéo à l'aide d'un modèle, d'une image initiale et de trois coordonnées du centre

15. méthodes introduites dans la sous-section II.B. du chapitre 2 et appliquées dans le chapitre 4 section VII

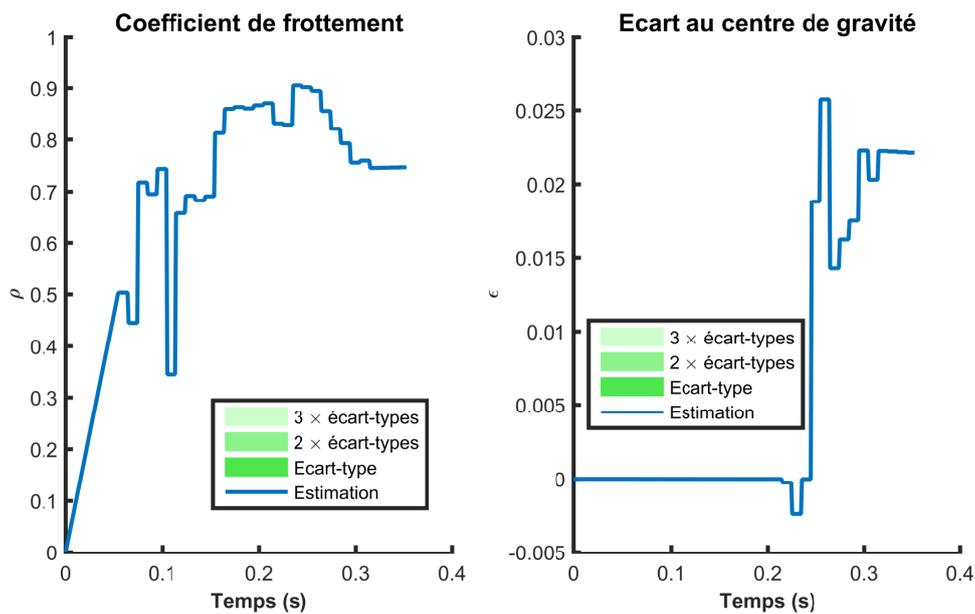


Figure 3.29 – Évolution des paramètres sur l'essai 1

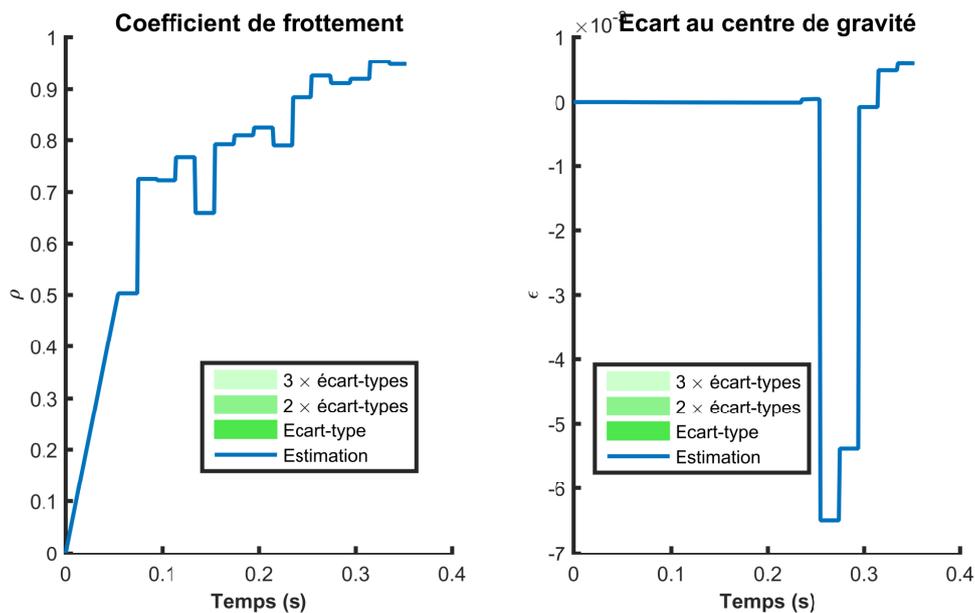


Figure 3.30 – Évolution des paramètres sur l'essai 2

de gravité apparent de l'objet. Cette compression comporte des pertes puisque les propriétés de filtrage de l'assimilation de données ne garantissent pas de repasser par les points observés. En revanche, les données dynamiques de la vidéo compressée seront filtrées selon modèle et minimisera les erreurs d'acquisition. La compression effectuée sera sémantique à la scène puisqu'elle s'effectue en connaissance d'un scénario de la scène. La reconstruction de la scène pourra se faire en utilisant là encore une méthode classique de compensation de mouvement comme dans le chapitre 4 section VII.

#### IV. Proposition d'un schéma d'assimilation de vidéos rapides sur des modèles dynamiques complexes et réduits

##### d. Réalité augmentée

L'utilisation conjointe d'un modèle complexe permet de se servir d'un nombre important de paramètres physiques internes sans que ces-derniers n'entrent dans le vecteur d'état. Dans le cas test de la règle glissante, c'est le cas des vitesses et des forces appliquées aux deux points de contact. Ces variables internes sont disponibles à chaque pas. La manipulation de données images nous offre un espace interactif pour disposer ces informations sur la scène. Dans cet exemple, nous avons choisi d'afficher les forces de frottement qui n'apparaissent pas dans le vecteur d'état mais interviennent dans le modèle.

##### e. Conclusion

L'ensemble des résultats proposés dans cette section montrent l'ensemble des possibilités de notre méthode d'assimilation de vidéos rapides en terme d'extrapolation à partir du modèle et des applications en réalité augmentée. La méthode testée sur un cas réel permet ainsi d'obtenir des résultats en prédiction, interpolation, et identification de paramètres en assimilant les données réelles observées par caméra et un modèle dynamique physique modélisant la scène observée : sur les vidéos présentées, nous avons choisi les forces de frottements.

## V Conclusion

Pour pouvoir atteindre les différents objectifs, nous avons développé une méthode permettant d'utiliser un modèle spécifique faisant intervenir de la physique tout en respectant un schéma de fonctionnement temps réel.

Nous avons choisi de développer une approche basée sur l'utilisation de filtres de Kalman. Ces filtres étant très populaires, nous avons souhaité montrer à quel point l'état de l'art était transverse au domaine d'étude, en terme d'applications mais également en terme d'innovations sur l'algorithme initial. Pour marquer ces multiples angles de vue, nous avons introduit ces filtres à travers une double approche.

**L'assimilation de données** regroupe l'ensemble des méthodes permettant de relier un modèle aux observations. En plus de montrer un cadre d'application sur des systèmes complexes et de grande taille, il semblait particulièrement adapté à notre problématique à travers son approche séquentielle utilisant des filtres de Kalman. Ces filtres utilisent un modèle dynamique pour proposer un triple cadre de filtrage, de lissage ou de prédiction.

**Le contexte bayésien** nous a permis d'éclairer sous un angle stochastique les filtres de Kalman. Une revue de l'état de l'art présentant un apport à travers une approche générale et transverse a été donnée : elle est structurée en trois parties.

Une première partie propose de regrouper de manière synthétique un ensemble d'extensions d'algorithmes de Kalman issus de plusieurs domaines. Nous avons cherché à faire ressortir les différentes versions existantes afin d'offrir au lecteur une vision large des possibilités proposées : fonctions non-linéaires, modèles continus, bruits de fonctionnement non-additifs, contraintes sur le système, système de grande taille.

Une deuxième partie propose quelques schémas de filtrage plus complexes utilisant des filtres de Kalman et permettant notamment de traiter des systèmes non-gaussiens. Enfin, des outils permettant de qualifier un filtre sont rappelés.

Nous avons développé notre approche en prenant appui sur cette revue bibliographique. La méthode que nous proposons permet d'assimiler un modèle dynamique lagrangien issu des équations de la physique du phénomène sous une forme de système différentiel. L'assimilation se sert des observations du système réel en position et en vitesse réalisées par caméra rapide. L'exploitation de la donnée vidéo se fait en transformant les observations eulériennes de la caméra avec une méthode de traitement vidéo développée dans le chapitre précédent. Notre schéma d'assimilation basé sur un filtre *UKF* fait correspondre le modèle aux données et nous permet une utilisation en extrapolation.

Cette méthode est illustrée sur une application d'une scène de règle glissant sur une table. La mise en place d'un modèle à deux points de contact nous permet de proposer des résultats sur la prédiction de la position finale de la règle, l'identification du coefficient de frottement et l'extrapolation de la règle permettant d'augmenter virtuellement la vitesse de la caméra d'un facteur 20. Toutes les variables internes au modèle telles que les forces ou les vitesses peuvent être affichées en réalité augmentée. La vidéo peut être compressée de manière sémantique à la scène en utilisant uniquement les coordonnées du centre de gravité.

## Bibliographie du chapitre 3

- Milena Anguelova : *Observability and Identifiability of Nonlinear Systems with Applications in Biology*. Thèse de doctorat, Chalmers Univ. of Technology [u.a.], Göteborg, 2007.
- Yaakov Bar-Shalom, X. Rong Li et Thiagalingam Kirubarajan : *Estimation with Applications to Tracking and Navigation : Theory Algorithms and Software*. John Wiley & Sons, 2004.
- Z. Bartosiewicz, M. Kosk et D. Mozyrska : Local observability and quasi-observers for nonlinear systems. *In Proceedings of MTNS*, volume 98, juillet 1998.
- Eric Blayo et Maëlle Nodet : Introduction à l'assimilation de données variationnelle. Rapport technique, mars 2012.
- Driss Boutat : La notion d'observateur pour les systèmes non-linéaires. *Supméca pour la Journée Mathématique (JMS)*, 2009.
- Pierre Brasseur et Jacques Verron : The SEEK filter method for data assimilation in oceanography : A synthesis. *Ocean Dynamics*, volume 56, pages 650–661, décembre 2006. ISSN 1616-7341, 1616-7228.
- Ye Chan, Chan Gook Park et Pho Kaung : Application of Unscented Kalman Filter with Non-symmetric Sigma Point Sampling on the Integrated Navigation System. 2014.
- Dominique Chapelle, M. Fragu, V. Mallet et Philippe Moireau : Fundamental principles of data assimilation underlying the Verdandi library : Applications to biophysical model personalization within euHeart. *Medical & Biological Engineering & Computing*, volume 51, pages 1221–1233, novembre 2013. ISSN 0140-0118, 1741-0444.
- Dominique Chapelle, Marina Vidrascu, Grégoire Derveaux, Philippe Moireau, Michel Bernadou, Frédéric Bourquin, Patrick Le Tallec et J. Sainte-Marie : Project-Team Macs : Modeling, Analysis and Control for Computational Structural Dynamics. Rapport technique, INRIA, Rocquencourt, 2006.
- A. Chouaki et P. Ladeveze : A posteriori constitutive relation error for dynamic model updating. *In PROCEEDINGS-SPIE THE INTERNATIONAL SOCIETY FOR OPTICAL ENGINEERING*, volume 1, pages 692–698. SPIE INTERNATIONAL SOCIETY FOR OPTICAL, 1998.
- Albert Alarcon Cot : *A Combined Kalman Filter and Error in Constitutive Relation Approach for System Identification in Structural Dynamics*. Thèse de doctorat, Ecole Polytechnique X, 2012.
- Nicolas Daget : Revue des méthodes d'assimilation. Rapport technique, Université de Toulouse, Université Toulouse III-Paul Sabatier, 2008.
- Frederica Darema : Dynamic Data Driven Applications Systems : A New Paradigm for Application Simulations and Measurements. *In Marian Bubak, Geert Dick van Albada, Peter M. A. Sloot et Jack Dongarra, éditeurs : Computational Science - ICCS 2004 : 4th International Conference, Kraków, Poland, June 6-9, 2004, Proceedings, Part III*, pages 662–669. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004. ISBN 978-3-540-24688-6.
- Gilles Debunne, Mathieu Desbrun, Marie-Paule Cani et Alan H. Barr : Dynamic real-time deformations using space & time adaptive sampling. *In Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, pages 31–36. ACM, 2001.

- J.R. Dormand et P.J. Prince : A family of embedded Runge-Kutta formulae. *Journal of Computational and Applied Mathematics*, volume 6, pages 19–26, 1980. ISSN 0377-0427.
- J. Dunik, O. Straka et M. Simandl : Stochastic Integration Filter. *IEEE Transactions on Automatic Control*, volume 58, pages 1561–1566, juin 2013. ISSN 0018-9286, 1558-2523.
- Jindrich Dunik, Ondrej Straka et Miroslav Šimandl : Sigma-Point Set Rotation in Unscented Kalman Filter : Analysis and Adaptation\*. 2014.
- Garry A. Einicke : *Smoothing, Filtering and Prediction. Estimating the Past, Present and Future*. INTECH Open Access Publisher, 2012.
- Karim El Mokhtari : *Estimation Circulaire Multi-Modèles Appliquée Au Map Matching En Environnement Contraint*. Thèse de doctorat, Dunkerque, 2015.
- Geir Evensen : The ensemble Kalman filter : Theoretical formulation and practical implementation. *Ocean dynamics*, volume 53, pages 343–367, 2003.
- Fredrik Gustafsson et Gustaf Hendeby : Some Relations Between Extended and Unscented Kalman Filters. *IEEE Transactions on Signal Processing*, volume 60, pages 545–555, février 2012. ISSN 1053-587X, 1941-0476.
- Jouni Hartikainen, Solin, Arno et Särkkä, Simo : Optimal Filtering with Kalman Filters and Smoothers. Rapport technique, août 11.
- Jindřich Havlík et Ondřej Straka : Performance evaluation of iterated extended Kalman filter with variable step-length. *Journal of Physics : Conference Series*, volume 659, page 012022, novembre 2015. ISSN 1742-6588, 1742-6596.
- Simon S Haykin : *Kalman Filtering and Neural Networks*. Adaptive and learning systems for signal processing, communications, and control. Wiley, New York, 2001. ISBN 978-0-471-46421-1 978-0-471-22154-8 978-0-471-36998-1.
- A. Hegyi, D Girimonte, R. Babuska et Bart De Schutter : A comparison of filter configurations for freeway traffic state estimation. In *Proceedings of the 2006 IEEE Intelligent Transportation Systems Conference*, Prentice-Hall information and system sciences series, pages 1029–1034, Toronto, Canada, septembre 2006.
- François Hild, Amine Bouterf, Ludovic Chamoin, Hugo Leclerc, Florent Mathieu, Jan Neggers, Florent Pled, Zvonimir Tomičević et Stéphane Roux : Toward 4D mechanical correlation. *Advanced Modeling and Simulation in Engineering Sciences*, volume 3, pages 1–26, 2016.
- Marine Jouin, Rafael Gouriveau, Daniel Hissel, Marie-Cécile Péra et Noureddine Zerhouni : Particle filter-based prognostics : Review, discussion and perspectives. *Mechanical Systems and Signal Processing*, volume 72-73, pages 2–31, mai 2016. ISSN 08883270.
- Simon J. Julier : The scaled unscented transformation. In *American Control Conference, 2002. Proceedings of the 2002*, volume 6, pages 4555–4559. IEEE, 2002.
- Simon J. Julier et Jeffrey K. Uhlmann : A new extension of the Kalman filter to nonlinear systems. In *Int. Symp. Aerospace/Defense Sensing, Simul. and Controls*, volume 3, pages 182–193. Orlando, FL, 1997.

### Bibliographie du chapitre 3

- Rudolf Emil Kalman : Contributions to the theory of optimal control. *Bol. Soc. Mat. Mexicana*, volume 5, pages 102–119, 1960a.
- Rudolph E. Kalman et Richard S. Bucy : New results in linear filtering and prediction theory. *Journal of basic engineering*, volume 83, pages 95–108, 1961.
- Rudolph Emil Kalman : A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, volume 82, pages 35–45, 1960b.
- Krishna Kumar Kottakki, Sharad Bhartiya et Mani Bhushan : State estimation of nonlinear dynamical systems using nonlinear update based Unscented Gaussian Sum Filter. *Journal of Process Control*, volume 24, pages 1425–1443, septembre 2014. ISSN 09591524.
- Roger RR Labbe : Kalman & Bayes Filters in Python. *Water Resources Research*, volume 40, décembre 2015.
- P Ladeveze : Nouvelle procédure d'estimation d'erreur relative à la méthode des éléments finis et applications. *Publications mathématiques et informatique de Rennes*, numéro S4, pages 1–19, 1977.
- Yann Lepoittevin : *Estimation de La Dynamique à Partir Des Structures Observées Dans Une Séquence d'images*. Thèse de doctorat, Université Pierre et Marie Curie, 2015.
- Andrew C Lorenc : Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, volume 112, pages 1177–1194, 1986.
- Basile Marchand, Ludovic Chamoin et Christian Rey : Coupling Modified Constitutive Relation Error, Model Reduction and Kalman Filtering Algorithms for Real-Time Parameters Identification. *Journal of Physics : Conference Series*, volume 657, page 012007, novembre 2015. ISSN 1742-6588, 1742-6596.
- S. Marchesseau, H. Delingette, M. Sermesant, R. Cabrera-Lozoya, C. Tobon-Gomez, P. Moireau, R.M. Figueras i Ventura, K. Lekadir, A. Hernandez, M. Garreau, E. Donal, C. Leclercq, S.G. Duckett, K. Rhode, C.A. Rinaldi, A.F. Frangi, R. Razavi, D. Chapelle et N. Ayache : Personalization of a cardiac electromechanical model using reduced order unscented Kalman filtering from regional volumes. *Medical Image Analysis*, volume 17, pages 816–829, octobre 2013a. ISSN 13618415.
- Stéphanie Marchesseau, Hervé Delingette, Maxime Sermesant et Nicholas Ayache : Fast parameter calibration of a cardiac electromechanical model from medical images based on the unscented transform. *Biomechanics and Modeling in Mechanobiology*, volume 12, pages 815–831, août 2013b. ISSN 1617-7959, 1617-7940.
- Stefano Mariani et Alberto Corigliano : Impact induced composite delamination : State and parameter identification via joint and dual extended Kalman filters. *Computer Methods in Applied Mechanics and Engineering*, volume 194, pages 5242–5272, décembre 2005. ISSN 00457825.
- Muhammad J. Mirza : A Modified Kalman Filter for Non-gaussian Measurement Noise. *In Communication Systems and Information Technology*, pages 401–409. Springer, 2011.
- Philippe Moireau : *Assimilation de Données Par Filtrage Pour Les Systèmes Hyperboliques Du Second Ordre. Applications à La Mécanique Cardiaque*. Thèse de doctorat, Citeseer, 2009.

### Chap. 3 : Analyse de vidéos par approche lagrangienne basée sur un modèle dynamique physique

- Philippe Moireau et Dominique Chapelle : Reduced-order Unscented Kalman Filtering with application to parameter identification in large-dimensional systems. *ESAIM : Control, Optimisation and Calculus of Variations*, volume 17, pages 380–405, avril 2011. ISSN 1292-8119, 1262-3377.
- Philippe Moireau, Dominique Chapelle et Patrick Le Tallec : Joint state and parameter estimation for distributed mechanical systems. *Computer Methods in Applied Mechanics and Engineering*, volume 197, pages 659–677, janvier 2008. ISSN 00457825.
- Philippe Moireau, Dominique Chapelle et Patrick Le Tallec : Filtering for distributed mechanical systems using position measurements : Perspectives in medical imaging. *Inverse Problems*, volume 25, page 035010, mars 2009. ISSN 0266-5611, 1361-6420.
- Naren Naik, R. M. O. Gemson, M. R. Ananthasayanam et others : Introduction to the Kalman Filter and Tuning its Statistics for Near Optimal Estimates and Cramer Rao Bound. *arXiv preprint arXiv :1503.04313*, 2015.
- Hong-Minh Nguyen, Olivier Allix et Pierre Feissel : A robust identification strategy for rate-dependent models in dynamics. *Inverse Problems*, volume 24, page 065006, décembre 2008. ISSN 0266-5611, 1361-6420.
- Dinh Tuan Pham, Jacques Verron et Marie Christine Roubaud : A singular evolutive extended Kalman filter for data assimilation in oceanography. *Journal of Marine systems*, volume 16, pages 323–340, 1998.
- Thomas Pottier : *Identification Paramétrique Par Recalage de Modèles Éléments Finis Couplée à Des Mesures de Champs Cinématiques et Thermiques*. Thèse de doctorat, Université de Savoie, Laboratoire SYMME, octobre 2010.
- Michael Roth, Carsten Fritsche, Gustaf Hendeby et Fredrik Gustafsson : The ensemble Kalman filter and its relations to other nonlinear filters. *In Signal Processing Conference (EUSIPCO), 2015 23rd European*, pages 1236–1240. IEEE, 2015.
- Simo Särkkä : On Unscented Kalman Filtering for State Estimation of Continuous-Time Nonlinear Systems. *IEEE Transactions on Automatic Control*, volume 52, pages 1631–1641, septembre 2007. ISSN 0018-9286.
- Simo Särkkä : Continuous-time and continuous–discrete-time unscented Rauch–Tung–Striebel smoothers. *Signal Processing*, volume 90, pages 225 – 235, 2010. ISSN 0165-1684.
- Simo Särkkä : *Bayesian Filtering and Smoothing*, volume 3. Cambridge University Press, 2013.
- Simo Särkkä et Juha Sarmavuori : Gaussian filtering and smoothing for continuous-discrete dynamic systems. *Signal Processing*, volume 93, pages 500–510, 2013.
- Simo Särkkä et Arno Solin : On Continuous-Discrete Cubature Kalman Filtering. *IFAC Proceedings Volumes*, volume 45, pages 1221–1226, juillet 2012. ISSN 14746670.
- Yoshikazu Sasaki : An objective analysis based on the variational method. *Journal of the Meteorological Society of Japan. Ser. II*, volume 36, pages 77–88, 1958.
- Stanley F. Schmidt et Leonard A. McGee : Discovery of the Kalman filter as a practical tool for aerospace and industry. Rapport technique 19860003843, NASA Ames Research Center, novembre 1985.

### Bibliographie du chapitre 3

- Dan Simon : *Optimal State Estimation : Kalman, H Infinity, and Nonlinear Approaches*. Wiley-Interscience, 2006. ISBN 0-471-70858-5.
- Dan Simon : Reduced order Kalman filtering without model reduction. *Control and Intelligent Systems*, volume 35, page 169, 2007.
- Dan Simon : Kalman filtering with state constraints : A survey of linear and nonlinear algorithms. *IET Control Theory & Applications*, volume 4, pages 1303–1318, août 2010. ISSN 1751-8644, 1751-8652.
- Dan Simon et Donald L. Simon : Constrained Kalman filtering via density function truncation for turbofan engine health estimation. *International Journal of Systems Science*, volume 41, pages 159–171, février 2010. ISSN 0020-7721, 1464-5319.
- Vincent Sircoulomb, Ghaleb Hoblos, Houcine Chafouk et José Ragot : State estimation under nonlinear state inequality constraints. A tracking application. *In Control and Automation, 2008 16th Mediterranean Conference On*, pages 1669–1674. IEEE, 2008.
- G.L. Smith, S.F. Schmidt et L.A. McGee : Application of statistical filter theory to the optimal estimation of position and velocity on board a circumlunar vehicle. Rapport technique R-135, National Aeronautics and Space Administration, 1962.
- Harold W. Sorenson : On the development of practical nonlinear filters. *Information Sciences*, volume 7, pages 253–270, 1974.
- Harold W Sorenson et Daniel L Alspach : Recursive Bayesian estimation using Gaussian sums. *Automatica*, volume 7, pages 465–479, 1971.
- Ondrej Straka, Jindrich Dunik et Miroslav Simandl : Scaling parameter in unscented transform : Analysis and specification. *In American Control Conference (ACC), 2012*, pages 5550–5555. IEEE, 2012a.
- Ondřej Straka, Jindřich Duník et Miroslav Šimandl : Truncation nonlinear filters for state estimation with nonlinear inequality constraints. *Automatica*, volume 48, pages 273–286, février 2012b. ISSN 00051098.
- Jeffrey K Uhlmann : *Dynamic Map Building and Localization : New Theoretical Foundations*. Theses, Oxford, 1995.
- Rudolph Van Der Merwe, Arnaud Doucet, Nando De Freitas et Eric A. Wan : The unscented particle filter. *In Advances in Neural Information Processing Systems*, pages 584–590, 2001.
- Eric A. Wan et Rudolph Van Der Merwe : The unscented Kalman filter for nonlinear estimation. *In Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, pages 153–158. IEEE, 2000.
- Li Wan-Chun, Wei Ping et Xiao Xian-Ci : A novel simplex unscented transform and filter. *In 2007 International Symposium on Communications and Information Technologies*, pages 926–931, octobre 2007.
- Yuanxin Wu, Dewen Hu, Meiping Wu et Xiaoping Hu : Unscented Kalman filtering for additive noise case : Augmented versus nonaugmented. *IEEE Signal Processing Letters*, volume 12, pages 357–360, mai 2005. ISSN 1070-9908.

Chap. 3 : Analyse de vidéos par approche lagrangienne basée sur un modèle dynamique physique

Zongbo Xie et Jiuchao Feng : Real-time nonlinear structural system identification via iterated unscented Kalman filter. *Mechanical Systems and Signal Processing*, volume 28, pages 309–322, avril 2012. ISSN 08883270.



# 4

## Application à l'analyse modale : une méthode temps réel pour l'identification de modes non linéaires

*Comme il est profond, ce mystère de l'Invisible ! Nous ne le pouvons sonder avec nos sens misérables, avec nos yeux qui ne savent apercevoir ni le trop petit, ni le trop grand, ni le trop près, ni le trop loin, ni les habitants d'une étoile, ni les habitants d'une goutte d'eau... [...] Ah ! si nous avions d'autres organes qui accompliraient en notre faveur d'autres miracles, que de choses nous pourrions découvrir encore autour de nous !*

– Guy de Maupassant

### Sommaire

---

I	Contexte de l'analyse modale . . . . .	160
II	Construction d'une fonction d'évolution pour l'assimilation séquentielle modale . . . . .	177
III	Un premier schéma d'assimilation utilisant un modèle théorique modal de la structure : le Kinematic multiSensors Assimilation Modal Identification (K-SAMI) . . . . .	186
IV	Formalisation de la chaîne complète . . . . .	187
V	Schéma d'assimilation développé sans connaissances sur la structure : Blind multiSensors Assimilation Modal Identification (B-SAMI) . . . . .	195
VI	Application au cas d'une poutre encastrée libre . . . . .	197
VII	Modifier la perception du système réel à l'aide de sa physique : proposition d'un autre algorithme de <i>motion magnification</i> . . . . .	215
VIII	Conclusion et perspectives . . . . .	218
	Bibliographie du chapitre 4 . . . . .	220

---

### Résumé

Le cadre que définit l'analyse modale utilise des signaux et des hypothèses de travail très propices à l'utilisation de méthodes d'assimilation de données. Ce chapitre propose d'appliquer les concepts développés dans les chapitres précédents associés à un modèle adéquat pour l'identification de paramètres modaux. Ces travaux introduisent deux schémas d'assimilation permettant de séparer les contributions modales dans un schéma temps réel, l'un utilisant un modèle des déformées modales théoriques de la structure, l'autre estimant directement les déformées modales expérimentales. Utilisés avec des mesures par caméra rapide, ils permettent un processus de mesure multipoint sans contact. Ils ouvrent d'autres possibilités : prédiction et interpolation spatiale et temporelle, fusion de données multicapteurs. Enfin, ils seront utilisés pour définir une nouvelle méthode de magnification de mouvement qui offre la possibilité de synthétiser de nouvelles vidéos enrichies par le modèle.



Dans la continuité des schémas développés dans le chapitre 3, nous proposons une application pour l'identification modale. La section I est une introduction pour comprendre les problématiques et l'état de l'art aussi bien sur le matériel utilisé que sur les méthodes d'identification modale existantes. L'utilisation d'un schéma d'assimilation avec un modèle adéquat sur des observations vidéos pourrait constituer un apport pour le domaine, mais avec quel modèle? La section II propose un tour d'horizon des modèles employés pour la démodulation et nous amène à la formulation de la fonction d'évolution. Nous développons dans la section III une fonction d'observation adaptée à des mesures par caméra ce qui nous conduit à proposer un premier schéma d'assimilation que nous appellerons **Kinematic-SAMI** (pour *Kinematic multiSensors Assimilation Modal Identification*). Ce schéma utilise des déformées théoriques de la structure pour identifier les paramètres modaux. L'assimilation se fait dans un schéma temps réel et offre la possibilité de faire de l'interpolation spatiale et temporelle en fonction des déformées modales. Il peut également fonctionner en extrapolation avec des observations partielles de la structure tout en estimant le comportement global de celle-ci. Une application sur une structure 3D sans stéréo-corrélation est proposée. Ce premier schéma permet de lisser les observations en fonction des déformées modales théoriques mais conduira à une estimation biaisée en cas de non correspondance, ce qui est par exemple le cas pour des non-linéarités géométriques. Nous proposons un second schéma d'assimilation que nous appellerons **Blind-SAMI** (pour *Blind multiSensors Assimilation Modal Identification*) dans la section V. Ce schéma permet d'identifier à la fois les paramètres modaux et les déformées modales expérimentales dans un schéma temps réel sans utiliser de déformées théoriques. L'absence de modèle spatial ne permettra de faire ni lissage, ni interpolation, ni prédiction spatiales. Une validation sera proposée dans la section VI sur un cas de poutre encastree libre. Avant de conclure, nous avons voulu proposer une application en synthèse vidéo. La section VII expose un nouvel algorithme de magnification de mouvement basé sur les SAMI. Il offre la possibilité de générer la vidéo d'une structure avec des propriétés modales modifiées. La figure 4.1 illustre le processus parcouru au cours de ce chapitre.

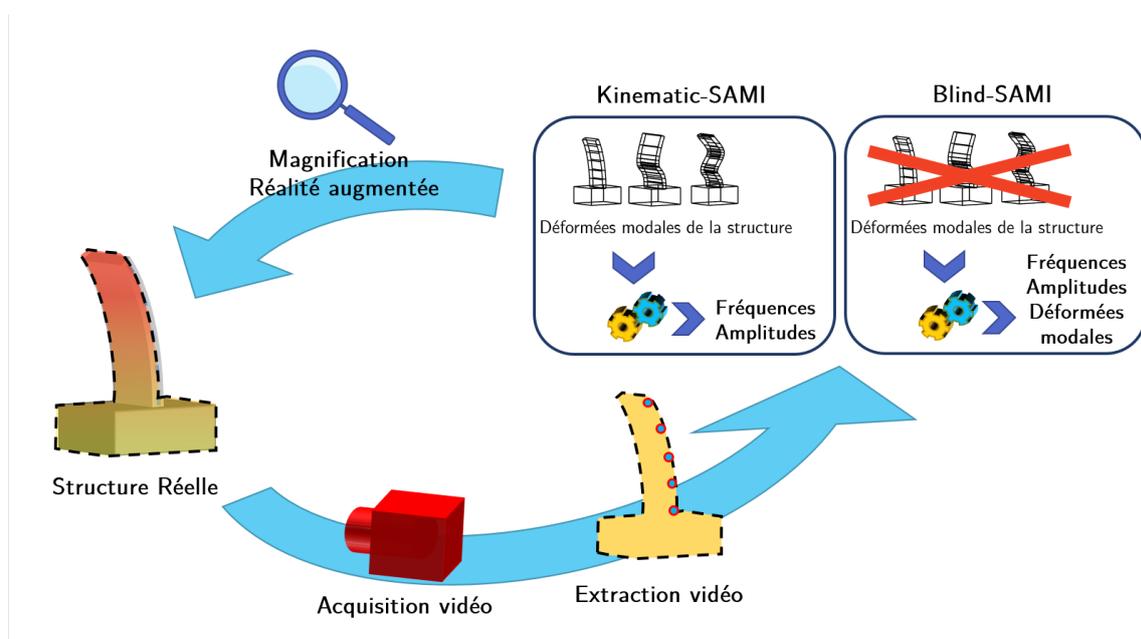


Figure 4.1 – Résumé graphique présentant la démarche et les résultats du chapitre 4

## I Contexte de l'analyse modale

Les travaux précédemment développés ont permis de valider une méthode permettant d'utiliser conjointement des données expérimentales et un modèle numérique dans un schéma temps réel. Aussi, dans le projet FAMOSYS, l'environnement de l'équipe en Vibration, Acoustique et Structures (VAST) au sein de Supméca a naturellement conduit à une application en analyse modale. La diversité des applications envisagées dans cette thèse montre le potentiel technique et industriel des méthodes d'assimilations étudiées.

Nous verrons dans un premier temps en quoi le cadre de l'analyse modale est un contexte favorable à l'assimilation de données. Le contexte bibliographique sera positionné dans un second temps précédant l'analyse des apports et limites de mesures par caméra rapide.

### I.A. Introduction : un cadre favorable

L'étude des vibrations d'un système peut être décomposée par bandes de fréquences. Trois domaines généraux sont distingués dans le tableau 4.1 issu de Garnier (2009). Ils se différencient notamment par l'amplitude vibratoire de réponse, par les systèmes concernés et par des phénomènes excitateurs. Ainsi, les fréquences extrêmement basses (EBF) concernent plutôt des ouvrages de grande hauteur, souvent statiques, tandis que les phénomènes excitateurs sont plutôt d'ordre naturel (séisme, vent, houle) mais également auto-générés par les systèmes mécaniques concernés (aéronefs, véhicules). Les très basses fréquences (TBF) constituent un domaine comportant beaucoup d'excitations, mais également le domaine où l'isolation vibratoire est maximale. Enfin les basses fréquences (BF) constituent la plage où la plupart des mécanismes mécaniques ont leur fréquences excitatrices principales, mais également celle où les structures mécaniques ont leur fréquences propres les plus gênantes, à la fois pour l'intégrité de la structure, mais également pour la stabilité du système. Les travaux réalisés porteront principalement sur cette gamme de fréquence.

Une vibration fait intervenir deux phénomènes dans un matériau ou une structure : une énergie cinétique qui est apportée au matériau et une énergie potentielle résultant de l'élasticité du matériau. Cet échange d'énergie va permettre au milieu élastique d'osciller autour d'une position d'équilibre. L'énergie cinétique apportée au matériau peut être engendrée de différentes manières. Une force externe oscillante peut par exemple être transmise au système ; cette force peut être également interne à la structure ou même interne au matériau par propagation d'une onde vibratoire ; la force excitatrice peut également ne pas être oscillante, comme dans le cas d'un choc ou d'un effort.

Les structures sont des résonateurs naturels : sous des conditions qui leurs sont propres, leur comportement peut atténuer ou amplifier une amplitude de vibration qui leur est imposée. Cette propriété

**Table 4.1** – Différents domaines fréquentiels

	Plage de fréquences
Fréquences extrêmement basses (EBF)	< 1 Hz
Très basses fréquences (TBF)	1-10 Hz
Basses fréquences (BF)	10-100 Hz
Moyennes fréquences (MF)	100-3000 Hz
Hautes fréquences (HF)	3-10 kHz

peut conduire à des amplitudes vibratoires pour lesquelles la structure peut être endommagée. Aussi l'identification des fréquences propres de la structure est primordiale pour le dimensionnement de structures afin de les éviter ou de les solliciter comme dans le cas des instruments de musique.

Appliquer un schéma d'assimilation n'a un sens que dans un domaine à vocation expérimentale, deux sous-domaines de l'analyse modale correspondent bien à ce cadre : l'analyse modale expérimentale et l'analyse modale opérationnelle. En reprenant les concepts développés dans cette thèse, une application dans le domaine de l'analyse modale a d'autant plus d'intérêt que l'évolution des systèmes vibratoires est régie par un modèle dynamique. D'autres hypothèses inhérentes à la dynamique des structures et à l'analyse modale facilitent l'usage de l'assimilation de données :

- Les signaux considérés sont centrés.
- Le cadre de travail induit un découplage spatiotemporel.
- L'évolution d'un signal peut s'écrire sous la forme d'un système d'état.

Au-delà du fait qu'il s'agit de systèmes dynamiques, la nature des signaux caractéristiques entre donc dans le cadre des hypothèses de l'assimilation de données.

## I.B. Analyse modale

L'analyse modale de structure s'est imposée comme une étape cruciale de la conception de systèmes mécaniques. À travers le développement de ces méthodes entrepris depuis les années 50, c'est la connaissance du modèle dynamique de la structure que l'on cherche à établir. Pour illustrer l'importance des vibrations basses fréquences sur la tenue mécanique des structures, citons la prévention des phénomènes de *flutter*. Le pont de Tacoma d'hier n'avait pas bénéficié de ces outils de dimensionnement contrairement à tous les engins volants d'aujourd'hui qui font l'objet d'études vibratoires au sol (*Ground Vehicle Testing - GVT*). La plupart de ces instabilités vibratoires (*POGO* pour les lanceurs spatiaux, *flutter* des ailes volantes, *wobble* des skateboards, *shimmy* pour les trains avant d'atterrissage, *judder* des systèmes de freinage...) sont autant de comportements qui peuvent conduire à une détérioration rapide de systèmes mécaniques. Pour éviter ces phénomènes, leur dimensionnement s'appuie sur une analyse modale expérimentale dont les fréquences d'intérêt sont dans la plupart des cas inférieures à 100 Hz. En amont l'analyse modale reste indispensable pour la détermination des paramètres modaux des structures : fréquence, amortissement, ainsi que la forme spatiale de chaque mode. Cette décomposition s'appuie sur deux hypothèses fortes de linéarité et d'invariance temporelle des modes de la structure. Plusieurs approches de l'analyse modale sont envisagées :

### L'analyse modale théorique

L'analyse modale théorique revient à trouver les paramètres modaux en résolvant le système aux valeurs propres défini par les équations générales du système étudié. On peut exprimer en toute généralité un système MDOF (*Multiple Degree Of Freedom*) en vibration sous la forme :

$$\begin{bmatrix} I & 0 \\ 0 & M \end{bmatrix} \begin{pmatrix} \dot{x} \\ \ddot{x} \end{pmatrix} = \begin{bmatrix} 0 & I \\ -K & -C \end{bmatrix} \begin{pmatrix} x \\ \dot{x} \end{pmatrix} + \begin{pmatrix} 0 \\ F_{ext} \end{pmatrix} + \begin{pmatrix} 0 \\ F_{NL}(x, \dot{x}) \end{pmatrix} \quad (4.1)$$

$x$  représente le déplacement du système et ses dérivées successives la vitesse et l'accélération.  $M$ ,  $K$ , et  $C$  les matrices de masse, de raideur et d'amortissement généralisés,  $F_{ext}$  les forces extérieures s'appliquant au système,  $F_{NL}$  les contributions non linéaires externes et internes du système exprimées sous forme de forces.

L'analyse modale théorique propose de résoudre le système comme un problème aux valeurs propres. Afin d'alléger les calculs, une approximation consiste à négliger les amortissements

## I. Contexte de l'analyse modale

et les termes de force, ce qui revient à :

$$(K - \omega_i^2 \cdot M)x_i = 0 \quad (4.2)$$

où  $\omega_i$  représente la pulsation propre et  $x_i$  le mode propre du mode  $i$ .

L'analyse modale théorique permet ainsi d'obtenir les modes de la structure sur des systèmes de plus en plus complexes grâce à l'augmentation de la puissance de calcul disponible pour les résolutions numériques. Cependant la complexité de certains systèmes impose des hypothèses simplificatrices qui peuvent conduire à des dérives par rapport au comportement modal de la structure réelle.

### Analyse Modale Expérimentale

L'analyse modale expérimentale est une famille de méthodes d'investigation visant à établir ou améliorer la connaissance du modèle dynamique des structures réelles. Elle propose d'utiliser des modèles paramétrés par un nombre réduit de variables : les paramètres modaux (fréquences propres, amortissements et déformées modales) pour décrire le comportement d'une structure dynamique complexe. L'identification de ces paramètres se fait sur des mesures de la structure réelle excitée sous des conditions particulières. Elle est primordiale pour les essais de systèmes aéronautiques ou automobiles afin de valider une conception. Les modes estimés de la structure réelle sont souvent comparés aux modes théoriques associés. Ces modes sont issus de l'analyse numérique du modèle réalisée en amont généralement à partir de structures simplifiées.

### Analyse Modale Opérationnelle

L'analyse modale opérationnelle s'attache également à identifier les paramètres modaux d'une structure en utilisant la structure réelle sans parvenir à maîtriser et mesurer l'excitation. Ces méthodes sont très utiles pour des ouvrages de génie civil qui sont difficiles à exciter de manière contrôlée, tout en étant excités en permanence par leur environnement.

Les concepts développés dans le projet FAMOSYS permettant de faire un pont entre des modèles et des données expérimentales, ils peuvent être appliqués dans les deux derniers domaines sus-cités.

### I.B.1. Analyse Modale Expérimentale

Si l'analyse modale expérimentale a permis de mieux comprendre le comportement des systèmes mécaniques et des structures, cela est aussi fortement lié au progrès des technologies de la chaîne de mesures. Les méthodes de mise en situation spécifique de la structure, de post-traitement et d'identification ont également contribué à cette évolution, parfois conjointement. Cette coévolution nous incite à introduire dans cette section l'analyse modale expérimentale sous ces deux différents éclairages.

### I.B.2. État de l'art matériel

L'avènement des accéléromètres a entièrement changé la vision expérimentale en mécanique vibratoire. Grâce à une étendue de mesure très large (165 dB de réponse dynamique) et un bruit crête à crête très faible (2  $\mu\text{g}$ ) (Cunha *et al.*, 2013), cette technologie s'est très largement imposée grâce à son rapport avantageux entre la précision, le coût, et l'encombrement. Dans beaucoup de domaines dont l'analyse modale, les accéléromètres sont établis comme instrument de référence, notamment chez les industriels. Comme souligné par la revue de Elies et Ebenhöch (2015), de nombreuses technologies d'accéléromètres existent, offrant des performances différentes. On retiendra qu'une

grande partie des accéléromètres permettent de couvrir la plage [0-1 kHz] voire [0-10 kHz]. Malgré leurs bonnes performances et leur utilisation pratique, quelques problèmes se posent concernant l'influence du capteur sur la structure. En effet, pour des petites structures l'influence de masse et de la liaison filaire peuvent s'avérer problématiques en terme de modification de la fréquence propre ou de l'amortissement. Des corrections sont alors appliquées à posteriori de la mesure. Cunha *et al.* (2013) oppose les structures de génie civile aux structures aéronautiques en soulignant la difficulté de séparation modale dans ce deuxième domaine dû à l'étalement des fréquences de résonance. Il note également la difficulté de faire de l'identification modale sur des câbles avec ce type de technologie. Les travaux récents de Régal *et al.* (2017) ont permis de comparer des mesures avec accéléromètres à des mesures sans contact pour une poutre. Les résultats montrent des variations des fréquences propres identifiées mais sont surtout significatifs au niveau de l'identification des amortissements dont les erreurs peuvent atteindre 100 %.

Ces résultats montrent l'intérêt d'utiliser des technologies sans contact pour préserver la précision des données d'un biais de mesure et permettre d'accroître la connaissance modale sur la structure réelle.

Pour permettre de passer outre cet obstacle, d'autres moyens de mesure existent. L'utilisation de techniques radar ou techniques interférométriques à base de micro-ondes permettent d'obtenir des déplacements ou des vitesses selon le mode d'utilisation (Gentile, 2011). Bien qu'utilisés pour les grandes structures, ces technologies ont une précision trop faible pour atteindre les niveaux de déplacements ou vitesses exigés pour l'AME (entre le mm et  $\mu\text{m}$ ). La limitation des technologies radar étant proportionnelle à la longueur d'onde, il a été naturel d'utiliser le laser comme source afin d'améliorer la précision. Les premières diodes laser des années 2000 et l'explosion des technologies qui a suivi ont permis progressivement de démocratiser les mesures laser dans d'autres domaines. Dans la dynamique de développement des télécommunications, la mesure par fibre optique est une des applications séduisantes par laser. Le principe des réseaux de capteurs à fibre optique est simple : la taille d'une fibre optique étant minime, elles sont incluses dans la structure au moment de sa fabrication. Une fibre est composée d'une succession de capteurs générant des échos reliés entre eux par une fibre classique. Les informations sont extraites dans la position temporelle de l'écho ou sur sa signature fréquentielle selon les technologies. Ces technologies sont idéales pour du *Structural Health Monitoring* (SHM) dans les domaines des grandes structures, malheureusement éloignées de notre cas d'étude. Avec une bande passante allant de valeurs inférieures au hertz à l'ordre du mégahertz, les lasers font partie des capteurs les plus justes et assurent une impartialité au niveau des résultats du fait de leur caractère sans contact (Berkovic et Shafir, 2012; von Martens, 2008). Cependant, deux désavantages majeurs empêchent les lasers de s'adapter à toutes les situations. Comme les accéléromètres, les lasers sont nativement des capteurs monopoints. Le coût d'un accéléromètre rend l'augmentation de la résolution spatiale abordable, ce qui est inenvisageable avec un laser. Certaines adaptations comme les lasers à balayage permettent d'avoir des mesures multipoints mais moyennant une hypothèse limitante de stationnarité du phénomène. Les cas d'observation transitoire, sollicitation aléatoire ou encore de rupture ou d'endommagement seront donc problématiques pour ce type de technologies. Néanmoins l'accent a été mis tout récemment sur la mise au point de technologies comportant plusieurs têtes lasers pour répondre aux demandes multipoints. D'autres défauts sont soulignés par Martin (2017) : par exemple les difficultés d'utilisation des technologies lasers sur des surfaces granuleuses. L'utilisation de lasers n'est par ailleurs pas sans conséquence sur la sécurité et des infrastructures spécifiques doivent être mises en place en plus de l'appareil. Le coût par point de mesure reste malgré tout très élevé, même si la technologie est juste tout en gardant une bande d'utilisation très large.

## I. Contexte de l'analyse modale

L'intérêt de pouvoir faire une mesure multipoint est de pouvoir accéder à une mesure du champ de déplacement ou du champ de vitesse à un instant donné. Ces mesures sont extrêmement riches en informations et s'abstraient de toute hypothèse d'interpolation spatiale ou temporelle. De par leur caractère matriciel et séquentiel, les caméras permettent d'acquérir ce type de mesures. On se réfèrera à l'état de l'art effectué sur les caméras rapides et ultra rapides (chapitre 1) pour s'apercevoir que si les mesures de champ sont très efficaces en statique ou faible dynamique, l'utilisation en dynamique nécessite l'emploi d'équipements de pointe. La raison vient de la relation spatiotemporelle qui régit les technologies actuelles de caméras rapides :

$$l_{champ} \propto \Delta t \quad (4.3)$$

pour une caméra donnée avec  $\Delta t$  le pas de temps, et  $l_{champ}$  la largeur du champ d'observation.

L'état de l'art du chapitre 1 propose des modèles allant jusqu'à la centaine de mégahertz sans toutefois garantir des résolutions ou des durées d'enregistrement suffisantes. Plus récemment s'est développé toute une série de caméras moyenne vitesse (100–1000 Hz) mais haute définition (4K soit  $4096 \times 3072$ ) basées sur une transmission de données améliorée. Ces caméras et leur futures évolutions ouvrent des perspectives d'utilisations dynamiques.

D'autre part, la précision des données acquises par caméra dépend toujours des hypothèses de mouvement plan de l'objet filmé, issues de la projection en 2D réalisée par le capteur. Avec l'aboutissement des recherches en stéréovision basées sur le modèle *pin-hole* et utilisant des coordonnées homogènes<sup>1</sup> (Szeliski, 2010), les mesures de champs 3D ont été très largement utilisées dans le domaine. Cependant elles utilisent essentiellement une technique : la corrélation croisée. Malgré les résultats sous pixeliques obtenus, l'algorithme n'est pas adapté à une utilisation temps réel malgré des codes de calcul très optimisés. Même si elle demeure séquentielle, l'emploi de la stéréocorrélation, plus lourde à traiter, n'apparaît pas envisageable dans un schéma temps réel pour les années à venir. Par ailleurs ces méthodes ont été abondamment employées ces dernières années en analyse modale à l'aide de caméras rapides, ceci malgré les faibles résolutions par rapport aux nouvelles possibilités 4K (tableau 4.2).

**Table 4.2** – Tableau présentant un état de l'art récent d'utilisation de caméras rapides dans des contextes vibratoires (avec l'aimable permission de Xavier Régat)

Référence	Année	Résolution	Vitesse	2D/3D	Objet
Wang <i>et al.</i> (2012)	2012	1280×800	300 Hz	3D	Capot de voiture
Helfrick <i>et al.</i> (2011)	2011	1024×1024	2000 Hz	3D	Plaque de sèche linge
Trebuña et Hagara (2014)	2014	-	2000 Hz	3D	Plaque de métal
Davis <i>et al.</i> (2015)	2015	480×288	5000 Hz	2D	Poutre et cylindre
Beberniss et Ehrhardt (2016)	2016	1024×1024	5000 Hz	3D	Poutre
Reu <i>et al.</i> (2016)	2016	800×800	3905 Hz	3D	Dalle

De nouvelles alternatives aux méthodes *DIC* (pour *Digital Inter Correlation*) émergent lors de l'utilisation de vidéos rapides. Deux approches sont actuellement développées : l'une algorithmique et l'autre matérielle. Le flot optique pour des petits mouvements devient maintenant une alternative crédible au niveau de la précision tout en étant moins coûteuse en temps de calcul (Liu *et al.*, 2015)<sup>2</sup>.

1. De nombreuses références sont disponibles pour expliquer ces notions comme sur le site de Mathworks pour la calibration de caméras : <https://fr.mathworks.com/help/vision/ug/camera-calibration.html>

2. Le sujet a été plus amplement développé dans le chapitre 2.

On note l'application de ces techniques de flot optique en identification modale dans certains travaux (Javh *et al.*, 2018). Enfin, il est intéressant de mentionner le laboratoire de Los Alamos qui développe ces dernières années des activités prospectives sur les mesures vibratoires par vidéo (Dorn *et al.*, 2017). Elles sont basées sur une technologie de rupture stable mais encore en développement pour atteindre des résolutions intéressantes. Ces caméras dites de vision artificielle (ou *Event-Based* ou encore *Neuromorphic*) sont des caméras bio-inspirées et basées sur le principe simple d'une acquisition événementielle pixel par pixel. Le potentiel de cette technologie est immense puisque les vitesses peuvent facilement atteindre 100 000i/s tout en ayant des mesures asynchrones et des flux de données très réduits. Ces caractéristiques en ferait une technologie idéale pour la mesure vibratoire si la résolution avait été plus grande. Le design particulier des pixels conduit à des grandes tailles de pixels ce qui aujourd'hui limite la résolution à des formats de l'ordre du QVGA. (chapitre 1)

On peut donc voir que les mesures de champs par caméra rapide représentent un intérêt non négligeable sur les recherches récentes. L'utilisation de ce type de technologie apporte des avantages en terme de mesure multipoint par rapport aux mesures accélérométriques ou par laser, cependant les performances en terme de fréquences d'échantillonnage ne sont pas encore tout à fait à la hauteur pour couvrir les moyennes fréquences mais suffisent pour les fréquences inférieures à 500 Hz (tableau 4.3). Quand à la précision, il est difficile de comparer, étant donné que la caméra est plutôt régie par le rapport  $\frac{\text{largeur du champ}}{\text{taille du pixel}}$ , rendant la précision dépendante du champ d'observation. De plus, les récentes avancées technologiques laissent entrevoir des améliorations significatives à une échelle de 5 ans, ce qui fait de la caméra rapide une technologie complète et prometteuse, applicable et concurrente à l'accélérométrie classique.

Accéléromètres	Lasers	Caméras rapides
<ul style="list-style-type: none"> <li>Précision</li> <li>Fréquence d'échantillonnage élevée</li> </ul>	<ul style="list-style-type: none"> <li>Précision</li> <li>Fréquence d'échantillonnage élevée</li> <li>Sans contact</li> <li>Multipoint sur des excitations stationnaires</li> </ul>	<ul style="list-style-type: none"> <li>Sans contact</li> <li>Multipoint sur des excitations transitoires / aléatoires</li> </ul>
<ul style="list-style-type: none"> <li>Installation sur la structure → longue et invasive</li> <li>Influence de la masse et du câble → Modification de la masse et de l'amortissement du système étudié.</li> <li>Capteur monopoint</li> </ul>	<ul style="list-style-type: none"> <li>Multipoint sur des excitations transitoires / aléatoires difficiles</li> </ul>	<ul style="list-style-type: none"> <li>La précision dépend du champ observé</li> <li>Fréquences d'échantillonnage moins élevées que pour les autres techniques</li> </ul>

**Table 4.3** – Tableau synthétisant les avantages et inconvénients des différents capteurs détaillés dans l'état de l'art

### I.B.3. État de l'art des méthodes d'identification modales

L'évolution des méthodes utilisées en analyse modale expérimentale est très dynamique en terme de publications novatrices. L'état de l'art des méthodes effectué ici est principalement basé sur les articles de Golinval et Kerschen (2014) et de Piranda (2001). Les méthodes d'identification peuvent être séparées en deux grandes familles : les méthodes temporelles et les méthodes fréquentielles.

## I. Contexte de l'analyse modale

L'analyse temporelle est bien plus ancienne avec des méthodes telles que la méthode d'Ibrahim (Ibrahim et Mikulcik, 1977). Le développement d'algorithmes rapides pour la transformée de Fourier discrète (*Fast Fourier Transform* ou *FFT*) a permis l'essor des méthodes fréquentielles. Elles permettent de gagner en précision en ajoutant des termes de résidu mais présentent quelques inconvénients comme la difficulté d'obtenir des mesures sur des courts intervalles temporels, le phénomène de fuite spectrale qui peut fortement impacter la précision des mesures, ou encore les difficultés de séparation de fréquences trop proches. Les hypothèses d'indépendance modale sont alors prépondérantes pour permettre une identification précise. Selon les différentes méthodes, on distingue les méthodes SDOF (pour *Single Degree Of Freedom*) et MDOF (pour *Multiple Degree Of Freedom*) selon le nombre de modes pris en compte dans l'analyse. Dans le domaine temporel, il n'est généralement pas possible de séparer les contributions modales, c'est pourquoi l'ensemble des méthodes temporelles repose sur une analyse multi degrés de liberté. Une distinction en terme d'entrées/sorties est faite selon les méthodes :

- SISO pour *Single Input Single Output*. Ces méthodes possèdent l'avantage d'être robustes mais elles reposent sur une hypothèse sous-jacente de faible résidu sur un mode des autres contributions modales. Elles permettent de n'identifier qu'une fréquence à la fois.
- SIMO pour *Single Input Multiple Output*. Contrairement aux méthodes SISO, plusieurs fréquences peuvent être identifiées à l'aide de ces méthodes.
- MIMO pour *Multiple Input Multiple Output*. Méthodes prenant en compte plusieurs types d'entrées et permettant d'identifier plusieurs fréquences en sortie.

Ces méthodes s'appuient souvent sur la mesure temporelle du signal de sortie ou sur la mesure de sa FRF (Fonction de Réponse en Fréquences), c'est-à-dire le spectre de la sortie du système normalisé par le spectre de l'entrée. Nous proposons de détailler quelques méthodes d'identification modale en commençant par les méthodes fréquentielles avant les méthodes temporelles.

### a. Les méthodes fréquentielles

#### Méthodes simple degré de liberté :

Sans s'attacher à une méthode en particulier, ces méthodes appliquent la forme analytique de la FRF aux données expérimentales. Elles sont robustes dans un cadre de séparation des contributions, mais possèdent des inconvénients majeurs. La fréquence modale utilisée est en réalité celle du pic de résonance. Elle ne prend donc pas en compte le décalage dû à l'amortissement et considère l'évolution en fréquence constante. L'amortissement est considéré comme constant. Il peut être mesuré à partir des fréquences correspondant à la règle des 3 dB (demi-puissance ou bande passante), grâce à la mesure du facteur de qualité, ou des informations issues des parties réels et imaginaires (prises indépendamment ou tracées dans le lieu de Nyquist). Enfin, l'obtention de la déformée modale peut se faire de différentes manières selon les grandeurs observées : soit à partir de la partie imaginaire de la FRF du déplacement ou de l'accélération, soit à partir de la partie réelle de la vitesse.

Les méthodes simple degré de liberté permettent une appropriation modale sous des hypothèses fortes de découplage modal et de linéarité.

#### Méthode LSFD (*Least Square Frequency Domain*) :

Travailler dans le domaine fréquentiel offre la possibilité de pouvoir borner en fréquence des intervalles d'analyse. Cette séparation permet ensuite d'utiliser la cohérence modale sur chacun d'entre eux afin d'identifier les modes de corps rigides des modes propres de la structure. Enfin travailler dans le domaine fréquentiel permet de prendre en compte de manière explicite les effets résiduels d'une mauvaise séparation modale. On utilise l'expression analytique de la

FRF qui donne :

$$H_{d,m}(\omega) = \sum_{k_{p\hat{o}l e}=1}^n \frac{A_{d,m}(k_{p\hat{o}l e})}{j\omega - \lambda_{k_{p\hat{o}l e}}} + \overline{\frac{A_{d,m}(k_{p\hat{o}l e})}{j\omega - \lambda_{k_{p\hat{o}l e}}} + u_{d,m} - \frac{l_{d,m}}{\omega^2}} \quad (4.4)$$

avec  $H_{d,m}$  la réponse spectrale pour le  $d^{\text{ème}}$  degré de liberté du  $m^{\text{ème}}$  mode. la réponse se modélise par une somme de  $n$  pôles complexes conjugués qui sont  $\lambda_{k_{p\hat{o}l e}}$ . On modélise en plus un résidu en masse  $u_{d,m}$  et en raideur  $l_{d,m}$ .

En posant  $A_{d,m}(k_{p\hat{o}l e}) = U_{d,m}(k_{p\hat{o}l e}) + j \cdot V_{d,m}(k_{p\hat{o}l e})$ , on peut mettre l'équation précédente sous la forme algébrique :

$$\begin{aligned} H_{d,m}(\omega) &= \sum_{k_{p\hat{o}l e}=1}^n U_{d,m}(k_{p\hat{o}l e}) \left( \frac{1}{j\omega + \lambda_{k_{p\hat{o}l e}}} + \frac{1}{j\omega - \lambda_{k_{p\hat{o}l e}}} \right) \\ &+ j \cdot V_{d,m}(k_{p\hat{o}l e}) \left( \frac{1}{j\omega + \lambda_{k_{p\hat{o}l e}}} - \frac{1}{j\omega - \lambda_{k_{p\hat{o}l e}}} \right) \\ &+ u_{d,m} - \frac{l_{d,m}}{\omega^2} \end{aligned} \quad (4.5)$$

La mesure de la FRF permet alors de connaître  $\lambda_{k_{p\hat{o}l e}}$  les pôles du système ainsi que la mesure de  $H_{d,m}(k)$ . Il reste  $2n + 2$  inconnues que l'on résout en respectant la condition sur la bande de fréquence choisie :

$$\frac{\omega_{max} - \omega_{min}}{\Delta\omega} + 1 > 2n + 2 \quad (4.6)$$

avec  $n$  le nombre de modes, et  $\Delta\omega$  la résolution fréquentielle.

#### L'estimateur PolyMAX :

PolyMAX est une des méthodes les plus utilisées actuellement (Peeters *et al.*, 2004). C'est une méthode MIMO qui s'appuie sur le formalisme de la LSCE définie dans la suite en se basant sur une résolution non-itérative à partir des FRF  $H_{FRF}(\omega) \in \mathbf{C}^{s \times e}$  pour un système à  $e$  entrées et  $s$  sorties.

Le modèle présenté précédemment peut s'écrire sous une forme de fraction polynomiale avec le numérateur  $Q(\omega) \in \mathbf{C}^{s \times e}$  et dénominateur  $P(\omega) \in \mathbf{C}^{e \times e}$  :

$$H(\omega) = Q(\omega) \cdot P(\omega)^{-1} \quad (4.7)$$

avec :

$$P(\omega) = \sum_{k=0}^n [\alpha_k] e^{-j\omega k \Delta t} \quad (4.8)$$

$$Q(\omega) = \sum_{k=0}^n [\beta_k] e^{-j\omega k \Delta t} \quad (4.9)$$

On introduit deux matrices  $[\alpha_k]$  et  $[\beta_k]$  à estimer et  $n$  l'ordre du modèle qui doit être supérieur au nombre de modes prévus dans la bande de fréquence d'étude puisque des modes numériques pourront intervenir.

## I. Contexte de l'analyse modale

L'estimation se fait non pas sur l'erreur  $E(\omega) = H(\omega) - H_{FRF}(\omega)$  mais sur la matrice linéarisée :

$$\begin{aligned} E^L(\omega) &= E(\omega) \cdot P(\omega) \\ &= Q(\omega) - H_{FRF}(\omega) \cdot P(\omega) \end{aligned} \quad (4.10)$$

Une solution pour les coefficients du dénominateur et du numérateur est ensuite obtenue via une minimisation au sens des moindres carrés de l'équation (4.10). On peut alors déterminer les pôles échantillonnés du système  $z_{k_{p\acute{o}le}}$  et en déduire les valeurs propres en temps continu du système :

$$\lambda_{k_{p\acute{o}le}} = -\frac{\log z_{k_{p\acute{o}le}}}{\Delta t} \quad (4.11)$$

Ce qui permet de déterminer les fréquences propres  $\omega_k$  et les amortissements  $\xi_k$  :

$$\lambda_{k_{p\acute{o}les}} = -\xi_k \omega_k \pm j \omega_k \sqrt{1 - \xi_k^2} \quad (4.12)$$

La FRF synthétisée peut être ensuite comparée à la FRF mesurée à l'aide d'indicateurs de corrélation que nous verrons par la suite (paragraphe c. ).

L'avantage de cette méthode est aussi dans l'utilisation de son diagramme de stabilisation. Durant cette procédure itérative, les pôles identifiés à l'ordre  $n$  sont comparés avec ceux d'ordre  $n - 1$ . Si la différence en fréquence et amortissement entre dans une marge d'erreur, le pôle est marqué comme stable. Ces diagrammes permettent de repérer la plupart des modes numériques et des parasites, ce qui est très pratique pour l'utilisateur.

### b. Les méthodes temporelles

#### La Méthode Exponentielle Complexe (LSCE) :

Cette méthode s'applique sur un domaine temporel sur des excitations impulsionnelles. Numériquement stable, elle fonctionne avec une majeure partie des modes. Elle a l'inconvénient d'utiliser plus de modes que dans le cas réel ce qui conduit à produire des alias de mode dans les résultats. Une étape de nettoyage de ces modes numériques doit alors être effectuée à l'aide de l'outil des diagrammes de stabilisation. La méthode utilise cette fois la transformée inverse de l'équation (4.4) afin de travailler dans le domaine temporel :

$$\begin{aligned} h_{d,m}(t) &= IFFT [H_{d,m}(\omega)] \\ &= IFFT \left[ \sum_{k_{p\acute{o}le}=1}^n \frac{A_{d,m}(k_{p\acute{o}le})}{j\omega - \lambda_{k_{p\acute{o}le}}} + \frac{\overline{A_{d,m}(k_{p\acute{o}le})}}{j\omega - \overline{\lambda_{k_{p\acute{o}le}}} } \right] \end{aligned} \quad (4.13)$$

En pratique, les pôles sont obtenus à l'aide de modèles auto-régressifs. Le passage dans le domaine temporel se fait en définissant le pas de temps par la fréquence d'échantillonnage (notée  $F_{ech}$ ) qui est généralement constante :

$$t = k \cdot \Delta t \quad (4.14)$$

$$\Delta t = \frac{1}{F_{ech}} \quad (4.15)$$

Ce qui nous donne en prenant le pôle échantillonné  $z_{k_{p\acute{o}le}} = e^{\lambda_{k_{p\acute{o}le}} \cdot \Delta t}$  :

$$h_{d,m}(k \cdot \Delta t) = \sum_{k_{p\acute{o}le}=1}^n \left( A_{d,m}(k_{p\acute{o}le}) \cdot z_{k_{p\acute{o}le}}^k + \overline{A_{d,m}(k_{p\acute{o}le})} \cdot \overline{z_{k_{p\acute{o}le}}^k} \right) \quad (4.16)$$

$$h_{d,m}(k \cdot \Delta t) = 2 \cdot \Re e \left[ \sum_{k_{p\acute{o}le}=1}^n \left( A_{d,m}(k_{p\acute{o}le}) \cdot z_{k_{p\acute{o}le}}^k \right) \right] \quad (4.17)$$

C'est en recalant les mesures expérimentales sur le modèle mathématique de l'équation (4.17) que l'identification va être faite. On aura alors  $A_{d,m}(k_{p\acute{o}le})$  le résidu définissant le déplacement du  $d^{\text{ème}}$  DDL du mode  $m$ .

On peut écrire en fréquentiel la fonction de transfert sous une forme de fraction polynomiale rationnelle :

$$\begin{aligned} H_{d,m}(\omega) &= \frac{Q_d(\omega)}{P_m(\omega)} \\ &= \frac{\sum_{i=0}^{o_p} a_i (\mathcal{J}\omega)^i}{\sum_{i=0}^{o_q} b_i (\mathcal{J}\omega)^i} \end{aligned} \quad (4.18)$$

ce qui donne :

$$\sum_{i=0}^{o_p} a_i \cdot P_m(\mathcal{J}\omega)^i = \sum_{i=0}^{o_q} b_i (\mathcal{J}\omega)^i \cdot Q_d(\omega) \quad (4.19)$$

On observe un modèle SISO en fréquence, qui peut se transformer en un modèle ARMA temporel :

$$\sum_{i=0}^{o_p} a_i \cdot p_m(t + \Delta t) = \sum_{i=0}^{o_q} b_i \cdot q_d(t + \Delta t) \quad (4.20)$$

Afin de simplifier la forme précédente (équation (4.20)), on utilise la réponse du système uniquement sur sa partie libre, ce qui nous permet de négliger la fonction d'entrée  $p_m$  pour  $t \geq 0$  :

$$\begin{cases} \sum_{i=0}^{o_q} b_i \cdot h_{d,m}(t + i\Delta t) = 0 & \text{pour } t > 0 \\ \sum_{i=0}^{o_q} b_i \cdot h_{d,m}(i\Delta t) = 0 & \text{pour } t = 0 \end{cases} \quad (4.21)$$

En reprenant l'équation (4.17), on obtient en temporel :

$$h_{d,m}(k \cdot \Delta t) = 2 \cdot \Re e \left[ \sum_{k_{p\acute{o}le}=1}^n \left( A_{d,m}(k_{p\acute{o}le}) \cdot z_{k_{p\acute{o}le}}^k \right) \right] \quad (4.22)$$

Soit en remplaçant dans l'équation (4.21) :

$$2 \cdot \Re e \left[ \sum_{k_{p\acute{o}le}=1}^n \left( A_{d,m}(k_{p\acute{o}le}) \sum_{k=0}^{o_q} b_k \cdot z_{k_{p\acute{o}le}}^k \right) \right] = 0 \quad (4.23)$$

## I. Contexte de l'analyse modale

On travaille dans une bande de fréquence donnée, ce qui nous permet de dire qu'il existe  $\{b_0, b_1, \dots, b_{o_q}\}$  vérifiant :

$$\sum_{k=0}^{o_q} b_k \cdot z_{k_{\text{pôle}}}^k = 0 \quad (4.24)$$

C'est un système linéaire que l'on résout à l'aide de  $o_q$  équations. À partir du moment où les coefficients  $b_k$  sont déterminés, ce que l'on fait à partir de l'équation (4.21), on peut l'utiliser pour trouver les pôles échantillonnés  $z_{k_{\text{pôle}}}$  qui nous permettent ensuite de trouver les valeurs  $\lambda_{k_{\text{pôle}}}$  propres continues à l'aide de l'équation (4.11).

La réalité expérimentale oblige une surdétermination du système, c'est-à-dire de prendre un nombre d'équation supérieur à  $o_q$ . Cette surdétermination permet d'augmenter la précision des résultats mais engendre un phénomène d'*aliasing*.

Ce problème est résolu dans une seconde étape à l'aide d'un diagramme de stabilisation.

### La méthode d'Ibrahim :

Cette méthode utilisant un cadre modal linéaire (Ibrahim et Mikulcik, 1977) est toujours d'actualité. Elle est basée sur l'équation générale d'un système dissipatif en réponse libre :

$$M \cdot \ddot{x}(t) + C \cdot \dot{x}(t) + K \cdot x(t) = 0 \quad (4.25)$$

avec  $x$ ,  $M$ ,  $K$  et  $C$  définis dans l'équation (4.1). Comme on l'a vu dans la méthode *LSCE*, la solution de la réponse de décroissance libre, en dehors de l'excitation est de la forme :

$$x(t) = \sum_{k=1}^{2n} a_k \cdot e^{\lambda_k \cdot t} \quad (4.26)$$

L'originalité est de construire un double système de  $2n$  équations, puisqu'il y a  $n$  pôles et  $n$  pôles conjugués en prenant les mesures aux instants  $\{t_1, t_2, \dots, t_{2n}\}$  :

$$[x(t_1) \dots x(t_{2n})] = [a_1 \dots a_{2n}] \cdot \begin{bmatrix} e^{\lambda_1 \cdot t_1} & \dots & e^{\lambda_1 \cdot t_{2n}} \\ \vdots & & \vdots \\ e^{\lambda_{2n} \cdot t_1} & \dots & e^{\lambda_{2n} \cdot t_{2n}} \end{bmatrix} \quad (4.27)$$

$$\iff \mathcal{X}_t = \mathcal{A}_t \cdot \Lambda_t \quad (4.28)$$

et aux instants décalés de  $\Delta t$  :

$$[x(t_1 + \Delta t) \dots x(t_{2n} + \Delta t)] = [a_1 \cdot e^{\lambda_1 \cdot \Delta t} \dots a_{2n} \cdot e^{\lambda_{2n} \cdot \Delta t}] \cdot \begin{bmatrix} e^{\lambda_1 \cdot t_1} & \dots & e^{\lambda_1 \cdot t_{2n}} \\ \vdots & & \vdots \\ e^{\lambda_{2n} \cdot t_1} & \dots & e^{\lambda_{2n} \cdot t_{2n}} \end{bmatrix} \quad (4.29)$$

$$\iff \mathcal{X}_{t+\Delta t} = \mathcal{A}_{t+\Delta t} \cdot \Lambda_t \quad (4.30)$$

On obtient ainsi :

$$\mathcal{X}_{t+2\Delta t} = \mathcal{A}_{t+2\Delta t} \cdot \Lambda_t \quad (4.31)$$

Ce qui donne le système d'équations linéaires

$$\begin{cases} \begin{bmatrix} \mathcal{X}_t \\ \mathcal{X}_{t+\Delta t} \end{bmatrix} = \begin{bmatrix} \mathcal{A}_t \\ \mathcal{A}_{t+\Delta t} \end{bmatrix} \cdot \Lambda_t \\ \begin{bmatrix} \mathcal{X}_{t+\Delta t} \\ \mathcal{X}_{t+2\Delta t} \end{bmatrix} = \begin{bmatrix} \mathcal{A}_{t+\Delta t} \\ \mathcal{A}_{t+2\Delta t} \end{bmatrix} \cdot \Lambda_t \end{cases} \quad (4.32)$$

(4.33)

$$\Leftrightarrow \begin{bmatrix} \mathcal{X}_{t+\Delta t} \\ \mathcal{X}_{t+2\Delta t} \end{bmatrix} \begin{bmatrix} \mathcal{X}_t \\ \mathcal{X}_{t+\Delta t} \end{bmatrix}^{-1} \begin{bmatrix} \mathcal{A}_t \\ \mathcal{A}_{t+\Delta t} \end{bmatrix} = \begin{bmatrix} \mathcal{A}_{t+\Delta t} \\ \mathcal{A}_{t+2\Delta t} \end{bmatrix} \quad (4.34)$$

Le système se résout alors en remarquant que pour chaque ligne, on a la relation suivante :

$$\begin{bmatrix} \mathcal{X}_{t+\Delta t} \\ \mathcal{X}_{t+2\Delta t} \end{bmatrix} \begin{bmatrix} \mathcal{X}_t \\ \mathcal{X}_{t+\Delta t} \end{bmatrix}^{-1} a_i = a_i \cdot e^{\lambda_i \cdot \Delta t} \quad (4.35)$$

En pratique, on considère un nombre d'échantillons supérieur à  $2n$  et l'on résout au sens des moindres carrés. Cette méthode a permis de faire une identification des paramètres modaux dans un contexte linéaire. L'aspect temporel est très important puisqu'il évite les phénomènes de fuite spectral et permet un traitement séquentiel des données.

Sous le nom de ITD (pour *Ibrahim Time Domain*), elle est très populaire dans le milieu de l'analyse modale expérimentale. Cependant Ibrahim a cherché à aller plus loin et a appliqué l'ITD sur des signaux aléatoires Ibrahim (1977). L'innovation majeure revient d'avoir utilisé un estimateur pertinent pour l'autocovariance d'un signal : le décrétement aléatoire proposé par Cole (1968) et peu connu comme le souligne Kopff (1990) une vingtaine d'années plus tard. Cette méthode a subi un certain désintérêt après sa sortie à cause de explosion de l'utilisation des techniques d'analyse spectrales. Elle a été utilisée dans des travaux récents (Lin et Tseng, 2014; Minko, 2005) pour ses avantages : transposition aisée sur des problématiques d'analyse modale opérationnelle, faible coût en calcul et faible coût en mémoire. Vincent (2016) le souligne en proposant même une utilisation sur 1 bit. Ces utilisations montrent la proximité de l'AME avec l'AMO.

#### Méthode des Sous-espaces Stochastiques :

Les méthodes SSI (pour *Stochastic Subspace Identification*) sont plus récentes que les précédentes (Van Overschee et De Moor, 1996) et proposent une approche temporelle différente basée sur un modèle d'état. Elles peuvent s'appliquer à la fois sur des réponses libres et sur des excitations aléatoires, sous une condition de non-corrélation. Elle introduit le cadre mathématique des séries temporelles, particulièrement bien adapté aux méthodes temporelles et au caractère discret des mesures.

Ainsi, on définit le déplacement à l'instant  $i$  du pas de temps  $\Delta t$  tel que  $x_k = x(k \cdot \Delta t)$ . La méthode SSI utilise alors une représentation d'état décrivant un système dynamique :

$$\begin{cases} s_{k+1} = A s_k + w_k \\ x_k = B s_k + v_k \end{cases} \quad (4.36)$$

Ce couple d'équation est composé d'une équation d'évolution et d'une équation d'observation (comme défini au chapitre 3). Il permet ainsi de faire coévoluer le modèle avec les mesures sur la structure réelle et de minimiser les écarts avec l'équation d'observation. Le système 4.1

## I. Contexte de l'analyse modale

dans sa version conservative en réponse libre correspond à ce formalisme. Aussi on définit  $s_k$  le vecteur d'état,  $w_k$  un bruit modélisant l'erreur du modèle et  $v_k$  celui de la mesure. Ces deux bruits sont supposés centrés gaussiens et non corrélés. On définit alors la covariance comme :

$$E \left[ \begin{pmatrix} w_k \\ v_k \end{pmatrix} \begin{pmatrix} w_{k+k} \\ v_{k+k'} \end{pmatrix}^T \right] = \begin{pmatrix} Q & S \\ S & R \end{pmatrix} \cdot \delta_{kk'} \quad (4.37)$$

Avec  $\delta_{kk'}$  la fonction de Kronecker. On pose  $P_0$  la covariance du vecteur d'état à l'instant  $k$ ,  $\bar{s}_k$  la moyenne du vecteur d'état et  $\Lambda_{k|k}$  la covariance de l'observation à l'instant  $k$ , on a alors les propriétés suivantes :

$$E[s_k \cdot s_k^T] = AP_0A^T + Q \quad (4.38)$$

$$E[s_{k+1} \cdot s_k^T] = G = AP_0B^T + S \quad (4.39)$$

$$\Lambda_0 = BP_0B^T + R = E[x_k \cdot x_k^T] \quad (4.40)$$

$$\Lambda_{k'} = BA^{k'-1}G = E[x_{k+k'} \cdot x_k^T] \quad (4.41)$$

Différentes méthodes ont été proposées dans l'état de l'art pour résoudre le système. Elles font appel au calcul de l'estimateur de la covariance de la mesure :

$$\Lambda_i = \frac{1}{N-i} \sum_{k=0}^{N-i-1} (x_{k+i} \cdot x_k^T) \quad (4.42)$$

Elles permettent alors de déterminer expérimentalement les pôles échantillonnés  $z_{k_{pôles}}$  et les vecteurs propres  $\varphi_m$  de  $A$ , qui permettent d'identifier les pôles  $\lambda_{k_{pôles}}$  et les déformées modales  $\Phi_m$  par les relations :

$$z_{k_{pôles}} = e^{\lambda_{k_{pôles}} \Delta t} \quad (4.43)$$

$$\Phi_m = B \cdot \varphi_m \quad (4.44)$$

Afin de compléter les méthodes présentées précédemment, nous proposons de rappeler quelques outils de caractérisation utilisés pour analyser les résultats.

### c. Un outil pour l'analyse modale expérimentale : le Critère d'Assurance Modale (MAC)

Les méthodes d'identification de l'analyse modale expérimentale offrent des résultats pour une excitation particulière de la structure. La vision expérimentale n'a pourtant de sens que dans la comparaison des résultats à la structure réelle. Le *MAC* est très largement utilisée pour comparer des familles de déformées modales entre elles. Soit  $\mathcal{B} = \{\Phi_1, \dots, \Phi_m\}$  et  $\mathcal{B}' = \{\Phi'_1, \dots, \Phi'_m\}$ , alors on définit le *MAC* comme étant :

$$\text{MAC}(\Phi_{(i)}, \Phi'_{(j)}) = \left( \frac{\Phi_{(i)}^T \cdot \Phi'_{(j)}}{\|\Phi_{(i)}\| \cdot \|\Phi'_{(j)}\|} \right)^2 \quad (4.45)$$

Il permet d'évaluer la corrélation entre deux modes  $\Phi$  et  $\Phi'$  appartenant à deux bases différentes. La valeur de ce critère varie de 0 à 1 et on considère que 0,9 constitue une valeur acceptable pour valider une correspondance entre deux modes.

Ces quelques grandes méthodes nous ont permis d'exposer les approches classiques d'identification modale expérimentale. Dans le prolongement de celles-ci, l'état de l'art plus récent propose d'autres méthodes d'identification. Nous proposons de nous éloigner de l'analyse modale expérimentale pour avoir une vision plus large des domaines ayant des problématiques similaires.

#### I.B.4. Autres stratégies d'identification

Les méthodes d'identification diffèrent selon l'utilisation d'un modèle et en fonction du niveau auquel se situe cette modélisation :

- Au niveau mécanique, par exemple en utilisant l'équation (4.1).
- Au niveau de la courbe de réponse du système comme par exemple pour les méthodes *LSCE* ou *ITD*.

Les avantages de l'utilisation d'un modèle physique sont nombreux : identification directe des paramètres, correction des mesures par l'utilisation du modèle, extraction d'autres paramètres physiques. Les méthodes présentées précédemment basées sur la modélisation des courbes obtenues sont plus robustes et plus éprouvées. Enfin, une nouvelle tendance s'est dégagée récemment par des techniques de décomposition du signal sans aucun modèle, dites empiriques : c'est l'*Empirical Modal Decomposition* ou *EMD* (Huang *et al.*, 1998; Schmitt *et al.*, 2014). Malgré l'avantage de pouvoir décomposer n'importe quel signal, ces méthodes sont très lourdes, et perdent totalement le sens physique du signal.

La méthode de *motion magnification* propose également d'extraire des informations de mouvement à partir de vidéos sans à priori sur l'observation. Cette méthode s'appuie sur différents outils mathématiques selon les publications : les *Complex Steerable Pyramids* (Wadhwa *et al.*, 2013) ou les pyramides de Riesz (Wadhwa *et al.*, 2014); il permet d'effectuer une extraction de la phase locale en l'isolant par bandes de fréquences. Cette variante de magnification est basée sur l'extraction de champs de déplacements d'une image à partir de ses informations de phase. Ainsi, pour une image 1D  $I(u)$ , lorsqu'on observe l'image déplacée d'un déplacement  $\delta_d(t)$  dans l'espace de Fourier, on a :

$$I(u + \delta_d(t)) = \sum_{-\infty}^{\infty} A_{\omega} e^{j\omega(u + \delta_d(t))} \quad (4.46)$$

Ce qui signifie que pour une bande de fréquence donnée, la phase  $\omega(u + \delta_d(t))$  contient des informations de déplacement. En filtrant la composante statique, on obtient aisément  $\omega\delta_d(t)$  que l'on peut alors facilement multiplier par un facteur d'amplification dans un espace fréquentiel local, avant de reconstruire l'image : on effectue ainsi une magnification.

Ces techniques obtiennent de très bons résultats mais reposent sur une chaîne de traitement aveugle en ne prenant pas en compte le modèle de ce qu'elles observent. L'avantage évident de ce type d'approche est l'adaptabilité de la méthode, mais on peut souligner qu'il est dommage de se priver de la connaissance physique du système observé.

Un grand nombre de méthodes proposent d'utiliser un modèle dans une approche plus orientée vibratoire que traitement du signal. Comme on a pu le voir dans la méthode SSI plus haut, les approches utilisant l'équation générale d'un système mécanique en vibration est très utilisée. Des approches similaires utilisant cette fois des processus séquentiels comme les filtres de Kalman sont utilisés avec des modèles physiques pour détecter les endommagements Seibold et Weinert (1996). On retrouve l'utilisation de modèles physiques associées aux filtres de Kalman dans le sens de la théorie de l'estimation à de multiples reprises pour de la détection en ligne (Ma *et al.*, 1999; Shao et Mechefske, 2009; Zhan et Mechefske, 2007) mais également dans le cadre du contrôle actif de vibration (Dong *et al.*, 2006). Ces implémentations reposent sur une mise en contexte du modèle de manière à pouvoir se ramener dans un cas linéaire.

La linéarité est le problème majeur de l'utilisation directe d'un modèle physique dans des méthodes d'identification. En effet, les méthodes telles que les filtres de Kalman peuvent s'adapter à des

## I. Contexte de l'analyse modale

fonctions d'évolution non-linéaires (voir chapitre 3 dans la partie III), mais conservent un formalisme dérivé du linéaire ; il permet certes d'être rapide au détriment d'erreurs d'approximations qui dépendent de la linéarisation de la fonction à chaque pas. Le cadre spécifique de l'analyse modale expérimentale autorisant des post-traitements « longs », l'étude de ces méthodes a été peu investiguée par rapport à d'autres domaines. Pour la sismique des sols ou des grands ouvrages ou encore le SHM (pour *Structure Health Monitoring*), la réalité expérimentale a poussé les chercheurs dans d'autres directions. Ainsi, Wu et Smyth (2007) montrent sur des cas simulés l'application d'identification de raideur et d'amortissement avec des filtres de Kalman. Ils insistent également sur l'amélioration de la précision par passage du filtre de Kalman étendu au filtre de Kalman Non Parfumé. Elle permet de fortement réduire les erreurs sur une identification SDOF. Papadimitriou *et al.* (2011) propose d'appliquer un filtre de Kalman linéaire pour déterminer la charge d'un système. Il utilise alors un filtre linéaire appliqué à un modèle mécanique conservatif mais ayant en entrée la sollicitation de la structure et utilisant le bruit gaussien de manière à déterminer des paramètres. Lourens *et al.* (2012) propose pour la première fois dans ce domaine d'appliquer un filtre de Kalman Augmenté, c'est-à-dire en composant un vecteur d'état de grandeurs observables et de paramètres optimisés par d'autres processus. Les résultats sont instables et Naets *et al.* (2015) formule un cadre plus robuste grâce à un filtre de Kalman linéaire et une détermination de la force par une modélisation de marche aléatoire. Cette approche sera encore améliorée par Eftekhari Azam *et al.* (2015) qui propose une structure de filtre différente et l'utilisation d'un UKF.

Naets *et al.* (2014) va également laisser entrevoir une application en réalité augmentée, en combinant capteurs de précision, caméra de visualisation et en exploitant la portée temps réel de l'EKF. L'application est faite sur le cas test académique de la poutre encastrée libre, mais développée par la suite dans un cadre plus général au sein du projet MODRIO qui s'est terminé fin 2015 (Tamarozzi *et al.*, 2014). Ce projet proposait de développer des schémas d'estimation afin d'intégrer des modèles réduits dans les tests réels et d'effectuer des mesures virtuelles dans un cadre modal. La solution proposée utilise directement l'équation de la dynamique dans le schéma d'estimation et propose d'estimer les forces entrantes sur le système en prenant en compte le comportement modal de ce dernier. Nous utiliserons une approche différente puisque nous nous concentrerons sur l'identification des paramètres modaux sans passer par l'intégration des équations de la dynamique. Par ailleurs, nous nous placerons dans le cadre plus élargi de l'assimilation de données. Les références récentes présentées dans la sous-sous-section II.B.1. du chapitre 3 nous montrent que les approches utilisant des méthodes d'assimilation de données sont en pleine expansion avec pour but d'allier connaissance du modèle mécanique et la réalité des essais. Aussi il est légitime de s'interroger sur la pertinence d'utiliser ces méthodes dans le contexte plus restreint de l'analyse modale expérimentale. Dans le cadre de l'analyse modale introduit précédemment, la linéarité est souvent présentée comme étant une hypothèse fondamentale de la théorie vibratoire. On a pu voir dans les différentes méthodes exposées à quel point les expérimentateurs essayaient de s'y ramener. Cependant la confrontation de l'expérience et du calcul montre que ce cadre d'étude est loin d'être applicable aux comportements observés sur des structures réelles. Les non-linéarités peuvent avoir des sources diverses : présence de frottements, de chocs, géométrie de la structure, assemblages entre sous-structures. Le constat de l'influence de la variation d'amplitude d'excitation sur la réponse de la structure est très explicite et il s'avère que des effets non-linéaires apparaissent très facilement. Par exemple sur le cas particulier de la poutre encastrée libre, des effets non-linéaires apparaissent à partir de déplacements verticaux de l'ordre du dixième de l'épaisseur de la poutre. En parallèle de ce cadre linéaire, une approche non-linéaire existe. Initiée dans les années 1960, elle a été relancée par Vakakis dans les

années 1990<sup>3</sup>. Cette approche purement non-linéaire n'entrera pas dans le cadre d'application des concepts développés dans cette thèse. En effet, les méthodes d'assimilation séquentielle peu coûteuses sont basées sur une approximation linéaire de fonctions non-linéaires, une application dans un cadre général non-linéaire ne serait pas judicieuse et pourrait amener à se placer aux limites du cadre des filtres de Kalman, dégradant ainsi les performances et limitant l'intérêt de la méthode. Foltête (1998) remarque avec humour que « les structures mécaniques classiques n'ont pas l'âme imaginative » et liste les non-linéarités communément rencontrées en analyse modale expérimentale. Il se place alors dans le cadre spécifique des faibles non-linéarités. Nous nous placerons dans la continuité de ce cadre théorique, ce qui nous permettra de travailler dans les conditions d'utilisations des méthodes non-linéaires d'assimilation de données tout en restant dans les hypothèses de l'analyse modale expérimentale.

### I.C. Apports et limites de la mesure par caméra rapide

Dans un contexte de grandes structures, d'applications automobiles ou de *GVT* (*Ground Vibration Tests*) en aéronautique, la taille des structures ne facilite pas la tâche de l'instrumentation. Les méthodes de mesures actuelles sont effectuées à l'aide d'accéléromètres, ce qui demande un coût important de temps d'installation, d'encombrement et une dégradation potentielle de la fidélité des mesures par un ajout de masse sur la structure, et surtout une modification des amortissements modaux dus à l'influence des câbles.

Aussi, les apports immédiats de la mesure par caméra rapide sont indéniablement matériels. Au niveau de l'encombrement, la possibilité qu'offrent les caméras de pouvoir faire de la mesure de champ permet une grande flexibilité au niveau du nombre de points capturés. Ce nombre de points peut être modulé selon les besoins en précision ou en vitesse d'exécution. L'aspect mesure de champ offre également intrinsèquement une synchronisation des points mesurés. Les caméras rapides ont l'avantage de faire de la mesure sans contact et d'offrir une mesure multi-points avec un impact faible sur la structure. Avantage ou inconvénient selon les usages, les caméras offrent une mesure native en déplacement ou en vitesse selon les méthodes de traitement vidéo utilisées mais ne permettent pas d'obtenir des mesures d'accélération.

Comme il a été vu dans le chapitre 1 les caméras rapides peuvent monter à des fréquences d'échantillonnage extrêmement élevées mais toujours au détriment de la précision spatiale. Elles sont victimes d'une certaine dualité spatio-temporelle dont la principale cause se trouve être le débit de transfert vers la mémoire voire la vitesse d'écriture sur la mémoire. Or, il se trouve que l'analyse modale nécessite à la fois des fréquences d'échantillonnage élevées et des mesures extrêmement précises. C'est un paramètre limitant de l'utilisation de caméras rapides en AME. Cette limitation peut être atténuée par différents procédés comme l'utilisation de méthodes d'extraction de points sous pixeliques et une utilisation réfléchie au niveau du choix d'objectifs. Les toutes récentes avancées matérielles ont permis de faire passer les résolutions HD (1920x1080) à des résolutions 4K (4096x3072) pour des vitesses allant jusqu'à 1000 Hz, améliorant un peu plus les possibilités pour l'étude des basses fréquences. Les problèmes de résolution spatiale étant propres aux systèmes numériques, il faut donc profiter de toute la partie analogique de la chaîne de mesure pour gagner en précision. Ainsi, l'utilisation d'un objectif spécifique couplé avec une calibration adéquate permettra de maximiser l'étalement du champ de vision sur le capteur et ainsi pouvoir accéder à des mesures très fines. Une ambiguïté se dessine : la caméra possède une résolution fixe en pixels, mais une résolution variable lorsqu'on

3. On se reportera aux travaux de Kerschen *et al.* (2013) pour l'état de l'art récent.

## I. Contexte de l'analyse modale

parle de pixel projeté sur la structure. En réalité, c'est le rapport entre la largeur du champ et la précision qui va caractériser la résolution d'une caméra. Un autre inconvénient majeur de la vidéo rapide se trouve dans la projection faite par la caméra d'une structure réelle en trois dimensions vers un capteur en deux dimensions. Pour pallier cette perte d'information, plusieurs moyens existent. Utiliser deux caméras afin de faire de la stéréo vision peut permettre de reconstruire une image en trois dimensions sur un champ défini en profondeur. Cette thématique est développée activement et son utilisation s'apprête à passer une importante marche normative, essentielle pour les applications aéronautique à la suite du projet DICCIT (Mathieu *et al.*, 2015). Pour une scène en 2D, on peut également négliger cette projection en plaçant la caméra de manière adéquate avec les hypothèses de l'expérience (projection selon l'axe  $\vec{z}$ , normal au mouvement plan  $(\vec{x}, \vec{y})$ ).

La synchronisation des points mesurés par caméras rapides offre la possibilité de réaliser l'acquisition de transitoires ou de sollicitations aléatoires particulièrement intéressantes pour l'AME ou l'AMO. On notera cependant l'intérêt de coupler ce capteur à des méthodes permettant de descendre sous la limite de résolution du pixel, ce que peuvent nous permettre les méthodes d'assimilation de données en travaillant au niveau du bruit des mesures. Enfin, il est nécessaire de compenser le problème majeur de projection de la structure en trois dimensions sur un capteur en deux dimensions, ce que nous verrons par la suite.

### I.D. Conclusion

Cette première partie introduit le domaine de l'analyse modale expérimentale. Après avoir constaté que le cadre de l'analyse modale était favorable pour l'utilisation de méthodes d'assimilation de données séquentielles, les principales techniques utilisées dans ce domaine ont été rappelées. Concernant les capteurs employés, les caméras rapides permettent de faire de l'analyse multipoints pour de la mesure de transitoires ou d'aléatoire sans problématiques de précision des mesures ou de coût supplémentaire par point observé. Concernant les méthodes d'identification modales nous avons décrit des méthodes fréquentielles qui peuvent poser certains problèmes de précision et qui ne sont pas applicables dans un schéma temps réel. Les méthodes temporelles possèdent l'avantage d'être séquentielles mais ne proposent pas la rapidité d'exécution des filtres de Kalman pour des évolutions non-linéaires des paramètres modaux. La recherche bibliographique a été ensuite complétée pour conclure que dans le domaine de la mécanique, les méthodes favorisant le dialogue essais-calculs étaient en pleine expansion. Parmi les différentes approches, les filtres de Kalman sont souvent employés pour une utilisation le temps réel.

Nous proposons ainsi d'utiliser les concepts développés dans ce manuscrit pour proposer une méthode d'assimilation de données dans le cadre de l'analyse modale expérimentale de structures. Cette méthode permettra de faire de l'identification modale faiblement non-linéaire tout en offrant d'autres avantages vus précédemment. La section suivante s'interroge sur un modèle d'évolution envisageable pour construire le modèle d'état de notre méthode.

## II Construction d'une fonction d'évolution pour l'assimilation séquentielle modale

### II.A. Contexte particulier de l'analyse modale

#### II.A.1. Propriétés des signaux étudiés

L'analyse vibratoire est un domaine suffisamment particulier pour pouvoir définir un cadre d'étude restreint à travers les types de signaux étudiés. De la même manière que Feldman (2011a) nous allons nous placer dans des conditions plus spécifiques en travaillant dans celui de l'analyse modale expérimentale.

Dans cette partie, les modèles des principaux signaux rencontrés sont présentés afin de définir des hypothèses sur la construction de notre modèle d'état.

L'analyse modale expérimentale offre le contrôle de l'excitation du système à l'expérimentateur, ce qui permet de modéliser de manière plus déterministe la réponse du système. On propose de rappeler quelques excitations caractéristiques de l'AME :

##### **Les signaux aléatoires avec des hypothèses de blancheur de spectre :**

Les signaux aléatoires peuvent être décrits par leur contenu fréquentiel. On utilise des outils comme la densité spectrale de puissance ou les moments statistiques pour les caractériser. Dans ces travaux nous n'utiliserons pas ces signaux mais ils feront partie de nos perspectives.

##### **La réponse libre à des excitations de type choc :**

Ce type de sollicitation n'est pas toujours facile à produire de manière à exciter les modes d'ordres élevés (problème d'une excitation de type Dirac imparfaite), surtout pour les grandes structures, mais possède l'avantage de permettre l'observation de l'amortissement de la structure à l'aide de la décroissance de l'amplitude. Le signal sur un point de mesure peut alors être modélisé par une somme de composantes modales :

$$s(t) = \sum_{i=1}^M a^i \cdot e^{-2\pi \int_0^t f^i(\tau) d\tau \xi^i} \cdot \cos \left( 2\pi \cdot \int_0^t f^i(\tau) d\tau + \phi^i \right) \quad (4.47)$$

$f^i$  représente alors la fréquence,  $\phi^i$  la phase,  $\xi^i$  l'amortissement et  $a^i$  l'amplitude pour le mode  $i$  parmi  $M$  modes.

##### **Les excitations de type quasi-sinus :**

Dans le cas des excitations de type quasi-sinus, l'évolution de la fréquence est définie différemment selon qu'elle soit constante ou variable. On parle alors de *chirp*. Le signal faisant intervenir l'intégrale de la fréquence, il est de la forme suivante pour un système SDOF :

$$s(t) = A^i \cos \left( 2\pi \cdot \int_0^t f^i(\tau) d\tau + \phi^i \right) \quad (4.48)$$

#### II.A.2. Cadre théorique

Parmi tous les modèles vus dans la section précédente, nous avons choisi de travailler sur des réponses libres à une sollicitation de type choc. Dans le cadre de l'utilisation d'une méthode stochastique

## II. Construction d'une fonction d'évolution pour l'assimilation séquentielle modale

comme un filtre de Kalman, on propose de rappeler l'expression générale d'un système vibratoire :

$$\begin{bmatrix} I & 0 \\ 0 & M \end{bmatrix} \begin{pmatrix} \dot{x} \\ \ddot{x} \end{pmatrix} = \begin{bmatrix} 0 & I \\ -K & -C \end{bmatrix} \begin{pmatrix} x \\ \dot{x} \end{pmatrix} + \begin{pmatrix} 0 \\ F_{ext} \end{pmatrix} + \begin{pmatrix} 0 \\ F_{NL}(x, \dot{x}) \end{pmatrix} + \begin{pmatrix} 0 \\ \mu \end{pmatrix} \quad (4.49)$$

Cette équation est identique à l'équation (4.1) sauf qu'on introduit  $\mu$  l'incertitude du modèle par rapport aux observations réelles. Dans notre cadre stochastique on suppose que  $\mu$  est gaussienne, centrée, de moyenne nulle et avec une covariance  $R_\mu$ .

On exprime alors la solution de ce système pour des **excitations de type choc**. Le cadre de l'analyse modale expérimentale nous autorise à effectuer un découplage spatiotemporel essentiel pour le modèle d'état. En posant la base modale :

$$\mathcal{B}(x, y, z) = \left\{ \Phi^1(x, y, z), \Phi^2(x, y, z), \dots, \Phi^n(x, y, z) \right\} \quad (4.50)$$

dépendant uniquement de l'espace et on peut exprimer la déformée du système en trois dimensions au temps  $t$  :

$$S(x, y, z, t) = \sum_{i=1}^n V^i(t) \cdot \Phi^i(x, y, z) + \nu \quad (4.51)$$

$$\text{avec } V^i(t) = A^i(t) \cdot \cos \left( 2\pi \cdot \int_0^t f^i(\tau) d\tau + \phi^i \right) \quad (4.52)$$

$$A^i(t) = a^i(t) \cdot e^{-2\pi \cdot \int_0^t f^i(\tau) d\tau \xi^i(t)} \quad (4.53)$$

$$\nu = \sum_{i=n+1}^{\infty} V^i(t) \cdot \Phi^i(x, y, z) + \mu \quad (4.54)$$

$f^i(t)$  représente alors la fréquence,  $\phi^i(t)$  la phase,  $\xi^i(t)$  l'amortissement et  $a^i(t)$  l'amplitude pour le mode  $i$ . L'introduction d'une incertitude sur le modèle  $\nu$  est particulièrement intéressante puisqu'elle nous permet d'opérer une réduction de modèle par l'intermédiaire d'une troncature modale.

Deux hypothèses importantes sont à souligner :

- La base modale utilisée est supposée purement réelle, c'est-à-dire sans phase entre les différents points d'observation pour chaque mode propre.
- On se place bien dans les hypothèses d'un système faiblement non-linéaire, puisque l'on considère que les vecteurs d'état sont invariants au cours du temps.

L'identification des paramètres du modèle passe par la séparation des contributions modales et en particulier par l'extraction des amplitudes et des fréquences. Étant donné la forme de la déformée du système, une méthode telle que la démodulation pourrait permettre de faire une identification. C'est cette direction que la partie suivante propose de développer.

### II.B. La démodulation comme modélisation d'un signal vibratoire

La modulation/démodulation est une technique très utilisée dans le domaine des télécommunications. Elle en a favorisé l'essor important depuis l'après-guerre jusqu'à nos jours en se basant sur un principe simple : mélanger de manière réversible deux signaux pour n'en faire qu'un. L'idée est déclinée sous diverses formes comme la modulation d'amplitude ou la modulation de phase. Ce sont ces techniques, utilisées en mono-bandes à l'époque qui contribuent toujours à de récents progrès en multi-bandes.

On rappelle le principe simple de la modulation : « d'inclure » un signal informatif  $m(t)$  dans une onde porteuse. Ce type de modulation est analogique. Les deux modulations les plus connues sont les modulations d'amplitude (AM) et de fréquence (FM) selon que le message informatif est inclus dans l'amplitude ou dans la phase de la porteuse. L'équation ci-dessous décrit le principe de la modulation d'amplitude d'un signal modulé  $s(t)$  pour un signal porteur  $s_p(t)$  et un signal modulant  $s_m(t)$  d'amplitude  $A$  et de pulsation  $\omega$  :

$$\text{Soit } \begin{cases} s_p(t) = A_p \cdot \cos(\omega_p t) \\ s_m(t) = A_m \cdot \cos(\omega_m t) \end{cases}$$

$$s(t) = s_p(t) + k \cdot s_p(t) s_m(t)$$

$$s(t) = A_p \cdot [1 + k \cdot A_m \cdot \cos(\omega_m t)] \cos(\omega_p t)$$

Dans la visée d'un schéma temps réel, toute utilisation de filtres induit différents phénomènes qui peuvent fausser une mesure d'information à un pas de temps donné. Outre le déphasage, le filtrage nécessite généralement la mise en place d'un tampon. Pourtant l'utilisation de ces filtres semble pouvoir répondre à l'utilisation que l'on recherche : séparer et caractériser les différentes composantes d'un signal multispectral ou non.

Lorsque l'on compare les signaux utilisés dans les télécommunications et ceux décrits dans la partie précédente, on s'aperçoit qu'il s'agit du même cadre théorique : signaux périodiques modulés, réponses harmoniques, sommes de sinus. Cependant les signaux utilisés en télécommunications sont générés pour respecter les hypothèses sur lesquelles s'appliquent les transformées. La démodulation est de plus utilisée sur des signaux dont la modulation est parfaitement connue, ce qui n'est pas le cas en AME.

Pourtant l'usage de techniques de démodulation sur des modèles d'évolution d'un signal vibratoire apparaît tentante et existe par ailleurs déjà dans la littérature.

### II.B.1. Différentes techniques de démodulation

La technique de démodulation que nous cherchons doit être adaptée aux signaux observés en analyse modale expérimentale et pouvoir fonctionner dans un schéma temps réel. C'est dans ce contexte que le domaine de l'analyse de signaux audio a été investigué. En particulier nous nous intéressons au domaine de la reconnaissance vocale. Dans ce domaine, la décomposition d'un signal multi-spectral en composantes mono-fréquences est essentielle pour faire une séparation spectrale.

Les applications de reconnaissance vocale ont souvent une forte coloration temps réel. La reconnaissance vocale sépare plusieurs type de sons. Dans le cas particulier des voyelles, elles sont composées des plusieurs composantes que l'on appelle formant. Chaque formant est un signal bande étroite multispectral modulé. Déterminer la voyelle peut se faire en obtenant les deux plus basses fréquences de chaque formant qui le composent. Autrement dit, chaque formant peut être modélisé par (Kvedalen, 2003) :

$$s_f(t) = A \cdot r^t \cos(t) \cos(\omega_p t + \beta \sin(\omega_m t)) \quad (4.55)$$

avec  $A$  l'amplitude,  $r$  le coefficient d'amortissement,  $\beta$  la déviation en fréquence, et  $\omega_p$  et  $\omega_m$  les

## II. Construction d'une fonction d'évolution pour l'assimilation séquentielle modale

porteuses et modulant. Une voyelle est modélisée par :

$$v(t) = \sum_{i=1}^N A_i r_i^t \cos(t) \cos(\omega_{p_i} t + \beta_i \sin(\omega_{m_i} t)) \quad (4.56)$$

L'identification faite en reconnaissance vocale est similaire par rapport à notre cadre d'application au niveau du temps réel et des modèles employés pour modéliser le signal. Quelques tendances se distinguent dans ce domaine pour traiter ce problème d'identification.

### L'opérateur d'énergie de Teager-Kaiser

Pour expliquer cet opérateur, on peut faire l'analogie avec l'énergie d'un système mécanique conservatif :

$$\ddot{x} + \frac{k}{m}x = 0 \quad (4.57)$$

On définit l'énergie du système comme la somme de l'énergie potentielle et de l'énergie cinétique :

$$E = \frac{1}{2}kx^2 + \frac{1}{2}m\dot{x}^2 \quad (4.58)$$

Les solutions du système étant et la forme  $x(t) = A \cos(\omega t + \phi)$ , en remplaçant, on a

$$E = \frac{1}{2}m\omega^2 A^2 \quad (4.59)$$

Ce résultat nous montre que l'énergie d'un système vibratoire est directement liée à son amplitude et sa fréquence.

La définition de l'opérateur d'énergie de Teager-Kaiser reprend ce principe pour un signal :

$$\Psi(x(t)) = (\dot{x}^2(t)) - x(t) \cdot \ddot{x}(t) \quad (4.60)$$

Ce qui donne pour un signal de la forme  $x(t) = A \cos(\omega t)$  :

$$\Psi(x(t)) = A^2 \omega^2 \quad (4.61)$$

$$\text{soit :} \quad (4.62)$$

$$\hat{\omega}^2 = \frac{\Psi(\dot{x}(t))}{\Psi(x(t))} \quad (4.63)$$

$$|\hat{A}^2(t)| = \frac{\Psi(x(t))^2}{\Psi(\dot{x}(t))} \quad (4.64)$$

$$(4.65)$$

Le principe de la démodulation à l'aide de l'opérateur de Teager-Kaiser est alors exact pour ce type de signaux et introduit des erreurs pour d'autres types de signaux. Par exemple pour un signal AM :

$$\Psi(A(t) \cos(\omega_p t + \theta)) = A^2(t) \cdot \omega_p^2 + \cos^2(\omega_p t + \theta) \cdot \Psi(A(t)) \quad (4.66)$$

Ce qui revient à considérer le second terme de l'équation ci-dessus comme un terme d'erreur, et nous ramène au cas idéal vu précédemment.

La grande force de cet opérateur réside dans sa version discrète qui nécessite très peu de ressources et peut être utilisée en temps réel, puisqu'elle ne requiert que trois mesures du signal consécutives (Potamianos et Maragos, 1994).

On pourra citer les algorithmes DESA qui utilisent cet opérateur pour estimer fréquence et amplitude. Cependant, le calcul de dérivées sur un signal numérique peut s'avérer délicat. Cette méthode est à utiliser avec précaution selon la qualité du signal à analyser. Il est intéressant de noter que c'est un opérateur utilisé avec succès par certains auteurs en analyse modale (Antoniadou *et al.*, 2012).

### La méthode de Shekel

Cette méthode utilise la même philosophie que l'opérateur de Teager, c'est-à-dire d'utiliser les dérivées successives, cette fois pour un signal AM. En passant le développement analytique, cette méthode permet d'obtenir :

$$\begin{cases} a(t) = \sqrt{x^2(t) - \frac{\dot{x}^2(t)x(t)}{\ddot{x}(t)}} \\ \omega(t) = \sqrt{-\frac{x(t)}{\ddot{x}(t)}} \end{cases} \quad (4.67)$$

Mêmes inconvénients pour cette transformée, qui utilise des dérivées secondes du signal, affectant la robustesse de l'estimation.

### Le signal analytique à l'aide de la transformée de Hilbert

La transformée de Hilbert est formulée de la manière suivante :

$$\mathcal{H}\{x(t)\} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(s)}{t-s} ds \quad (4.68)$$

La démodulation par transformée de Hilbert est basée sur le concept de signal analytique :

$$s_a(t) = x(t) + j\mathcal{H}\{x(t)\} \quad (4.69)$$

Elle permet de former un signal en quadrature de phase par rapport au signal initial observé (voir figure 4.2).

La démodulation est alors immédiate en prenant le module et la phase du signal analytique considéré :

$$\begin{cases} |\hat{a}(t)| = |s_a(t)| \\ \hat{\omega} = \arg(s_a(t)) \end{cases} \quad (4.70)$$

L'utilisation du signal analytique présente également des défauts. La démodulation ne propose que d'identifier la valeur absolue de l'amplitude du signal et invoque une hypothèse limitante sur la nature du signal de bande étroite. Par ailleurs le calcul des transformées de Hilbert n'est pas trivial et peut introduire des erreurs.

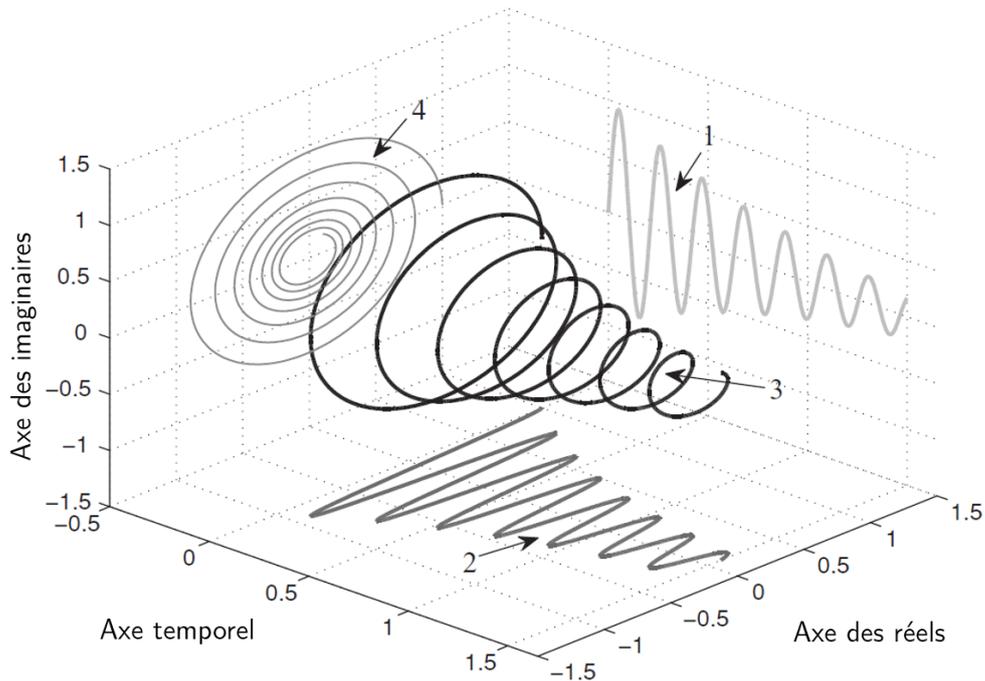
### La méthode de Prony

La méthode de Prony a été développée en 1795 par Gaspard de Prony. Elle s'applique sur des signaux de la forme :

$$s(t) = \sum_{i=1}^N A_i e^{\sigma_i t} \cos(\omega_i t + \phi) \quad (4.71)$$

$$(4.72)$$

## II. Construction d'une fonction d'évolution pour l'assimilation séquentielle modale



- (1) Projection effectuée par la transformée de Hilbert
- (2) Signal réel
- (3) Signal analytique
- (4) Phase dans le plan complexe

Feldman, © 2011 Elsevier

**Figure 4.2** – Figure expliquant la transformée de Hilbert tirée de Feldman (2011a) p13

soit en complexe :

$$s_z(t) = \sum_{i=1}^N \frac{1}{2} A_i e^{\pm j\phi_i} e^{\lambda_i t} \quad (4.73)$$

Cette méthode est proche de la méthode d'Ibrahim (ITD) présentée précédemment (paragraphe b. ).

Elle possède beaucoup de paramètres à régler ce qui contraint à beaucoup de mesures pour permettre une identification fidèle et qui de fait réduit la précision d'une estimation pour des fréquences variables.

Cette introduction aux techniques de démodulation est loin d'être exhaustive et d'autres méthodes existent, le but de ce paragraphe étant de présenter deux grandes familles de démodulation, l'une basée sur une vision géométrique complexe et l'autre axée sur une vision plus différentielle.

Néanmoins ces méthodes pourraient faire des modèles efficaces à utiliser dans des schémas d'assimilation de données. La démodulation possède toutes les caractéristiques pour modéliser l'évolution du signal, puisqu'elle permet d'identifier les paramètres modaux pour une composante modale. On propose dans la suite d'appliquer les concepts d'assimilation de données dans le cadre particulier du signal analytique.

## II.B.2. Une démodulation selon Hilbert pour multiplier les observations

Les propriétés de la transformée de Hilbert sur la détection d'enveloppe ont été très largement utilisées en mécanique comme dans d'autres domaines. Ainsi, il est difficile de ne pas citer Feldman (Feldman, 2011a,b) tant il a contribué au développement et à la démocratisation de l'utilisation de la transformée de Hilbert. La transformée possède de nombreux avantages pour faire de l'estimation de l'amplitude et de la fréquence instantanée pour les signaux rencontrés dans la mécanique vibratoire décrits dans la sous-sous-section II.A.1.. Elle est toujours très largement utilisée et de nombreux auteurs cherchent à la mettre en œuvre dans des thématiques récentes telles que les réseaux neuronaux (Ondra *et al.*, 2017). Il est important de mentionner que différents schémas d'identification de démodulation ont été testés au cours de ces travaux :

- Un modèle réel de la forme classique  $s(t) = \sum_{i=1}^M A_i \cos(\omega_i t + \phi_i)$ . Sans qu'une preuve formelle ait été apportée, nous n'avons pas réussi à faire converger l'assimilation de donnée pour ce modèle.
- Un modèle de Hilbert classique de la forme  $s(t) = \sum_{i=1}^M A_i (\cos(\omega_i t) + j \sin(\omega_i t))$ . Ce modèle est celui que nous allons utiliser par la suite.
- Un modèle de Prony a également été testé, mais l'identification de l'amortissement s'est avérée infructueuse.

Le choix que nous avons fait d'utiliser la transformée de Hilbert est lié à la fois à l'exigence du schéma temps réel, au cadre de l'assimilation de donnée, mais également à la nature des informations identifiées. Il apparaît clairement que l'amplitude et la fréquence modale sont naturellement présentes dans le domaine fréquentiel, tandis que l'amortissement est plus facile à identifier dans le domaine temporel. Il est intéressant d'observer dans l'état de l'art de domaines plus éloignés (électronique de puissance) que des filtres de Kalman utilisant le signal analytique avaient été proposés. Le but de ces filtres était différent, puisqu'il s'agissait d'obtenir des informations spectrales dans un schéma séquentiel, de manière à pouvoir filtrer certaines composantes. Ainsi, Girgis *et al.* (1991) proposent un modèle d'état basé sur le signal analytique capable de suivre des composantes spectrales à l'aide d'un Kalman linéaire, sans toutefois pouvoir suivre les fréquences. Ces pistes utilisant la transformée de Hilbert à travers le signal analytique pour suivre ou détecter des harmoniques sont d'ailleurs toujours d'actualité dans ce domaine (de la O Serna et Rodríguez-Maldonado, 2012; Will et Cardoso, 2012). Par contre, les modèles proposés ne permettent pas de suivre en même temps la fréquence puisqu'elle n'intervient pas dans le vecteur d'état. Dion *et al.* (2012) introduisent des modèles similaires dans le domaine de l'analyse modale mais proposent par contre de suivre également la fréquence à travers la version non-linéaire de Kalman *EKF*. Ce schéma permet un calcul rapide des signaux analytiques utilisés pour la détection d'enveloppes sans passer par le domaine de Fourier. C'est cette approche utilisant un espace d'observation mono-dimensionnel que nous avons choisie afin de la généraliser dans le cadre d'un espace d'observation N-dimensionnel. En reprenant le cadre décrit par Dion *et al.*, la solution générale donnée dans l'équation (4.51) pour un mode  $i$  s'exprime sous la forme :

$$V^i(t) = A^i(t) \cdot \cos \left( 2\pi \cdot \int_0^t f^i(\tau) d\tau + \phi^i \right) \quad (4.74)$$

L'utilisation d'une transformée de Hilbert pour construire le signal analytique permet d'extraire ensuite l'enveloppe du système. Cependant, l'utilisation de cette transformée ainsi que l'implémentation numérique que l'on fait n'est qu'une approximation qui n'est correcte que si l'on respecte la première condition du théorème de Bedrosian (Bedrosian, 1963) plus connue sous l'hypothèse du signal à bande étroite. Cette hypothèse se traduit sur le signal par une variation lente de l'amplitude par

## II. Construction d'une fonction d'évolution pour l'assimilation séquentielle modale

rapport à la variation du cosinus. Comme indiqué précédemment, cette configuration est utilisée dans les télécommunications avec succès, étant donné que les signaux sont construits en amont de manière adéquat puis filtrés en aval pour respecter cette hypothèse. Dans notre cas, on fait une identification d'un signal non maîtrisé, sur des signaux non plus périodiques mais harmoniques (Pachaud *et al.*, 2013). Malgré l'ancienneté de l'utilisation du signal analytique dans le domaine de l'analyse vibratoire, l'extension de son utilisation en dehors de la condition de Bedrosian reste un sujet actif (Shahriar *et al.*, 2017).

Dans notre cas, les conditions d'excitation de la structure nous permettent de considérer pour chaque mode que l'évolution de l'amplitude est lente par rapport à la fréquence d'intérêt, ce qui nous assure des hypothèses de bande étroite. Cependant les signaux caractéristiques que nous sommes susceptibles d'observer sont harmoniques, ce qui va engendrer des distorsions dans les signaux identifiés.

Ainsi notre modèle d'identification n'est pas idéal, mais cette erreur de modèle pourra être prise en compte dans la structure du filtre de Kalman. Aussi, on exprime le signal analytique sous la forme :

$$V_z^i(t) = A^i(t) \cdot e^{2\pi j \left( \int_0^t f^i(\tau) d\tau + \phi^i \right)} \quad (4.75)$$

On peut alors introduire le schéma classique de modulation en bande étroite pour réécrire l'équation sous la forme :

$$V^i(t) \simeq A^i(t) \cdot \cos \left( 2\pi \cdot \int_0^t f_p^i(\tau) + f_\Delta^i \cdot m(\tau) \cdot d\tau \right) \quad (4.76)$$

$$\text{avec } f^i(t) \simeq f_p^i + f_\Delta^i \cdot m^i(t) \quad (4.77)$$

$$f_\Delta^i \ll f_p^i \quad (4.78)$$

$$|m_i(t)| < 1 \quad (4.79)$$

avec  $f_p^i$  la fréquence de la porteuse,  $f_\Delta^i$  la déviation en fréquence et  $m_i(t)$  le signal modulant pour le mode  $i$ . Les conditions de bande étroite se traduisent donc par :

$$\frac{\partial A^i(t)}{\partial t} \ll 2\pi \cdot (f_p^i(t) + f_\Delta^i \cdot m(t)) \quad (4.80)$$

### II.B.3. Modèle d'état

Les conditions d'utilisation définies, il reste à décliner la transformée de Hilbert sous forme de modèle d'état, ce qu'on reprend dans l'article de Dion *et al.* (2012). Posons  $t = k \cdot \Delta t$  et  $\Delta t$  le pas de temps d'échantillonnage, et posons  $x_k^i$  l'approximation du signal analytique discrétisé :

$$x_k^i = \alpha_k^i \cdot e^{2\pi j f_k^i k \cdot \Delta t} \quad (4.81)$$

En considérant une évolution lente des paramètres  $\alpha^i$  et  $f^i$ , on les suppose constants entre deux pas de temps.

$$x_{k+1}^i \simeq \alpha_k^i \cdot e^{2\pi j f_k^i (k+1) \Delta t} \quad (4.82)$$

Ce qui nous donne le système d'état suivant :

$$x_{k+1}^i \simeq x_k^i \cdot e^{2\pi j f_k^i \Delta t} \quad (4.83)$$

On exprime alors le signal analytique sous sa forme algébrique, de manière à rendre plus robuste notre vecteur d'état à l'aide d'une autre variable aléatoire circulaire : sa transformée de Hilbert.

$$x_k^i = x_k^{1,i} + j \cdot x_k^{2,i} \quad (4.84)$$

$$\text{avec } x_{k+1}^{1,i} = x_k^{1,i} \cos(2\pi f_k^i \Delta t) - x_k^{2,i} \sin(2\pi f_k^i \Delta t) \quad (4.85)$$

$$x_{k+1}^{2,i} = x_k^{1,i} \sin(2\pi f_k^i \Delta t) + x_k^{2,i} \cos(2\pi f_k^i \Delta t) \quad (4.86)$$

Soit dans une forme vectorielle :

$$\begin{pmatrix} x_{k+1}^{1,i} \\ x_{k+1}^{2,i} \end{pmatrix} = \begin{bmatrix} \cos(2\pi f_k^i \Delta t) & -\sin(2\pi f_k^i \Delta t) \\ \sin(2\pi f_k^i \Delta t) & \cos(2\pi f_k^i \Delta t) \end{bmatrix} \begin{pmatrix} x_k^{1,i} \\ x_k^{2,i} \end{pmatrix} \quad (4.87)$$

Afin de permettre au filtre d'identifier également la fréquence, on exprime la fonction d'évolution et le vecteur d'état de la manière suivante :

$$X_{k+1}^i = F_k^i \cdot X_k^i \quad (4.88)$$

$$\text{avec } X_k^i = \begin{pmatrix} x_k^{1,i} \\ x_k^{2,i} \\ x_k^{3,i} \end{pmatrix} \quad x_k^{3,i} = 2\pi f_k^i \Delta t \quad (4.89)$$

$$\text{et } F_k^i = \begin{pmatrix} \cos(x_k^{3,i}) & -\sin(x_k^{3,i}) & 0 \\ \sin(x_k^{3,i}) & \cos(x_k^{3,i}) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (4.90)$$

Le vecteur d'état ainsi défini nous donne accès à l'amplitude instantanée par la relation  $\alpha_k^i = \sqrt{x_k^{1,i^2} + x_k^{2,i^2}}$ .

Cette représentation d'état permet de suivre le signal analytique d'un système SDOF. Afin d'appliquer ce modèle pour décomposer la structure sur plusieurs modes, on construit le vecteur d'état en concaténant ce vecteur d'état « unitaire ». On fait de même avec la fonction d'évolution que l'on construit par bloc, ce qui donne un vecteur d'état pour  $n$  modes :

$$X_k = \begin{pmatrix} X_k^1 \\ X_k^2 \\ \dots \\ X_k^n \end{pmatrix} \quad \text{et} \quad F_k = \begin{pmatrix} F_k^1 & 0 & \dots & 0 \\ 0 & F_k^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & F_k^n \end{pmatrix} \quad (4.91)$$

Par la suite nous utiliserons ce modèle d'évolution dans un filtre de Kalman permettant d'utiliser des modèles d'état non linéaire : l'EKF ou l'UKF. On pose dans ce contexte  $W_k$  le bruit gaussien et centré de ce modèle, de covariance  $Q_k$ .

#### II.B.4. Remarque sur une formulation purement complexe

Comme le montre Dash *et al.* (1999), il est possible de poser un modèle similaire sous une formulation complexe en utilisant l'ECKF (pour *Extended Complex Kalman Filter*). La version complexe

### III. Un premier schéma d'assimilation utilisant un modèle théorique modal de la structure : le Kinematic multiSensors Assimilation Modal Identification (K-SAMI)

des filtres de Kalman existe et ne diffère que dans les formules de changement de base où la matrice transposée est également conjuguée. L'intérêt en utilisant notre modèle d'état est de limiter les variables dans le vecteur d'état, tout en diminuant le nombre d'opérations par le fait que la variable d'état est complexe et la fonction d'évolution est diagonale. Les tests effectués n'ont pas permis de mettre en évidence une quelconque amélioration que ce soit en vitesse d'exécution ou en stabilité. Cependant, on peut simplement remarquer que cette structure de filtre permet de répartir la covariance sur deux matrices plus petites ce qui limite les problèmes de stabilité dus à de trop grandes matrices de covariance.

### III Un premier schéma d'assimilation utilisant un modèle théorique modal de la structure : le Kinematic multiSensors Assimilation Modal Identification (K-SAMI)

Le modèle que nous avons présenté dans la section précédente nous permet de travailler dans un espace d'état fréquentiel tout en traitant séquentiellement des données temporelles. Cet espace fréquentiel est intrinsèquement tronqué d'où la nature réduite de notre modèle.

Nous devons maintenant faire le lien avec notre espace d'observation. Nos observations sont des ensembles de points reliés par des propriétés particulières. En effet, si on pose l'espace discret tel que :

$$x = p_x \cdot \Delta x \quad (4.92)$$

$$y = p_y \cdot \Delta y \quad (4.93)$$

$$z = p_z \cdot \Delta z \quad (4.94)$$

$$\text{et } p = (p_x, p_y, p_z) \quad (4.95)$$

Alors l'équation (4.51) permet de décomposer la déformée mesurée sur l'ensemble des modes propres de la structure :

$$S_{k,p} = \sum_{i=1}^n V_k^i \cdot \Phi_p^i + \nu \quad (4.96)$$

Or  $V_k^i$  est la partie réelle du signal analytique, qui est définie à partir du vecteur d'état :

$$V_{k,p}^i = \text{Re} \left( x_{1,k}^i + j x_{2,k}^i \right) \quad X_k^i = \begin{pmatrix} x_k^{1,i} \\ x_k^{2,i} \\ x_k^{3,i} \end{pmatrix} \quad (4.97)$$

Aussi, la question de l'estimation de la partie modale se pose. Dans cette partie, on s'appuie sur l'hypothèse forte de l'AME de l'invariabilité temporelle des déformées modales. On peut alors s'appuyer sur la base modale  $\mathcal{B} = \{\Phi^1, \Phi^2, \dots, \Phi^n\}$  définie à l'équation (4.50) pour proposer une fonction d'observation. L'utilisation de cette base des déformées modales de la structure implique donc une connaissance à priori de la structure et de ses déformées modales. Nous poserons plus tard (Section V) le problème sans cette contrainte.

Les différentes déformées modales permettent alors de faire un lien entre les différents points observés de la structure, offrant une redondance d'observation à l'estimation du filtre du Kalman. On peut

alors définir la fonction d'observation  $H$  comme une matrice de taille  $((p_x + p_y + p_z) \times 3n)$  telle que :

$$H = \begin{pmatrix} \Phi^1 & 0 & 0 & \Phi^2 & 0 & 0 & \dots & \Phi^n & 0 & 0 \end{pmatrix} \quad (4.98)$$

La construction du modèle d'état a été choisie de manière à concentrer les non-linéarités sur la seule fonction d'évolution. Enfin, afin de pouvoir exprimer la fonction dans le formalisme stochastique, on définit le  $\nu$  le bruit de mesure gaussien centré et  $R_\nu$  la covariance associée.

Avec le système d'état établi dans cette section, un modèle de la structure est introduit par la fonction d'observation, tandis que la fonction d'évolution permet d'identifier les amplitudes et fréquences. C'est cette nouvelle méthode que nous appelons *Kinematic-SAMI* pour *Kinematic multiSensors Assimilation Modal Identification*.

## IV Formalisation de la chaîne complète

### IV.A. Méthodes de vision

Comme il a été mentionné plus haut, les méthodes de visions utilisées sont interchangeable. Plusieurs méthodes présentées dans le chapitre 2 peuvent être appliquées ici, cependant nous avons fait le choix arbitraire d'utiliser des méthodes de détection et suivi de points déterministes sur la structure. Ce choix a été fait pour trois raisons :

- Il permet de s'affranchir d'erreurs d'interpolation lors de la projection des points détectés sur l'espace d'observation.
- Il va permettre de gagner en précision en utilisant des marqueurs spécifiques.
- Il va permettre de directement comparer nos performances avec des capteurs monopoints.

Deux méthodes seront utilisées pour de la détection de marqueurs afin de proposer des comparaisons.

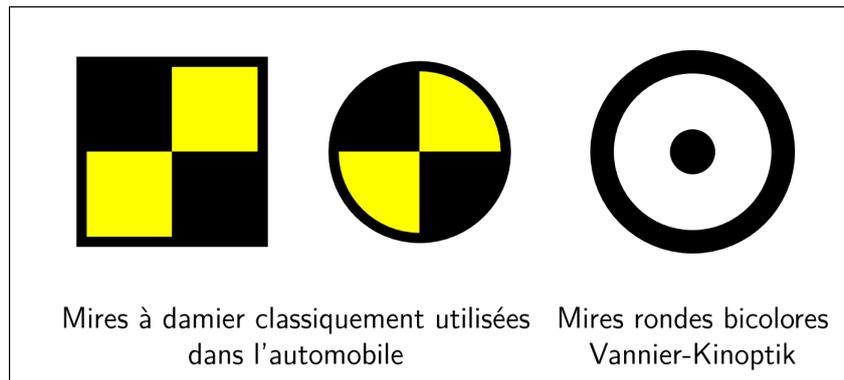
Étant donné la nature 2D des signaux, les précautions nécessaires ont été prises pour garantir que la projection implicite de la caméra sur le plan de son capteur reflète un minimum d'erreur. Cela peut se traduire par plusieurs conditions dans la mise en place expérimentale :

- L'expérience doit se dérouler dans un plan 2D.
- Le plan de l'expérience et celui du capteur doivent idéalement être parallèles.

L'utilisation de marqueurs déterministes revient à faire du suivi de mires. Plusieurs types de mires existent, comme les célèbres damiers de l'automobile (figure 4.3). Le choix de ces mires a été déterminé à l'origine par les dépouillements manuels qui consistaient à faire coïncider un réticule avec le centre apparent du damier. Avec les antécédents de Vannier-Kinoptik sur ce sujet, nous avons choisi des mires rondes bicolores (figure 4.3). Ces mires utilisent l'effet moyenneur d'une ellipse pour gagner en précision sur leurs centres ; elles nous permettent d'avoir une erreur sur la position de  $\pm 0,1$  pixels tandis que les mires classiques à damier permettent d'obtenir une erreur sur la position de  $\pm 0,5$  pixels (Grémillet, 2010).

Nous allons exploiter l'invariant local spatial que ces mires présentent au niveau de l'intensité. Les méthodes présentées ci-après sont basées sur cette invariance et sur le caractère centré des petits mouvements des mires. Notamment, on dégage une hypothèse d'observation :

$$\forall (C_{i,k}, C_{j,k}) \in \mathbb{R}^2 \times \mathbb{R}^2, \quad \|C_{i,k}, C_{i,k+1}\| < \|C_{i,k}, C_{j,k}\| \quad (4.99)$$



**Figure 4.3** – Les fameuses mires à damier connues pour leur utilisation en crash-test et les mires utilisées chez Vannier-Kinoptik

avec  $C_{i,k}$  les coordonnées du centre de la mire  $i$  à l'instant  $k$ .

Les méthodes utilisées ne pouvaient pas être de type extraction de fond comme dans les parties précédentes étant donné l'hypothèse de mouvement. Certains points sont amenés à avoir un déplacement très faible et ce type de méthodes les incorporerait progressivement avec l'arrière plan. On a donc choisi trois autres approches :

**Méthode manuelle et statique : le seuillage en intensité**

La première méthode utilisée est aussi la plus simple à mettre en œuvre, mais également la moins robuste aux changements de luminosité, d'où un paramétrage manuel à l'initialisation. Cette méthode est statique puisque qu'elle n'utilise pas de cohérence temporelle.

Il s'agit d'une première étape de seuillage suivie d'une ouverture. Cette première étape nous permet d'avoir une image binaire de l'ensemble des masques des mires. Les moments statistiques 2D de Hu permettent de les séparer et d'obtenir leurs propriétés. Le centre de chaque mire nous donne leur position tandis que l'aire et l'excentricité de l'ellipse entourant le masque de chaque mire nous permettent d'avoir des informations sur la fiabilité de la mesure. Dans le cas d'un éclairage constant sur une mire sans reflet, ces paramètres sont stables. Grâce à sa simplicité, cette méthode est de loin la plus rapide mais aussi la moins robuste.

**Méthode automatique et dynamique : utilisation de *tracker***

Cette méthode est basée sur les détecteurs de points d'intérêts. Étant donné l'invariant local introduit par les mires, on utilise un détecteur MSER (chapitre 2 sous-sous-section III.B.2.) qui extrait les zones localement invariantes en intensité (Krig, 2014). Elle va permettre d'identifier une série d'ellipses  $\{\mathcal{E}_1, \dots, \mathcal{E}_j, \dots, \mathcal{E}_e\}$  dans l'image décrites par leur centre  $C_j$  et leurs deux demi-axes. La détection de mire se fait alors en calculant les scores paramétrés par le seuil  $C_0^+$  :

$$\forall J \in [1, e] \text{ et } C_0^+ \in \mathbb{R}_+^*, \quad \sigma_J = \sum_{j=1}^e \|C_j, C_J\|_2 < C_0^+ \quad (4.100)$$

Ce test permet de regrouper les ellipses par centre et ainsi déterminer les positions des mires. Des conditions sur l'écart de diamètre entre le plus petit cercle de chaque groupe et le grand cercle permettent de limiter les fausses détections.

Cette méthode étant relativement longue à exécuter (de l'ordre de la seconde), elle semble

moins en adéquation pour une utilisation temps réel. C'est pourquoi nous avons choisi d'utiliser un schéma de suivi basé sur un autre détecteur : celui de Shi-Tomasi. Il détecte les angles dans l'image et le schéma de suivi compare ensuite les angles de l'image courante avec la précédente afin d'établir une correspondance et estimer le mouvement. Cette méthode est robuste à l'éclairage et au bruit de mesure. Cependant, contrairement au détecteur MSER qui est adapté à l'invariance des mires, celui de Shi-Tomasi ne pourra pas déterminer quels sont les points d'intérêts correspondants à une mire. La partie traitement vidéo de la figure 4.5 illustre le fonctionnement de cette méthode. On choisit donc d'initialiser le détecteur Shi-Tomasi à l'aide du détecteur MSER.

Cette méthode est néanmoins problématique au niveau conceptuel. La mire est une zone stable en intensité donc de basse fréquence spatiale et ronde tandis que le détecteur Shi-Tomasi recherche des coins, plus susceptibles d'apparaître sur des zones de haute fréquence spatiale. Même si la méthode fonctionne bien à cause du bruit introduit par le support d'impression et l'acquisition, nous proposons une amélioration de la mire que nous appellerons **mires HBF** (pour Hautes et Basses Fréquences). Afin d'améliorer sa robustesse, du contenu hautes fréquences est ajouté au schéma initial des mires rondes, en ajoutant des motifs gris (voir figure 4.4). Le détecteur MSER sera initialisé sur des images filtrées en passe-bas tandis que le détecteur Shi-Tomasi travaillera sur les images originales.

#### Méthode de Filtrage en Cascade

Dernière méthode utilisable : les filtres en cascade proposés par Viola et Jones (2001). Étant donné que les mires sont déterminées à l'avance, on peut entraîner un filtre afin d'identifier les mires dans les images. Cette méthode est longue sur la phase d'entraînement, mais rapide lors de l'exécution. Un algorithme du Hongrois est ajouté pour permettre une correspondance entre les points. Par contre la précision de ce type de méthode est limitée au pixel, ce qui modère son utilisation comparée aux résolutions sous pixeliques obtenues par ailleurs.

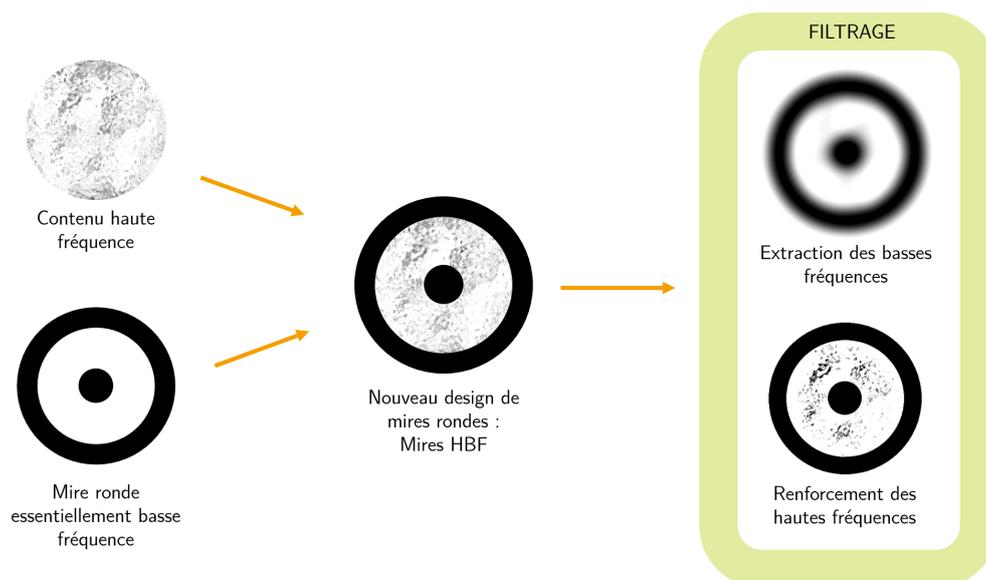


Figure 4.4 – Proposition d'un nouveau design de mires comprenant à la fois des hautes et des basses fréquences : les mires HBF.

## IV. Formalisation de la chaîne complète

Ces trois méthodes nous permettent d'obtenir les coordonnées de points déterminés sur une image dans des conditions expérimentales particulières. Ces points constituent les données entrantes utilisées dans le cadre des filtres de Kalman. Comment utiliser le filtre développé section III de ce chapitre dans ce cadre ? La suite de cette section présente comment employer le Kinematic SAMI en 2D puis en 3D sans passer par une coûteuse méthode de reconstruction.

### IV.B. Formalisation pour un cas 2D

Afin de mieux décrire la structure complète de la chaîne de mesure, nous avons préféré utiliser un schéma afin de mieux comprendre les jonctions entre les différentes parties sur la figure 4.5. Cette figure synthétise l'implémentation que nous avons effectuée. Dans le cadre bleu, le processus d'extraction des informations vidéos permet d'identifier les mires dans l'image grâce à un détecteur MSER et initialise un algorithme de suivi de Shi-Tomasi. Ce bloc fournit en sortie les positions des différentes mires à chaque instant. L'assimilation ensuite utilise un schéma de Kalman non parfumé. L'espace d'état est modal et il nous permet à tout instant une identification décrite dans le bloc du bas. Il est alors possible d'utiliser les modèles pour faire de la prédiction, de l'interpolation et de la compression. Enfin le bloc gris schématise une étape de réalité augmentée. Les parties d'assimilation de données et de traitement vidéo ont été décrites jusqu'ici mais il a été choisi de présenter également dans ce schéma les traitements effectués à partir de l'espace d'état afin d'illustrer l'extraction des différentes informations, puis l'utilisation de ces informations dans le flux vidéo. L'espace d'état est également un espace modal où à une transformation près, les paramètres modaux sont décomposés sur la base modale. La décomposition de ce formalisme nous permet ainsi de réexprimer les différents objectifs du projet FAMOSYS appliqués dans ce cadre.

- Une identification des paramètres modaux de la structure pourra être réalisée.
- À l'aide de ces paramètres modaux identifiés, nous pourrions faire de la prédiction en utilisant la fonction d'évolution pour calculer l'état futur du système. On notera que bien que dans un cas non-linéaire, la fonction d'évolution ne modélise pas d'évolution de la fréquence du système. Ainsi, lorsqu'on calcule un état futur, on se met dans les hypothèses du modèle, c'est-à-dire que la fréquence est constante au cours du temps, ce qui peut s'avérer faux.
- La compression de donnée n'a pas beaucoup d'intérêt dans le cadre de l'analyse modale expérimentale, sauf que lors d'essais en caméra rapide, la quantité de données à stocker est très importante : on pourra la réduire en proposant de reconstituer l'objet d'intérêt à partir des positions de ses mires en fonction du temps.
- Enfin ce schéma 2D permet d'effectuer une interpolation temporelle ou spatiale au sens du modèle mais également une extrapolation spatiale en estimant l'état d'une structure complète observée partiellement.

Les concepts du projet FAMOSYS appliqués au contexte de l'analyse modale expérimentale peuvent être utilisés pour faire de l'interpolation, de l'identification, de la compression de données, et de la prédiction. Cette application vient d'être décrite sur un cas d'application 2D, à l'aide d'un suivi de cibles spécifiques. Différentes méthodes de suivi de cibles ont été développées et restent entièrement interchangeables.

Cet exemple en 2D va permettre de valider la méthode sur différents cas tests. Cependant, il est important de souligner que les applications industrielles sont quasi exclusivement traitées en 3D. Nous utilisons ainsi le même formalisme pour proposer une nouvelle méthode de mesure de vibrations 3D à l'aide des caméras 2D sans reconstruction.

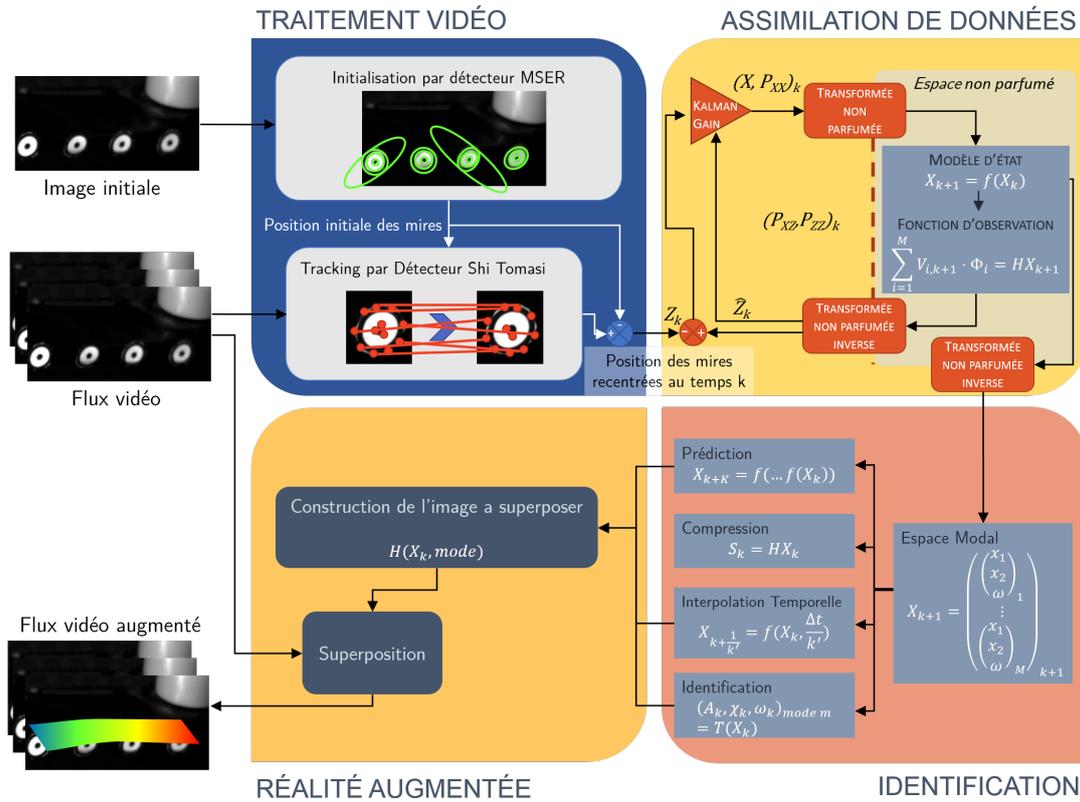


Figure 4.5 – Schéma présentant l'assimilation de données associée à des mesures vidéo pour des scènes 2D

#### IV.C. Formalisation pour un cas 3D

Les mesures par stéréocorrélation sont les mesures les plus précises actuelles en vidéo, mais elles souffrent de plusieurs défauts :

- Un temps de traitement relativement long, en raison de la reconstruction en trois dimensions basée sur la correspondance entre deux images de deux points de vue.
- Une mise en place du dispositif expérimental encore complexe due à la calibration nécessaire aux méthodes de stéréovision.
- Afin de permettre une correspondance entre les deux différentes caméras et afin de minimiser l'erreur, le champ de ces caméras est contraint.

La mesure de champ 3D a donc actuellement un certain prix. Il apparaîtrait judicieux de pouvoir travailler dans des images 2D calibrées, sans jamais effectuer de passage à la 3D tout en faisant correspondre un nombre de données réduites. On effectue un étalonnage des repères caméras non pas entre elles mais vis-à-vis d'un repère lié à la structure.

Nous proposons pour répondre à ces problématiques d'utiliser un filtre de Kalman pour assimiler les différentes données des images 2D sur un modèle 3D. La figure 4.6 synthétise notre approche. On propose alors de définir une série d'espaces d'observations dépendant de chaque caméra  $O_{cam}$  et un espace d'affichage du modèle  $O_{model}$ . Les espaces d'observations utilisés par le filtre sont en deux dimensions tandis que l'espace d'affichage du modèle sera en trois dimensions. Le concept est donc le même que dans la partie précédente : assimiler les informations dans un espace d'état modal

#### IV. Formalisation de la chaîne complète

$\mathcal{M}$  pour ensuite projeter ces informations dans l'espace d'observation. La différence réside dans le nombre d'espaces d'observation : ici, on a plusieurs espaces d'observations reliés par des relations particulières.

On définit l'application de projection  $p_{monde \rightarrow 2D} : \mathbb{R}^4 \rightarrow \mathbb{R}^3$  la matrice permettant de passer d'un espace en coordonnées homogènes en quatre dimensions tel que l'espace monde  $\mathcal{E}_{monde}$  que constitue le laboratoire, au repère de la caméra. On remarque que cette même application s'applique également à l'espace du modèle  $\mathcal{O}_{model}$ .

On précisera d'ailleurs que ces deux espaces coïncident pour un changement de coordonnées qu'on définit par l'application  $g_{R,t} : \mathbb{R}^4 \rightarrow \mathbb{R}^4$  paramétrée par une rotation  $R$  et une translation  $t$ . On utilise le cadre des conditions d'application du modèle pin-hole<sup>4</sup>. Ces relations de passage, que ce soit le changement de coordonnées ou les matrices de projection sont déterminées lors d'une phase d'étalonnage.

On commence par établir une relation d'équivalence entre l'espace monde et l'espace modèle. Par définition, la famille de vecteurs des déformées modales de la structure  $\mathcal{B} = \{\Phi^1, \dots, \Phi^M\}$  est une base de l'espace modèle  $\mathcal{O}_{model}$ . On peut alors exprimer la réduction entre l'espace modèle et l'espace monde par la relation suivante :

$$\forall S_{3D} \in \mathcal{E}_{monde}, \quad \exists (\nu, V^i) \in (\mathbb{R}^4, \mathbb{R}) \mid S_{3D} = g \left( \sum_{i=1}^M V^i \cdot \Phi^i + \nu \right) \quad (4.101)$$

On définit alors les espaces images  $\mathcal{E}_{cam_m}$  comme suit pour une caméra  $cam_m$  par rapport à une déformée observée  $S_{3D}$  à l'aide de l'application de projection  $p_{cam_m}$  :

$$\text{Soit } S_{2D} \in \mathcal{E}_{cam_m}, \quad \forall S_{2D} \in \mathcal{E}_{cam_m}, \quad \exists (p_{cam_m} : \mathbb{R}^4 \rightarrow \mathbb{R}^3, \quad g_{R,t} : \mathbb{R}^4 \rightarrow \mathbb{R}^4) \quad (4.102)$$

$$S_{2D} = p_{cam_m} \circ g_{R_{cam_m}, t_{cam_m}}(S_{3D}) \quad (4.103)$$

De l'autre côté de l'assimilation de données, on obtient un espace d'état modal  $\mathcal{M}$  relié aux espaces d'observation caméras par :

$$\begin{aligned} \forall X \in \mathcal{M}, \quad \exists S_{3Dobs} \in \mathcal{O}_{model} \mid S_{3Dobs} = H(X) \\ \forall X \in \mathcal{M}, \quad \exists S_{2Dobs,m} \in \mathcal{O}_{cam_m} \mid S_{2Dobs,m} = \text{MAP}_{cam_m} [H(X)] \end{aligned} \quad (4.104)$$

avec la fonction d'observation  $H$  utilisée dans l'assimilation composée avec une fonction de conversion  $\text{MAP}_{cam_m}$  définie pour chaque caméra, permettant d'avoir une correspondance entre les points observés sur le modèle 3D et les points dans le plan capteur de la caméra  $cam_m$ .

De cette manière, on observe la structure à l'aide de multiples mesures en 2D et la fonction d'observation nous permet de construire la structure en 3D. Grâce à l'assimilation de données cet assemblage d'espaces nous donne un espace modèle réduit par rapport à l'espace monde. On obtient bien des mesures en 3D tout en utilisant le modèle pour assimiler les acquisitions. La solution présentée ici comporte beaucoup d'avantages :

- On ne travaille que sur des images 2D, minimisant ainsi le surcout de calcul et les erreurs liés à la reconstruction 3D.
- La méthode offre la possibilité de n'avoir qu'une caméra sur chaque zone.

4. De nombreuses références sont disponibles pour expliquer ces notions comme sur le site de Matlab<sup>®</sup> pour l'étalonnage de caméras : <https://fr.mathworks.com/help/vision/ug/camera-calibration.html>

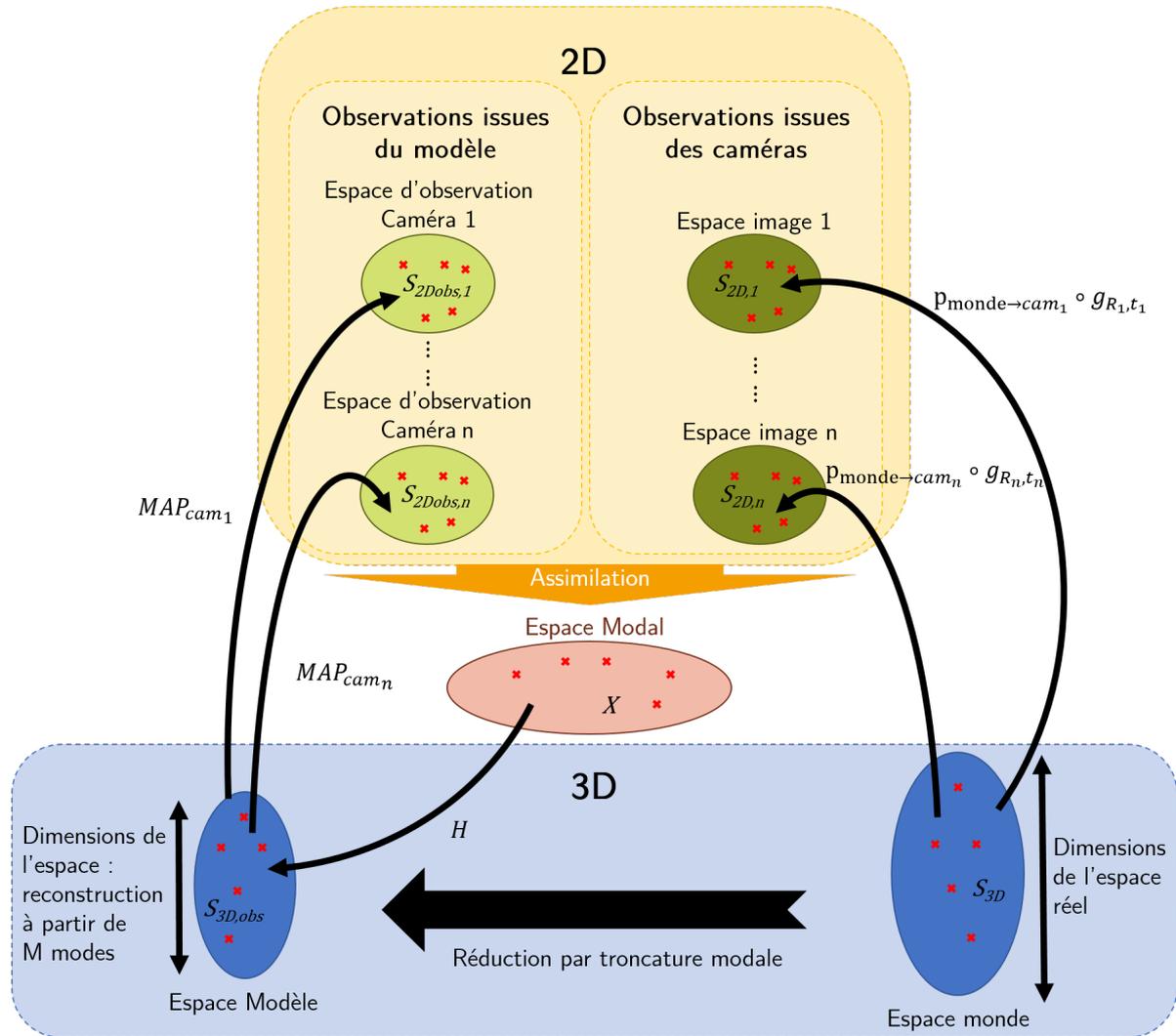


Figure 4.6 – Schéma présentant les espaces présentés pour l'assimilation 3D

- On peut travailler avec des points totalement masqués grâce à la fonction  $MAP_{cam_m}$  en reconstruisant une information 3D à l'aide de points non observés par toutes les caméras, permettant une interpolation spatiale selon le modèle.
- Enfin on peut utiliser plusieurs caméras de caractéristiques différentes.

Il faut cependant bien préciser que la reconstruction 3D n'a de sens que dans une vision modale et les avantages précédemment cités ne sont valables que dans les mêmes conditions d'application de toute cette section, c'est-à-dire que les caméras doivent être placées en connaissance des mouvements observés dans la scène de telle sorte que les informations perdues par la projection faite par chaque caméra soient négligeables par rapport à l'ordre de  $\nu$ , le résidu de la réduction.

#### IV.D. Fusion d'informations : utilisation de capteurs complémentaires

Bien que centrés sur une utilisation exclusive de mesures vidéos, ces travaux ne se veulent pas pour autant complaisants envers les caméras rapides numériques. Aussi, les caméras numériques comme

## IV. Formalisation de la chaîne complète

les autres capteurs vus dans la sous-sous-section I.B.2. ont chacun leur points forts, et leurs points faibles. Il serait judicieux de profiter des avantages de chaque technologie.

L'assimilation de données a été développée dans des domaines comme la météorologie ou l'océanographie et ces méthodes sont utilisées depuis longtemps pour agglomérer des données issues de différents types de capteurs. Nous proposons donc également de coupler plusieurs capteurs de technologies différentes afin d'utiliser les propriétés du filtre de Kalman pour fusionner ces informations.

L'intérêt en 3D est immédiat, puisque l'ajout de capteurs permettrait de fiabiliser certains points de mesures moins bien définis ou gagner en précision spatiale et temporelle en utilisant des capteurs avec une fréquence d'échantillonnage plus élevée.

Le sujet ne sera pas approfondi dans ce manuscrit, et on pourra se référer à Bloch (1996) pour avoir un aperçu des opérateurs disponibles pour effectuer une fusion de données, ou Castanedo (2013) pour une avoir une vision plus globale des problématiques de ce domaine. On proposera tout de même ici un exemple dans la partie VI.B.4..

### IV.E. Fonctionnement en prédiction, en lissage, en interpolateur (cf partie 3)

Le schéma de filtrage proposé dans cette partie introduit un modèle spatial et temporel du système. Ainsi, une fois que le filtre a recalé le modèle par rapport aux observations, il est possible d'exploiter au maximum le modèle d'état.

#### Temporellement

La fonction d'évolution permet de modéliser le système à un temps futur, moyennant l'erreur qu'il introduit. Il nous est donc possible de prédire le comportement du système, ce qui peut être utile en cas de détection précoce de défauts. Plus intéressant encore dans notre cas, il nous permet d'augmenter virtuellement la fréquence d'échantillonnage en considérant les observations comme événementielles. Ainsi, au sens du modèle, une fréquence d'échantillonnage virtuelle pourra être choisie comme  $F'_{ech} \propto F_{ech}$  si on souhaite repasser par des mesures. Elle pourra être choisie quelconque sinon. Ce schéma est évidemment à utiliser avec la plus grande précaution puisqu'il ne peut qu'interpoler temporellement les comportements observables à la fréquence initiale. Il sera impossible d'observer un mode absent dans l'observation mais par contre on pourra faire ressortir des modes très évanescents. Cette interpolation au sens du modèle est directement limitée par l'erreur de celui-ci : une interpolation sur un nombre de pas de temps trop grand fera diverger le filtre. On peut éventuellement augmenter l'erreur sur le modèle pour gagner sur l'intervalle d'interpolation, ce qui a également pour inconvénient de détériorer la précision.

#### Spatialement

Dans des conditions de fonctionnement normales du filtre, ce modèle spatial vient enrichir l'observation discrète. Il permet donc de faire une interpolation spatiale au sens des déformées modales analytiques ou de fonctionner avec des observations partielles. L'utilisation d'un modèle spatial permet en outre d'effectuer un lissage des acquisitions.

Une application de ces fonctionnements en extrapolation sera présentée dans la sous-sous-section VI.B.4..

Malgré les extrapolations qu'il offre, l'utilisation d'un modèle spatial est cependant une source d'approximation du comportement réel de la structure. La méthode d'identification filtre les observations par rapport aux déformées modales analytiques. Par exemple, une non-linéarité ponctuelle comportementale sur la structure n'existant pas sur le modèle, celle-ci sera intégrée dans le résidu alors

qu'il serait judicieux de l'intégrer dans les modes propres de la structure. C'est l'objet de la partie suivante dans laquelle nous proposons une version du filtre identifiant également la base modale.

## V Schéma d'assimilation développé sans connaissances sur la structure : Blind multiSensors Assimilation Modal Identification (B-SAMI)

Le développement du Kinematic SAMI proposé dans la partie précédente s'appuie sur une hypothèse forte de l'analyse modale expérimentale : celle de l'invariance temporelle d'un mode. Cependant ce schéma possède un inconvénient majeur : celui d'utiliser des modes théoriques de la structure précalculés par simulation ou mesurés par d'autres méthodes.

Dans la plupart des mesures expérimentales actuelles, des études numériques amont sont menées et fournissent ces modes. Cependant, des hypothèses simplificatrices sont faites au niveau des amortissements et de la structure afin de rester autour d'un domaine linéaire acceptable pour pouvoir découpler les modes.

La méthode présentée dans la section III a l'avantage de pouvoir suivre des variations de la fréquence de résonance, ainsi que du coefficient d'amortissement, en temps réel. La précision du K-SAMI dépendra donc de l'exactitude des modes théoriques par rapport à la structure réelle. Pour fixer les idées, prenons l'exemple courant d'une non-linéarité comportementale. Cette dernière sera interprétée par le filtre comme une erreur de mesure et donc causera dans le meilleur des cas une distorsion si elle n'est pas comprise dans l'écart-type du bruit de mesure. Le cas échéant, elle sera absorbée dans le bruit, augmentant de fait la valeur du résidu.

Pour le Blind-SAMI, nous proposons de reprendre la même fonction d'évolution que précédemment, en l'utilisant sans connaissances sur la structure afin d'éviter cette approximation implicite des modes expérimentaux de la structure sur les modes propres théoriques dans l'assimilation.

Dans ce cas, il faut repenser les espaces d'état et d'observation. En effet, l'espace d'état correspondait à un espace modal dans le paragraphe précédent parce que les modes propres de la structure permettaient d'avoir une base de l'espace d'observation. On choisit de ne pas avoir d'a priori sur les modes de la structure : l'espace d'observation se décompose désormais sur la base des modes propres expérimentaux que l'on cherche à déterminer également. Il est donc nécessaire de modifier le vecteur d'état de manière à faire figurer la base des déformées modales en plus des fréquences. De manière à conserver le formalisme de la transformée de Hilbert on considère donc pour chaque point son signal analytique et pour chaque mode sa fréquence propre.

La fonction est alors construite par blocs de manière à faire évoluer chaque point indépendamment des autres et ce pour chaque mode. En reprenant les notations de la partie précédente, on aurait pour un point  $pt$  appartenant au mode  $i$  à l'instant  $k$  :

$$\begin{pmatrix} x_{k+1}^{1,i} \\ x_{k+1}^{2,i} \end{pmatrix}_{pt} = \begin{bmatrix} \cos(2\pi f_k^i \Delta t) & -\sin(2\pi f_k^i \Delta t) \\ \sin(2\pi f_k^i \Delta t) & \cos(2\pi f_k^i \Delta t) \end{bmatrix} \begin{pmatrix} x_k^{1,i} \\ x_k^{2,i} \end{pmatrix}_{pt} \quad (4.105)$$

En construisant alors le vecteur d'état, on pose  $x_k^{3,i} = 2\pi f_k^i \Delta t$  et on définit la fonction d'évolution

## V. Schéma d'assimilation développé sans connaissances sur la structure : Blind multiSensors Assimilation Modal Identification (B-SAMI)

pour un mode  $i$  de la structure pour les  $N$  points de la structure :

$$\begin{pmatrix} \begin{pmatrix} x_{k+1}^{1,i} \\ x_{k+1}^{2,i} \\ \vdots \\ x_{k+1}^{1,i} \\ x_{k+1}^{2,i} \\ \vdots \\ x_{k+1}^{N,i} \end{pmatrix} \\ x_{k+1}^{3,i} \end{pmatrix}_i = \begin{pmatrix} \begin{bmatrix} \cos(x_k^{3,i}) & -\sin(x_k^{3,i}) \\ \sin(x_k^{3,i}) & \cos(x_k^{3,i}) \end{bmatrix} & \dots & 0 \\ 0 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \begin{bmatrix} \cos(x_k^{3,i}) & -\sin(x_k^{3,i}) \\ \sin(x_k^{3,i}) & \cos(x_k^{3,i}) \end{bmatrix} \\ 0 & \dots & 0 \end{pmatrix}_i \cdot \begin{pmatrix} \begin{pmatrix} x_k^{1,i} \\ x_k^{2,i} \\ \vdots \\ x_k^{1,i} \\ x_k^{2,i} \\ \vdots \\ x_k^{N,i} \end{pmatrix} \\ x_k^{3,i} \end{pmatrix}_i \quad (4.106)$$

Il reste encore à concaténer l'ensemble des modes souhaités. On remarquera d'ailleurs que rien n'oblige à prendre les modes des structures ni dans l'ordre ni successifs comme proposé dans ce manuscrit.

Dans un but de simplification de l'écriture et d'amélioration de la précision, nous avons choisi d'utiliser pour cette version un filtre de Kalman non parfumé (*UKF*). Cela nous évite une formalisation des fonctions sous forme matricielle ainsi que les calculs de Jacobiennes.

En posant le nouvel espace d'état  $\mathcal{M}_B$  on définit alors la fonction d'observation comme une application de conversion telle que :

$$\forall X \in \mathcal{M}_B, \quad h : \mathbb{R}^{M \times 2(N+1)} \rightarrow \mathbb{R}^N \mid h(X) = \sum_{i=1}^M \begin{pmatrix} \begin{pmatrix} x_k^{1,i} \\ x_k^{2,i} \\ \vdots \\ x_k^{N,i} \end{pmatrix} \end{pmatrix}_i \quad (4.107)$$

On utilisera également l'application suivante pour extraire non pas les déformées modales, mais leur valeur absolue :

$$\forall X \in \mathcal{M}_B, \quad h_D : \mathbb{R}^{M \times 2(N+1)} \rightarrow \mathbb{R}^{M \times N} \mid h_D(X) = \left( \begin{pmatrix} \left( \|x_k^{1,i} + j \cdot x_k^{2,i}\| \right)_1 \\ \left( \|x_k^{1,i} + j \cdot x_k^{2,i}\| \right)_2 \\ \vdots \\ \left( \|x_k^{1,i} + j \cdot x_k^{2,i}\| \right)_N \end{pmatrix}_1, \dots, \begin{pmatrix} \left( \|x_k^{1,i} + j \cdot x_k^{2,i}\| \right)_1 \\ \left( \|x_k^{1,i} + j \cdot x_k^{2,i}\| \right)_2 \\ \vdots \\ \left( \|x_k^{1,i} + j \cdot x_k^{2,i}\| \right)_N \end{pmatrix}_M \right) \quad (4.108)$$

En contrepartie cette formulation est beaucoup plus lourde à traiter, à cause de la taille du vecteur d'état. Sur un exemple classique de structure à 25 points avec une identification sur 5 modes, le vecteur d'état du système comporterait 255 variables.

Nous avons proposé un schéma du filtre de Kalman avec le modèle de Hilbert afin de pouvoir identifier simultanément les fréquences et les déformées modales dans un schéma temps réel. Ce modèle sera validé et appliqué sur un cas test réel dans la sous-section VI.C..

## VI Application au cas d'une poutre encastrée libre

Afin de valider les schémas d'assimilation présentés dans la section précédente, nous avons choisi le cas académique de la poutre encastrée libre. Nous détaillerons dans une première partie notre cadre expérimental pour les simulations et pour les cas tests réels, puis nous testerons successivement le K-SAMI et le B-SAMI.

### VI.A. Cadre expérimental

#### VI.A.1. Considérations sur la notion de schéma temps réel

Comme dans l'ensemble de ces travaux, l'un des objectifs est de travailler dans une logique séquentielle afin de pouvoir mettre en place un schéma temps réel de traitement dans la chaîne de mesure. En ce sens, c'est le cas pour les déclinaisons présentées dans ce chapitre. Cependant il est important de préciser que pour plusieurs raisons notamment technologiques, le but de ces travaux n'était pas d'atteindre le temps réel :

- Le code utilisé dans la suite a été conçu sous MatLab® et n'est pas optimisé outre mesure. Le temps d'exécution pourra être fortement réduit par le passage dans un langage plus industriel. En l'occurrence, la partie la plus lente de la chaîne étant la partie vision, l'utilisation de la librairie OpenCV spécialement optimisée pour processeurs Intel® et utilisant les capacités de traitements parallèles des cartes graphiques NVIDIA® supportant CUDA permet de réduire considérablement les temps d'exécutions en déchargeant les processeurs principaux.
- Dans la plupart des caméras haute vitesse et plus particulièrement pour la totalité d'entre elles au-dessus d'un certain niveau de performance, le fonctionnement de la caméra se fait en deux temps (chapitre 1) : une étape d'enregistrement et une seconde étape d'export de la vidéo. Néanmoins l'évolution du marché fortement influencée par l'industrie 4.0 accroît ses efforts pour le développement de technologies d'imagerie permettant de faire du flux à haute vitesse et à haute résolution. L'évolution technologique du matériel va dans le sens de l'utilisation des concepts développés dans ces travaux.

On utilisera dans cette partie expérimentale une caméra moyenne vitesse avec une résolution étendue (voir tableau 4.4). La propriété intéressante de diminuer le nombre de pixels verticaux pour augmenter la vitesse sera exploitée pour atteindre les fréquences nécessaires.

**Table 4.4** – Propriétés de la caméra utilisée

Capteur	CMOS
Vitesse en pleine résolution	181 Hz
Résolution	4096x3072
Vitesse maximum	3000 Hz

## VI. Application au cas d'une poutre encastrée libre

### VI.A.2. Cas tests présentés

Le cas test de la poutre encastrée libre sera d'abord décliné en simulation, permettant de tester les limites des SAMI. Une validation expérimentale permettra ensuite de valider l'ensemble de la chaîne depuis la caméra jusqu'à la synthèse vidéo augmentée.

Ce cas test présente l'avantage de montrer l'intérêt d'une mesure sans contact. On travaillera dans un plan 2D de manière à considérer uniquement les modes de flexion. On considère les hypothèses de faibles amplitudes, ce qui nous permet d'utiliser les hypothèses Euler-Bernoulli. Cependant, ce cadre n'exclut pas des effets non-linéaires sur les paramètres modaux (Cumunel, 2008). On va donc tester les différentes versions des schémas SAMI décrites précédemment sur deux cas tests :

1. Une simulation non-linéaire d'une poutre encastrée-libre comportant trois modes afin de valider les filtres sur des variations de fréquence.
2. Un cas test réel sur une règle encastrée-libre excitée au choc proche de l'encastrement afin d'énergiser un maximum de modes.

Enfin on utilisera dans la dernière partie un cas test réel de la spatule avant d'un ski encastré à son milieu. Ce cas test plus complexe sera utilisé comme exemple d'application pour la magnification développée dans la section VII.

### VI.A.3. Caractéristiques et montage expérimental

Cette section décrit les différentes caractéristiques de notre cas test dans ses deux déclinaisons : numérique et réel.

#### a. Simulation d'une poutre encastrée-libre

La simulation doit nous permettre de valider le modèle, c'est pourquoi nous avons choisi de la rendre moins réaliste au niveau des non-linéarités, de manière à se placer dans des cas défavorables, bien qu'éloignés du comportement réel de la structure. Parmi les trois modes simulés, deux sont choisis assouplissants et un rigidifiant<sup>5</sup>. Les valeurs des trois premiers modes de la structure seront simulées selon les caractéristiques présentées dans le tableau 4.5. Afin de simuler la mesure, on ajoute un bruit gaussien et centré de covariance  $R_{sim} = 200 \times 10^{-6}$  m (soit 10 % de l'amplitude du troisième mode). Cette gaussianité est une hypothèse sujette à controverse, notamment parce que le bruit du capteur a plutôt une tendance poissonnienne. Cependant nous supposons que les autres incertitudes de la chaîne de mesure majorent les bruits non gaussiens. On vérifiera cette hypothèse sur les cas tests réels. La fréquence de la simulation a été choisie de 200 Hz et la simulation se fera sur 25 points régulièrement espacés le long de la poutre.

**Table 4.5** – Fréquences théoriques de la poutre encastrée libre simulée

Mode	1	2	3	4	5	6	7	8	9	10
Fréquence (Hz)	1,676	10,51	29,41	57,64	95,29	142,3	198,8	264,7	340,0	424,7

5. La simulation de ces modes est obtenue en utilisant une évolution de la fréquence en  $1 \pm \frac{1}{\tau}$ .

### b. Poutre encastrée-libre réelle

Afin de pouvoir avoir des amplitudes conséquentes, la poutre choisie est un réglet standard d'aluminium de 50 cm avec une section angulaire de 1 mm encastré dans un bloc d'acier sur 5 cm. Rapidement, on peut constater que la déformée  $S(x, t)$  à l'instant  $t$  pour l'abscisse  $x$  peut s'exprimer dans l'espace monde sur la base de ses déformées modales  $\Phi^i(x)$  comme indiqué dans l'équation (4.109) :

$$\forall (x, t) \in [0, L] \times \mathbb{R}_+, S(x, t) = \sum_{i=1}^{\infty} V^i(t) \Phi^i(x) \quad (4.109)$$

Cette structure classique respecte bien nos hypothèses d'applications. Les propriétés mécaniques du réglet ont été résumées dans le tableau 4.6.

Il est important de souligner qu'étant donné l'écart entre les modules d'Young de l'aluminium et de l'acier, l'encastrement n'est pas autant idéal que le suppose le modèle d'Euler-Bernoulli. L'hypothèse de déplacements négligeables près de l'origine pourra être observée dans l'analyse des résultats.

La structure étant très fine, nous avons préféré utiliser pour cette structure des marqueurs physiques mats dont l'intensité sur l'image se situe au milieu de l'échelle de gris. On utilisera donc uniquement une technique de seuillage pour cette structure qui nous permettra d'obtenir la position à chaque instant de 45 points espacés d'un centimètre sur le réglet. Après calibration, on obtient une taille de pixel projetée sur la structure (appelée *Ground Sample Distance*) de  $108 \times 10^{-6}$  m. La caméra a été paramétrée pour une fenêtre de  $4096 \times 512$  pixels avec une fréquence de 350 Hz. Par contre, les séquences utilisées pour comparer les méthodes de vision, mais également la fusion ont été effectuées sur cette même structure, cette fois munie des mires décrites dans l'annexe IV.A.. Ces mires de 0,5 cm ont été disposées de la même manière que pour les marqueurs physiques mats. Les mires pourront ainsi être traitées à la fois par les algorithmes de suivi introduits précédemment, mais également par des techniques de seuillage.

**Table 4.6** – Propriétés de la poutre encastrée libre

Longueur (m)	0,450
Hauteur (m)	0,020
Largeur (m)	0,0012
Densité ( $\text{kg m}^{-3}$ )	2700
Module d'Young (GPa)	74

### c. Montage utilisé

Le montage expérimental est composé de trois éléments :

- une caméra haute vitesse ;
- la structure ;
- l'éclairage.

Malgré un choix réfléchi des cas tests, les structures observées ont toutes le mauvais goût d'être spéculaires, à cause de la présence de métal dans le champ. L'utilisation d'un éclairage rasant ne permettant pas pour autant de collecter assez de lumière pour avoir des résultats corrects. Il a fallu jouer sur la préparation des structures pour éviter ces réflexions autant locales que parasites.

Une photographie du montage est présentée figure 4.7 pour le réglet.

En accord avec les suppositions faites précédemment, on place la caméra face au réglet et l'axe optique normal au plan des déplacements. De plus, on se place dans de bonnes conditions d'éclairage, malgré les remarques sur les reflets spéculaires. Un exemple d'acquisition est donné dans la figure 4.8.

## VI. Application au cas d'une poutre encastrée libre

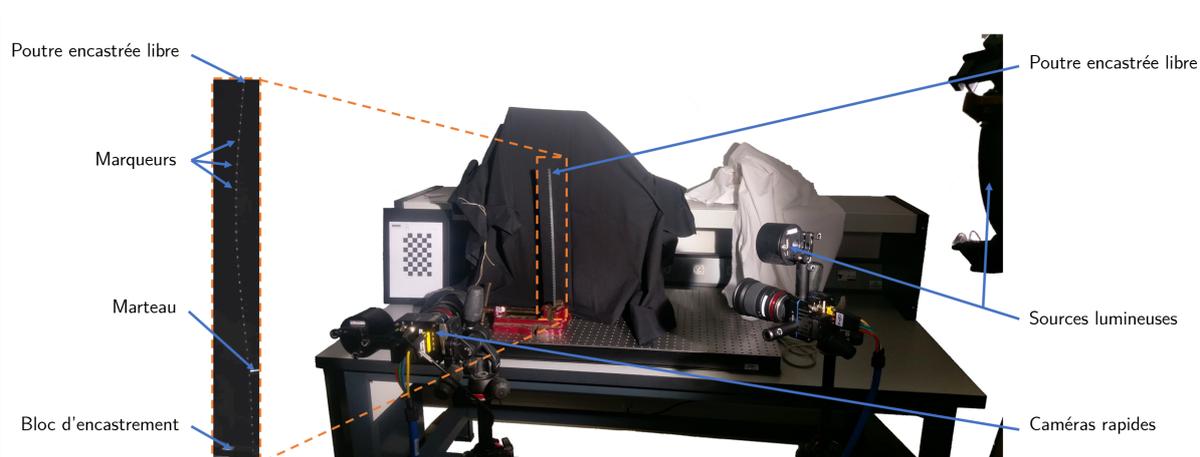


Figure 4.7 – Photographie du montage expérimental pour le réglage



Figure 4.8 – Exemple d'acquisition de la poutre sollicitée au choc. À noter que les points rouges au centre de chaque mire sont issus du suivi par seuillage

### d. Outils utilisés pour l'analyse des résultats

Afin de pouvoir évaluer les performances de la méthode, on utilise des indicateurs permettant de quantifier des écarts.

- Un indicateur d'écart classique RMSE (la racine carrée de l'erreur quadratique moyenne) sera utilisé pour estimer la norme du résidu. Bien que dans les simulations nous pourrions avoir les modes exacts, il nous sera difficile de comparer mode par mode sur les cas réels. Cette mesure de RMSE est très largement utilisée dans le domaine de l'assimilation de données ou de l'estimation stochastique.
- Étant donné la réduction effectuée par le modèle (équation (4.51)), le résidu contient des modes. Nous proposons d'étudier la forme spatiale du résidu comme un indicateur.

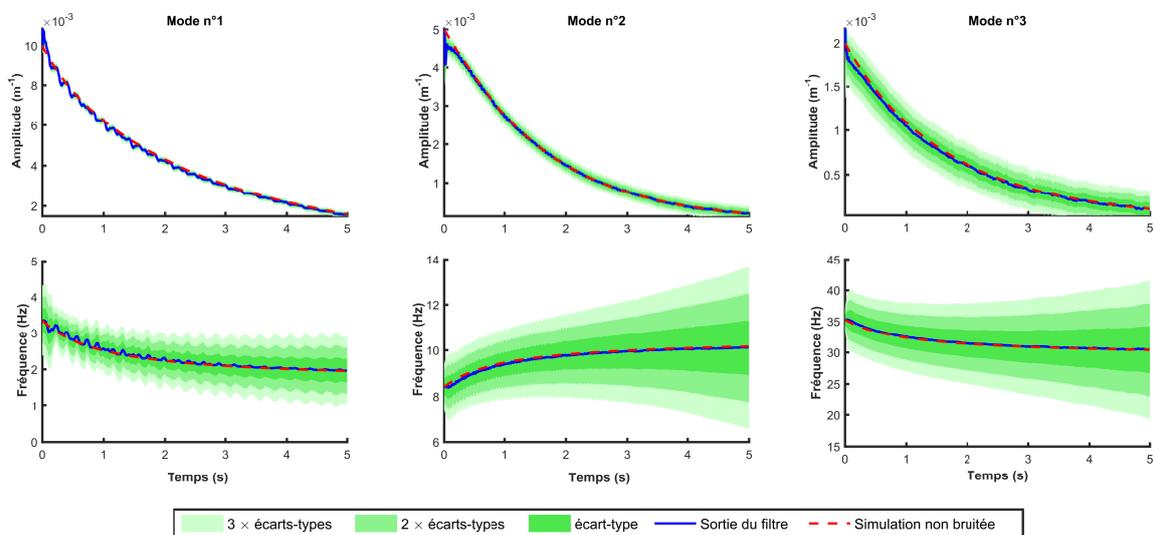
## VI.B. Validation du schéma Kinematic-SAMI

Les résultats générés par les cas tests décrits dans la partie précédente vont être analysés pour chaque méthode. Dans cette partie, on n'utilisera que le K-SAMI.

### VI.B.1. Simulations

Les résultats sont présentés sur la figure 4.9. Toutes les informations disponibles dans l'espace d'état de l'assimilation de données sont condensées dans cette figure. Les informations cachées en pointillés rouges sont superposées aux sorties du filtre en bleu. Le signal entrant n'a pas été affiché. Afin de pouvoir comprendre le fonctionnement du filtre, les covariances ont été affichées également en dégradés de verts selon la largeur choisie.

Le filtre arrive à suivre l'évolution de la fréquence en fonction du temps pour des comportements rigidifiants et assouplissants selon les modes, ce qui était le comportement attendu. L'amplitude est également correctement suivie. L'augmentation de la covariance résulte d'une information modale moins présente dans le signal : avec l'atténuation, les signaux sont de plus en plus proches du bruit, ce qui amène le filtre à augmenter le niveau de covariance. Cette impression est très influencée par l'échelle de courbe. L'identification semble osciller autour de la valeur cachée, toujours en respectant sa plage de covariance. En regardant l'évolution de l'erreur au cours du temps sur la figure 4.10a on constate que malgré les faibles erreurs sur l'amortissement et la fréquence, ces oscillations parasites sont toujours présentes, d'autant plus fortement que le mode est faible. Ces oscillations sont dues à la structure du filtre et aux approximations faites sur la transformée de Hilbert. On valide cette méthode qui, malgré un niveau de bruit élevé sur l'entrée, arrive à identifier les différents paramètres modaux sur une séparation modale figure 4.10b.



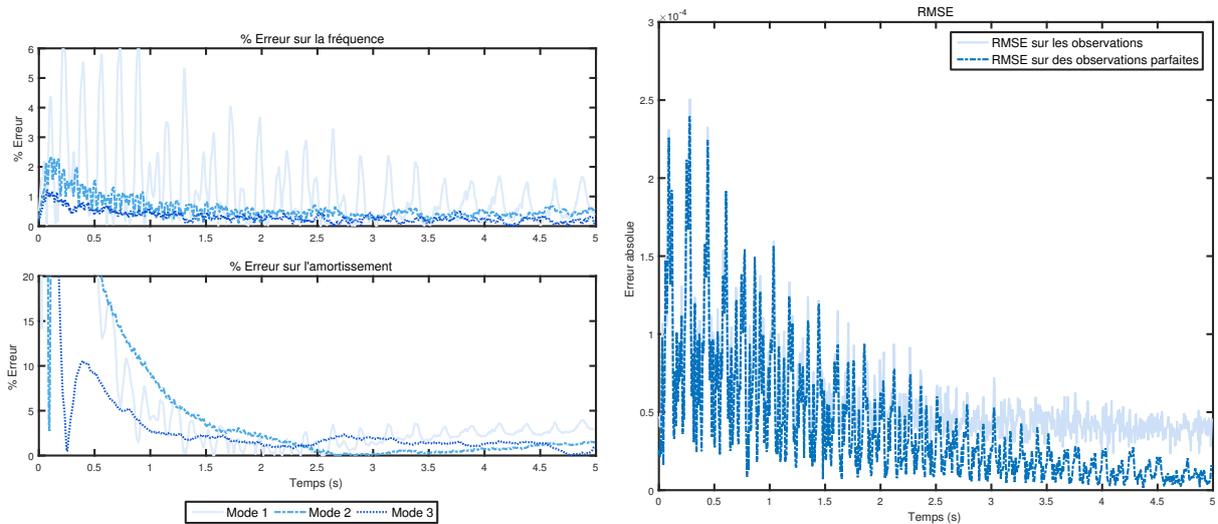
**Figure 4.9** – Résultats de l'identification en fréquence et amplitude pour la simulation non-linéaire d'une poutre encastree-libre

**Remarque : analyse des erreurs d'identification**

Afin de comprendre d'où viennent les variations de l'identification, une analyse spectrale de l'amplitude identifiée des différents modes est proposée. Pour plus de clarté, on choisit une simulation sans variation de fréquence avec un bruit plus faible de covariance  $R_{bruit} = 2 \times 10^{-5} m$ . On choisit les fréquences telles que :  $f_1 = 1,67$  Hz,  $f_2 = 6,67$  Hz et  $f_3 = 26,67$  Hz, et on propose d'effectuer une analyse spectrale sur l'amplitude identifiée du premier mode. Très facilement, on identifie les fréquences sur le gain de la FRF de l'amplitude identifiée normalisée par rapport à son niveau initial (figure 4.11).

On observe bien les effets décrits par Pachaud *et al.* (2013) sur la transformée de Hilbert d'un signal harmonique. Dans notre cas, nous avons modulation des fondamentales et apparition des harmoniques de rang 1. Cette analyse est primordiale puisque les distorsions s'avérant entièrement déterminées, il sera possible de proposer ultérieurement une version du filtre corrigeant ces effets.

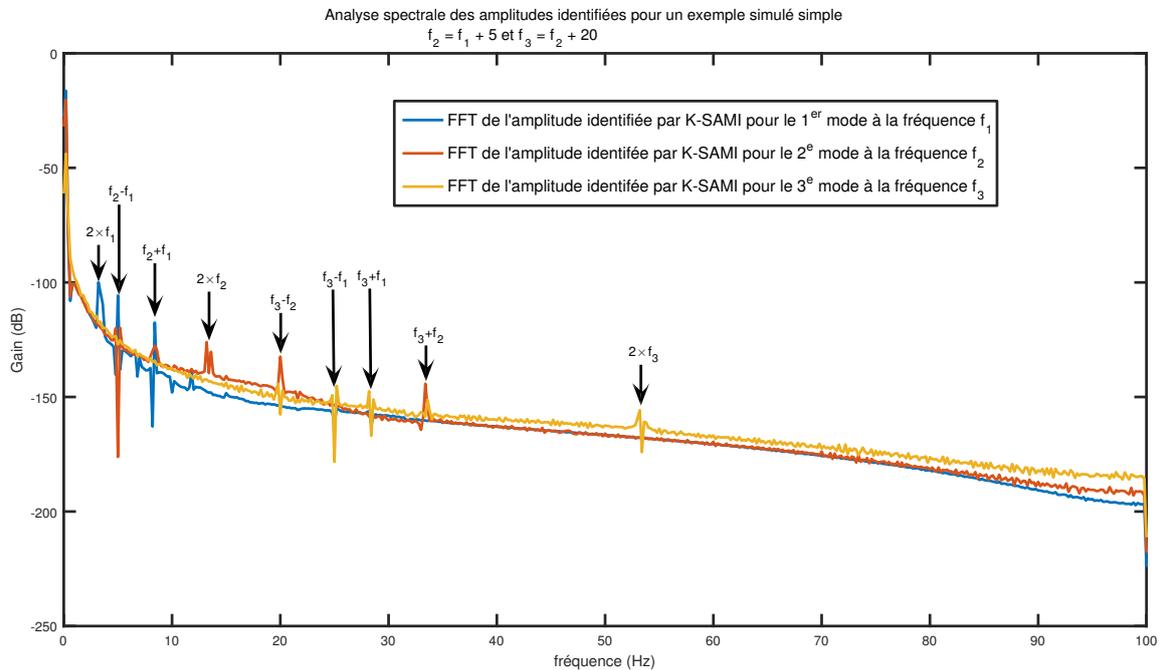
## VI. Application au cas d'une poutre encastrée libre



(a) Résultats de l'identification en fréquence et en amplitude pour la simulation non-linéaire d'une poutre encastrée-libre : erreur sur les différents paramètres identifiés en fonction du temps

(b) RMSE pour l'identification modale sur un cas test simulé non-linéaire : comparaison au cas non-bruité et au cas bruité

**Figure 4.10** – Résultats de l'identification en fréquence et en amplitude pour la simulation non-linéaire d'une poutre encastrée-libre : RMSE



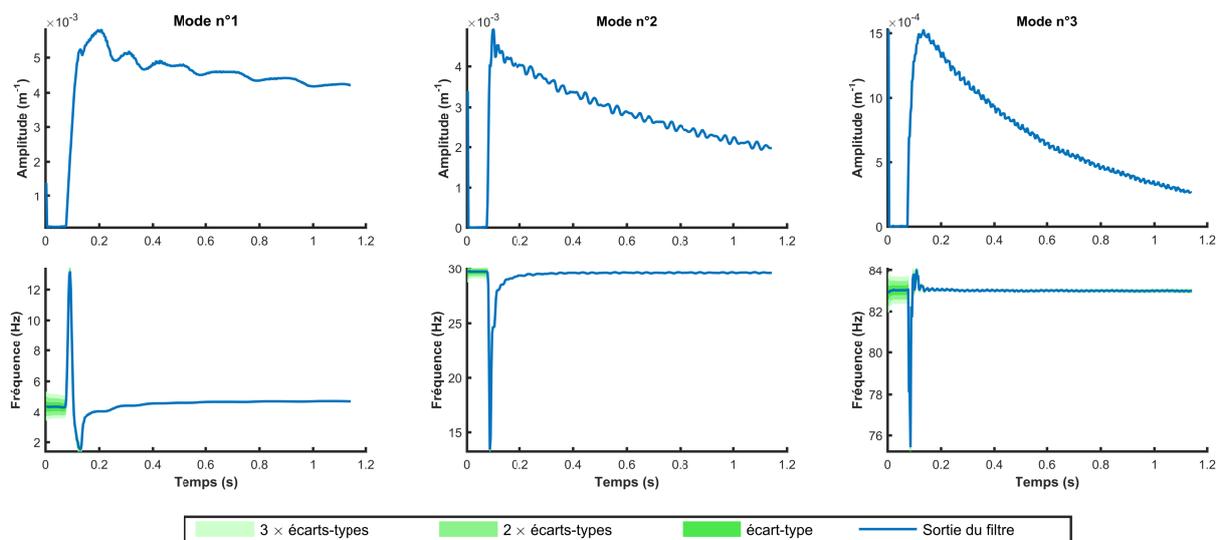
**Figure 4.11** – Analyse Spectrale des amplitudes identifiées des 3 différents modes

## VI.B.2. Cas réel

Dans le cadre d'un essai réel, le caractère expérimental nous impose de présenter les résultats différemment. Afin de vérifier les valeurs des fréquences dans le signal, on effectue une analyse spectrale sur le déplacement de chaque point d'acquisition dont les résultats sont présentés sur le tableau 4.7.

Les valeurs de pic et les largeurs à mi-hauteur sont données afin de quantifier l'étalement fréquentiel dû à l'amortissement ou à un comportement non-linéaire. La résolution de cette transformée de Fourier vaut 0,87 Hz. Elle est plutôt faible mais difficilement améliorable, à cause de la nature transitoire du signal. Nos résultats concordent et les écart-types donnés par le filtre semblent plus précis que la résolution de l'analyse de Fourier. Soulignons que K-SAMI parvient à identifier avec précision le quatrième mode, pourtant à un niveau d'amplitude proche du bruit.

La figure 4.13 présente la FFT du signal en amplitude afin d'avoir les valeurs de chaque contribution. Les résultats de l'identification de 3 modes présentés sur la figure 4.12 montrent que le filtre réalise une bonne identification et qu'il converge rapidement. Ce cas test réel donne également un aperçu de la robustesse du filtre puisque ce-dernier parvient à faire l'identification même lorsqu'il est initialisé avant le choc.



**Figure 4.12** – Résultats de l'identification pour le cas test réel d'une poutre encastrée-libre

En comparant le RMSE de cet essai (figure 4.15a) avec celui du cas simulé, on comprendra que le cas simulé était choisi volontairement difficile à identifier, que ce soit au niveau des variations de paramètres ou du bruit de simulation. Le RMSE de l'essai converge vers une valeur de l'ordre de  $1 \times 10^{-5}$  m, qui est un ordre de grandeur en-dessous de la taille du pixel. Ce résultat est issu à la fois des propriétés des filtres de Kalman qui lui permettent de travailler au niveau du bruit, mais également grâce à la méthode de traitement vidéo qui permet de descendre à des précisions sous pixeliques.

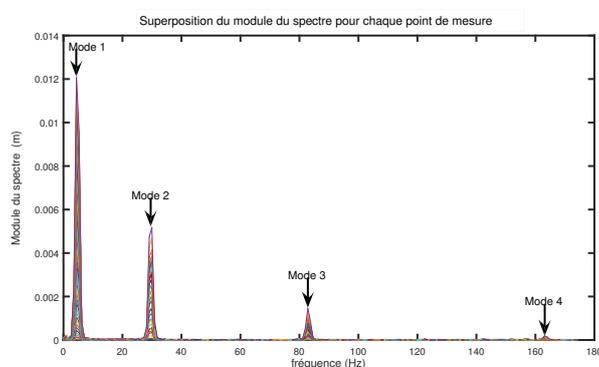
Le filtre de Kalman repose sur une hypothèse forte de gaussianité, que nous proposons de vérifier à travers une analyse de résidus.

## VI.B.3. Analyse de Résidus

Afin de mieux comprendre le fonctionnement du filtre, nous proposons d'observer l'évolution du résidu du filtre en fonction des modes identifiés. Le résidu contient la valeur absolue de la somme des modes non identifiés (équation (4.51)) et il est facile d'afficher le résidu sous une forme spatiale pour faire ressortir les modes qu'il contient.

Les résultats présentés figure 4.16 pour une identification de 2 modes dessinent une forme spatiale du résidu correspondant à la valeur absolue du troisième mode. L'expérience est concluante puisque lorsqu'on identifie 3 modes, la forme du résidu change et prend plutôt la forme du quatrième mode. Les résultats sont moins évidents pour le quatrième mode qui se trouve beaucoup plus proche du niveau de bruit. L'analyse spectrale nous permet d'affirmer que le résidu restant après le quatrième mode n'est constitué que de bruit et de fréquences parasites. On notera que le modèle est spatialement moins juste aux deux extrémités, ce qui reste logique, étant donné que ce sont les endroits où les hypothèses sont le moins respectées : l'encastrement parfait d'un côté et l'hypothèse d'Euler-Bernoulli de faible non-linéarité de l'autre due aux amplitudes plus importantes du bout de la règle.

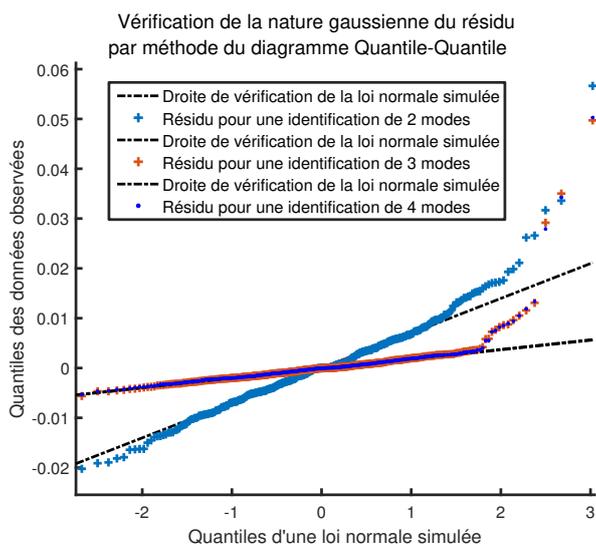
Pour compléter cette analyse, nous proposons de vérifier la nature gaussienne du résidu en utilisant un diagramme quartile-quartile. La figure 4.14 montre les résultats respectivement pour deux modes identifiés et trois modes identifiés. La nature gaussienne semble acceptable sur ces courbes puisque le résidu observé suit la même pente qu'une loi normale simulée. La fonction de densité de probabilité des erreurs centrées en fonction des modes identifiés (voir figure 4.15b) montre également une forme gaussienne avec une variance qui diminue avec le nombre de modes identifiés qui augmente.



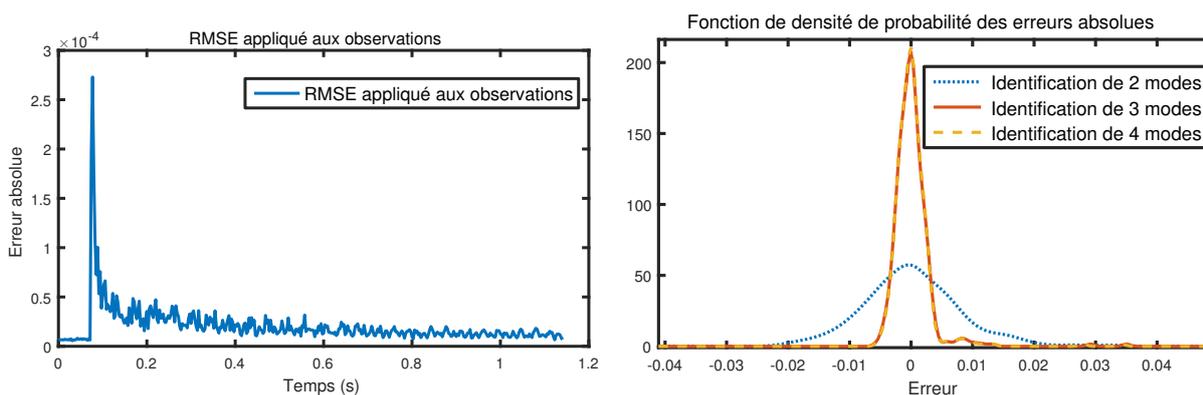
**Figure 4.13** – Gain de l'analyse de Fourier pour chaque point de mesure sur le cas test réel d'une poutre encastrée libre

**Table 4.7** – Analyse modale : comparaison entre une analyse de Fourier et notre méthode

Mode n°	Transformée de Fourier (Hz)		Notre méthode (Hz)	
	Pic	Largeur à mi-hauteur	Moyenne	6 × Écart-type
1	4,408	1,8	4,720	0,06
2	29,97	1,9	29,60	0,10
3	82,87	1,7	82,92	0,18
4	163,1	1,8	163,5	0,35



**Figure 4.14** – Diagramme Q-Q pour vérification de la nature gaussienne du résidu de l'identification pour différents nombres de modes identifiés



**(a)** RMSE de l'identification sur le cas test réel de la poutre encastrée-libre par rapport aux observations **(b)** Fonction de densité de probabilité des erreurs centrées en fonction du nombre de modes identifiés

**Figure 4.15** – Deux manières différentes de présenter les erreurs : le RMSE en fonction du temps et la fonction de densité de probabilité centrée en fonction du nombre de modes identifiés

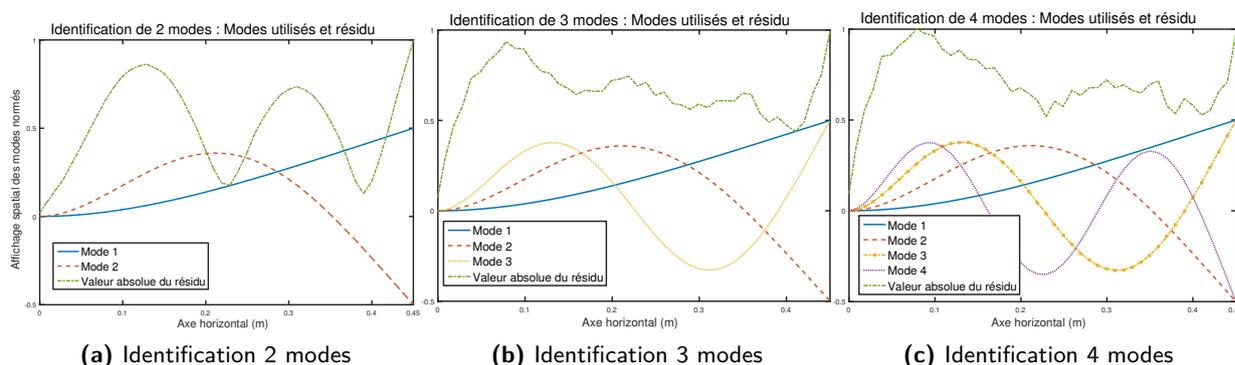
#### VI.B.4. Utilisation des modèles d'assimilation pour la prédiction

Comme exposé dans la sous-section IV.D. et la sous-section IV.E., il est possible d'utiliser le filtre pour faire de la prédiction dès sa convergence. On propose dans cette partie un aperçu d'ensemble des différentes possibilités qu'offrent ces utilisations. On introduira la fusion multicapteurs, l'interpolation temporelle et spatiale, ainsi que la compression de données.

##### a. Fusion de données

On propose de reprendre les données issues de l'expérience présentée dans l'annexe C.. On a donc l'observation de l'ensemble des points donnée par une caméra principale, l'observation d'un point (N°8) par un capteur de déplacement, et l'observation d'un point (N°43) donnée par une deuxième

## VI. Application au cas d'une poutre encastrée libre



**Figure 4.16** – Résidu spatial et modes normés identifiés avec l’affichage des modes théoriques

caméra de caractéristiques identiques à la première. Il existe de multiples manières de fusionner les mesures Bloch (1996). Nous présentons ici une manière simple en augmentant l’espace d’observation de deux points supplémentaires côté filtre et en concaténant les mesures côté acquisition. L’espace des observations ne correspond plus à l’observation réelle de la structure, ce qui nous oblige d’utiliser une seconde fonction d’observation uniquement pour l’affichage.

**Table 4.8** – Résolution des capteurs utilisés

Capteur	Résolution	Niveau de Bruit après méthodes d’extractions	Covariance choisie (R)
Caméra principale (1)	109,7 $\mu\text{m}$	$1 \times 10^{-7}$ m	$1 \times 10^{-12}$
Caméra secondaire (2)	19,84 $\mu\text{m}$	$5 \times 10^{-8}$ m	$2 \times 10^{-13}$
Capteur de déplacement	10 $\mu\text{m}$	$1 \times 10^{-7}$ m	$5 \times 10^{-11}$

Notre mesure de déplacement étant plus précise que la mesure réalisée par notre caméra principale et possédant des erreurs de mesures sur des amplitudes importantes de l’excitation, nous avons préféré choisir volontairement une valeur moins élevée pour la covariance de manière à englober ces erreurs (tableau 4.8).

Contrairement à ce que l’on pourrait attendre, la figure 4.17 nous montre que la fusion de deux mesures de précisions différentes ne permet pas d’améliorer ici de manière significative la précision de l’estimation. Plusieurs raisons peuvent expliquer cette absence d’amélioration dans notre cas.

- La raison la plus probable est que le RMSE est effectué ici en utilisant comme données d’acquisition les mesures de la caméra principale. Il est donc logique que la fusion avec d’autres capteurs soit moins exacte.
- Au-delà de la résolution et du bruit des mesures, nous avons montré que les capteurs ne mesuraient pas la même quantité, entachant la mesure d’approximations, qui se retrouvent dans le RMSE (voir Annexe C.).

### b. Observation partielle et interpolation spatiale

On propose dans cette partie d’effectuer une prédiction spatiale à partir des déformées modales. Il est possible de réduire l’espace d’observation de la structure tout en extrapolant sur le modèle

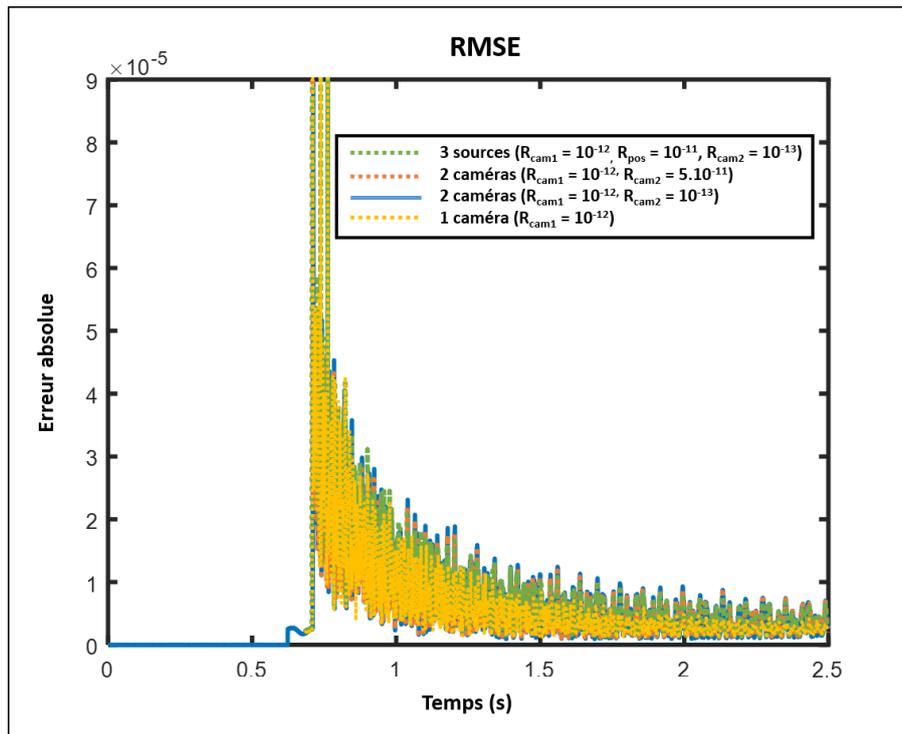
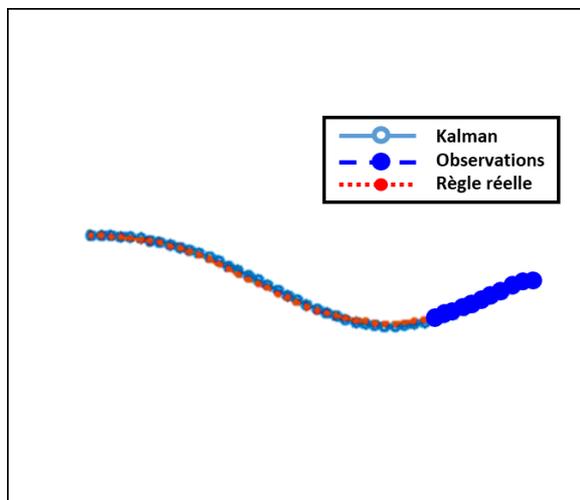
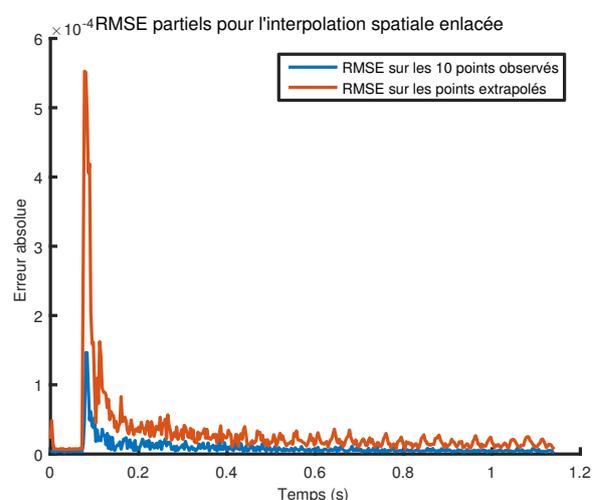


Figure 4.17 – Fusion de plusieurs sources de données avec les réglages sur la covariance



(a) Position des observations et points extrapolés pendant le mouvement de la structure

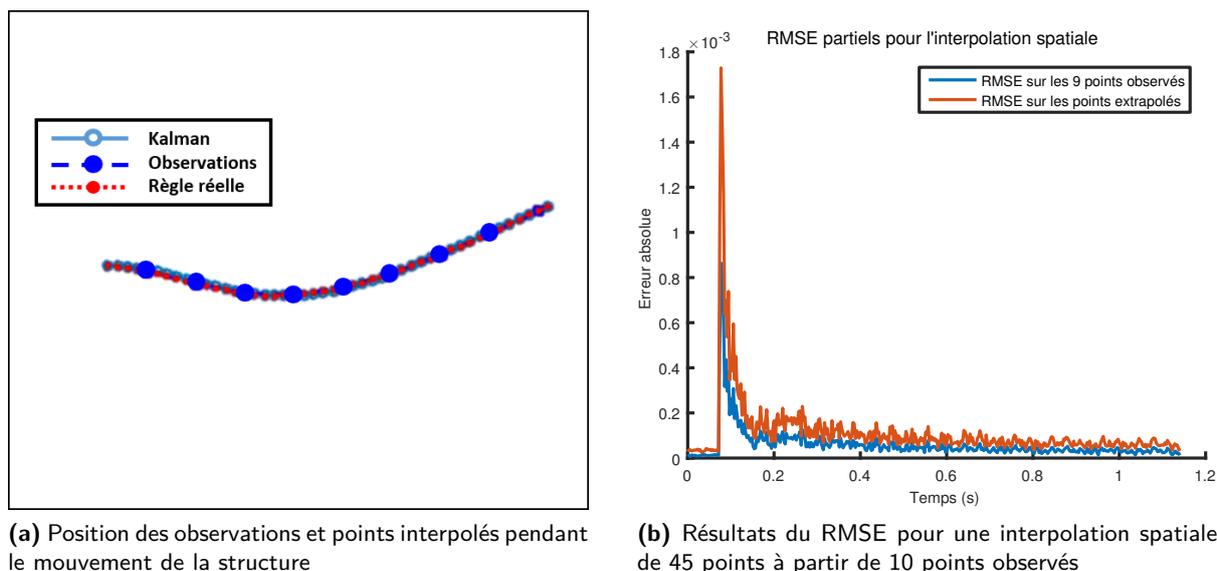


(b) Résultats du RMSE pour une extrapolation spatiale

Figure 4.18 – Configuration de l'observation et résultat temporel du RMSE pour l'extrapolation spatiale

d'observation le reste de la structure. Ce type de fonctionnement se retrouve dans la littérature sous le nom d' « expansion modale » . En reprenant cette fois l'exemple de la règle encadrée libre utilisé pour l'identification, on propose d'observer les dix derniers points et de prédire les autres. La figure 4.18 compare les résultats sur deux RMSE partiels : l'un sur les points observés et l'autre sur

## VI. Application au cas d'une poutre encastrée libre



**Figure 4.19** – Configuration de l'observation et résultat temporel du RMSE pour l'interpolation

les points prédits.

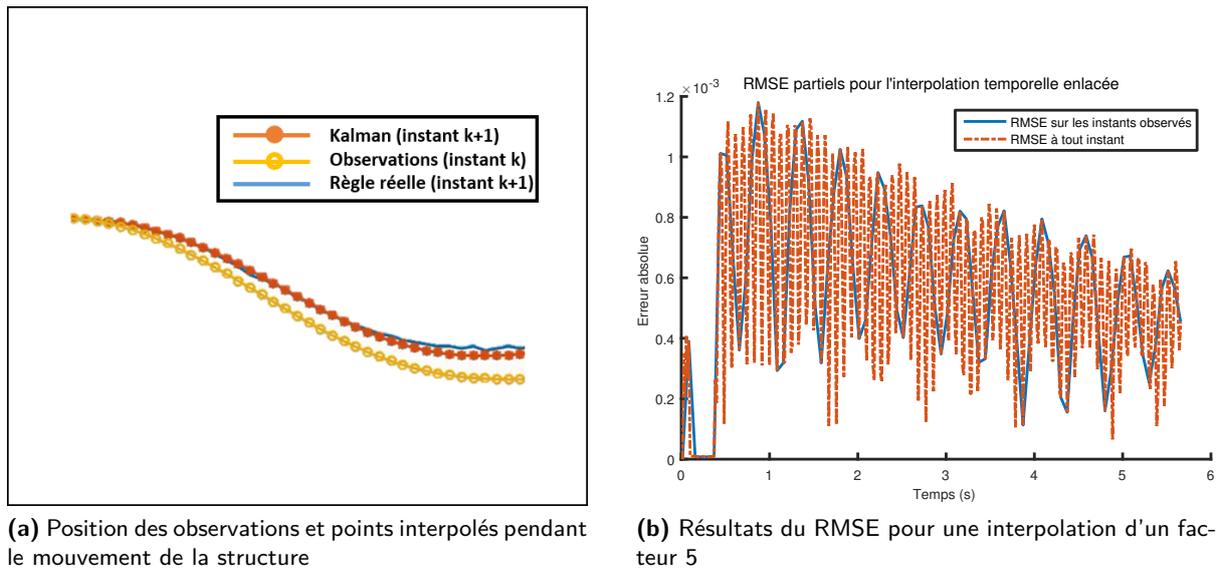
La correspondance entre les données réelles « cachées » au filtre et les données prédites est satisfaisante. La grande partie de l'erreur réside spatialement dans l'encastrement, ce qu'on ne peut pas voir sur le RMSE. Le résultat du RMSE est tout à fait acceptable, compte-tenu qu'il opère en aveugle.

Il est également possible d'interpoler au sens du modèle des points situés entre chaque position mesurée. L'avantage d'utiliser le modèle par rapport à une interpolation classique réside dans l'absence de limitation sur le nombre de points interpolés. Les résultats de la figure 4.19 sont satisfaisants puisque les données interpolées du modèle ont une erreur du même ordre de grandeur que les données observées. Les résultats vidéos des interpolations sont disponibles dans l'annexe D.

### c. Interpolation temporelle

Intéressons-nous maintenant aux informations apportées par la fonction d'évolution. Comme présenté dans la sous-section IV.E., la fonction d'évolution propose un modèle temporel des observations. L'utilisation de cette fonction va permettre de faire l'interpolation temporelle. La fréquence d'échantillonnage de l'essai que l'on traite figure 4.20 étant de  $F_{ech} = 350$  Hz, nous proposons d'observer les sorties de K-SAMI sur une fréquence cinq fois plus grande. La figure 4.20a montre pour une observation en jaune à un pas donné, comment le Kinematic-SAMI en rouge parvient à suivre la règle réelle en bleu alors qu'il ne l'observe pas.

Les erreurs constatées sont d'évidence plus élevées que sans interpolation, mais le filtre a bien un comportement minimisant le RMSE en fonction du temps. La vitesse de convergence vers une erreur minimale semble plus faible puisqu'elle dépend principalement du nombre d'observations reçues. Nous pourrions réduire cette convergence en initialisant plus proche de la première observation.



**Figure 4.20** – Configuration de l'observation et résultat temporel du RMSE pour l'interpolation temporelle d'un facteur 5

#### d. Compression

Le filtrage modal que nous appliquons nous permet d'extraire les différents paramètres modaux de la structure présents dans l'acquisition. En prenant la compression au sens défini dans la sous-section IV.E., on a ici une scène à fond fixe dont les informations pertinentes sont les informations modales. Une manière de compresser ces données revient donc à effectuer une troncature de la base modale puis à reconstituer l'image. L'affichage peut être fait à partir du modèle, mais également en générant une vidéo avec la méthode proposée dans la section VII. Si  $n$  est le nombre d'images de la séquence, le taux de compression obtenu sera de l'ordre de  $\frac{1}{n}$ . Les informations restituées par la suite offriront l'avantage d'être plus précises que l'acquisition originale grâce à la précision sous-pixélique des méthodes utilisées. Le rendu vidéo restera dans la définition d'origine et ne pourra pas être compressé, à moins de sur-échantillonner l'image et ainsi d'utiliser la précision de l'identification.

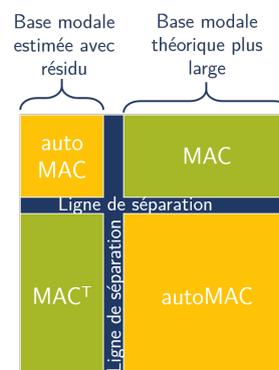
### VI.C. Validation du schéma Blind-SAMI

L'application du schéma *Blind-SAMI* sur la structure observée permet une analyse sous la seule hypothèse que chacun de ces points se décompose sur une base modale commune. Autrement dit, il n'y a aucun lien spatial entre les différents points. De manière à pouvoir d'abord valider ce modèle, puis le comparer avec l'autre version, on propose d'utiliser les mêmes cas tests que pour la partie précédente, c'est-à-dire de valider dans un premier temps le modèle sur le cas simulé de la poutre encastree libre pour ensuite l'appliquer au même cas réel. Les indicateurs utilisés dans la section précédente ne seront pas conservés pour l'amplitude car moins pertinents. Cette méthode permettant de récupérer les déformées modales. Elle nous donne en réalité non plus l'amplitude d'un mode théorique, mais les amplitudes de chaque point de ce mode. Nous utiliserons également le critère de MAC pour pouvoir caractériser les déformées modales estimées avec les déformées modales théoriques. Ainsi, nous utiliserons le MAC et l'autoMAC (c'est-à-dire le critère de MAC

## VI. Application au cas d'une poutre encastrée libre

d'une base sur elle-même) au sein d'une même représentation (cf. figure 4.21).

Le MAC nous renseignera sur la similarité entre les modes estimés et la base modale théorique, ce qui nous permettra de quantifier la qualité de notre estimation. Nous avons volontairement séparé les blocs de MAC et d'autoMAC par des lignes et colonnes bleues pour renforcer la lisibilité. Notre matrice compare les valeurs absolues des trois modes estimés et le résidu avec la valeur absolue des modes théoriques. Nous afficherons ce MAC dynamiquement, de manière à observer la disparition des modes. Enfin, nous enrichirons la base des modes propres estimés avec le résidu de cette estimation. Cet enrichissement nous permettra de détecter en temps réel la présence de composantes modales dans le résidu.



**Figure 4.21** – Explication pour les matrices de MAC exposées dans la suite

### VI.C.1. Simulations

Pour la validation de ce second schéma de filtrage, nous reprenons le même cas test que précédemment. Sans plus s'étendre sur les conditions de simulations déjà décrites auparavant, on présente les mêmes résultats que précédemment. La figure 4.22 présente les estimations des amplitudes et fréquences modales, tandis que la figure 4.23a présente l'identification de l'amortissement. L'estimation est satisfaisante au niveau des déformées modales comme de la fréquence. En revanche cette méthode semble moins précise que la précédente pour estimer le coefficient d'amortissement.

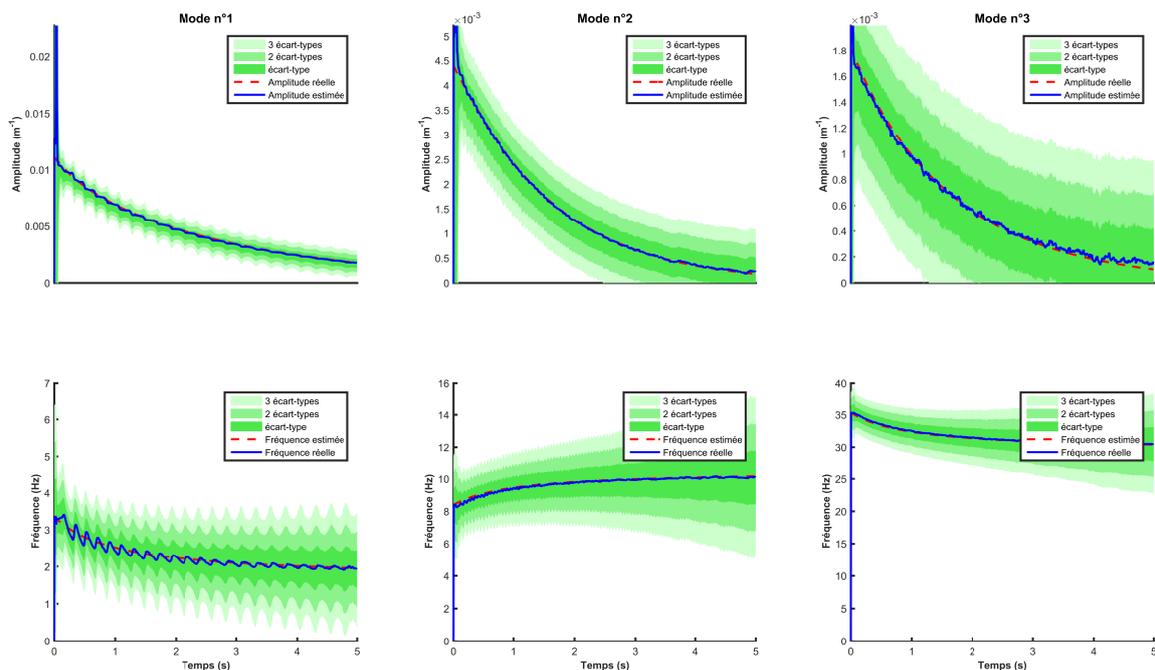
La comparaison des RMSE (figure 4.23b) de cette version du filtre avec la précédente fait ressortir une plus grande sensibilité au bruit ce qui découle de l'absence de modèle spatial permettant de lisser les données de chaque point. C'est également cette plus grande sensibilité au bruit qui rend l'identification du coefficient d'amortissement plus bruitée.

### VI.C.2. Cas réel

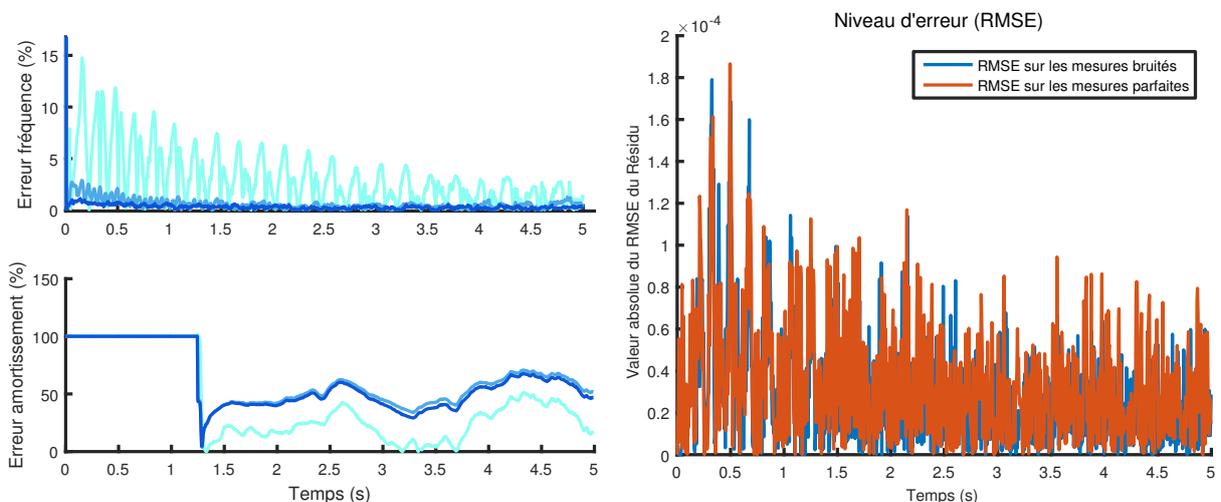
Le cas test réel de la poutre encastrée libre est très adapté pour tester cette méthode : il possède quatre modes dans son contenu spectral dont le quatrième est très évanescent. L'estimation sera effectuée sur trois modes uniquement afin de laisser un mode dans le résidu pour l'analyse qui suivra. Les résultats sont donnés figure 4.26. Les amplitudes moyennes correspondent à celles trouvées pour l'analyse du même cas test par la méthode K-SAMI comme on peut le voir sur la figure 4.25. Même s'il existe des différences de niveau dues au fait que la valeur observée diffère d'un facteur, les pentes sont similaires.

Le filtre est donc capable d'identifier en temps réel les modes propres expérimentaux. On vérifie à l'aide du RMSE présenté figure 4.24 que les niveaux de résidus sont similaires à ceux obtenus dans la partie précédente, c'est-à-dire que la précision est la même par rapport à la valeur du pixel projeté sur la structure. On remarquera sur la figure 4.24c que l'erreur ne peut être caractérisée gaussienne que sur la partie centrale de la fonction de probabilité. Nous n'avons pas poussé plus loin notre investigation sur la question.

## Chap. 4 : Application à l'analyse modale : une méthode temps réel pour l'identification de modes non linéaires



**Figure 4.22** – Résultats de l'estimation des paramètres modaux par le schéma Blind-SAMI et comparaison avec les vraies valeurs de la simulation



**(a)** Erreur en % de l'estimation sur la fréquence et l'amortissement

**(b)** Valeur du RMSE pour des mesures bruitées ou des mesures parfaites

**Figure 4.23** – Résultats de l'identification en fréquence et amplitude pour la simulation non-linéaire d'une poutre encastree-libre : RMSE

### VI.C.3. Discussions et analyses

De la même manière que dans la partie VI.B., nous allons procéder à une analyse de résidu. Cependant les déformées modales étant estimées, l'analyse spatiale se portera également sur une appréciation de la corrélation entre ces modes estimés et les modes théoriques.

## VI. Application au cas d'une poutre encastrée libre

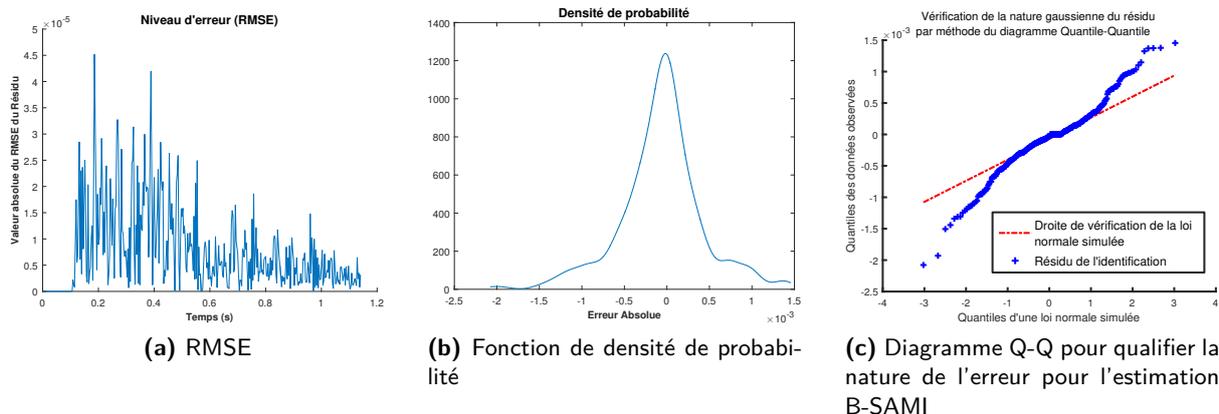


Figure 4.24 – Caractérisation de l'erreur pour le schéma B-SAMI

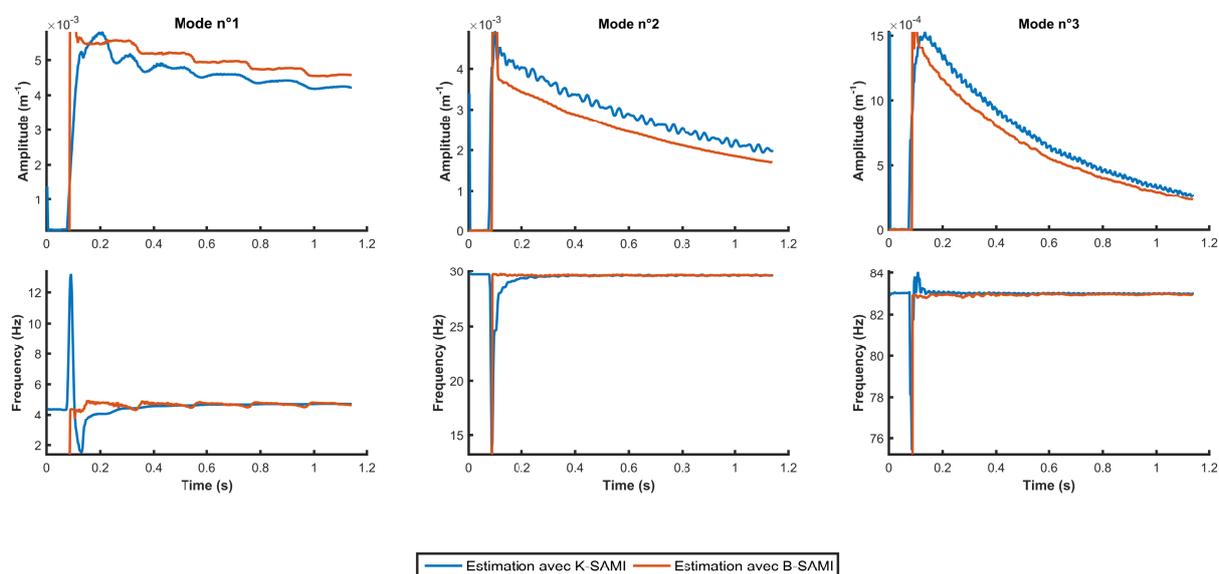
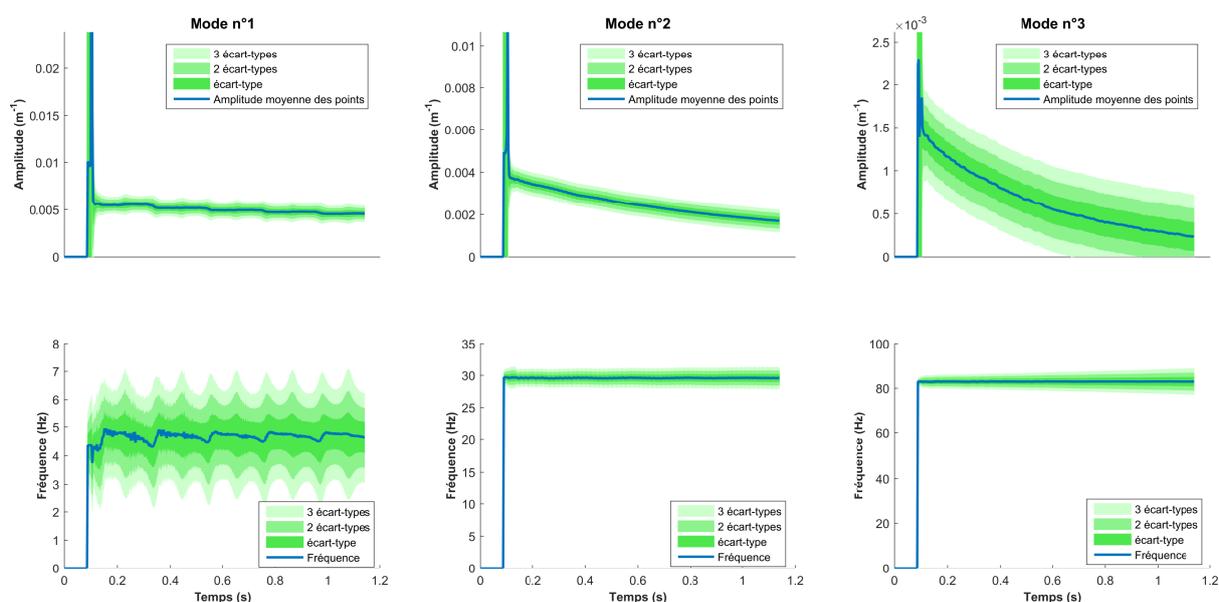


Figure 4.25 – Comparaison entre les estimations de K-SAMI et de B-SAMI

Afin de montrer le bon fonctionnement du filtre, nous avons choisi de comparer deux instants de l'essai : le premier proche du choc et le second à la fin de notre acquisition. La figure 4.27 présente ainsi deux manières de voir les choses : le MAC et l'autoMAC des modes estimés et du résidu, et leur représentation spatiale normalisée. Une valeur bleu représente une corrélation nulle et rouge bordeaux une corrélation maximale. L'autoMAC nous renseigne sur les dépendances potentielles entre les modes d'une même base. Nous avons choisi d'afficher le critère de MAC des modes estimés sur une base modale théorique beaucoup plus grande de manière à pouvoir mettre en évidence des modes inattendus.

On remarque que les modes propres estimés sont bien corrélés avec les modes propres théoriques correspondants. Les autoMAC des deuxièmes et troisièmes modes montrent une moins bonne définition que le premier, ce que l'on retrouve dans la partie MAC. On vérifie bien sur les représentations spatiales cette correspondance. Une remarque intéressante peut être soulevée pour le premier mode. En effet, il s'agit du mode dont la distorsion est la plus grande lors de l'estimation et il aurait pu être

## Chap. 4 : Application à l'analyse modale : une méthode temps réel pour l'identification de modes non linéaires



**Figure 4.26** – Résultats de l'estimation des paramètres modaux par le schéma B-SAMI sur le cas test réel de la poutre encastree libre

attendu d'avoir une « énergie » étalée sur plusieurs modes. On s'aperçoit avec le MAC qu'il ne s'agit bien que de distorsions dues à la transformée de Hilbert, faisant apparaître des fréquences parasites sans aucun sens modal : on ne retrouve pas d'autre corrélation de ce premier mode dans la base modale théorique. Au niveau du résidu, on remarquera que le mode 4 est bien présent au début de l'enregistrement puisqu'il y a une forte corrélation sur le MAC. On s'aperçoit en observant le résidu spatial que le résidu a bien une forme proche du quatrième mode. Sur la deuxième colonne de la figure 4.27 présentant le MAC au pas de temps 368, on note que la corrélation avec le quatrième mode a disparu, ce qui s'explique bien par le caractère plus évanescent du quatrième mode. Ce mode 4 montre un des avantages de notre schéma SAMI quelle que soit sa version : être capable d'extraire les informations modales à chaque instant grâce à son schéma séquentiel. L'annexe E. présente une animations de cette figure statique.

### VI.D. Conclusion sur l'application des SAMI

Cette partie applicative a permis de tester et valider nos deux méthodes. Basées sur un schéma d'assimilation de données vidéos et permettant d'intégrer des modèles d'états non-linéaires, ces schémas SAMI nous permettent de multiple applications.

La première version développée, *Kinematic-SAMI* offre tous les avantages de l'utilisation d'un modèle dans des observations décrites dans les objectifs du projet FAMOSYS (chapitre 1). La structure découplée du modèle d'état nous autorise même d'avoir deux modèles : l'un spatial et l'autre temporel. Nous avons pu valider l'identification des paramètres modaux avec des précisions inférieures à 1% pour la fréquence et 5% pour l'amortissement, l'erreur sur l'amplitude convergeant vers  $1 \times 10^{-5}$  m pour la simulation. L'enrichissement des mesures à l'aide du modèle spatial a ensuite été testée et nous avons pu obtenir des *RMSE* tout à fait satisfaisant pour des valeurs interpolées ou des valeurs extrapolées en fonction du modèle spatial. L'enrichissement des mesures avec le modèle temporel nous a permis d'effectuer une

## VI. Application au cas d'une poutre encastrée libre

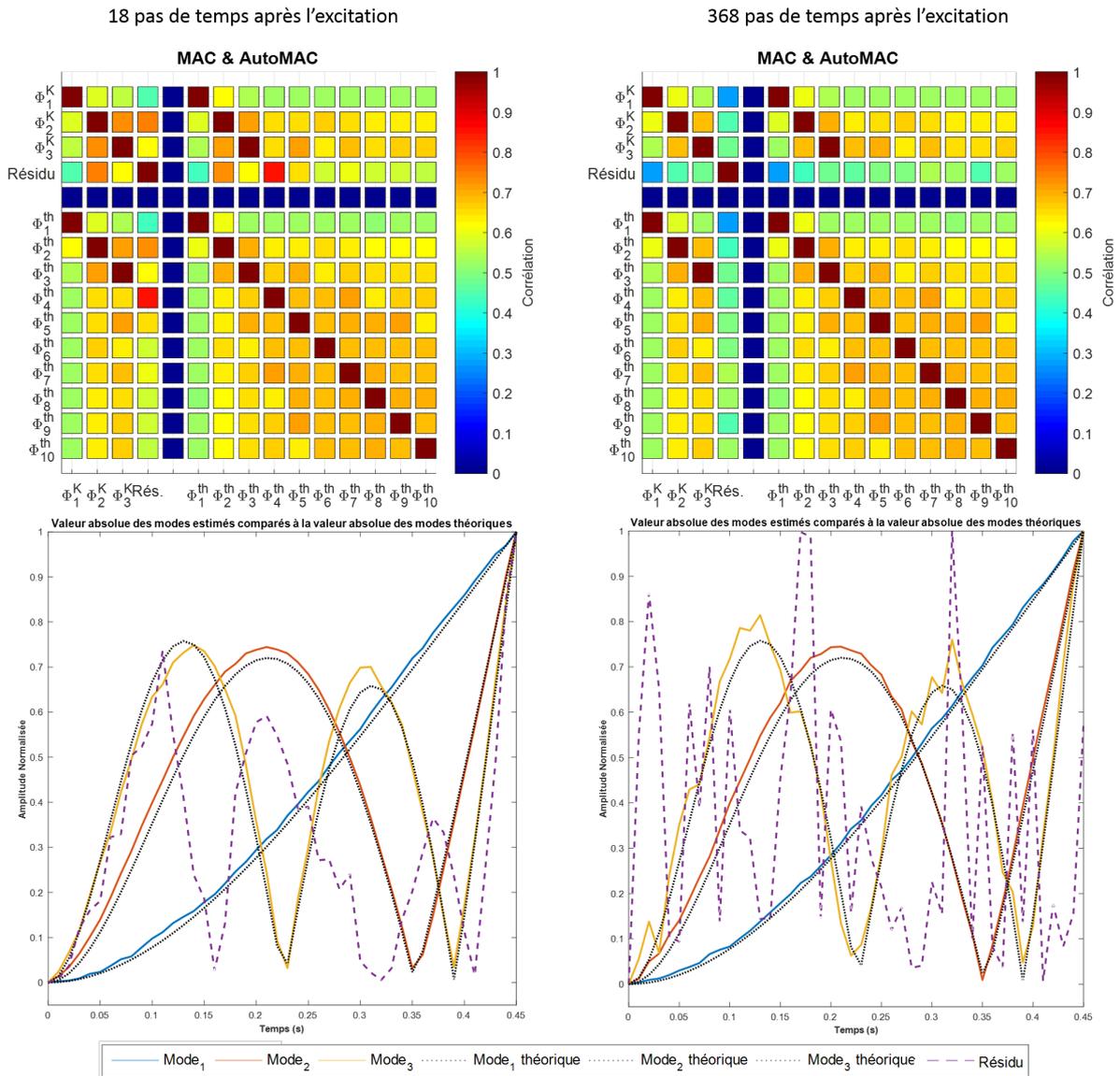


Figure 4.27 – MAC et déformées pour deux instants de l'estimation

interpolation temporelle et pourrait servir à de la prédiction même s'il n'y pas beaucoup d'intérêt dans notre cadre d'étude.

La seconde version développée, **Blind-SAMI** fonctionne sans modèle spatial et propose d'estimer, en plus des fréquences, les déformées expérimentales pour chaque mode. Si l'absence de modèle spatial limite ses capacités à enrichir les mesures, il obtient des résultats sur l'estimation de modes propres expérimentaux sur le même cas test que précédemment. Nous avons proposé une nouvelle utilisation du critère de MAC dynamique pour mieux caractériser le fonctionnement de ce filtre que nous avons reliée à l'observation spatiale des déformées modales.

Nous avons pu observer que ces méthodes possèdent l'avantage de pouvoir utiliser des technologies de capteurs différentes en entrée à travers les résultats en fusion de données. Elles peuvent être utilisées au sein d'autres algorithmes comme nous avons pu le proposer sur un cas 3D dans la sous-

section IV.C.. Afin d'enrichir la perception du système réel que nous observons, nous proposons de développer dans une dernière partie un algorithme réutilisant les sorties d'un K-SAMI pour synthétiser une vidéo différente de celle acquise.

## VII Modifier la perception du système réel à l'aide de sa physique : proposition d'un autre algorithme de *motion magnification*

L'approche développée dans cette thèse propose d'utiliser des méthodes d'assimilation de données, et plus particulièrement des filtres de Kalman pour les appliquer dans un contexte non-linéaire dans un schéma temps réel. En l'appliquant à l'analyse modale expérimentale, nous avons pu faire de l'identification, de la prédiction, de la fusion mais également de la compression et de l'interpolation au sens du modèle physique. Cependant, lorsqu'on se réfère aux concepts développés dans les parties précédentes, nous ne développons ici que de la réalité augmentée par superposition d'informations à la vidéo originale mais pas de vidéo synthétisée selon le modèle. Dans cette dernière partie, nous appliquerons les concepts de la compensation de mouvement développés dans le chapitre 2 afin de pouvoir créer une vidéo en fonction des paramètres identifiés. Nous proposerons ainsi une nouvelle méthode pour faire de la magnification avec les avantages de notre approche temps réel.

### VII.A. Modification d'une vidéo à partir de l'espace d'observation du K-SAMI

Cette partie diffère des précédentes sur sa position dans la chaîne de mesure. On trouvera ici en entrée les images à l'instant  $k - 1$  et  $k$  et les vecteurs d'états correspondants. On propose dans cette partie de développer une nouvelle méthode de magnification, c'est-à-dire un outil d'affichage permettant d'amplifier les contributions modales d'une structure donnée.

L'activité concernant cette technique de *motion magnification* que nous avons déjà développée<sup>6</sup> est très concentrée autour d'une équipe du MIT comportant W.T.Freeman et F.Durand. Les approches utilisant les *Steerable Complex Pyramids* faisant intervenir du filtrage par bande dans un espace spectral, elle ne permettent pas de travailler dans un schéma séquentiel propre à pouvoir s'adapter dans un schéma temps réel. Par contre les derniers articles basés sur une transformée de Riesz permettent une utilisation en schéma temps réel.

Les concepts développés dans cette thèse permettent d'effectuer de l'identification schéma temps réel, recalant ainsi un modèle sur les observations qui sont extraites des images. Faire de la magnification dans ce schéma reviendrait à concevoir une méthode permettant de générer une image modifiée à l'instant  $k$  à partir :

- des images autour de l'instant  $k$  d'une vidéo ;
- des vecteurs d'état modifiés par un facteur de magnification  $M_\alpha$ .

Il s'avère que l'état de l'art développé dans le chapitre 2 sous-section II.B. décrit des stratégies de reconstruction puisque c'est un processus essentiel dans la compression de données vidéo. Les standards MPEG ont introduit des algorithmes permettant d'approximer toute une série d'images en calculant par estimation de mouvement un champ de déplacements des pixels de l'image. On peut alors supprimer les informations pour lesquelles l'approximation est concluante. Les algorithmes permettant d'effectuer cette transformation comme le *block matching* ou le flot optique n'ont de sens que si l'image peut être reconstruite, c'est-à-dire d'obtenir l'image  $k + 1$  à partir d'une carte

6. Dans la sous-sous-section I.B.3. de ce chapitre.

## VII. Modifier la perception du système réel à l'aide de sa physique : proposition d'un autre algorithme de *motion magnification*

de champ de déplacement de pixels et de l'image au temps  $k$ . C'est la famille des algorithmes dits de *motion compensation* ou compensation de mouvement.

Plusieurs approximations sont faites dans cette famille d'algorithmes qui imposent des hypothèses d'applications particulières. Tout d'abord, les mouvements des pixels sont supposés linéaires, ce qui implique :

- de pouvoir effectuer une linéarisation des trajectoires de chaque pixel, soit pour l'intensité  $I_k(u)$  de pixels repérés par le vecteur  $u$  et  $a$  un vecteur de déplacement dans le repère capteur :

$$I_{k+1}(u+a) \simeq I_k(u) + \vec{\nabla} I_k(u) \cdot a \quad (4.110)$$

- de supposer l'intensité de chaque pixel constante sur un voisinage temporel.

Enfin ces algorithmes s'appuient sur une hypothèse fautive de conservation de la quantité de pixels dans l'image. En effet, pour une estimation/compensation de mouvement, il est difficile de gérer les apparitions/disparitions d'objets pourtant intrinsèques à la projection sur le plan 2D du capteur. Cette hypothèse est souvent vérifiée localement à la fois temporellement et spatialement.

Ces conditions d'application de l'estimation/compensation de mouvements sont difficiles à appliquer sans dégradation de la vidéo pour des scènes quelconques. Cependant, notre cadre de vidéos rapides sur des scènes 2D est bien choisi. En reprenant les hypothèses déjà formulées dans le chapitre 2 sur les vidéos rapides, on rappelle les conditions intéressantes ici :

- avoir des déplacements 2D ce qui évite les effets de perspective et nous place dans un cadre plus que favorable pour une conservation globale des quantités de pixels ;
- avoir des conditions d'éclairage idéales qui garantissent que l'intensité des pixels est conservée au cours du temps en fonction de la position sur la structure. À noter que cette hypothèse n'est pas respectée localement au niveau des reflets ;
- avoir des petits mouvements de pixels découlant directement du caractère rapide des vidéos.

Ces hypothèses renforcent le cadre d'application de la linéarisation des trajectoires de pixels. Nos vidéos ont des caractéristiques particulières qui permettent de mieux satisfaire le cadre des hypothèses de l'estimation/compensation de mouvements que pour l'utilisation classique qui est faite sur des vidéos quelconques.

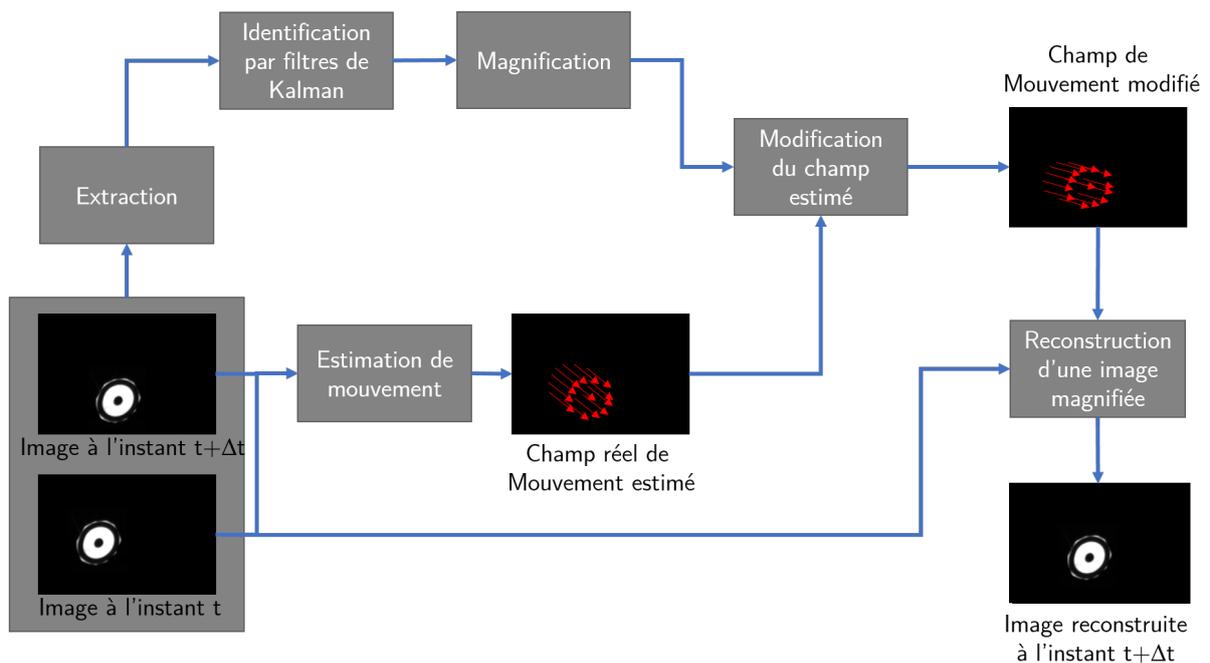
Nous proposons ici de reprendre cet algorithme et de détourner son utilisation. L'idée est de prendre les sorties du filtre présenté précédemment, d'attribuer un gain par mode sur les amplitudes identifiées puis de projeter dans l'espace d'observation. À un instant donné, l'estimation de mouvement fournit un champ de déplacements à partir des deux dernières images. On génère un champ virtuel à partir de la différence entre les positions extraites et les positions magnifiées ramenées dans l'espace d'observation. Ce champ virtuel vient modifier le champ réel afin d'obtenir l'image magnifiée à partir des pixels de l'image réelle. Le schéma proposé figure 4.28 résume l'algorithme modifié pour afficher la magnification.

Nous avons proposé ici d'utiliser notre méthode d'assimilation au sein d'un schéma de magnification. Une méthode d'affichage a été développée à partir des images réelles et de la projection dans l'espace d'observation des variables identifiées et magnifiées de manière à ce que l'algorithme puisse agir comme une « loupe » dans le domaine modal.

### VII.B. Application sur le cas réel d'un ski

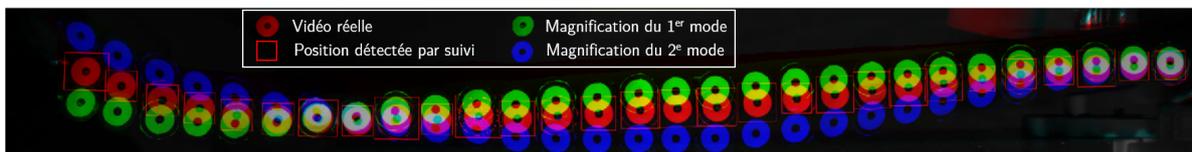
À titre d'exemple d'application de l'algorithme présenté dans cette partie, on propose d'utiliser la vidéo d'un ski excité au choc. À l'aide du schéma *Kinematic-SAMI* nous avons obtenu une

Chap. 4 : Application à l'analyse modale : une méthode temps réel pour l'identification de modes non linéaires



**Figure 4.28** – Une nouvelle méthode de magnification basée sur de l'identification modale et sur de la compensation de mouvements

identification des deux premiers modes. L'utilisation d'un algorithme de magnification va permettre d'utiliser la séparation modale afin d'augmenter artificiellement l'amplitude de vibration pour des modes dont les contributions ne sont pas visibles. Afin de pouvoir comparer la magnification sur une



**Figure 4.29** – Image 160 de la séquence du ski avec une excitation de type choc (canal rouge) - Magnification du 1er mode uniquement (canal vert) - Magnification du second mode (canal bleu) en gardant la contribution du premier mode

image statique, nous avons choisi d'utiliser les différents canaux de couleur pour afficher plusieurs informations sur la même image présentée figure 4.29. Une image de la vidéo réelle avec le suivi vidéo des mire encadrées sur le canal rouge et deux images magnifiées différemment au même instant à partir de cette image réelle et du modèle utilisé dans l'assimilation. Ainsi, sur le canal vert, on a choisi de retirer complètement le second mode et de ne pas amplifier le premier mode, tandis que sur le canal bleu, on a magnifié le second mode avec un facteur 2,5 tout en conservant la contribution initiale du premier mode, de manière à garder l'image magnifiée centrée par rapport à l'image originale.

Le résultat de cette synthèse est probant cependant quelques limitations sont à souligner :

- Pour éviter trop de distorsions, on doit rester dans un cadre de petits mouvements, ce qui doit être pris en compte lors du choix des coefficients.

## VIII. Conclusion et perspectives

- Bien que nous ne l'ayons pas mentionné, le changement de direction de pixels induit des doublons : plusieurs pixels peuvent ainsi se retrouver au même endroit. Ainsi, lorsqu'un objet quitte sa position initiale, très peu de pixels viennent combler sa place dans sa position initiale. Ce problème est récurrent dans le domaine de la synthèse de vidéos, et il existe des solutions qui ne seront pas développées dans ce manuscrit.
- Quelle que soit la méthode d'extraction vidéo choisie, les méthodes actuelles imposent d'avoir une discrétisation spatiale inférieure à celle des pixels. Ainsi, lorsque l'on va utiliser ces informations sur la vidéo réelle, une interpolation sera nécessaire.

La magnification permet d'afficher directement dans une vidéo les résultats du K-SAMI. En prenant un gain unitaire pour l'ensemble des modes de la structure filmée, elle permet simplement d'observer le résultat de l'assimilation de données. Cette vidéo synthétisée peut être enrichie avec les modèles spatiotemporels du *K-SAMI*, ce qui permettrait d'augmenter la résolution spatiale de la vidéo ou d'insérer des images pour ralentir le mouvement.

## VIII Conclusion et perspectives

Afin d'appliquer les concepts proposés dans cette thèse au domaine de l'analyse vibratoire et plus particulièrement dans le domaine de l'analyse modale expérimentale, nous avons repris et développé un modèle basé sur la transformée de Hilbert. Ce modèle a été adapté à une extension non-linéaire d'un filtre de Kalman (EKF et UKF dans notre cas). Nous l'avons ensuite conformé à une mesure multipoint dans le but de pouvoir utiliser des extractions de mesures par vidéo rapide.

Deux filtres différents utilisant ce modèle ont été proposés dans un contexte de faibles non-linéarités pour des mesures au choc.

**Un schéma d'assimilation *Kinematic-SAMI* utilisant un modèle de la structure** a été développé et validé sur des cas tests non linéaires simulés et réels. Ce filtre permet d'extraire sur un schéma temps réel les différents paramètres modaux. Sa structure permet en outre d'enrichir la mesure par les modèles utilisés en permettant de faire du lissage, de l'interpolation et de la prédiction spatiale mais également de l'interpolation temporelle. Ce filtre modal va ainsi « linéariser » spatialement les observations en les assimilant à une base modale de la structure qui lui est fournie.

**Un schéma d'assimilation aveugle *Blind-SAMI* sans connaissance de la structure** a alors été proposé afin de répondre au besoin de mesurer les non-linéarités. Toujours en schéma temps réel, il est néanmoins plus gourmand en calcul puisqu'il impose de travailler avec des systèmes d'état de grande taille. Si toute extrapolation spatiale est impossible avec cette version, l'interpolation temporelle est toujours applicable. Un dernier avantage de ce filtre est qu'il permet de décomposer avec un schéma temps réel une structure sur ses modes expérimentaux avec toutes ses non-linéarités tout en identifiant les différentes fréquences modales.

Ces filtres ont été déclinés et validés sur une chaîne de mesures vidéo en deux dimensions en utilisant des méthodes d'extraction vidéo de suivi de points d'intérêt. Cette chaîne permet de proposer une solution innovante de capteurs multipoints sans contact permettant d'effectuer de l'identification non-linéaire de paramètres modaux. Une application à d'autres types d'excitations est envisagée ainsi qu'un développement théorique dans le domaine de l'analyse modale opérationnelle. La structure offre la possibilité de mélanger des capteurs de technologies différentes ou d'utiliser d'autres méthodes d'extraction vidéo en utilisant le filtre de manière à faire de la fusion de données. Une application sur des sources de données à différentes fréquences d'échantillonnage est en cours de préparation.

#### Chap. 4 : Application à l'analyse modale : une méthode temps réel pour l'identification de modes non linéaires

Un algorithme d'une application du schéma temps réel Kinematic-SAMI en 3D a été proposé avec l'avantage de ne pas utiliser de stéréocorrélation, coûteuse en temps et en calcul. Une application sur la maquette GARTEUR (Balmes et Wright, 1997) devrait suivre pour valider cette méthode. Enfin, un nouveau schéma de magnification de mouvement a été proposé et validé sur un cas réel. Il utilise les méthodes SAMI avec des méthodes de compensation de mouvement pour synthétiser une vidéo de structure aux propriétés modales modifiées ou simplement enrichie par le modèle.

Ce chapitre a permis de montrer comment les méthodes d'assimilation de données pouvaient être appliquées aux vidéos rapides en offrant la possibilité de passer d'une simple grille discrète de pixels eulériens à une vision continue d'un modèle lagrangien. Du pixel on extrait ainsi les informations modales, de manière à superposer un modèle spatio-temporel aux données acquises, on augmente la vidéo, on fusionne plusieurs sources d'informations, jusqu'à retourner au pixel en synthétisant une vidéo selon les paramètres de l'utilisateur, concluant ainsi le tour d'horizon des possibilités qu'offrent ces méthodes.

## Bibliographie du chapitre 4

- I. Antoniadou, G. Manson, N. Dervilis, T. Barszcz, W. J. Staszewski et K. Worden : Use of the Teager-Kaiser energy operator for condition monitoring of a wind turbine gearbox. *In Proceedings of the International Conference on Noise and Vibration Engineering ISMA 2012*, pages 4255–4268, 2012.
- Etienne Balmes et J. R. Wright : GARTEUR group on Ground Vibration Testing. Results from the test of a single structure by 12 laboratories in Europe. *OFFICE NATIONAL D'ETUDES ET DE RECHERCHES AEROSPATIALES ONERA-PUBLICATIONS-TP*, 1997.
- Timothy J. Beberniss et David A. Ehrhardt : High-speed 3D digital image correlation vibration measurement : Recent advancements and noted limitations. *Mechanical Systems and Signal Processing*, volume 86, pages 35–48, avril 2016. ISSN 08883270.
- E. Bedrosian : A product theorem for Hilbert transforms. *Proceedings of the IEEE*, volume 51, pages 868–869, mai 1963. ISSN 0018-9219.
- Garry Berkovic et Ehud Shafir : Optical methods for distance and displacement measurements. *Advances in Optics and Photonics*, volume 4, page 441, décembre 2012. ISSN 1943-8206.
- Isabelle Bloch : Information combination operators for data fusion : A comparative review with classification. *IEEE Transactions on Systems, Man, and Cybernetics-Part A : Systems and Humans*, volume 26, pages 52–67, 1996.
- Federico Castanedo : A Review of Data Fusion Techniques. *The Scientific World Journal*, volume 2013, pages 1–19, 2013. ISSN 1537-744X.
- JR. Cole, H. : On-the-line analysis of random vibrations. *In 9th Structural Dynamics and Materials Conference, Structures, Structural Dynamics, and Materials and Co-located Conferences*. American Institute of Aeronautics and Astronautics, avril 1968.
- Gwendal Cumunel : *Extensomètres à Fibre Optique Longue Base (ELB) Pour l'évaluation Dynamique Des Structures*. Thèse de doctorat, 2008.
- A. Cunha, E. Caetano, F. Magalhães et C. Moutinho : Recent perspectives in dynamic testing and monitoring of bridges : RECENT PERSPECTIVES IN DYNAMIC TESTING AND MONITORING OF BRIDGES. *Structural Control and Health Monitoring*, volume 20, pages 853–877, juin 2013. ISSN 15452255.
- P. K. Dash, A. K. Pradhan et G. Panda : Frequency estimation of distorted power system signals using extended complex Kalman filter. *IEEE Transactions on Power Delivery*, volume 14, pages 761–766, 1999.
- Abe Davis, Katherine L. Bouman, Justin G. Chen, Michael Rubinstein, Frédo Durand et William T. Freeman : Visual vibrometry : Estimating material properties from small motions in video. *In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5335–5343. IEEE, 2015.
- J. A. de la O Serna et J. Rodríguez-Maldonado : Taylor-Kalman-Fourier Filters for Instantaneous Oscillating Phasor and Harmonic Estimates. *IEEE Transactions on Instrumentation and Measurement*, volume 61, pages 941–951, avril 2012. ISSN 0018-9456, 1557-9662.

- J.-L. Dion, I. Tawfiq et G. Chevallier : Harmonic component detection : Optimized Spectral Kurtosis for operational modal analysis. *Mechanical Systems and Signal Processing*, volume 26, pages 24–33, janvier 2012. ISSN 08883270.
- Xing-Jian Dong, Guang Meng et Juan-Chun Peng : Vibration control of piezoelectric smart structures based on system identification technique : Numerical simulation and experimental study. *Journal of Sound and Vibration*, volume 297, pages 680–693, novembre 2006. ISSN 0022460X.
- Charles Dorn, Sudeep Dasari, Yongchao Yang, Garrett Kenyon, Paul Welch et David Mascareñas : Efficient Full-Field Operational Modal Analysis Using Neuromorphic Event-Based Imaging. In Julie M. Harvie et Javad Baqersad, éditeurs : *Shock & Vibration, Aircraft/Aerospace, Energy Harvesting, Acoustics & Optics, Volume 9 : Proceedings of the 35th IMAC, A Conference and Exposition on Structural Dynamics 2017*, pages 97–103. Springer International Publishing, Cham, 2017. ISBN 978-3-319-54735-0.
- Saeed Eftekhari Azam, Eleni Chatzi et Costas Papadimitriou : A dual Kalman filter approach for state estimation via output-only acceleration measurements. *Mechanical Systems and Signal Processing*, volume 60-61, pages 866–886, août 2015. ISSN 08883270.
- Stephan Elies et Stephan Ebenhöch : Performance Analysis of Commercial Accelerometers of Different Technologies. Venice, Italy, août 2015. IARIA.
- M. Feldman : *Hilbert Transform Applications in Mechanical Vibration*. Wiley, 2011a. ISBN 978-1-119-99152-6.
- Michael Feldman : Hilbert transform in vibration analysis. *Mechanical Systems and Signal Processing*, volume 25, pages 735–802, avril 2011b. ISSN 08883270.
- Emmanuel Foltête : *Identification Modale de Structures Linéaires et Faiblement Non-Linéaires*. Thèse de doctorat, Université de Franche-Comté, Besançon, novembre 1998.
- Bernard Garnier : Vibrations des structures industrielles Notions de physique des vibrations. *Techniques de l'ingénieur Vibrations en milieu industriel, mesures, surveillance et contrôle*, volume base documentaire : TIB424DUO., 2009.
- Carmelo Gentile : Structural health monitoring of stay cables using the microwave interferometry. *Proceedings of the IOMAC*, 2011.
- Adly A. Girgis, W. Bin Chang et Elham B. Makram : A digital recursive measurement scheme for online tracking of power system harmonics. *IEEE transactions on Power Delivery*, volume 6, pages 1153–1160, 1991.
- GKJC Golinval et Gaetan Kerschen : *Experimental Modal Analysis*. 2014.
- Philippe Grémillet : Note relative à la précision de mesure de vitesse avec MotionTrack, mai 2010.
- Mark N. Helfrick, Christopher Niezrecki, Peter Avitabile et Timothy Schmidt : 3D digital image correlation methods for full-field vibration measurement. *Mechanical Systems and Signal Processing*, volume 25, pages 917–927, avril 2011. ISSN 08883270.
- Norden E. Huang, Zheng Shen, Steven R. Long, Manli C. Wu, Hsing H. Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung et Henry H. Liu : The empirical mode decomposition and the Hilbert

## Bibliographie du chapitre 4

- spectrum for nonlinear and non-stationary time series analysis. *In Proceedings of the Royal Society of London A : Mathematical, Physical and Engineering Sciences*, volume 454, pages 903–995. The Royal Society, 1998.
- S.R. Ibrahim : Random Decrement Technique for Modal Identification of Structures. *Journal of Spacecraft and Rockets*, volume 14, pages 696–700, novembre 1977. ISSN 0022-4650, 1533-6794.
- SR Ibrahim et EC Mikulcik : A method for the direct identification of vibration parameters from the free response. 1977.
- Jaka Javh, Janko Slavi ? et Miha Bolte ?ar : High frequency modal identification on noisy high-speed camera data. *Mechanical Systems and Signal Processing*, volume 98, pages 344–351, janvier 2018. ISSN 08883270.
- G. Kerschen, M. Peeters, J. C. Golinval et C. Stéphan : Nonlinear Modal Analysis of a Full-Scale Aircraft. *Journal of Aircraft*, volume 50, pages 1409–1419, septembre 2013. ISSN 0021-8669, 1533-3868.
- Paul Kopff : Renaissance d'une technique de l'analyse modale d'avant les transformées rapides : Le décrétement aléatoire. *Recherche-Traitement du signal*, volume 8, pages 345–354, 1990.
- Scott Krig : *Computer Vision Metrics : Survey, Taxonomy, and Analysis*. Apress, Berkely, CA, USA, 1st édition, 2014. ISBN 1-4302-5929-9 978-1-4302-5929-9.
- Eivind Kvedalen : Signal processing using the Teager energy operator and other nonlinear operators. *Master, University of Oslo Department of Informatics*, volume 21, 2003.
- Chang-Sheng Lin et Tse-Chuan Tseng : Response-only modal identification using random decrement algorithm with time-varying threshold level. *Journal of Mechanical Science and Technology*, volume 28, pages 2099–2109, juin 2014. ISSN 1738-494X, 1976-3824.
- Tianshu Liu, Ali Merat, M. H. M. Makhmalbaf, Claudia Fajardo et Parviz Merati : Comparison between optical flow and cross-correlation methods for extraction of velocity fields from particle images. *Experiments in Fluids*, volume 56, août 2015. ISSN 0723-4864, 1432-1114.
- E. Lourens, E. Reynders, G. De Roeck, G. Degrande et G. Lombaert : An augmented Kalman filter for force identification in structural dynamics. *Mechanical Systems and Signal Processing*, volume 27, pages 446–460, février 2012. ISSN 08883270.
- Chih-Kao Ma, Dong-Cherng Lin et Jer-Ming Chang : Estimation of forces generated by a machine mounted upon isolators under operating conditions. *Journal of the Franklin Institute*, volume 336, pages 875–892, 1999.
- Guillaume Martin : *Etude Du Crissement de Frein Par Intercorrélation*. Thèse de doctorat, 2017.
- Florent Mathieu, François Hild et Jérémie Bellec : Le projet DICCIT : Des outils pour la comparaison essai/calcul et l'identification. Lyon, France, août 2015. AFM, Association Française de Mécanique.
- Ignace Davy Mendoume Minko : *Identification Des Systèmes Dynamiques Stochastiques. Application à l'évaluation Dynamique Des Ponts Sous Sollicitations Ambiantes*. Thèse de doctorat, Université Blaise Pascal-Clermont-Ferrand II, 2005.

- F. Naets, F. Cosco et W. Desmet : An extended Kalman filter approach for augmented strain/stress visualization in mechanical systems. *In 2014 IEEE/ASME 10th International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, pages 1–6, septembre 2014.
- F. Naets, J. Cuadrado et W. Desmet : Stable force identification in structural dynamics using Kalman filtering and dummy-measurements. *Mechanical Systems and Signal Processing*, volume 50–51, pages 235–248, janvier 2015. ISSN 0888-3270.
- V. Ondra, I.A. Sever et C.W. Schwingshackl : A method for detection and characterisation of structural non-linearities using the Hilbert transform and neural networks. *Mechanical Systems and Signal Processing*, volume 83, pages 210–227, janvier 2017. ISSN 08883270.
- Christian Pachaud, Timothée Gerber, Marcin Firla, Nadine Martin et Corinne Mailhes : Consequences of non-respect of the Bedrosian theorem when demodulating. *In The Tenth International Conference on Condition Monitoring and Machinery Failure Prevention Technologies (CM2013)*, 2013.
- Costas Papadimitriou, Claus-Peter Fritzen, Peter Kraemer et Evangelos Ntotsios : Fatigue predictions in entire body of metallic structures from a limited number of vibration sensors using Kalman filtering. *Structural Control and Health Monitoring*, volume 18, pages 554–573, août 2011. ISSN 15452255.
- Bart Peeters, Herman Van der Auweraer, Patrick Guillaume et Jan Leuridan : The PolyMAX frequency-domain method : A new standard for modal parameter estimation? *Shock and Vibration*, volume 11, pages 395–409, 2004.
- Jean Piranda : Analyse vibratoire expérimentale. *TDI*, décembre 2001.
- Alexandros Potamianos et Petros Maragos : A comparison of the energy operator and the Hilbert transform approach to signal and speech demodulation. *Signal processing*, volume 37, pages 95–120, 1994.
- Xavier Régal, Gwendal Cumunel, Michel Bornert et Marc Quiertant : Analyse modale d'une poutre par corrélation d'images numériques. Lille, septembre 2017.
- Phillip L. Reu, Daniel P. Rohe et Laura D. Jacobs : Comparison of DIC and LDV for practical vibration and modal measurements. *Mechanical Systems and Signal Processing*, volume 86, pages 2–16, mars 2016. ISSN 08883270.
- Jeremy Schmitt, Nelly Pustelnik, Pierre Borgnat, Patrick Flandrin et Laurent Condat : 2D Prony-Huang Transform : A New Tool for 2D Spectral Analysis. *IEEE Transactions on Image Processing*, volume 23, pages 5233–5248, décembre 2014. ISSN 1057-7149, 1941-0042.
- S. Seibold et K. Weinert : A time domain method for the localization of cracks in rotors. *Journal of Sound and Vibration*, volume 195, pages 57–73, 1996.
- Md Rifat Shahriar, Pietro Borghesani, R.B. Randall et Andy C.C. Tan : An assessment of envelope-based demodulation in case of proximity of carrier and modulation frequencies. *Mechanical Systems and Signal Processing*, volume 96, pages 176–200, novembre 2017. ISSN 08883270.
- Yimin Shao et Chris K. Mechefske : Gearbox vibration monitoring using extended Kalman filters and hypothesis tests. *Journal of Sound and Vibration*, volume 325, pages 629–648, août 2009. ISSN 0022460X.

- Richard Szeliski : *Computer Vision : Algorithms and Applications*. Springer Science & Business Media, 2010.
- Tommaso Tamarozzi, Roland Pastorino, Martijn Vermaut, Bert Pluymers, Wim Desmet, Karl Janssens et Herman Van der Auweraer : MODRIO Project : Reduced Flexible Model for twistbeam and input/stress monitoring strategies. Rapport technique 11004, KU Leuven, septembre 2014.
- F. Trebuña et M. Hagara : Experimental modal analysis performed by high-speed digital image correlation system. *Measurement*, volume 50, pages 78–85, avril 2014. ISSN 02632241.
- Peter Van Overschee et B. L. De Moor : *Subspace Identification for Linear Systems : Theory—Implementation—Applications*. Springer Science & Business Media, 1996.
- Rémy Vincent : *Identification Passive En Acoustique : Estimateurs et Applications Au SHM*. Thèse de doctorat, Université Grenoble Alpes, 2016.
- Paul Viola et Michael Jones : Rapid object detection using a boosted cascade of simple features. *In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference On*, volume 1, pages I–I. IEEE, 2001.
- Hans-Jürgen von Martens : Metrology of vibration measurements by laser techniques. page 709802, juin 2008.
- Neal Wadhwa, Michael Rubinstein, Frédo Durand et William T. Freeman : Phase-based video motion processing. *ACM Transactions on Graphics (TOG)*, volume 32, page 80, 2013.
- Neal Wadhwa, Michael Rubinstein, Frédo Durand et William T. Freeman : Riesz pyramids for fast phase-based video magnification. *In Computational Photography (ICCP), 2014 IEEE International Conference On*, pages 1–10. IEEE, 2014.
- Weizhuo Wang, John E. Mottershead, Thorsten Siebert et Andrea Pipino : Frequency response functions of shape features from full-field vibration measurements using digital image correlation. *Mechanical Systems and Signal Processing*, volume 28, pages 333–347, avril 2012. ISSN 08883270.
- Newton C. Will et Rafael Cardoso : Comparative analysis between FFT and Kalman filter approaches for harmonic components detection. *In Industry Applications (INDUSCON), 2012 10th IEEE/IAS International Conference On*, pages 1–7. IEEE, 2012.
- Meiliang Wu et Andrew W. Smyth : Application of the unscented Kalman filter for real-time nonlinear structural system identification. *Structural Control and Health Monitoring*, volume 14, pages 971–990, novembre 2007. ISSN 15452255, 15452263.
- Yimin Zhan et Chris K. Mechefske : Robust detection of gearbox deterioration using compromised autoregressive modeling and Kolmogorov–Smirnov test statistic. Part II : Experiment and application. *Mechanical Systems and Signal Processing*, volume 21, pages 1983–2011, juillet 2007. ISSN 08883270.

# Conclusion générale

Cette thèse se place à l'intersection entre plusieurs domaines puisqu'au fil des chapitres, nous évoluons d'un contexte vidéo vers un contexte de dynamique multicorps grâce aux méthodes d'assimilation de données que nous avons utilisées d'un point de vue plutôt stochastique. Ces travaux aboutissent à des applications en analyse modale expérimentale où deux schémas temps réel d'assimilation de données pour l'identification modale sont proposés. Ils sont également intégrés dans des algorithmes de magnification et d'identification sur des structures 3D. Le cadre de l'étude à la fois industriel et centré sur les vidéos rapides est présenté dans le premier chapitre. Il présente les objectifs et la problématique de l'étude qui consiste à développer des méthodes permettant de relier un modèle à un système réel perçu à l'aide de caméras rapides dans un schéma temps réel. L'utilisation de ce modèle permettrait de faire de l'interpolation, de la prédiction et de l'identification dans un but de réalité augmentée, de compression sémantique et de synthèse de vidéos.

Les apports de ces travaux se présentent également sous plusieurs aspects. Dans un contexte vidéo, nous avons d'abord élaboré une méthode permettant de faire de l'interpolation et de la compression sémantique de séquences. Cette méthode est issue de la philosophie couramment utilisée en traitement vidéo, c'est-à-dire d'utiliser des modèles cinématiques élémentaires appliqués à un niveau pixélique. Travailler dans ces conditions assure une certaine généralité aux algorithmes en regard de la scène. En revanche, le modèle ne permet pas d'enrichir suffisamment les observations ni d'introduire des modèles physiques plus élaborés décrivant la scène de la séquence.

Notre méthode utilise une estimation de mouvement pour pouvoir déduire des trajectoires des pixels et ainsi assurer une interpolation. Les limitations actuelles de l'estimation de mouvement aux « petits mouvements » pour des temps de calculs raisonnables nous ont conduit à élaborer un algorithme permettant d'étendre cette estimation aux « grands mouvements ». Constatant que ces limitations sont dues essentiellement au caractère pixélique des traitements qui ne permettent pas d'utiliser la notion d'objet, nous avons introduit une double description d'une scène à la fois eulérienne et lagrangienne. En distinguant alors l'objet au sens vidéo du sens mécanique, nous avons posé des hypothèses permettant de les faire correspondre. Il est alors possible de décrire le mouvement d'un objet d'une séquence par un modèle simple de cinématique du solide indéformable. Nous estimons alors la cinématique de grands mouvements en effectuant une composition de mouvements : dans le repère de l'image l'objet est considéré indéformable entre deux images consécutives et sa trajectoire est décrite par la cinématique de son centre tandis que dans le repère de l'objet on se trouve dans un cadre de « petits mouvements ». Les méthodes d'estimation de mouvements vidéos permettent de déterminer les déformations du contour et de l'intérieur de l'objet.

L'originalité de cette méthode réside dans l'utilisation d'un double formalisme à la frontière de la mécanique et de la vidéo. Elle est applicable pour nos objectifs de compression sémantique et d'interpolation tout en étant générique. Cette méthode aurait de nombreuses applications par exemple lors de la synchronisation de différents flux vidéos en permettant de faire une interpolation et synthétiser

les images au même instant. Par ailleurs, les utilisateurs de vidéos rapides apprécieront l'opportunité d'avoir des séquences compressées sémantiquement, c'est-à-dire non-compressées pour les zones d'intérêt de la vidéo. Cependant ces résultats partiels nous ont poussés à élaborer d'autres méthodes permettant d'utiliser des modèles physiques plus complets.

L'utilisation de modèles physiques avec les observations par caméra rapide nous ont amené à proposer un schéma d'assimilation de données adéquat. Dans la méthode présentée, nous utilisons un algorithme d'assimilation de données séquentiel et non-linéaire permettant de relier un modèle physique dynamique sous forme de modèle d'état avec des informations en déplacement et en vitesse. Cette méthode nous permet ainsi d'enrichir les observations à l'aide du modèle spécifiquement construit pour décrire la scène filmée. L'utilisation du modèle assimilé nous permet de faire de l'identification de paramètres physiques, de la prédiction et de l'interpolation ainsi que de la compression sémantique et de la réalité augmentée.

Le développement de cette méthode s'appuie sur une approche bibliographique multi-domaines apportant un éclairage issu de l'assimilation de données et un autre issu de l'estimation stochastique. Le premier éclairage nous permet de détailler l'évolution de différentes méthodes utilisées par des domaines comme la météorologie pour recaler des modèles à la fois complexes et physiques. L'aspect temps réel n'intervient que pour les méthodes séquentielles qui n'utilisent que les informations passées pour déterminer un état présent : les filtres de Kalman. Nous avons souhaité apporter une contribution sur les filtres de Kalman en présentant une synthèse didactique d'un ensemble de versions du filtre. Cette synthèse n'apporte aucun nouveau développement mais détaille exhaustivement les avantages et les hypothèses d'utilisation de ces filtres et les différentes extensions de ces hypothèses. Ce contexte nous permet d'utiliser avantageusement un modèle pour lisser les données observées, faire du filtrage ou de la prédiction.

Notre approche s'applique pour le cas particulier de modèles continus sous forme différentielle. Le schéma que nous proposons utilise un schéma d'intégration associé à un modèle d'état continu afin de le discrétiser et utilise un filtre de Kalman non-linéaire de la catégorie *derivative-free* évitant d'effectuer une linéarisation. L'espace d'évolution contraint par la nature des informations issues des caméras ne comporte que des positions et des vitesses. Les applications envisagées regroupent les modèles analytiques permettant de décrire les objets de la scène avec un nombre réduit de variables et de paramètres, assurant une utilisation rapide.

Ces résultats peuvent être employés pour augmenter le champ d'application des mesures par caméra sur des grandeurs déduites à l'aide du modèle. Dans le cadre de surveillance de chaînes de production, il serait ainsi possible d'obtenir des informations météorologiques en utilisant de simples caméras.

Parallèlement au développement d'un schéma d'assimilation appliqué à la mécanique, nous avons élaboré un schéma d'assimilation dans le cadre de l'analyse modale expérimentale pour faire de l'identification modale non-linéaire en temps réel. Nous avons nommé cette méthode SAMI (pour *multiSensors Assimilation for Modal Identification*) puisqu'elle autorise le traitement de plusieurs sources d'observation issues de capteurs différents. L'utilisation de caméras rapides offre ainsi une chaîne de mesure sans contact permettant une acquisition multipoint pour de l'estimation transitoire ou aléatoire.

Le développement de cette chaîne s'est appuyé sur les récents résultats sur l'influence de la masse et des câbles sur la précision des fréquences et amortissements identifiés. L'utilisation de caméras rapides est limitée en fréquence d'acquisition mais offre la possibilité d'acquérir un champ en deux dimensions à chaque instant. L'extraction d'informations de ces séquences vidéos a été effectuée en concevant une mire spécifique et une méthode de suivi associée aux caractéristiques géométriques de cette mire. L'utilisation de ces mires permet de gagner en précision puisque le centre peut être

## Conclusion générale

mesuré au dixième de pixel. La méthode SAMI se base sur l'hypothèse de découplage spatio-temporel caractéristique de l'analyse modale de systèmes faiblement non-linéaires. Deux versions sont alors proposées utilisant le même modèle d'évolution basé sur une démodulation multi-porteuses du signal. La première version utilise un modèle spatial théorique de la base des déformées modales de la structure étudiée : c'est le Kinematic-SAMI. La présence d'un modèle spatial et temporel assimilé permet de faire de l'interpolation mais également de la prédiction : le Kinematic-SAMI offre la possibilité d'observer une partie d'une structure tout en estimant par expansion modale la position de la totalité de la structure. Nous proposons ainsi un algorithme d'utilisation avantageuse pour une identification modale 3D puisque celui-ci n'utilise pas de stéréo-reconstruction, réduisant ainsi son temps de calcul. Contrairement aux autres apports, nous n'avons pas présenté d'implémentation de cet algorithme dans ce document. La deuxième version utilise une fonction d'observation différente afin de projeter le vecteur d'état dans l'espace d'observation sans utiliser de modèle spatial à priori : c'est le Blind-SAMI. L'utilisation d'un tel filtre permet d'accéder à l'identification des déformées modales expérimentales. Nous avons pu montrer sur un cas réel d'une structure 2D que ce filtre était plus sensible au bruit étant donné qu'il ne bénéficie pas du lissage par un modèle spatial, mais qu'il prend en compte le comportement réel de la structure. Les applications effectuées avec le Kinematic-SAMI utilisant le modèle spatial ne sont par ailleurs plus envisageables dans ce cas.

Afin de donner un aperçu d'applications utilisant la synthèse de vidéo, nous avons présenté un nouvel algorithme de magnification basé sur les SAMI présentés plus haut. Cet algorithme fonctionne comme une loupe modale puisqu'il permet de modifier le mouvement de la structure selon des coefficients sur chaque mode de celle-ci. La synthèse de la vidéo a été élaborée en détournant le mécanisme de compression/reconstruction par estimation de mouvement puis compensation de mouvements. À partir du champ initial calculé, il est proposé de le modifier selon le champ interpolé issu de la décomposition modale identifiée magnifiée. Les pixels de l'image réelle sont alors déplacés par compensation de mouvement pour représenter le nouveau comportement modal de la structure. Ces différents apports pourraient être utilisés dans le cadre notamment de *GVT*<sup>7</sup>. Ils apporteraient un protocole expérimental moins lourd à mettre en place et surtout des résultats en temps réel en fonction des modèles théoriques avec le Kinematic-SAMI et en fonction des déformées modales expérimentales avec le Blind-SAMI. La magnification permettrait alors de venir visualiser les résultats en amplifiant ou atténuant certaines composantes, de manière à vérifier chaque mode individuellement.

## Perspectives

### Méthode d'interpolation de grands mouvements

L'implémentation proposée dans ces travaux est déjà fonctionnelle et permet de traiter des cas d'application concrets. Un projet industriel est en cours de lancement pour un prototype d'implémentation matérielle de la méthode. Comme il est par ailleurs précisé à plusieurs reprises, l'extraction d'objets est effectuée pour des images 2D. La conception modulaire de l'algorithme nous permet d'envisager l'utilisation de capteurs permettant d'obtenir des informations 3D sur la scène. L'extraction des objets et le développement théorique sera différent, mais la méthode pourra être transposée et sera d'autant plus avantageuse que la quantité de données en 3D est plus importante qu'en 2D.

### Assimilation de données vidéos avec un modèle mécanique

Le schéma d'assimilation de données proposé dans le chapitre 3 n'utilise pas d'outils concer-

7. *Ground Vibration Test* qui sont les tests de validation en vibration opérés sur les aéronefs.

nant les performances des modèles. Comme indiqué dans l'état de l'art sur les filtres de Kalman, le sujet existe et il serait intéressant de compléter notre approche. Lors de l'implémentation, nous avons fait le choix d'utiliser la représentation réduite d'un solide indéformable qui correspondait bien au cas test réel présenté. Cependant ce cas n'est pas généralisable et il serait judicieux d'utiliser une représentation à l'aide de descripteurs locaux en modifiant la fonction d'observation choisie. Un plus grand nombre d'observations pourra par ailleurs faciliter la détermination de l'état de l'objet. Un passage en 3D est envisagé pour faciliter l'extraction de l'objet. Le modèle mécanique utilisé est généralement déjà adapté à trois dimensions. Ces travaux ont été centrés sur l'utilisation de modèles analytiques. Une application sur des modèles éléments finis pourrait ouvrir d'autres perspectives. L'utilisation de cette méthode à des fins de capteurs logiciels est déjà en développement dans le cadre du projet FUI EUGENE. Ce projet vise une application sur des chaînes de production en allant dans une direction plus orientée contrôle.

### **Schémas d'assimilation de données vidéos dans le contexte de l'analyse modale**

Les schémas SAMI constituent un premier résultat solide qu'il est nécessaire de développer pour obtenir des applications industrielles intéressantes. D'un côté, il est nécessaire de continuer à perfectionner le schéma. Dans cette perspective, il est possible de proposer une fonction d'évolution plus précise en partant du constat fait dans la figure 4.11 sur les imperfections de la transformée de Hilbert. Compte-tenu de l'état de l'art évoqué sur le théorème de Bedrosian, une autre perspective consistera à étudier les limites et les effets de modes doubles sur la démodulation par transformée de Hilbert. Si la partie concernant l'extraction de positions à partir de mires est fonctionnelle, elle n'en est pas moins perfectible. Tout d'abord, nous n'avons pas pu mener une étude rigoureuse sur l'amélioration effective de nos nouvelles mires par rapport aux anciennes. Leur utilisation pose question par rapport aux techniques de mesures de champ, permettant de récupérer plus d'informations. Une des perspectives de ces travaux consisterait à adapter la fonction d'observation pour utiliser des méthodes de mesure de champ. Un schéma hybride utilisant la covariance pour prendre la décision de figer les modes estimés pourrait être proposé. Ce schéma permettrait alors d'avoir à la fois les avantages du modèle spatial de Kinematic-SAMI tout en estimant les déformées modales expérimentales comme Blind-SAMI. Nous avons proposé une application intéressante de ces schémas pour la magnification et une application à l'identification d'une structure 3D pour le Kinematic-SAMI. Une perspective proche consiste à valider cette méthode sur un cas test. Nous utiliserons la maquette GARTEUR (Balmes et Wright, 1997) pour faire cette validation. Une application du Blind-SAMI sur une structure possédant une non-linéarité comportementale est en cours dans le cadre du projet FUI CLIMA. Elle permettra de valider les performances d'estimation des déformées modales expérimentales.

# Publications associées à ces travaux

Adrien Goeller, Jean-Luc Dion, Ronan Le Breton et Thierry Soriano : Real-time contactless identification of nonlinear modal parameters using a Kalman filter in a vision context. *Mechanical Systems and Signal Processing (Soumis)*.

Adrien Goeller, Jean-Luc Dion, Ronan Le Breton et Thierry Soriano : Identification modale temps réel par caméra rapides et assimilation de données : Une nouvelle méthode de magnification. *In Actes Du 23e Congrès Français de Mécanique*, Lille, août 2017a.

Adrien Goeller, Jean-Luc Dion, Ronan Le Breton et Thierry Soriano : Real time modal identification of vibrating beam using Kalman filters and high speed cameras. *In 35th Society For Experimental Mechanics IMACS*, Garden Grove, California USA, février 2017b.

Adrien Goeller, Jean Luc Dion, Ronan Le Breton, Thierry Soriano et Bernard Roux : Dynamic system stochastic identification mixed with video processing : Validation on a real case. *In Mechatronics (MECATRONICS)/17th International Conference on Research and Education in Mechatronics (REM), 2016 11th France-Japan & 9th Europe-Asia Congress On*, pages 146–151. IEEE, 2016.

Adrien Goeller, Jean-Luc Dion, Thierry Soriano et Bernard Roux : Real Time Dynamic System Stochastic Identification in Video Capture for Data Compression, Image Interpolation, Prediction, and Augmented Reality. volume 8, Boston, Massachusetts, USA, août 2015. ASME. ISBN 978-0-7918-5718-1.



# Annexes

## A. Application dans le cadre de l'interpolation vidéo

Parmi les applications visées, la méthode développée dans le paragraphe précédent utilise un modèle de déplacement des objets élémentaires. Ce modèle ne permet pas de faire une prédiction très précise spatialement et sera de plus en plus faussée avec l'éloignement temporel de la dernière valeur observée. Le cas de l'interpolation est plus particulier. L'encadrement entre deux mesures permet de minimiser l'erreur à l'abord de ces deux bornes. Cependant, il nous a semblé intéressant de développer sur la limite qu'il pouvait exister entre une interpolation soumise aux limites imposées par le théorème d'échantillonnage et l'extrapolation du modèle cinématique que nous employons pour caractériser cette trajectoire.

### 1. Trajectoire et échantillonnage

Lors de la mesure d'une trajectoire sur le plan du capteur, il est primordial de savoir quelle est la précision sur la mesure afin de pouvoir distinguer les phénomènes visibles des phénomènes trop évanescents pour être vus. Cette problématique est d'importance dans le domaine des vidéos rapides où une mauvaise appréciation peut engendrer des surcoûts non négligeables.

Aussi, il s'agit ici de relier le pas d'échantillonnage temporel  $\Delta t$  avec le pas d'échantillonnage, que nous noterons  $\Delta s$  défini sur l'abscisse curviligne de la trajectoire.

Le cadre imposé par la vidéo rapide est particulier puisqu'il vient modifier notre perception des événements en changeant une chaîne de perception complètement intégrée par notre cerveau. En effet dans une chaîne vidéo classique, les événements réels continus sont discrétisés selon une fréquence d'échantillonnage  $f_e$  définie en images par secondes (*frame per second - fps*) et lors de la reconstruction du signal discrétisé, la fréquence d'affichage utilisée ( $f_{aff}$ ) est la même. Ce processus nous permet de poser la relation simple entre la vitesse réelle  $v_r(t)$  du phénomène et la vitesse apparente  $v_v(t)$  :

$$v_v(t) = \frac{f_{aff}}{f_e} \cdot v_r(t) \quad (111)$$

ce qui dans le cas classique revient à :

$$\begin{cases} v_v(t) &= \frac{f_{aff}}{f_e} \cdot v_r(t) \\ f_{aff} &= f_e \end{cases} \quad (112)$$

## A. Application dans le cadre de l'interpolation vidéo

On voit ainsi rapidement comment ralentir ou accélérer la perception des phénomènes : en venant modifier ce rapport de fréquences. La modification de la fréquence reste limitée. L'accélération est aisée à produire, que ce soit en augmentant la fréquence d'affichage ou en diminuant la fréquence d'échantillonnage. Ralentir un phénomène est en revanche beaucoup plus difficile puisqu'elle nécessite d'avoir des informations supplémentaires.

C'est ce que proposent les caméras rapides en venant échantillonner les phénomènes avec une plus grande fréquence. Ainsi on peut définir un facteur de ralentissement par le rapport des fréquences d'affichage et des fréquences d'échantillonnage

$$v_v(t) = \nu \cdot v_r(t) \quad (113)$$

ce qui dans le cas classique revient à :

$$\begin{cases} v_v(t) &= \nu \cdot v_r(t) \\ \nu &= \frac{f_{aff}}{f_e} \end{cases} \quad (114)$$

Or la relation entre le pas de temps  $\Delta t = \frac{1}{f_e}$  et le pas spatial entre deux positions mesurées  $\Delta s$  dépend de la vitesse, puisque :

$$\Delta s = |v_v(t)| \cdot \Delta t \iff \Delta s = f_{aff} \cdot |v_r(t)| \cdot \Delta t^2 \quad (115)$$

Cette relation fait intervenir le module de la vitesse du phénomène mesuré : l'échantillonnage des positions est en réalité variable en fonction de la vitesse du phénomène. Ainsi une vitesse constante engendrerait un échantillonnage constant, une augmentation de la vitesse dégraderait l'échantillonnage tandis qu'une diminution de la vitesse diminuerait le pas spatial entre deux positions successives.

Cette considération sur l'échantillonnage, ajoutée au fait que les trajectoires des objets observés ne sont pas toujours stationnaires, ne permettent pas d'invoquer le théorème d'échantillonnage. Pourtant il existe bien une limite à ces acquisitions, que l'on va détailler dans un cas local. Dans l'espace paramétré du plan capteur, on pose la trajectoire  $s(t)$  définie par le couple  $(x(t), y(t))$ . Supposons que cette trajectoire est parcourue à une vitesse  $\vec{v}(t)$  suffisamment lente pour que la discrétisation reflète le phénomène continu de manière acceptable. On introduit une perturbation transitoire de trajectoire  $(\delta x(t), \delta y(t))$  parcourue à une vitesse localement constante  $\vec{v}_\delta$  telle que  $\vec{v}_\delta \gg \max(\vec{v}(t))$ , que l'on caractérise par un temps caractéristique  $\Delta t_\delta$  (voir figure 30). La question revient à trouver la condition sur les paramètres pour pouvoir observer le phénomène. Ici le pas de temps minimum pour pouvoir reconstruire le phénomène se résume à avoir  $\Delta t$  tel que :

$$\Delta t \geq \frac{1}{2} \cdot \Delta t_\delta, \text{ soit pour une vitesse variant peu } \Delta s \geq |v_\delta| \cdot \frac{1}{2} \cdot \Delta t_\delta \quad (116)$$

Comme attendu, l'échantillonnage ne permettra pas de capturer des phénomènes dont le temps caractéristique se trouve inférieur à deux fois le pas de temps séparant deux acquisitions. Cependant, il est important de noter qu'au niveau global de la trajectoire, il faut considérer l'influence de la variation du module de la vitesse réelle du phénomène observé, pas forcément connue. Aussi, toute reconstruction par interpolation ne peut se faire que dans la limite d'un facteur 2 par rapport au pas de temps initial.



Cette forme de l'erreur montre l'influence de la variation du rayon de courbure sur la qualité de l'interpolation. Ce modèle sera donc particulièrement peu performant dans les situations où la vitesse est discontinue (rebond, mauvaise projection d'un mouvement de tourbillon...).

### 3. Cas d'application sur une simulation de balle rebondissante

Afin de fixer les idées, nous avons choisi un cas simple d'une scène à fond constant avec un objet dont les mouvements mettent à mal les hypothèses d'application de notre modèle d'interpolation.

Nous avons ainsi choisi de modéliser une balle rebondissante. Notre modèle est indéformable, étant donné que les hypothèses à tester sont au niveau de la trajectoire. Le rebond a été modélisé par l'inversion de la composante verticale de la vitesse en-deçà d'une position limite du centre de gravité représentant le sol en plus du facteur de rebond. Ce modèle nous permet à la fois d'avoir des variations du module de la vitesse, mais également des variations continues (sommet des rebonds) et discontinues (rebonds) de la direction.

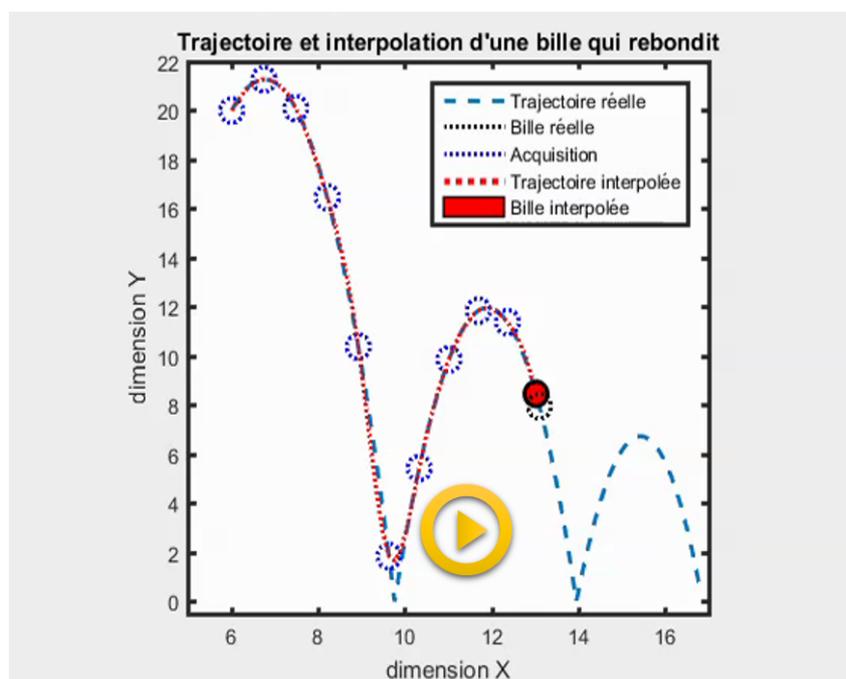
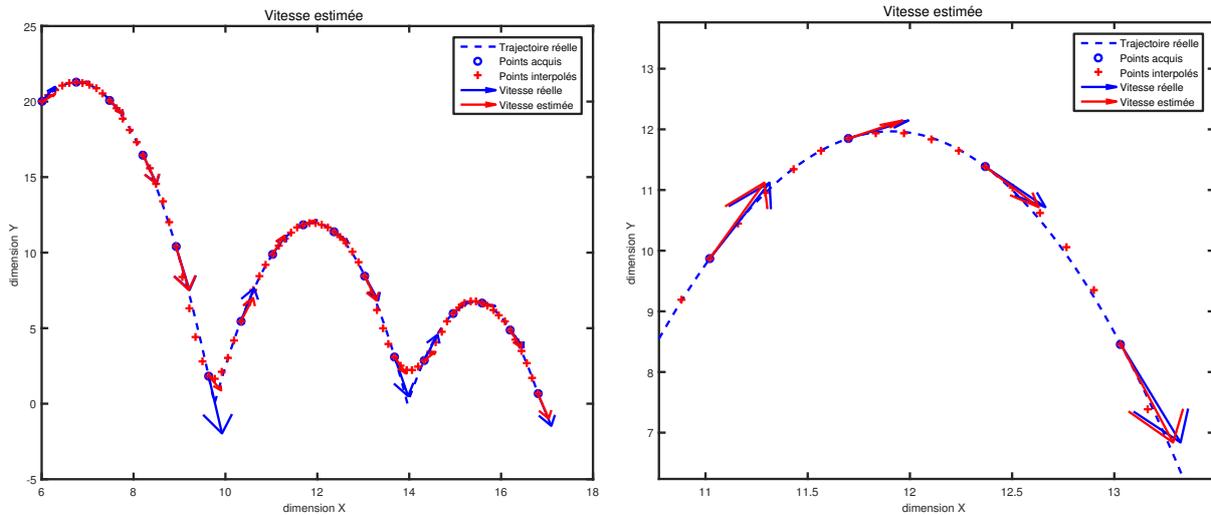


Figure 31 – Vidéo présentant la vidéo interpolée avec notre méthode.

La vidéo de la figure 31 montre la discrétisation choisie pour la caméra. Deux phénomènes peuvent être vus ici :

- l'influence du module de la vitesse sur l'échantillonnage spatiale de la trajectoire ;
- la difficulté des splines de modéliser des vitesses non continues.

Afin de mieux mettre en évidence les difficultés du modèle sur ce cas test, nous avons comparé sur la figure 32 la vitesse réelle en terme de module et de direction avec la vitesse estimée de l'interpolation en spline du déplacement. La figure 32b nous montre alors que l'estimation des vitesses **sur les points d'acquisition favorables au modèle** est déjà entachée d'une erreur. La figure souligne également que les erreurs en terme de module et de direction sont maximisées dans les zones de rebond.



(a) Comparaison des vitesses aux points interpolés sur l'ensemble de la trajectoire

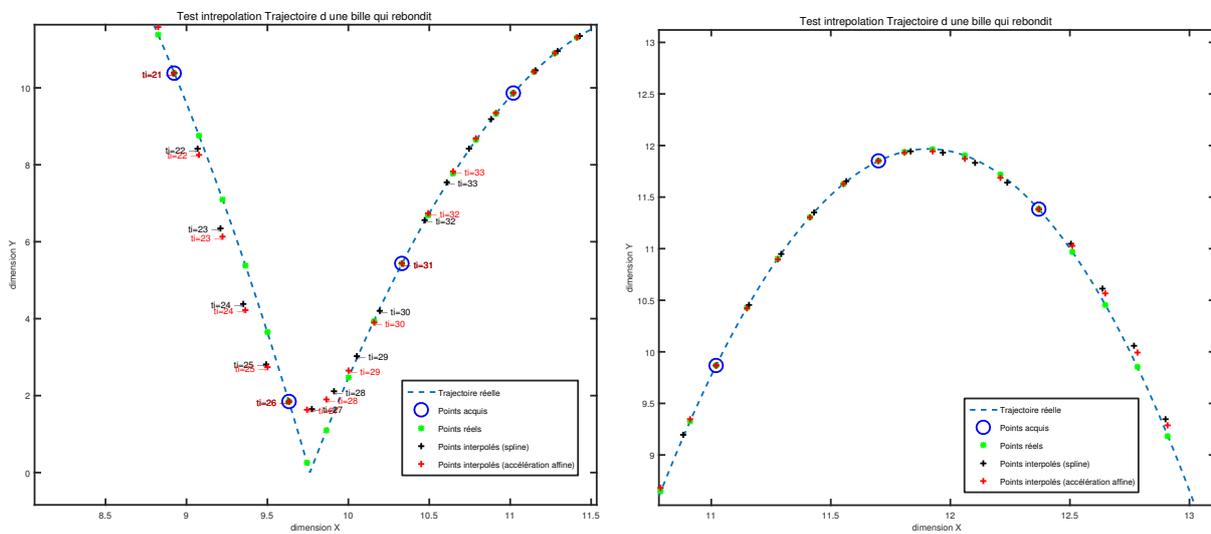
(b) Comparaison des vitesses aux points interpolés : zoom sur une partie de la trajectoire favorable au modèle de spline qu'est le sommet du rebond

**Figure 32** – Comparaison des vitesses aux points interpolés

Bien que les problématiques de dérivation numérique de données limiteraient l'utilisation de la vitesse et de surcroît de l'accélération, nous profitons de l'utilisation d'un cas simulé pour avoir accès à une mesure non bruitée de l'accélération et ainsi pouvoir comparer une interpolation de l'accélération avec celle du déplacement. La figure 33a nous montre que dans le cas d'une vitesse discontinue, aucune des deux grandeurs n'offre une estimation des positions interpolées significativement meilleure. Cependant, on remarquera que l'accélération interpolée permet de mieux rendre compte de la dynamique du mouvement que l'interpolation par spline qui compte de la trajectoire tout en ayant un décalage temporel (figure 33b).

Ce cas d'application nous permet de bien mettre en évidence les difficultés de notre modèle d'interpolation par spline par rapport à une trajectoire quelconque. Son extrapolation nous permet d'avoir une bonne approximation de la trajectoire réelle lorsque les vitesses sont continues tandis qu'il est peu performant en dehors de ses hypothèses d'application optimale notamment sur des portions où la vitesse est discontinue. Comme attendu, nous avons pu également mettre en évidence que la dynamique de la scène n'était pas restituée par le modèle cinématique. L'interpolation/extrapolation ne permettra pas de rendre compte de la véritable dynamique de la scène. La qualité de l'extrapolation que nous obtenons avec ce modèle élémentaire reflète le peu d'informations qu'il apporte. Les erreurs sont engendrées par un comportement du modèle inadéquat face aux données rencontrées : c'est l'inconvénient des propriétés adaptatives de cette méthode. En revanche, elle est très efficace pour faire de la compression sémantique ou de l'interpolation dans les limites de la dynamique capturée par l'acquisition.

## A. Application dans le cadre de l'interpolation vidéo



(a) Comparaison de l'interpolation de positions par spline et de celle linéaire de l'accélération avec la trajectoire réelle de la balle : zoom sur l'endroit de discontinuité de la vitesse qu'est le rebond

(b) Comparaison de l'interpolation de positions par spline et de celle linéaire de l'accélération avec la trajectoire réelle de la balle : zoom sur le sommet du rebond

**Figure 33** – Comparaison de l'interpolation de positions par spline et de celle linéaire de l'accélération avec la trajectoire réelle de la balle

## B. Modèle de la Règle à 2 points de contacts

La règle glisse sur le plan  $(O; \vec{x}_0, \vec{y}_0)$  confondu avec le plan  $(G; \vec{x}_r, \vec{y}_r)$ . Le glissement de la règle sur une table est modélisé par un glissement ponctuel en 2 points de contact  $O_1$  et  $O_2$ . La modélisation adoptée et les notations utilisés sont représentés sur la figure 34. On note  $\vec{GO}_1 = l_1 \cdot \vec{x}_r$  et  $\vec{GO}_2 = -l_2 \cdot \vec{x}_r$ .

L'ensemble est soumis à l'accélération de la pesanteur suivant  $-\vec{z}_0$  dans le repère  $\mathcal{R}_0 (O_0; \vec{x}_0, \vec{y}_0, \vec{z}_0)$  réputé galiléen. On note  $\mathcal{R}_r$  le repère lié à la règle  $(G; \vec{x}_r, \vec{y}_r, \vec{z}_r)$ .

$\mu$  désigne le coefficient de frottement de glissement de la règle sur le plan.

On note  $\alpha$  l'angle entre  $(\vec{x}_0, \vec{x}_r)$ .

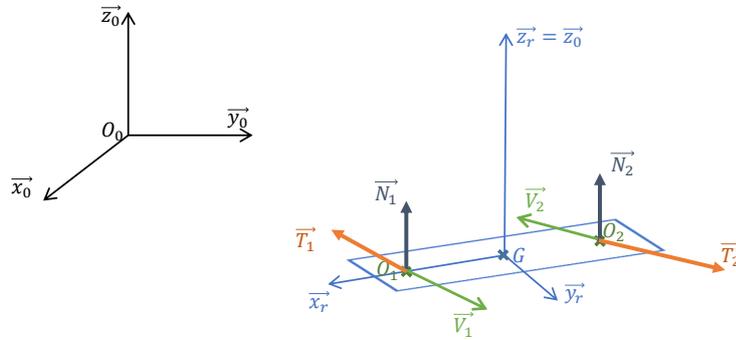


Figure 34 – Modèle de la règle glissant sur la table

### 1. Bilan des actions mécaniques

La règle est soumise :

— au poids

$$\{\mathcal{T}_{pes}\} = \left\{ \begin{array}{c} -m \cdot g \cdot \vec{z}_0 \\ \vec{0} \end{array} \right\}_G$$

— à l'action de contact en  $O_1$

$$\{\mathcal{T}_{O_1}\} = \left\{ \begin{array}{c} \vec{N}_1 + \vec{T}_1 \\ \vec{0} \end{array} \right\}_{O_1}$$

avec :  $\vec{N}_1 = N_1 \cdot \vec{z}_0$  et  $\vec{T}_1 = -\mu \cdot N_1 \frac{\vec{V}_1}{\|\vec{V}_1\|}$

— à l'action de contact en  $O_2$

$$\{\mathcal{T}_{O_2}\} = \left\{ \begin{array}{c} \vec{N}_2 + \vec{T}_2 \\ \vec{0} \end{array} \right\}_{O_2}$$

avec :  $\vec{N}_2 = N_2 \cdot \vec{z}_0$  et  $\vec{T}_2 = -\mu \cdot N_2 \frac{\vec{V}_2}{\|\vec{V}_2\|}$

## B. Modèle de la Règle à 2 points de contacts

$\vec{T}_1$  et  $\vec{T}_2$  tout comme  $\vec{V}_1$  et  $\vec{V}_2$  appartiennent au plan de glissement ( $O; \vec{x}_0, \vec{y}_0$ ).

### 2. Cinématique du système

Torseur cinématique de la règle :

$$\{\mathcal{V}_{(r/\mathcal{R}_0)}\} = \left\{ \begin{array}{c} \dot{\alpha} \cdot \vec{z}_0 \\ \overrightarrow{V_{G \in r/\mathcal{R}_0}} \end{array} \right\}_G$$

Avec :  $\overrightarrow{V_{G \in \mathcal{R}_r/\mathcal{R}_0}} = \overrightarrow{V_{G/\mathcal{R}_0}} = \begin{pmatrix} v_x \\ v_y \\ 0 \end{pmatrix}_{(\vec{x}_0, \vec{y}_0, \vec{z}_0)}$

d'où :

$$\begin{aligned} \vec{V}_1 &= \overrightarrow{V_{O_1 \in r/\mathcal{R}_0}} = \overrightarrow{V_{G/\mathcal{R}_0}} + \overrightarrow{O_1 G} \wedge \dot{\alpha} \cdot \vec{z}_0 \\ &= \overrightarrow{V_{G/\mathcal{R}_0}} - l_1 \cdot \vec{x}_r \wedge \dot{\alpha} \cdot \vec{z}_0 \\ &= \overrightarrow{V_{G/\mathcal{R}_0}} + l_1 \dot{\alpha} \cdot \vec{y}_r \\ &= \overrightarrow{V_{G/\mathcal{R}_0}} + l_1 \dot{\alpha} \cdot (-\sin \alpha \vec{x}_0 + \cos \alpha \vec{y}_0) \\ &= \begin{pmatrix} v_x - l_1 \dot{\alpha} \sin \alpha \\ v_y + l_1 \dot{\alpha} \cos \alpha \\ 0 \end{pmatrix}_{(\vec{x}_0, \vec{y}_0, \vec{z}_0)} \end{aligned}$$

de même :

$$\begin{aligned} \vec{V}_2 &= \overrightarrow{V_{O_2 \in r/\mathcal{R}_0}} = \overrightarrow{V_{G/\mathcal{R}_0}} + \overrightarrow{O_2 G} \wedge \dot{\alpha} \cdot \vec{z}_0 \\ &= \overrightarrow{V_{G/\mathcal{R}_0}} + l_2 \cdot \vec{x}_r \wedge \dot{\alpha} \cdot \vec{z}_0 \\ &= \overrightarrow{V_{G/\mathcal{R}_0}} - l_2 \dot{\alpha} \cdot \vec{y}_r \\ &= \begin{pmatrix} v_x + l_2 \dot{\alpha} \sin \alpha \\ v_y - l_2 \dot{\alpha} \cos \alpha \\ 0 \end{pmatrix}_{(\vec{x}_0, \vec{y}_0, \vec{z}_0)} \end{aligned}$$

d'où

$$\begin{aligned} \|\vec{V}_1\| &= \sqrt{(v_x - l_1 \dot{\alpha} \sin \alpha)^2 + (v_y + l_1 \dot{\alpha} \cos \alpha)^2} \\ \|\vec{V}_2\| &= \sqrt{v_x^2 + v_y^2 + l_1^2 \dot{\alpha}^2 + 2l_1 \dot{\alpha} (-v_x \sin \alpha + v_y \cos \alpha)} \end{aligned}$$

### 3. Cinétique

$m$  représente la masse de la règle répartie de façon homogène.

On donne le torseur cinétique du système :

$$\{\mathcal{C}_{(r/\mathcal{R}_0)}\} = \left\{ \begin{array}{c} m \cdot \overrightarrow{V_{G/\mathcal{R}_0}} \\ \overrightarrow{\sigma_{G \in r/\mathcal{R}_0}} \end{array} \right\}_G$$

L'opérateur d'inertie de la règle est donné par  $\bar{I}_{(G,r)} = \frac{m}{12} \begin{pmatrix} l^2 + h^2 & 0 & 0 \\ 0 & L^2 + h^2 & 0 \\ 0 & 0 & L^2 + l^2 \end{pmatrix}_{(\vec{x}_r, \vec{y}_r, \vec{z}_r)}$

où  $L$ ,  $l$  et  $h$  sont respectivement la longueur, la largeur et la hauteur de la règle ( $h$  est négligeable).  
D'où l'expression du moment cinétique :

$$\overrightarrow{\sigma_{G \in r / \mathcal{R}_0}} = \bar{I}_{(G,r)} \cdot \dot{\alpha} \vec{z}_0 + \vec{0} = \dot{\alpha} \cdot \frac{m(L^2 + l^2)}{12} \vec{z}_0$$

#### 4. Dynamique

Le torseur dynamique s'exprime sous la forme :

$$\{\mathcal{D}_{(r/\mathcal{R}_0)}\} = \left\{ \begin{array}{c} m \cdot \overrightarrow{\Gamma_{G/\mathcal{R}_0}} \\ \overrightarrow{\delta_{G \in r / \mathcal{R}_0}} \end{array} \right\}_G$$

avec :

$$\overrightarrow{\Gamma_{G/\mathcal{R}_0}} = \left. \frac{d\overrightarrow{V_{G/\mathcal{R}_0}}}{dt} \right|_{\mathcal{R}_0} = \begin{pmatrix} \dot{v}_x \\ \dot{v}_y \\ 0 \end{pmatrix}_{(\vec{x}_0, \vec{y}_0, \vec{z}_0)}$$

et

$$\overrightarrow{\delta_{G \in r / \mathcal{R}_0}} = \left. \frac{d\overrightarrow{\sigma_{G \in r / \mathcal{R}_0}}}{dt} \right|_{\mathcal{R}_0} + \vec{0} = \ddot{\alpha} \cdot \frac{m(L^2 + l^2)}{12} \vec{z}_0$$

#### 5. Principe Fondamental de la Dynamique

Le PFD s'écrit :

$$\{\mathcal{D}_{(r/\mathcal{R}_0)}\}_G = \{\mathcal{T}_{pes}\}_G + \{\mathcal{T}_{O_1}\}_G + \{\mathcal{T}_{O_2}\}_G$$

$$\left\{ \begin{array}{c} m \cdot \overrightarrow{\Gamma_{G/\mathcal{R}_0}} \\ \overrightarrow{\delta_{G \in r / \mathcal{R}_0}} \end{array} \right\}_G = \left\{ \begin{array}{c} -mg \cdot \vec{z}_0 \\ \vec{0} \end{array} \right\}_G + \left\{ \begin{array}{c} \vec{N}_1 + \vec{T}_1 \\ l_1 \cdot \vec{x}_r \wedge (\vec{N}_1 + \vec{T}_1) \end{array} \right\}_G + \left\{ \begin{array}{c} \vec{N}_2 + \vec{T}_2 \\ -l_2 \cdot \vec{x}_r \wedge (\vec{N}_2 + \vec{T}_2) \end{array} \right\}_G$$

soit :

$$m \cdot \begin{pmatrix} \dot{v}_x \\ \dot{v}_y \\ 0 \end{pmatrix}_{(\vec{x}_0, \vec{y}_0, \vec{z}_0)} = \begin{pmatrix} -\mu \left( \frac{N_1}{\|\vec{v}_1\|} (v_x - l_1 \dot{\alpha} \sin \alpha) + \frac{N_2}{\|\vec{v}_2\|} (v_x + l_2 \dot{\alpha} \sin \alpha) \right) \\ -\mu \left( \frac{N_1}{\|\vec{v}_1\|} (v_y + l_1 \dot{\alpha} \cos \alpha) + \frac{N_2}{\|\vec{v}_2\|} (v_y - l_2 \dot{\alpha} \cos \alpha) \right) \\ -mg + N_1 + N_2 \end{pmatrix}_{(\vec{x}_0, \vec{y}_0, \vec{z}_0)}$$

## B. Modèle de la Règle à 2 points de contacts

et l'équation du moment dynamique :

$$\begin{pmatrix} 0 \\ 0 \\ \ddot{\alpha} \cdot \frac{m(L^2+l^2)}{12} \end{pmatrix}_{(\vec{x}_r, \vec{y}_r, \vec{z}_0)} = \begin{pmatrix} 0 \\ -l_1 \cdot N_1 + l_2 \cdot N_2 \\ -\mu \left[ \frac{N_1 \cdot l_1}{\|\vec{V}_1\|} [\cos \alpha (v_y + l_1 \dot{\alpha} \cos \alpha) - \sin \alpha (v_x - l_1 \dot{\alpha} \sin \alpha)] + \dots \right. \\ \left. - \frac{N_2 \cdot l_2}{\|\vec{V}_2\|} [\cos \alpha (v_y - l_2 \dot{\alpha} \cos \alpha) - \sin \alpha (v_x + l_2 \dot{\alpha} \sin \alpha)] \right] \end{pmatrix}_{(\vec{x}_r, \vec{y}_r, \vec{z}_0)}$$

$$\begin{pmatrix} 0 \\ 0 \\ \ddot{\alpha} \cdot \frac{m(L^2+l^2)}{12} \end{pmatrix}_{(\vec{x}_r, \vec{y}_r, \vec{z}_0)} = \begin{pmatrix} 0 \\ -l_1 \cdot N_1 + l_2 \cdot N_2 \\ -\mu \left[ \frac{N_1 \cdot l_1}{\|\vec{V}_1\|} [v_y \cos \alpha - v_x \sin \alpha + l_1 \dot{\alpha}] + \dots \right. \\ \left. - \frac{N_2 \cdot l_2}{\|\vec{V}_2\|} [v_y \cos \alpha - v_x \sin \alpha - l_2 \dot{\alpha}] \right] \end{pmatrix}_{(\vec{x}_r, \vec{y}_r, \vec{z}_0)}$$

## C. Comparaison des performances des méthodes de vision

Les méthodes de vision présentées dans la sous-section IV.A. permettent d'extraire des informations sur l'image. Elles utilisent des marqueurs spécifiques conçus pour maximiser leur précision. Nous utilisons une séquence test effectuée sur la poutre encastree libre préparée avec des marqueurs tous les centimètres. Une image de la séquence est présentée dans la figure 35.



**Figure 35** – Image issue de la séquence de la poutre munie de mire HBF traitée par seuillage : le centre déterminé est en rouge.

Les mire utilisées sont les mire HBF présentées dans le sous-section IV.A.. Elles sont choisies d'un diamètre de 0,5 cm. Afin de pouvoir exploiter plus finement les résultats, l'essai est séparé en trois parties. La partie avant le choc permet d'observer les sources de bruits présentes sur le montage et de comparer l'exactitude des signaux extraits entre eux. L'étude pendant le choc nous permettra à l'aide d'une comparaison de spectres de voir les niveaux de bruit et de quantifier le contenu spectral des différents signaux. La comparaison sur la fin de l'acquisition permet de vérifier les dérives de la mesure. On remarquera sur les données qu'il reste encore un mouvement résiduel de la structure au début de la séquence qui correspond au premier mode. On utilise également un capteur de déplacement placé à 8 cm de l'encastrement afin de comparer les mesures obtenues.

Cette comparaison va permettre :

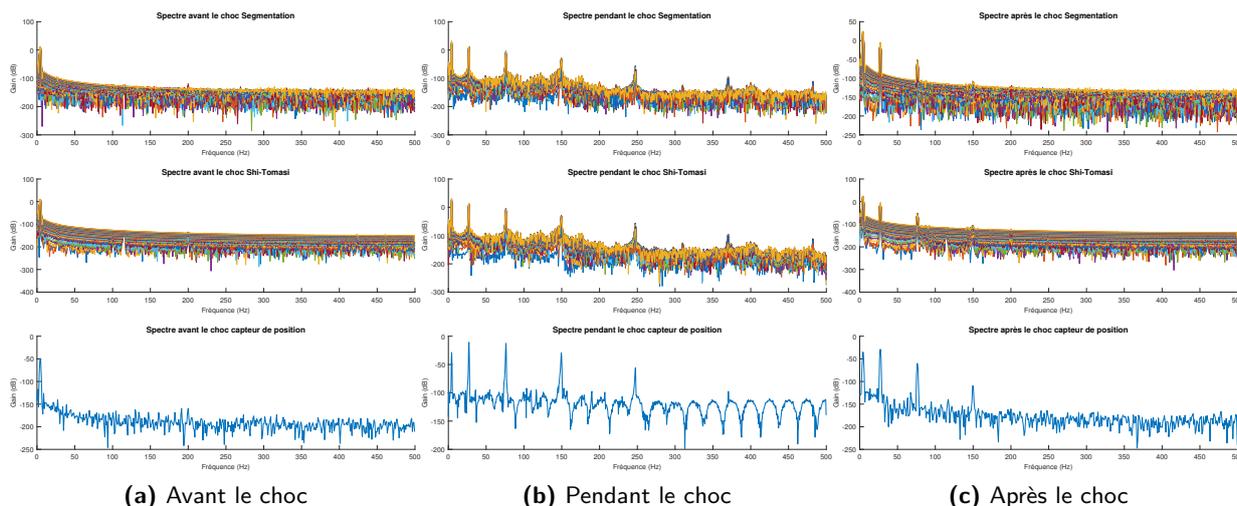
- d'obtenir les niveaux de bruit et les erreurs que génèrent les différentes méthodes ;
- de pouvoir caractériser les sources de bruits qui perturbent la mesure sur ce montage ;
- d'utiliser les spectres de fréquences propres de la structure pour comparer les informations spectrales contenues dans les différents signaux.

En regardant les figures temporelles données sur la figure 37, on peut noter que les valeurs correspondent bien entre les différents signaux. On remarque cependant un léger déphasage du capteur de position qui correspond certainement à une erreur de positionnement du capteur par rapport à la mire. Les signaux sont moyennés selon leur valeur moyenne sur la totalité de la plage d'enregistrement et il est étonnant de constater que la méthode de segmentation accuse un décalage vertical non négligeable au départ de l'acquisition, décalage qui disparaît à la fin de l'enregistrement. On a donc une dérive de la mesure pour la segmentation, ce qui n'est pas le cas pour le suivi. En regardant attentivement la fin de l'acquisition, on s'aperçoit que les déplacements issus de mesures par des méthodes vidéo semblent osciller autour du capteur de déplacement. Nous proposons de regarder les spectres afin de valider le contenu spectrale de cette séquence test.

Les mesures spectrales sont présentées figure 36. D'une manière générale, le spectre des données extraites par le suivi de Shi-Tomasi est plus lisse. Les mesures faites par le capteur de position ont un bruit légèrement supérieur aux mesures par vidéo. On remarquera également que le bruit de mesure pendant le choc est similaire pour les deux méthodes d'estimation, les différences ne se voyant qu'avant et après. Les fréquences propres identifiées sur les différents spectres sont regroupées dans le tableau 9. Cette analyse en trois temps nous permet d'identifier les sources de bruit pouvant intervenir sur ce montage. Les fréquences 99,89 Hz et 199,8 Hz correspondent à la fréquence de l'éclairage que l'on utilise. Cette fréquence parasite pourra être supprimée en utilisant un éclairage

## C. Comparaison des performances des méthodes de vision

diffus synchronisé, ce qui n'est pas notre cas. La fréquence de 115,4 Hz semble elle plus inhérente au montage qu'à la mesure puisqu'elle est présente également sur le capteur de position. Le suivi de Shi-Tomasi permet d'obtenir des harmoniques jusqu'à la fréquence de coupure (500 Hz, les mesures sont faites à 1000 Hz), tandis que la segmentation est trop bruitée pour pouvoir aller au-delà de 400 Hz. La segmentation laisse transparaître deux harmoniques de la fréquence d'alimentation de l'éclairage, montrant ainsi une plus grande sensibilité aux variations de luminosité. Deux autres fréquences sont ainsi visibles en dehors du choc uniquement sur le spectre de Shi-Tomasi, ne permettant pas de conclure sur l'origine de ces fréquences.

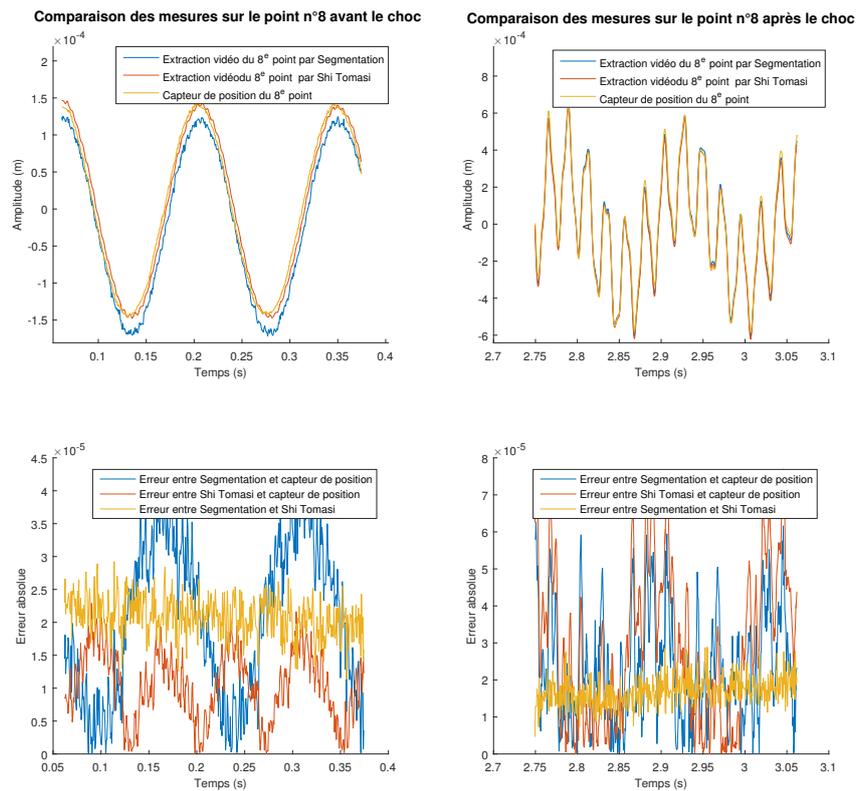


**Figure 36** – Spectre avant, pendant et à la fin du choc pour comparer les mesures par extraction vidéo et les mesures de position

**Table 9** – Fréquences propres extraites des spectres

Spectre Segmentation (Hz)			Spectre Shi-Tomasi (Hz)			Spectre Capteur de Position (Hz)		
Avant	Pendant	Après	Avant	Pendant	Après	Avant	Pendant	Après
4,44	4,19	4,21	4,44	4,19	4,21	4,44	4,19	4,21
27,75	27,23	27,34	27,75	27,23	27,34	27,75	27,23	27,34
-	76,09	76,76	-	76,09	76,76	-	76,09	76,76
99,89	-	-	99,89	-	-	-	-	-
115,4	-	114,6	115,4	-	115,7	-	-	116,7
-	149,4	149,3	-	149,4	-	-	149,3	-
199,8	200,1	199,8	199,8	-	199,8	-	-	-
-	247,5	-	-	247,5	247,1	-	247,5	-
-	370,5	-	-	370,5	-	-	370,5	-
-	400,2	-	-	-	-	-	-	-
-	-	-	424	-	424	-	-	-
-	-	-	460	-	460,4	-	-	-
-	482	-	-	482	-	-	-	-

Les méthodes d'extraction vidéo permettent d'obtenir des données exploitables, exactes et sur une large bande en fréquence. En déplacement, il semblerait que l'on ait des dérives de mesures. Nous avons souhaité investiguer ce point. La caméra principale donne la position des mires dans le plan

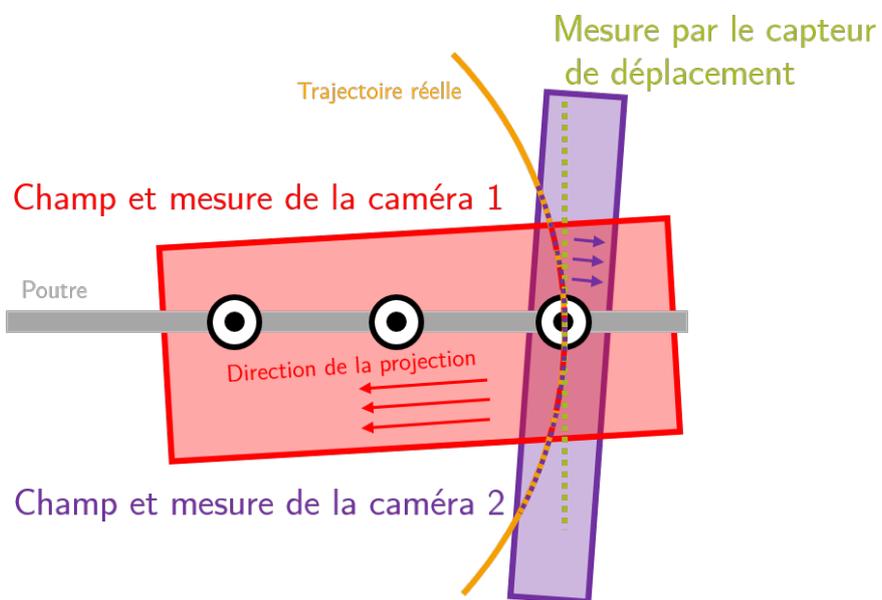


**Figure 37** – Comparaison des mesures par caméra et par capteur de position avant et après un choc : valeurs et erreurs en fonction du temps

de la tranche de la règle avec une projection que nous avons supposée négligeable. Ce n'est pas le cas du capteur de déplacement qui mesure la hauteur de la règle pour une section fixe, sans tenir compte de la courbure de la règle. Nous avons représenté schématiquement les données acquises par les différents capteurs sur la figure 38 en ajoutant également une deuxième caméra en champ proche utilisée pour les résultats sur la fusion de données (chapitre 4 paragraphe a. ). Le problème de la caméra secondaire est différent. Bien que l'hypothèse de projection 2D de la scène soit respectée, la projection n'est cependant pas la même que la caméra principale. Ce cas montre les limitations de cette hypothèse, mais aussi l'importance de la calibration des caméras afin de travailler dans le même repère.

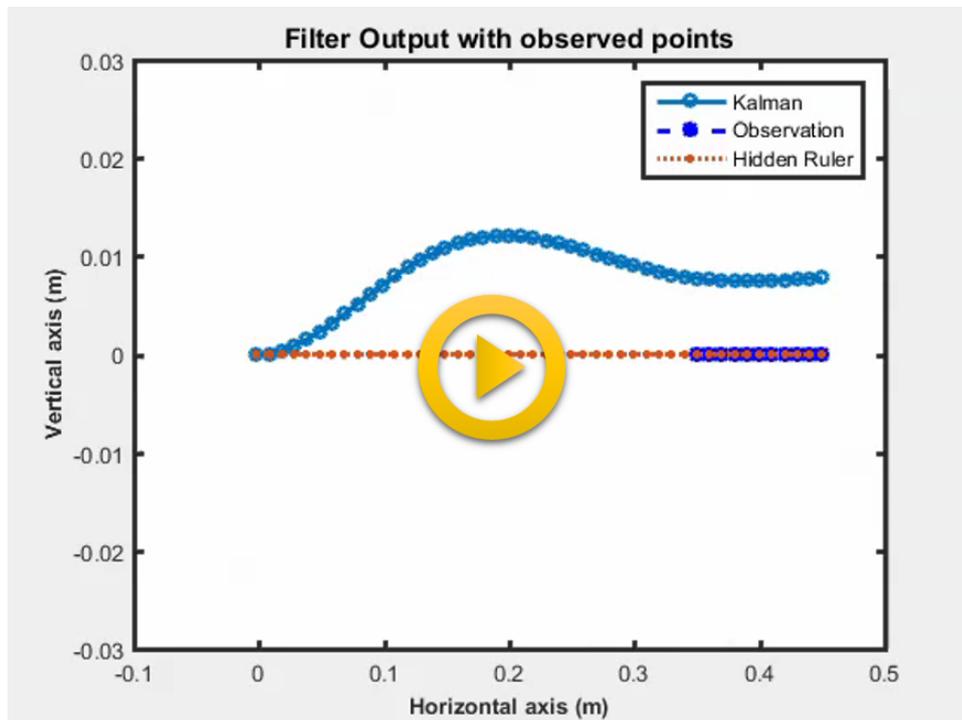
On remarquera également que le niveau d'erreur moyen augmente également en fonction de l'amplitude et reste en dessous de la valeur du pixel (ici  $1,01 \times 10^{-4}$  m/pixel). Un rapide calcul nous permet d'estimer qu'à 8 cm il existe un décalage géométrique entre la mire et le capteur de position dû à la courbure de la règle. Ce décalage est estimé à  $2,3 \mu\text{m}$  pour les amplitudes avant le choc mais à  $135 \mu\text{m}$  après le choc, ce qui explique l'augmentation de l'erreur.

### C. Comparaison des performances des méthodes de vision



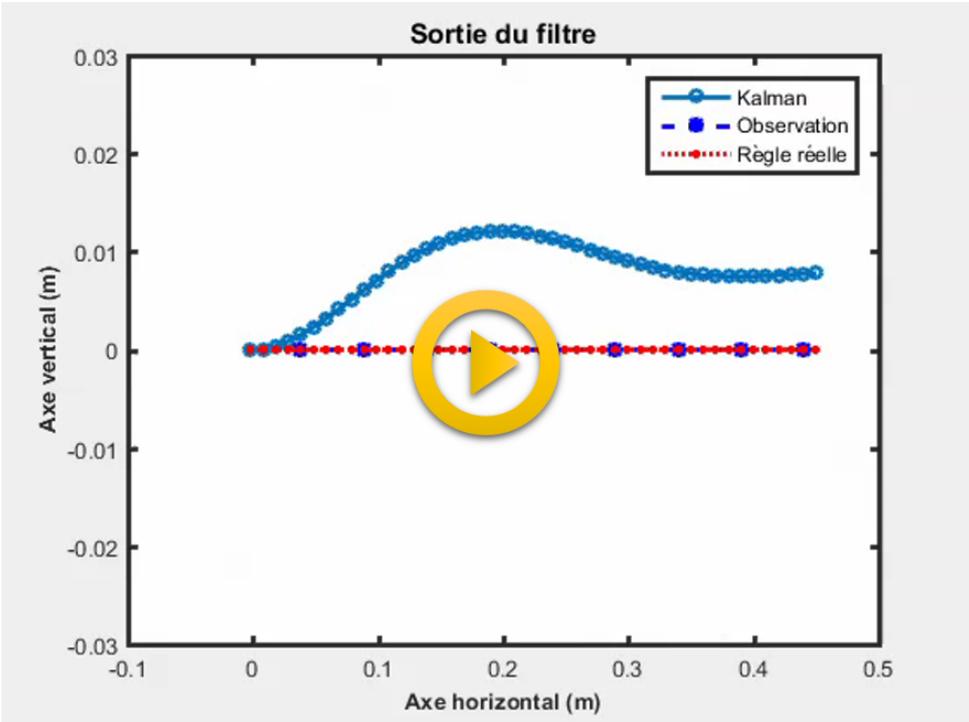
**Figure 38** – Schéma explicatif sur la provenance des erreurs de mesures entre les capteurs

**D. Résultats animés sur le cas test réel de la poutre encastree libre pour l'identification avec le schéma Kinematic-SAMI : applications à l'interpolation et l'extrapolation spatiale**

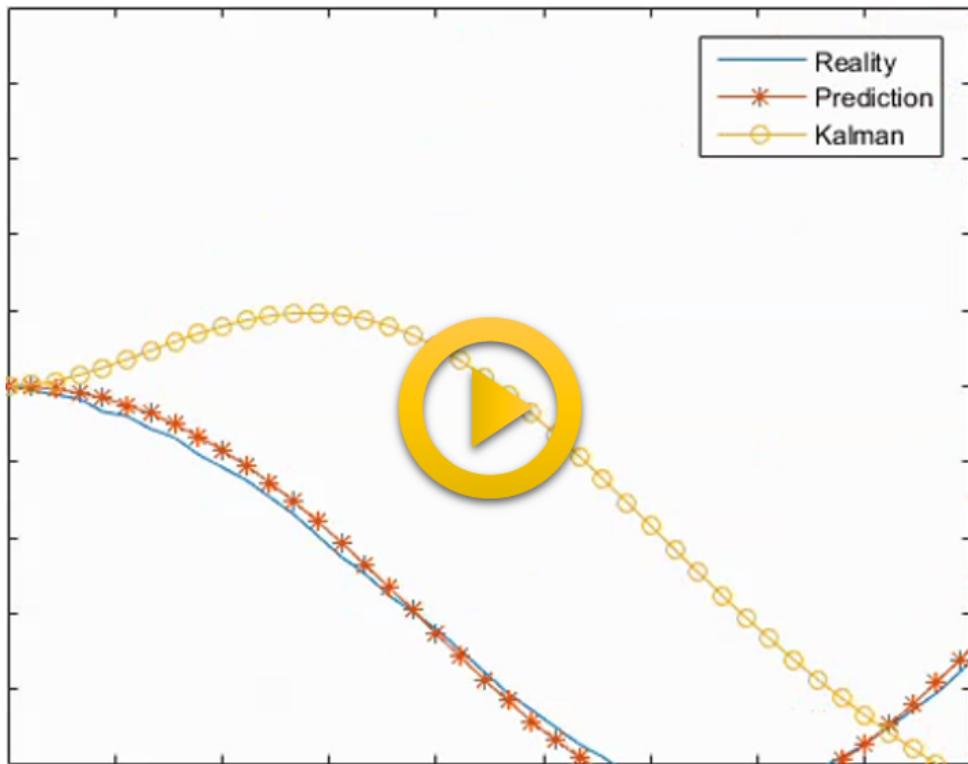


**Figure 39** – Résultats dynamiques de la comparaison entre déformées modales extrapolées estimées et déformées modales réelles pour des observations partielles sur l'extrémité de la règle sur le cas test réel de la poutre encastree libre pour les trois premiers modes

D. Résultats animés sur le cas test réel de la poutre encastree libre pour l'identification avec le schéma Kinematic-SAMI : applications à l'interpolation et l'extrapolation spatiale



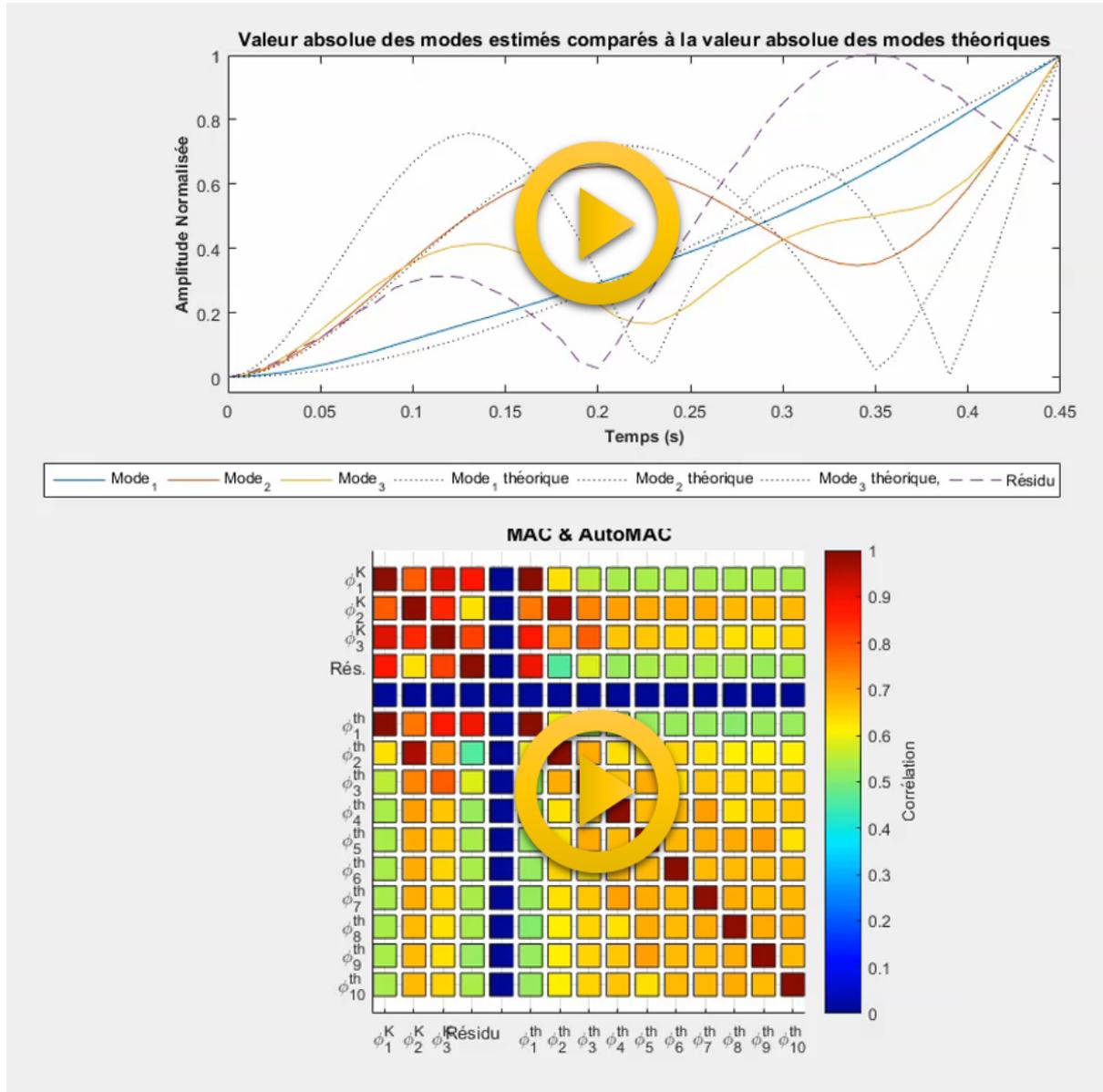
**Figure 40** – Résultats dynamiques de la comparaison entre déformées modales interpolées estimées et déformées modales réelles pour des observations équiréparties sur la règle sur le cas test réel de la poutre encastree libre pour les trois premiers modes



**Figure 41** – Résultats dynamiques de la comparaison entre déformées modales interpolées estimées et déformées modales réelles pour une interpolation temporelle d'un facteur 10 sur le cas test réel de la poutre encastree libre pour les trois premiers modes



**E. Résultats animés sur le cas test réel de la poutre encastrée libre pour l'identification avec le schéma Blind-SAMI**



**Figure 42** – Résultats dynamiques de l'évolution du critère de MAC et de l'identification des déformées modales correspondantes sur le cas test réel de la règle encastrée libre pour les trois premiers modes



## Bibliographie des annexes

Ernst Hairer : Interpolation et Approximation. *In Analyse Numérique*. Université de Genève, juin 2005.

I.J. Schoenberg : *Cardinal Spline Interpolation*. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, 1973. ISBN 978-0-89871-009-0.



**Titre :** Contribution à la perception augmentée de scènes dynamiques : schémas temps réels d'assimilation de données pour la mécanique du solide et des structures

**Mots clés :** Traitement vidéo, Identification non-linéaire, Assimilation de Données, Filtrés de Kalman, Magnification, Analyse Modale.

**Résumé :** Dans le monde industriel comme dans le monde scientifique, le développement de capteurs a toujours répondu à la volonté d'observer l'inobservable. La caméra rapide fait partie de ceux-là puisqu'elle permet de dévoiler des dynamiques invisibles, de la formation de fissure au vol du moustique. Dans un environnement extrêmement concurrentiel, ces caméras sont principalement limitées par le nombre d'images acquises par seconde. Le but de cette thèse est d'augmenter la capacité de dévoiler la dynamique invisible en enrichissant l'acquisition initiale par des modèles dynamiques. La problématique consiste alors à élaborer des méthodes permettant de relier en temps réel un modèle et la perception d'un système réel. Les bénéfices de cette utilisation offrent ainsi la possibilité de faire de l'interpolation, de la prédiction et de

l'identification. Cette thèse est composée de trois parties. La première est axée sur la philosophie du traitement vidéo et propose d'utiliser des modèles élémentaires et génériques. Un algorithme d'estimation de grands mouvements est proposé mais l'approche actuellement proposée n'est pas assez générique pour être exploitée dans un contexte industriel. La deuxième partie propose d'utiliser des méthodes d'assimilation de données séquentielle basées sur la famille des filtres de Kalman afin d'associer un modèle avec des observations par caméras rapides pour des systèmes mécaniques. La troisième partie est une application à l'analyse modale expérimentale non linéaire. Deux schémas d'assimilation temps réel multicapteurs sont présentés et leur mise en oeuvre est illustrée pour de la reconstruction 3D et de la magnification.

**Title:** Contribution to augmented observation of dynamic scenes : real time data assimilation schemes for solid and structure mechanics

**Keywords:** Video processing, nonlinear identification, data assimilation, Kalman Filter, motion magnification, modal analysis.

**Abstract:** The development of sensors has always followed the ambition of industrial and scientific people to observe the unobservable. High speed cameras are part of this adventure, revealing invisible dynamics such as cracks formation or subtle mosquito flight. Industrial high speed vision is a very competitive domain in which cameras stand out through their acquisition speed. This thesis aims to broaden their capacity by augmenting the initial acquisition with dynamic models. This work proposes to develop methods linking in real time a model with a real system. Aimed benefits are interpolation, prediction and identification. Three parts are developed. The first one is

based on video processing and submits to use kinematic elementary and generic models. An algorithm of motion estimation for large movements is proposed but the generic nature does not allow a sufficient knowledge to be conclusive. The second part proposes using sequential data assimilation methods known as Kalman filters. A scheme to assimilate video data with a mechanical model is successfully implemented. An application of data assimilation in modal analysis is developed. Two multi sensors real time assimilation schemes for nonlinear modal identification are proposed. These schemes are integrated in two applications on 3D reconstruction and motion magnification.

