



HAL
open science

Utilisation d'une caméra Temps-de-vol pour la gestion des mouvements en radiothérapie externe

Thomas Wentz

► **To cite this version:**

Thomas Wentz. Utilisation d'une caméra Temps-de-vol pour la gestion des mouvements en radiothérapie externe. Traitement du signal et de l'image [eess.SP]. Télécom Bretagne, Université de Bretagne Occidentale, 2013. Français. NNT: . tel-00908839

HAL Id: tel-00908839

<https://theses.hal.science/tel-00908839>

Submitted on 25 Nov 2013

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.

Sous le sceau de l'Université européenne de Bretagne

Télécom Bretagne

En habilitation conjointe avec l'Université de Bretagne Occidentale

Ecole Doctorale – sicma

Utilisation d'une caméra Temps-de-Vol pour la gestion des mouvements en radiothérapie externe

Thèse de Doctorat

Mention : Sciences et Technologies de l'Information et de la Communication (STIC)

Présentée par **Thomas Wentz**

Département : Image et Traitement de l'Information

Laboratoire : LaTIM

Directeurs de thèse : Dimitris Visvikis

Soutenue le 04 Juillet 2013

Jury :

Mme. Jocelyne Troccaz Directrice de Recherche au Laboratoire TIMC-IMAG (Rapporteur)
M. George Kontaxikis Professeur à l'Universidad Politécnica de Madrid (Rapporteur)
M. Dimitris Visvikis Directeur de Recherche au Laboratoire LATIM (Directeur de thèse)
M. Frédéric Lamare Ingénieur de recherche au CHU de Bordeaux (Examineur)
M. Olivier Pradier Professeur des universités – Praticien Hospitalier à Brest (Examineur)
M. Hadi Fayad Maître de conférence à l'université de Brest (Examineur)
Mme. Emmanuelle Le Fur Chef de Clinique en Radiothérapie (Invitée)



Table des matières

Introduction	19
Etat de l'art	25
1 Contexte médical	25
1.1 Le cancer	26
1.1.1 Définition et description	26
1.1.2 Epidémiologie	26
1.1.3 Le diagnostic du cancer	27
1.1.4 Le traitement du cancer	29
1.2 La radiothérapie externe	30
1.2.1 Historique	30
1.2.2 Principe et étapes du traitement par radiothérapie	31
1.2.3 Erreur due aux mouvements	34
1.3 Prise en charge du mouvement en radiothérapie externe	36
1.3.1 Le mouvement respiratoire	36
Description	36
Mécanique de l'appareil respiratoire	36
Amplitude du mouvement respiratoire	37
Prise en compte du mouvement respiratoire en radiothérapie	38
1.3.2 Repositionnement entre les fractions	43
Lasers	43
Imageries de contrôle	44
1.3.3 Synthèse et enjeux pour la recherche scientifique	45
2 Contexte scientifique	47
2.1 Systèmes commerciaux de suivi tumoral	48
2.2 Observation du mouvement surfacique	51
2.2.1 Systèmes acquérant un signal 1D	51
Le spiromètre	51

Table des matières

La ceinture de pression	52
Le système infrarouge passif de Varian	52
2.2.2 Marqueurs ponctuels	53
Marqueurs infrarouges	53
Marqueurs optiques	53
Marqueurs électromagnétiques	55
2.2.3 Les scanners tridimensionnels	55
La lumière structurée	56
Le scanner 3D par laser	57
2.2.4 Les caméras Temps-de-Vol	58
Etudes techniques	60
Applications	62
2.3 Les modèles de respiration	64
2.3.1 Les modèles basés sur un point	65
2.3.2 Les modèles basés sur un champ de déformation	65
2.4 Le repositionnement par l'analyse de la surface du patient	66
2.5 Synthèse	68

Apports méthodologiques et études menées **73**

3 Comparaison des caméras ToF et de la Kinect **75**

3.1 Principe de fonctionnement et caractéristiques	76
3.2 Expérimentations sur fantôme statique	78
3.3 Expérimentations sur fantôme dynamique	80
3.4 Estimation du mouvement respiratoire	82
3.5 Synthèse et conclusion	85

4 Extraction du mouvement respiratoire externe **89**

4.1 Modélisation surfacique des données	91
4.1.1 Les fonctions B-Splines	92
4.1.2 L'algorithme MBA	93
4.1.3 Extraction du mouvement	95
4.2 Caractérisation de la qualité de mesure	96
4.2.1 Paramètres de la caméra	99
4.2.2 Impact de la modélisation	100
4.3 Comparaison avec un examen 4D-TDM comme vérité terrain	102
4.3.1 L'examen 4D-TDM	103
4.3.2 Conditions d'observation cliniques	104
4.3.3 Mise en correspondance des données	106

4.3.4	Traitement des données	108
4.3.5	Résultats	109
	Influence de la taille de la ROI	109
	Concordance spatiale et temporelle	109
	Analyse quantitative	112
4.4	Prédiction du mouvement interne	115
4.5	Synthèse et conclusion	117
5	Estimation du mouvement surfacique 3D	121
5.1	Le flot optique	122
5.2	Le recalage élastique	124
5.3	Description de la méthodologie	127
5.4	Evaluation de la méthode	128
	5.4.1 Tests effectués	128
	5.4.2 Respect de la contrainte d'illumination constante	130
	5.4.3 Estimation du déplacement 2D	134
	5.4.4 Estimation du déplacement 3D	138
	5.4.5 Comparaison avec un recalage élastique	139
5.5	Synthèse et conclusion	140
6	Le repositionnement inter-fraction avec une caméra ToF	145
6.1	Création d'un système stéréo d'observation : calibrage de 2 caméras ToF . .	147
	6.1.1 Etat de l'art	147
	6.1.2 Proposition de méthode et comparaison avec l'état de l'art	148
	Calibrage avec un plan	149
	Calibrage avec une sphère	150
	Calibrage avec un cube	150
	6.1.3 Expérimentations	152
	6.1.4 Impact sur le repositionnement	157
6.2	Prise en compte du mouvement respiratoire	160
6.3	Etude Clinique	167
	6.3.1 Objectifs, moyens et déroulement	167
	6.3.2 Dispositif d'observation	168
	6.3.3 Système avec 1 seule caméra	170
	6.3.4 Système avec 2 caméras	171
6.4	Synthèse et conclusion	174
	Synthèse, conclusion générale et perspectives	179

Table des figures

1.1	Incidence et taux de survie à 5 ans pour différents types de cancer	27
1.2	Evolution de la mortalité dûe au cancer en France	28
1.3	Lors de la chimiothérapie, l'administration des substances chimiques bloquent la mitose (division du noyau des cellules) et ainsi ciblent les cellules se divisant rapidement, ce qui est le cas des cellules cancéreuses. La plupart des substances actuelles ne différencient pas les cellules saines des cellules cancéreuses et détruisent aussi des cellules à division rapide tels que les cellules sanguines, de l'épithélium intestinal ou des cheveux.	29
1.4	Image du premier patient traité pour un retinoblastome par radiothérapie administrée par un accélérateur linéaire en 1957	30
1.5	Capture d'écran d'un système de planification de traitement qui permet la simulation des lignes de tirs afin d'analyser les conséquences sur la tumeur et les organes à risque.	31
1.6	Différentes marges sont ajoutées successivement au volume décrivant la tumeur vue par imagerie (GTV). Le CTV prend compte les extensions microscopiques de la tumeur invisibles à l'image. Le PTV ajoute deux autres types de marges au CTV, les marges internes (IM) décrivant le mouvement de la tumeur, et les marges dues au Setup (SM) afin de compenser les incertitudes liées au systèmes.	32
1.7	Les collimateur multilames permettent d'administrer des faisceaux à géométrie variables afin de cibler celle du volume tumoral.	33
1.8	La VMAT change la conformation de l'accélérateur lors de son déplacement en arc de cercle autour du patient	34
1.9	Description du processus permettant la planification du traitement avant de pouvoir traiter le patient	34
1.10	Déplacement du GTV associé à une tumeur lors de la respiration (M.Rains (2010))	35
1.11	Biomécanique respiratoire chez l'homme	37
1.12	Indication des différentes directions anatomiques utilisées dans ce manuscrit	38

Table des figures

1.13	Tableau tiré du rapport du groupe de travail 76 récapitulant le mouvement (moyenne (min-max)) en millimètre des tumeurs situées dans le poumon lors de la respiration dans différentes études (Keall et al., 2006). BP = bas des poumons. MP = milieu des poumons. HP = haut des poumons.	39
1.14	Tableau tiré du rapport du groupe de travail 76 récapitulant le mouvement (moyenne (min-max)) en millimètre des organes de l'abdomen durant la respiration (Keall et al., 2006).	40
1.15	Exemple du système d'asservissement avec blocage respiratoire volontaire avec le feedback vu à l'aide de lunettes.	41
1.16	A gauche : capture d'écran du logiciel associé au système RPM. A droite : schéma décrivant le principe de fonctionnement du Gating.	42
1.17	Le système commercial Cyberknife estime la position d'une tumeur en temps réel par suivi de la position de marqueurs externes fixés sur la veste du patient (Bondiau et al., 2007).	43
1.18	Exemple clinique de ciblage d'une tumeur du sein. On y distingue les trois lasers qui permettent au manipulateur d'aligner le patient à la position de référence	44
1.19	Capture d'écran du logiciel permettant la comparaison visuelle entre deux images portales avec surimpression du PTV préalablement prévu	45
2.1	Illustration du dispositif Calypso permettant de suivre des tumeurs de la prostate en temps réel (SiteCalypso (2013)).	49
2.2	Illustration du dispositif Cyberknife qui est constitué de : (1) l'accélérateur linéaire guidé par un bras robotisé ; (2) deux plaques radiographiques scellées au sol ; (3) une veste couverte de marqueurs infrarouges suivies par (4) trois caméras CCD.	50
2.3	Le dispositif ExacTrac est constitué de deux imageurs par rayons X fixés obliquement par rapport au patient. Conçu initialement pour le repositionnement, un suivi des marqueurs optiques infrarouge permet le suivi d'une tumeur.	50
2.4	Le système infrarouge passif RPM commercialisé par Varian est composé d'un émetteur de lumière infrarouge (à gauche), d'un bloc plastique transparent sur lequel sont fixés deux réflecteurs infrarouges visible sur l'image infrarouge à droite.	52
2.5	En haut, illustration des deux types de marqueurs infrarouges, passifs et actifs. Le système Polaris (NDI Company) est présenté en bas avec des marqueurs réfléchissants associés au système.	53
2.6	Principe de la détection stéréoscopique.	54

2.7	Exemple de patient sur lequel sont fixés 31 marqueurs optiques (Hostettler, 2008).	54
2.8	Le système commercial AlignRT de VisionRT utilise la lumière structurée pour reconstruire les surfaces en 3D (SiteVisionRT (2013))	56
2.9	La mesure par temps de vol consiste à émettre puis mesurer de la lumière infrarouge réfléctée par les surfaces présentes dans la scène.	58
2.10	Principe de mesure de la profondeur par mesure de l'offset	59
2.11	Propriétés de la caméra ToF SR4000 de SwissRanger	60
2.12	Principe de mesure de la profondeur par mesure de l'offset	61
2.13	Un flou apparait lorsque l'objet observé se déplace pendant l'acquisition (Foix et al., 2011)	62
2.14	Illustrations de différentes applications utilisant une caméra ToF en vision. En haut à gauche, le suivi dynamique de la marche par fitting d'un modèle articulé (Jensen et al., 2008). En haut à droite, ajout d'un objet virtuel dans une scène réelle (Bartczak et al., 2008). En bas, détection en temps réel de points d'intérêt tels que le nez et une main (Boehme et al., 2008).	64
3.1	Le motif de points infrarouge projeté permet à la Kinect d'estimer une carte de profondeur associant à chaque pixel un point 3D.	76
3.2	Tableau répertoriant les principales caractéristiques des trois caméras comparées dans ce chapitre.	77
3.3	Erreur moyenne d'estimation du mouvement surfacique nul du fantôme pour les trois caméras en fonction de la distance.	79
3.4	Influence de l'angle d'observation sur la mesure de mouvement pour les trois caméras.	79
3.5	Résolution en profondeur de la Kinect sur une gamme allant de 0.6 à 2 mètres.	80
3.6	Distribution du déplacement estimé pour chaque point d'une ROI pour les trois caméras. Le déplacement réel est de 2 mm	81
3.7	Erreur moyenne d'estimation du mouvement linéaire du fantôme en suivant le point 3D correspondant au pixel au centre du champ de vue	82
3.8	Illustration du suivi du mouvement linéaire du fantôme en suivant le point 3D correspondant au pixel au centre du champ de vue. La vitesse est d'un centimètre par seconde.	83
3.9	Illustration des deux ROIs prises en compte pour l'extraction du mouvement respiratoires. Leur taille est respectivement de 1x5 cm ² et 5x8 cm ²	83
3.10	Récapitulatif du nombre approximatif de points par ROI et par type de caméra.	84
3.11	Mouvement respiratoire observé par la caméra ToF SR4000 et la Kinect à 0.9 m et 1.5 m avec à gauche une ROI abdominale de 1x5 cm ² et à droite une ROI abdominale de 5x8 cm ²	85

Table des figures

3.12	Le déplacement en profondeur pour les points d'une large région d'intérêt est montré ici par un code couleur allant de 0 mm (bleu) à 5 mm (rouge) avec à gauche les données de la caméra ToF SR4000 et à droite celles de la Kinect. En haut, observations à 0.9 m et en bas à 1.5 m.	86
4.1	Mouvement de la peau selon l'axe antéro-postérieur le long d'un cycle respiratoire (Hostettler, 2008). Le mouvement est mesuré à l'aide de marqueurs optiques, plus précisément des ARTags dont la position est déterminée par un système de caméras stéréoscopiques. L'abscisse représente la direction craniale-caudale, le long de laquelle les marqueurs sont situés, avec à gauche la région abdominale et dans la partie droite la région thoracique.	91
4.2	Illustration du phénomène de Runge où un jeu de données simple est interpolé par un polynôme de degré 4	93
4.3	Modélisation d'un nuage de points de mesure (en rouge) par une fonction B-Splines (en bleu).	93
4.4	Exemple de l'influence de la résolution de la grille de contrôle sur la modélisation des données de mesures et l'intensité de lissage associée. Une grille hautement résolue aura tendance à modéliser parfaitement les points de mesure au détriment du réalisme.	94
4.5	Les deux points de mesure z_1 et z_2 influencent chacun leur 16 points de contrôle proches. Réciproquement, les points de contrôle sont influencés par tout points de mesure dans l'espace délimité par les 24 points de contrôle voisins : les points verts reçoivent l'influence de z_1 et de z_2	95
4.6	Extrait de Lee et al. (1997) qui illustre l'optimisation du temps de calcul induit par le rééchantillonnage des fonctions entre chaque itération plutôt qu'avec la résolution finale.	96
4.7	La modélisation des surfaces permet de les manipuler aisément et de permettre une extraction extrêmement rapide d'un champ de vecteurs 1D décrivant la déformation locale d'une surface.	97
4.8	Vue du fantôme anthropomorphique utilisé pour simuler le type de relief caractéristique du torse humain, ainsi que de l'actionneur linéaire utilisé pour le déplacer.	98
4.9	Erreur moyenne d'estimation d'un déplacement nul en fonction de la distance d'observation.	101
4.10	A 1m, distribution des erreurs moyennes le long de la surface (gauche), puis à l'échelle d'un pixel ayant l'erreur médiane (droite)	101
4.11	Erreur d'estimation d'un mouvement triangulaire de 2cm d'amplitude pour différents angles d'observation. L'estimation est effectuée avec les points de contrôle sans (gauche) et avec (droite) filtrage spatial gaussien.	102

4.12	Le système RPM de Varian est constitué d'une caméra fixée au bout du lit de traitement qui observe le déplacement de deux marqueurs réfléchissants infrarouges, visibles sur cette image. Le logiciel intégré fournit alors en temps réel une estimation du déplacement dans les trois directions.	103
4.13	Le Binning des données TDM acquises dynamiquement est effectué en sélectionnant les projections pour un même instant respiratoire. Cet instant peut être considéré par une même amplitude du signal respiratoire ou par une même phase à chaque cycle de respiration	103
4.14	Le système de fixation de la caméra ToF est constitué d'un bras girafe prolongé d'une rotule afin de permettre un maximum de degrés de liberté tant en hauteur, profondeur et orientation. La matériel est de la marque Manfrotto (Vitec Company, Italie).	105
4.15	Vue du dispositif complet depuis l'arrière du tunnel d'acquisition. Une des difficultés est l'impossibilité de monter plus haut la caméra afin d'avoir un meilleur angle d'observation puisque le tunnel du scanner est arrondi vers l'extérieur.	105
4.16	Un exemple de la surface d'un patient lors d'un examen 4D-TDM utilisé dans cette étude. On voit que la restriction du champ à trente centimètres dans la direction cranio-caudale ne permet d'observer qu'une bande d'environ dix centimètres en dessous des seins.	106
4.17	La comparaison des données Scanner avec les données ToF nécessite une mise en correspondance spatiale des données qui est effectuée en déterminant la matrice de passage reliant le repère de la caméra à la table d'examen.	107
4.18	Le nuage de points 3D décrivant la table permet la détermination du plan associé (en bleu) qui fournit le vecteur normal associé (violet). Après analyse par ACP, les deux vecteurs propres associés aux deux valeurs propres ayant le plus de poids fournissent un repère complet décrivant la table et son orientation complète, $\mathbf{R}_{table\ d'\ examen}$	107
4.19	Le traitement des données est découpé en trois phases : la transformation des données de la caméra afin de correspondre aux données TDM (Section 4.3.3), la correction du mouvement du patient et l'extraction de la trace respiratoire, puis la sélection et la modélisation des surfaces décrivant un cycle respiratoire moyen.	108
4.20	Les traces respiratoires sont extraites de régions d'intérêt de taille différentes.	110
4.21	Traces associées au ROIS illustrées figure 4.20.	110

4.22	Illustration du cas 1 présenté d'une patiente avec une forte poitrine. En haut à droite, la représentation 3D de la surface du volume TDM. En haut à gauche, une carte de couleur représentant la différence en profondeur entre inspiration et expiration complète. Les deux cartes du bas montrent pour la ROI exploitée, les résultats à partir des données TDM (gauche) et de la caméra ToF (droite).	112
4.23	Illustration du cas 2 présenté d'une patiente avec une poitrine de taille standard. En haut à droite, la représentation 3D de la surface du volume TDM. En haut à gauche, une carte de couleur représentant la différence en profondeur entre inspiration et expiration complète. Les deux cartes du bas montrent pour la ROI exploitée, les résultats à partir des données TDM (gauche) et de la caméra ToF (droite).	113
4.24	Illustration du cas 3 présenté d'un patient avec une morphologie maigre. En haut à droite, la représentation 3D de la surface du volume TDM. En haut à gauche, une carte de couleur représentant la différence en profondeur entre inspiration et expiration complète. Les deux cartes du bas montrent pour la ROI exploitée, les résultats à partir des données TDM (gauche) et de la caméra ToF (droite).	114
4.25	Les différentes ROIs utilisées pour estimer le mouvement surfacique ont une taille allant de 4x3 cm ² (vert) à 20x3 cm ² (bleu).	115
4.26	Mouvement d'une ROI de 20x3 cm ³ lors d'un cycle respiratoire, estimé à partir des données d'une caméra ToF et d'un Scanner 4DCT.	115
4.27	Différence moyenne d'estimation du mouvement par une caméra ToF par rapport à celui extrait des données TDM pour un cycle respiratoire complet, pour 7 patients et 3 tailles de ROIs	116
4.28	Les 13 points d'intérêt sont des points caractéristiques des poumons. Leur trajectoire est déterminée par analyse des 22 volumes TDM par un expert médical.	116
4.29	Les 10 régions servant à l'extraction des 10 signaux décrivant le mouvement externe surfacique sont des bandes horizontales mesurant 2x20 cm ²	117
5.1	Le flot optique fournit un vecteur de déplacement pour chaque pixel de l'image initiale. La contrainte de continuité permet d'éviter des aberrations en partant de l'hypothèse que deux pixels voisins ne peuvent pas avoir des déplacements opposés.	124
5.2	Illustration extraite de Nickolls and Dally (2010) qui récapitule les gains en temps de calcul atteignables par transfert sur GPU en fonction du type de traitement à effectuer. Le gain potentiel varie fortement, allant de 9 pour de la reconnaissance vocale à 137 pour une application de simulation moléculaire.	126

5.3	Le pixel $P(u,v-1)$ est analysé comme ayant bougé à la localisation $(u,v+1)$ entre les observations i et $i+1$: son mouvement en 3D est le vecteur formé par les points 3D correspondants $P(u,v-1)$ et $P(u,v+1)$. De même, le pixel $P(u-1,v-1)$ s'est déplacé à une position non-définie à l'instant $i+1$: le point 3D correspondant à cette localisation est déterminé à l'aide de la fonction B-splines décrivant la surface, à l'emplacement ayant comme coordonnées x et y interpolé des 4 voisins les plus proches.	128
5.4	Illustration du champ de déplacement estimé par flot optique entre inspiration complète et expiration complète. En rouge, la localisation des marqueurs dessinés sur le patient.	129
5.5	Une même scène permet d'avoir deux images : celle de gauche représente l'intensité lumineuse alors que celle de droite représente la carte de profondeur convertie en niveaux de gris.	130
5.6	Illustration de la pyramide simulée pour test, avec la vue selon la direction z . En c), le flot optique entre les deux images.	131
5.7	La détection d'un mouvement fantôme par le flot optique vient du fait qu'un pixel qui se déplace selon z change de niveau de gris. La flot optique le détecte donc comme ayant bougé latéralement, vers le pixel voisin qui lui, prend la valeur initiale de niveau de gris.	131
5.8	Flot optique entre les images de la scène avant et après translation de 4 mm vers la surface. La taille des vecteurs est normalisée.	132
5.9	Distribution des composantes des vecteurs de déplacement selon les directions u et v entre -1 et 1 pixel. La translation est de 4 mm.	133
5.10	Distribution des composantes des vecteurs de déplacement selon les directions u et v entre -1 et 1 pixel. La translation est de 8 mm.	133
5.11	Distribution des composantes des vecteurs de déplacement selon les directions u et v entre -1 et 1 pixel. La translation est de 12 mm.	133
5.12	Deux images d'un même patient à expiration et inspiration complète.	134
5.13	Erreur d'estimation du flot optique entre $Im(0)$ et $Im(1/3)$	135
5.14	Erreur d'estimation du flot optique entre $Im(0)$ et $Im(2/3)$	136
5.15	Erreur d'estimation du flot optique entre $Im(0)$ et $Im(1)$	136
5.16	Moyenne des erreurs d'estimation des composantes u et v du flot optique pour 10 cas d'étude et deux méthodes d'estimation, basées sur Horn and Schunck (1981) et Lucas et al. (1981), en fonction du déplacement maximal (au sens de la distance de Manhattan). En vert, le meilleur résultat pour chaque cas.	137

Table des figures

5.17	Erreur moyenne d'estimation absolue du flot optique pour 5 cas d'étude et deux méthodes d'estimation, basées sur Horn and Schunck (1981) et Lucas et al. (1981), pour les cinq cas où le déplacement maximal est inférieur à 4 pixels.	138
5.18	Erreur moyenne d'estimation en millimètres du mouvement 3D des marqueurs optiques pour les cinq cas dont le déplacement maximal est inférieur à 4 pixels.	139
5.19	Illustration des données fournies par une caméra ToF (en haut) et une Kinect (en bas) avec : a) une image d'intensité et b) un nuage de points 3D. L'image d'intensité de la Kinect est bien plus résolue que celle de la caméra ToF (640*480 contre 176*144) ce qui améliore la précision de la détection du mouvement des marqueurs.	140
5.20	Illustration du champ de déplacement estimé par flot optique entre inspiration complète et expiration complète.	141
5.21	Détail des résultats pour le patient dont le mouvement surfacique est estimé par la fusion des données 2D/3D et l'algorithme de recalage élastique PCD. Les données sont fournies par une caméra Kinect.	141
6.1	Description des points clef des méthodes de calibrage	148
6.2	Régression linéaire d'un jeu de données (croix) à l'aide de la méthode des moindres carrés (pointillés) et de la méthode des moindres médians carrés (ligne pleine)	151
6.3	Les deux objets physiques utilisés sont (a) une sphère en polystyrène de 15 cm de diamètre et (b) un cube en bois peint en blanc de diamètre 30x30x30 cm qui sert aussi de plan à détecter. En c) et d), les scènes 3D des objets en question.	153
6.4	Représentation du champ de vue commun aux deux caméras qui est divisé en deux sous-espaces : Le sous-espace Calibrage et le sous-espace Monde. Les croix rouges indiquent les positions ayant servi au calibrage et/ou à la validation alors que les croix vertes désignent des positions ayant seulement servi pour validation.	154
6.5	Variabilité de l'estimation du centre de la sphère fixe.	155
6.6	Estimation de la translation de la sphère.	156
6.7	Variabilité sur l'estimation des angles d'Euler associés à un plan.	156
6.8	Erreur de calibrage des différentes méthodes au centre et à la périphérie du champ de vue.	157
6.9	Configuration du dispositif d'observation par rapport à la table d'examen. Les caméras sont centrées sur l'isocentre de traitement.	158

6.10 Régions d'intérêt pour l'évaluation du repositionnement sur volontaire (région 1 à gauche, région 2 à droite).	158
6.11 Fusion des observations latérales du volontaire avec en bleu les données de la caméra A, et en rouge celles de la caméra B.	159
6.12 Erreur d'estimation du mouvement latéral du fantôme quand l'observation est effectuée par la caméra A, la B ou les deux.	159
6.13 Erreur d'estimation du mouvement en profondeur du fantôme quand l'observation est effectuée par la caméra A, la B ou les deux.	160
6.14 Erreur d'estimation de la translation d'un humain quand l'observation est effectuée par la caméra C (en rouge) ou le système stéréo (bleu). La région d'intérêt est la première de la figure 6.10.	161
6.15 Erreur d'estimation de la translation d'un humain quand l'observation est effectuée par la caméra C (en rouge) ou le système stéréo (bleu). La région d'intérêt est la deuxième de la figure 6.10.	161
6.16 Erreur moyenne d'estimation du mouvement en millimètre pour toute la gamme de translations.	162
6.17 Trois régions d'intérêt sont délimitées afin de caractériser l'instant respiratoire associé à la surface acquise.	163
6.18 Signaux respiratoires correspondant aux trois régions d'intérêt illustrées figure 6.17.	163
6.19 Trois acquisitions sont sélectionnées, la deuxième pour sa concordance avec la première selon les trois signaux, la troisième pour sa concordance d'après le signal abdominal uniquement.	164
6.20 Résultats du recalage surfacique d'un patient ayant subi une translation de 8 mm dans la direction cranio-caudale. La surface pré-déplacement est sélectionnée à expiration complète. Elle est ensuite recalée avec chaque acquisition d'un cycle respiratoire équivalent après le déplacement. Les meilleurs recalages sont avec la première et la vingt-sixième images (en vert).	165
6.21 Résultats du recalage surfacique d'un patient ayant subi une translation de 8 mm dans la direction antérieure. La surface pré-déplacement est sélectionnée à expiration complète. Elle est ensuite recalée avec chaque acquisition d'un cycle respiratoire équivalent après le déplacement. Les meilleurs recalages sont avec la deuxième et la vingt-troisième images (en vert).	166
6.22 Vue d'ensemble du dispositif d'observation composé de deux caméras ToF fixés au plafond à l'aide de systèmes amovibles. Leur position est limitée par le champ de déplacement de l'accélérateur ainsi que par le système laser de repositionnement. Elle est optimisée avec comme critère la plus faible distance d'observation par rapport à l'isocentre et un champ de vue maximal afin d'englober l'ensemble du thorax et de l'abdomen.	169

6.23	Le système de fixation de la caméra située à gauche par rapport à l'accélérateur est amovible, avec une barre métallique qui s'enfile dans un support visé au plafond. La caméra de droite est elle tout le temps fixée à un système mécanique permettant une position de repos non-gênante dans le faux plafond grâce à une fixation magnétique. La position baissée est assurée par une butée.	170
6.24	Erreur moyenne d'estimation du déplacement d'un patient en fonction de l'amplitude de déplacement	171
6.25	La surface fournie par les deux caméras permet une description complète du tronc du patient malgré une plus faible densité de points en haut du thorax. La zone bleue correspond à la région prise pour estimer le déplacement . .	172
6.26	Résultats de la phase 2 de l'étude, sur trois patients, avec les erreurs de détection et l'estimation de la conformité ainsi que la prédiction du déplacement. A droite, les résultats de Moser et al. (2012) pour les localisations pulmonaires	173
6.27	Résultats de la phase 3 de l'étude, sur trois patients, avec les erreurs de détection et l'estimation de la conformité ainsi que la prédiction du déplacement. A droite, les résultats de Moser et al. (2012) pour les localisations pulmonaires	175

Introduction

Le traitement par radiothérapie externe est aujourd'hui un maillon essentiel dans le traitement du cancer. Utilisé conjointement avec la chirurgie et/ou la chimiothérapie, elle permet d'améliorer le pronostic de survie de manière significative. L'irradiation par des faisceaux de rayons X de haute énergie et des faisceaux d'électrons permet de détruire les cellules présentes dans le volume ciblé. Les dernières évolutions cliniques ont concerné la forme et la configuration de ces faisceaux afin de permettre l'irradiation de volumes cibles décrivant précisément la géométrie de la tumeur : l'utilisation des collimateurs multilames associés aux avancées dans l'aide informatique à la planification rend possible une radiothérapie avec modulation d'intensité (IMRT ou Intensity Modulated RadioTherapy). La deuxième grande avancée est l'utilisation de l'image comme aide à la décision et à la vérification de la bonne conformité de la position du volume cible lors de chaque fraction du traitement.

Les défis actuels dans ce domaine médical concernent la prise en charge du mouvement afin d'assurer un ciblage optimal de la tumeur. L'appareillage standard en pratique clinique est assez performant pour délivrer une dose avec une précision très importante, mais des facteurs potentiels d'erreur viennent entacher la précision finale de la délivrance de dose. Un facteur d'erreur commun à toutes les localisations tumorales est le repositionnement du patient entre chaque fraction afin de faire correspondre sa position spatiale avec celle de référence, fixée en accord avec un examen de tomodensitométrie (TDM) utilisé pour la planification. La stratégie actuelle est composée de deux degrés d'action : le premier est un repositionnement journalier effectué à l'aide de repères tatoués sur la peau qui sont alignés sur un système de lasers calibrés dans la salle de traitement. Le second degré d'action va consister à imager le patient afin de vérifier que les structures internes d'intérêt sont alignées par rapport à une image de référence, la fréquence de cette vérification dépendant de la localisation tumorale. Différents types d'imageurs sont développés afin de permettre une visualisation de bonne qualité tout en limitant la dose induite qui doit être retirée du traitement réel administré. Cependant cette dose a tendance à augmenter lorsque les imageurs fournissent des images de qualité supérieure. L'autre grand facteur d'incertitude de la localisation tumorale concerne principalement les cancers du poumons et du sein : le mouvement respiratoire déplace la tumeur par la variation du volume pulmonaire, ce mouvement pouvant atteindre plusieurs centimètres. La réponse clinique classique à ce phénomène est d'élargir le volume à irradier afin de garantir un délivrance de dose complète, provoquant en contrepartie l'irradiation de tissus et organes sains à la périphérie de la tumeur ou le long des lignes de tir.

Des solutions technologiques sont actuellement développées afin de prendre en charge ces mouvements et ainsi limiter le recours aux solutions irradiantes tels que les imageurs pour le repositionnement et les marges supplémentaires en présence de mouvement respiratoire. Au delà de l'irradiation à un seul moment respiratoire (le "Gating") ou des systèmes de contention, la recherche se dirige vers le développement de systèmes asservissant l'irradiation au mouvement respiratoire par l'intermédiaire de la prédiction du mouvement de la tumeur. Des solutions commerciales existent mais malgré les avancées significatives qu'elles apportent,

de l'imagerie irradiante est nécessaire au bon fonctionnement du système. Le développement récent de modèles de respiration permettant de déduire le mouvement respiratoire interne par l'intermédiaire du mouvement externe sans imagerie complémentaire est une alternative intéressante comme outil prédictif du déplacement tumoral. Pour le repositionnement, des systèmes de visualisation 3D du patient à haute vitesse sont évalués afin d'assurer un repositionnement complémentaire au laser, pouvant aussi à moyen terme proposer la gestion du mouvement pendant la fraction.

Ces deux domaines d'innovation ont en commun l'observation de la surface d'un patient afin, dans un cas, de caractériser le mouvement respiratoire externe, et dans l'autre de pouvoir estimer la position corporelle par rapport à une position de référence. C'est dans ce cadre de recherche que s'inscrit ce travail de thèse lors duquel nous avons étudié l'utilisation d'une caméra Temps-de-Vol (ToF) pour ces applications. Cette technologie est assez récente et les derniers développements techniques améliorent considérablement la qualité de mesure, amélioration qui rend envisageable leur utilisation pour mesurer des déplacements d'un ordre de grandeur de quelques millimètres, ce qui est le cas du mouvement respiratoire externe et de la précision désirée pour le repositionnement.

Ce manuscrit est organisé de la façon suivante : la première partie décrit le contexte médical et scientifique qui ont justifié la tenue de ces travaux. Après un rappel sur le cancer et la radiothérapie externe à travers son histoire et ses évolutions, la prise en compte des mouvements en pratique clinique est présentée avec en premier lieu le mouvement respiratoire. Les technologies proposées afin de le circonvenir sont détaillées. De même, la gestion du repositionnement inter-fraction est aussi présentée. Le second chapitre de cette partie présente l'état actuel de la recherche pour la prise en compte de ces mouvements. Les systèmes commerciaux prédisant le mouvement tumoral sont présentés, suivis par les différentes solutions techniques pour suivre le mouvement respiratoire externe, que ce soit à l'aide d'appareils fournissant un signal 1D, de marqueurs ponctuels ou de scanners 3D. Un focus particulier est fait sur la caméra Temps-de-Vol : après description de la technologie, nous présentons les facteurs d'erreur de mesures, puis le champ d'application de cette caméra. Ce chapitre se conclut par la présentation de deux domaines d'application pour lesquels les scanners 3D peuvent apporter une véritable plus-value, soit la création de modèles de respiration reliant mouvement respiratoire interne et mouvement surfacique externe, ainsi que le repositionnement inter-fraction par observation de la totalité de la surface du patient.

La deuxième partie de ce manuscrit détaille les différents travaux effectués tout au long de cette thèse, répartis en quatre chapitres distincts. Le chapitre 3 cherche à caractériser la précision de mesure d'un mouvement surfacique des deux modèles de caméra ToF les plus répandus, la SR4000 (Mesa Imaging) et la CamCube 3.0 (PMD[Vision]). Nous avons ajouté la Kinect de Microsoft dans la comparaison, cette caméra étant elle aussi apparue récemment. Malgré le fait que ce ne soit pas une caméra ToF, ses performances sont encore peu

étudiées et sa place dans l'état de l'art reste encore à déterminer.

Le quatrième chapitre présente les travaux effectués afin d'utiliser une caméra ToF pour la caractérisation du mouvement surfacique respiratoire en temps réel. Nous montrons qu'une telle analyse peut être effectuée avec une précision locale de l'ordre du millimètre à la vitesse de 10 Hz, grâce à la modélisation des données 3D et une méthode d'extraction du mouvement basée sur les points de contrôles issues de cette modélisation. La validation est effectuée sur fantôme anthropomorphique et sur patients par comparaison avec des données issues d'examen 4D-TDM. Pour finir, la faisabilité de la prédiction du mouvement interne est montrée en utilisant le modèle de respiration précédemment développé au sein du laboratoire.

Le cinquième chapitre présente une méthodologie permettant l'extraction de la déformation 3D de la surface entière du patient à haute vitesse. Cette méthode propose la fusion des informations 2D et 3D fournies par la caméra pour contraindre les déformations entre deux acquisitions à l'aide du flot optique connu à partir des images d'intensité de la scène. La validation est effectuée par comparaison avec d'une part, le mouvement de repères visuels sur la peau du patient et d'autre part, un algorithme de recalage élastique. Enfin, une discussion souligne les développements ultérieurs à apporter à la méthode, tant au niveau méthodologique qu'à celui de la validation et du temps de calcul.

Le chapitre 6 contient les développements effectués pour évaluer l'utilisation de cette caméra pour le repositionnement inter-fraction. Après une première partie méthodologique où est présentée une méthode de calibrage spatial permettant la création d'un système stéréoscopique de deux caméras ToF ainsi qu'une analyse de la prise en compte du mouvement respiratoire lorsque le repositionnement concerne des régions thoraciques ou abdominales, une étude clinique effectuée au service de radiothérapie du CHRU de Brest est présentée. Divisée en trois parties, elle présente l'utilisation directe d'une seule caméra, de deux caméras sans puis avec prise en compte du mouvement respiratoire.

Enfin, la dernière partie rappelle les différents travaux effectués ainsi que leurs apports tant méthodologiques que cliniques par rapport à l'état de l'art. Les perspectives ouvertes sont examinées ainsi que les travaux futurs envisagés afin de confirmer et/ou transférer en clinique certains résultats ou méthodes.

Etat de l'art

Chapitre 1

Contexte médical

Sommaire

1.1 Le cancer	26
1.1.1 Définition et description	26
1.1.2 Epidémiologie	26
1.1.3 Le diagnostic du cancer	27
1.1.4 Le traitement du cancer	29
1.2 La radiothérapie externe	30
1.2.1 Historique	30
1.2.2 Principe et étapes du traitement par radiothérapie	31
1.2.3 Erreur due aux mouvements	34
1.3 Prise en charge du mouvement en radiothérapie externe	36
1.3.1 Le mouvement respiratoire	36
1.3.2 Repositionnement entre les fractions	43
1.3.3 Synthèse et enjeux pour la recherche scientifique	45

Ce chapitre a pour but de présenter le contexte médical dans lequel s’inscrit ce travail de thèse. Après avoir présenté le diagnostic et le traitement du Cancer, nous nous focaliserons plus spécifiquement sur le traitement par radiothérapie qui forme à l’heure un des trois piliers de l’action contre le cancer avec la chirurgie et la chimiothérapie. Son histoire à travers le vingtième siècle est marquée par une évolution technique et technologique permettant une délivrance de dose de bonne qualité. Après avoir décrit les différentes étapes du traitement, la dernière section de ce chapitre se concentrera sur la gestion des différents mouvements dégradant la qualité du traitement en diminuant la précision de la localisation tumorale, c’est-à-dire le mouvement respiratoire et l’erreur de repositionnement entre chaque fraction.

1.1 Le cancer

1.1.1 Définition et description

Le cancer se caractérise par une altération génétique qui conduit à une dérégulation cellulaire, généralement entraînée par la modification du gène influençant le contrôle de la multiplication. La cellule ainsi touchée n'est plus soumise au mécanisme de mort cellulaire programmée (ou apoptose) et peut se multiplier à l'infini. Cette prolifération dans un tissu sain met en danger la survie de celui-ci, et peut s'étendre aux tissus environnants et plus loin, jusqu'à la formation de métastases.

On considère le cancer comme une maladie multigénique car 10 à 20 gènes altérés sont nécessaires pour sa formation. Il est montré que ces altérations se produisent en chaîne, l'une favorisant la suivante, et ceci sur une durée allant de 5 à 20 ans. Les facteurs de mutations peuvent être soit endogènes (hérédité), soit exogènes, par l'exposition à des substances cancérogènes, aux radiations, à des agents infectieux... Son dépistage passe soit par l'examen physique, soit par l'observation par imagerie ou par dosage biologique. Son traitement vise à détruire les cellules cancéreuses.

1.1.2 Epidémiologie

L'Institut National du Cancer estime qu'il y a eu en France en 2011 365 500 nouveaux diagnostics de cancers, soit respectivement 207 000 hommes et 158 500 femmes. On considère que 1 200 000 personnes vivent avec ou après un cancer. Les cancers les plus fréquents chez l'homme sont situés dans la prostate, le poumon et le colon/rectum (avec respectivement 71 000, 27 500 et 21 500 cas). Chez la femme, le cancer le plus fréquent est celui du sein avec 53 000 cas, suivi du colon/rectum (19 000 cas) et du poumon (12 000 cas). Le taux de survie à 5 ans moyen tout cancer confondu est de 52 % avec de très fortes disparités en fonction des localisations, il est toutefois admis que le principal facteur associé à ce taux de survie est la rapidité de détection du cancer : dans le cas des cancers dits profonds, les taux sont les plus bas : 13 % chez l'homme (respectivement 18 % chez la femme) pour le poumon, 11 % (14 %) pour l'oesophage.

Les données complètes pour la France en 2005 sont présentées figure 1.1 avec l'incidence, la mortalité et le taux de survie à 5 ans pour chaque localité. Les chiffres indiquent une diminution du taux de mortalité global sur la période 1983-1987 comparée à 2003-2007 avec une forte diminution chez les hommes : -22 % de 208.7 décès à 162.6 décès pour 100 000 hommes. La tendance est moins forte chez les femmes : -14 % de 92.8 à 79.9 décès pour 100 000 femmes. La figure 1.2, extraite d'un rapport de l'Institut National du Cancer montre cette évolution au fil des ans.

La lutte contre le cancer est cause nationale depuis une vingtaine d'années et l'état français lance régulièrement des plans globaux de lutte, le dernier s'étalant sur la période

1.1. Le cancer

Incidence (1980-2005) et mortalité (2005) par cancer en France

Cancer	Incidence	Mortalité
Tous les cancers	319 380	145 762
Prostate	62 245	9 202
Sein	49 814	11 201
Côlon-rectum	37 413	16 865
Poumons	30 651	26 624
Cavité buccale et pharynx	12 270	4 000
Lymphome Malin Non Hodgkinien	10 224	4 203
Vessie	9 679	4 482
Rein	7 949	3 684
Mélanome (peau)	7 401	1 440
Pancréas	7 218	7 787
Estomac	6 794	4 756
Thyroïde	6 672	403
Foie	6 433	7 390
Utérus (corps)	5 774	1 800
Œsophage	4 721	3 850
Myélocome	4 516	2 789
Ovaire	4 375	3 180
Système nerveux	4 120	3 019
Larynx	3 735	1 406
Leucémie lymphoïde chronique	3 224	1 059
Leucémies aiguës	3 082	2 733
Col de l'utérus	3 068	1 067
Testicule	2 002	82
Maladie de Hodgkin	1 544	218
Plèvre	906	1 090

Survie à cinq ans par localisation et par sexe

Hommes		Femmes	
75 % et +			
Lèvre	96 %	Thyroïde	95 %
Testicule	95 %	Maladie de Hodgkin	92 %
Thyroïde	88 %	Mélanome de la peau	89 %
Maladie de Hodgkin	85 %	Lèvre	88 %
Mélanome de la peau	83 %	Sein	85 %
Prostate	80 %	Corps utérin	76 %
50 à 74 %			
Pénis	67 %	Col utérin	70 %
Rein	63 %	Rein	64 %
Vessie	60 %	Os, articulations et cartilages articulaires	62 %
Os, articulations et cartilages articulaires	59 %	Larynx	59 %
Côlon	58 %	Rectum	58 %
Rectum	55 %	Côlon	57 %
Larynx	54 %	Vulve et vagin	52 %
Lymphome malin non hodgkinien	54 %	Vessie	50 %
25 à 49 %			
Nasopharynx	45 %	Langue	45 %
Intestin grêle	45 %	Myélomes multiples	43 %
Myélomes multiples	42 %	Intestin grêle	42 %
Cavité orale	38 %	Ovaire	40 %
Langue	33 %	Fosses nasales, sinus annexes de la face, oreille moyenne et oreille interne	37 %
		Hypopharynx	34 %
		Leucémie aiguë lymphoblastique	29 %
		Estomac	28 %
25 % et -			
Estomac	23 %	Système nerveux central	19 %
Système nerveux central	20 %	Poumon	18 %
Poumon	13 %	Œsophage	14 %
Œsophage	11 %	Mésotéliome de la plèvre	13 %
Foie	7 %	Foie	9 %
Mésotéliome de la plèvre	6 %	Pancréas	7 %
Pancréas	5 %		

FIGURE 1.1 – Incidence et taux de survie à 5 ans pour différents types de cancer

2009-2013. Il prévoit 730 millions d'euros afin d'agir sur trois points : améliorer la qualité des soins par la recherche et le développement, lutter contre les inégalités sociales et géographiques face au cancer, et améliorer la qualité de vie post-cancer (SitePlanCancer (2013)).

1.1.3 Le diagnostic du cancer

Lors de la présence de suspicion de cancer, l'examen histo-pathologique confirme ou infirme sa présence, puis précise la nature de celui-ci, son degré d'évolution et cherche à détecter la présence de métastases. La biopsie, l'analyse des marqueurs tumoraux ainsi que l'imagerie médicale sont les trois piliers de ce diagnostic.

Le prélèvement du tissu suspect est suivi d'une analyse au microscope permettant la détection de cellules tumorales. C'est cette analyse qui détermine si la tumeur est bénigne ou maligne, déterminant ainsi la stratégie thérapeutique. Un bilan sanguin permet ensuite

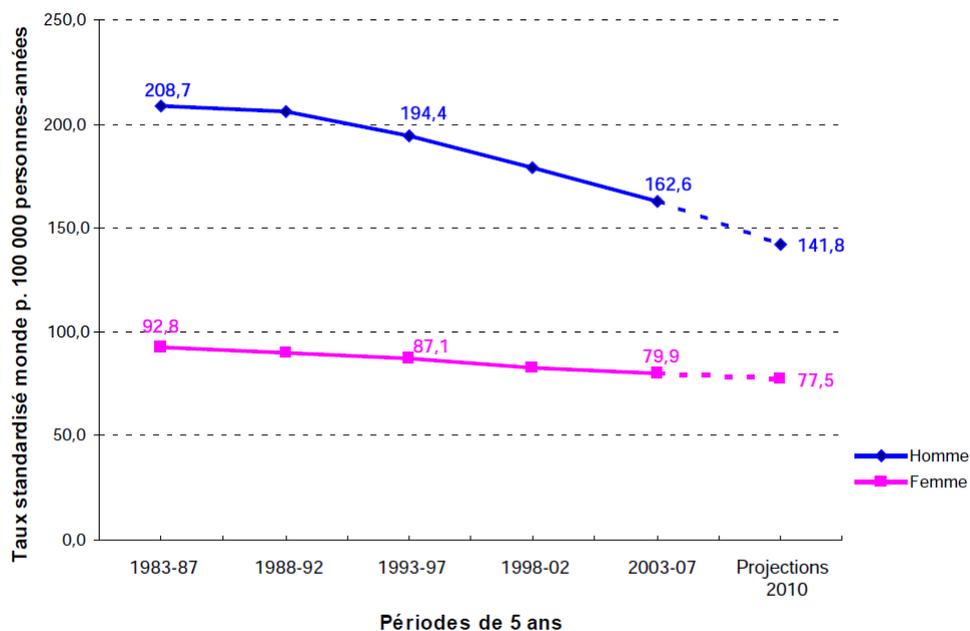


FIGURE 1.2 – Evolution de la mortalité due au cancer en France

de détecter et doser la présence en quantité anormale de marqueurs tumoraux caractéristiques. Certaines cellules saines produisent aussi ces marqueurs, mais en très faible quantité : l'augmentation de la concentration de ces molécules peut correspondre à la présence d'une tumeur, même si des pathologies différentes peuvent expliquer cette augmentation. L'évolution de la concentration d'un marqueur au fil du traitement est aussi généralement un facteur quantitatif reflétant l'évolution de la taille de la tumeur.

Enfin, l'imagerie médicale sert à visualiser les organes internes ainsi que les volumes tumoraux, ce qui servira à la mise en place d'une stratégie thérapeutique visant à leur destruction. Différentes modalités sont utilisées en fonction des localisations. On peut citer l'échographie qui permet l'examen des organes de l'abdomen par ultrason, ainsi que la radiographie et le scanner TDM (pour TomoDensitoMétrie) utilisant les rayons X pour diagnostiquer les localisations thoraciques, abdominales et mammaires. Le cerveau, organe mou et riche en eau, est lui observé par modalité IRM (Imagerie par Résonance Magnétique), qui utilise un champ magnétique pour exciter et relaxer les atomes d'hydrogènes présents dans l'eau. Ces différentes modalités sont dites anatomiques, car elle décrivent la structure des organes, leur forme et leur densité. L'imagerie TEP (Tomographie par Emission de Positons) est dite fonctionnelle car la détection de la désintégration d'un élément radioactif associé à des molécules de glucose permet d'obtenir une carte de concentration. Sachant que les tumeurs consomment une forte quantité de cette substance, l'imagerie TEP donne bien une information fonctionnelle sur la tumeur. La fusion des modalités TDM et TEP permet finalement d'obtenir dans un même repère l'information anatomique et fonctionnelle d'une tumeur, permettant de caractériser au mieux celle-ci.

1.1.4 Le traitement du cancer

Après diagnostic et caractérisation du cancer, le traitement consiste généralement à l'administration et/ou la combinaison de trois interventions différentes qui sont la chirurgie, la radiothérapie et la chimiothérapie.

L'intervention chirurgicale consiste à retirer la tumeur cancéreuse tout en cherchant à préserver l'intégrité physique du patient. La qualité de l'intervention permet une diminution des traces laissées par la chirurgie : par exemple, la majorité des femmes atteintes d'un cancer du sein ne subissent plus d'ablation partielle ou totale à l'issue de l'opération. Lors de l'intervention, le chirurgien peut en plus analyser des ganglions proches afin de détecter la présence de métastases et ainsi potentiellement étendre la zone retirée.

Dans tous les cas, la chirurgie n'assure pas le traitement complet car il peut rester des cellules tumorales soit aux alentours de la zone retirée, soit sous forme de métastases. L'oncologue peut alors administrer des médicaments anti-cancéreux lors d'une chimiothérapie afin de prévenir une récurrence et traiter les métastases. Une troisième méthode consiste à brûler un volume anatomique afin d'en éliminer les cellules avec des radiations de haute énergie. C'est la radiothérapie qui est dite néo-adjuvante si elle est pratiquée avant la chirurgie afin de réduire la taille de la tumeur et/ou diminuer le risque de récurrence ; elle est dite adjuvante si elle suit la chirurgie avec comme objectif de détruire les éventuelles cellules cancéreuses à proximité de l'ablation. La section suivante présente plus en détail la radiothérapie en commençant par son histoire. Le protocole complet est alors présenté pour finir par montrer l'influence du mouvement sur la qualité du traitement de radiothérapie.

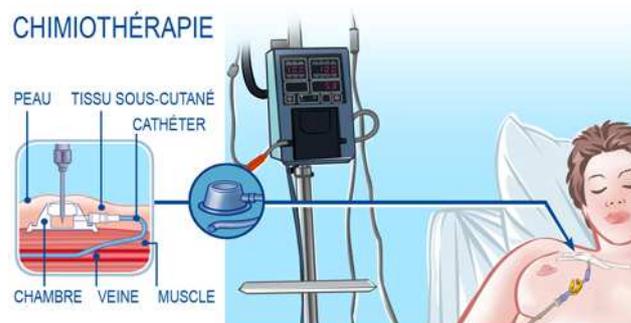


FIGURE 1.3 – Lors de la chimiothérapie, l'administration des substances chimiques bloque la mitose (division du noyau des cellules) et ainsi ciblent les cellules se divisant rapidement, ce qui est le cas des cellules cancéreuses. La plupart des substances actuelles ne différencient pas les cellules saines des cellules cancéreuses et détruisent aussi des cellules à division rapide tels que les cellules sanguines, de l'épithélium intestinal ou des cheveux.

1.2 La radiothérapie externe

1.2.1 Historique

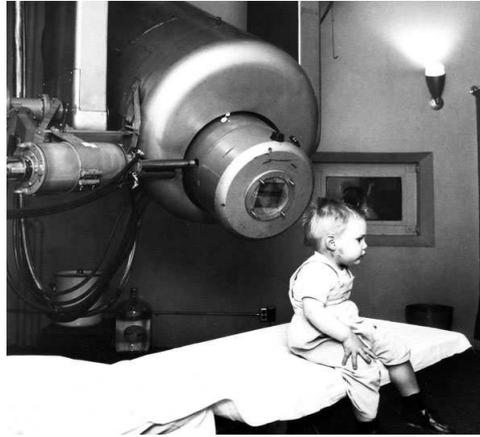


FIGURE 1.4 – Image du premier patient traité pour un retinoblastome par radiothérapie administrée par un accélérateur linéaire en 1957

La genèse de la radiothérapie se trouve dans la découverte des rayons X par le physicien allemand W.C.Roentgen (1845-1923) en 1895, suivie l'année suivante par celle de la radioactivité naturelle par H. Becquerel. Après avoir observé que la radiation créait des brûlures cutanées, des médecins tentèrent très rapidement de soigner des cancers périphériques en les brûlant. Après une courte période durant laquelle la radiation fut prescrite et utilisée sans conscience d'un danger potentiel, les effets néfastes d'une exposition prolongée sont découverts, la radiation endommageant les cellules, entraînant l'apparition de phénomènes cancéreux. L'irradiation n'est aujourd'hui utilisée qu'en radiothérapie et dans un cadre extrêmement réglementé pour éviter des débordements.

L'enjeu principal du traitement par radiothérapie est d'arriver à cibler le mieux possible le volume anatomique potentiellement touché par la prolifération cellulaire tumorale et de lui appliquer une dose optimale tout en évitant l'irradiation des tissus sains environnants. Le 20^{ème} siècle est marqué par une évolution de la qualité du traitement due en grande partie à l'évolution matérielle qui a permis au fil des ans un meilleur contrôle de la production des rayons X. La première moitié du siècle est marquée par la production de rayons X grâce au Radium avec une énergie associée allant jusqu'à 200 kilovolts, ce qui interdisait la pénétration des rayons dans le corps à cause de leur absorption par les tissus. Cependant, des tumeurs superficielles ou proches de la surface de la peau étaient traitées. Un saut technologique s'opère avec l'invention de l'accélérateur linéaire qui permet de générer des ions à haute énergie sans source radioactive, tout en étant assez puissant pour atteindre l'anatomie

profonde. Rolf Widerøe brevète l'invention en 1928 et produit le premier appareil. Le premier patient fut traité en 1953 avec un accélérateur générant des ions à 8 MegaVolts (Figure 1.4). Les technologies étant de plus en plus performantes, c'est l'arrivée de l'informatique qui permet un contrôle et une systématisation des traitements avec l'apparition des premiers logiciels de dosimétrie (Figure 1.5). La radiothérapie et l'oncologie médicale se structurent en France avec la création en 1985 de la Société Française de Radiothérapie Oncologique (SFRO).

On considère que de nos jours plus de 50% des patients atteints de cancer sont traités par radiothérapie à un moment du traitement.

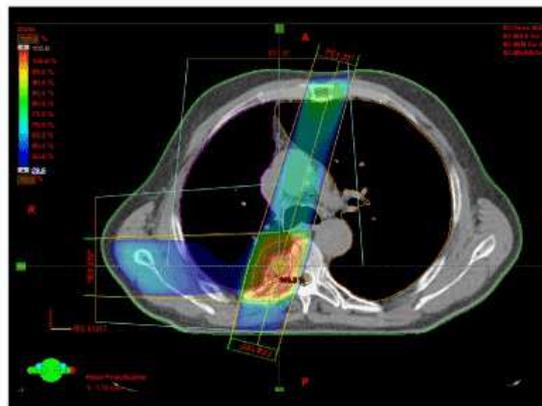


FIGURE 1.5 – Capture d'écran d'un système de planification de traitement qui permet la simulation des lignes de tirs afin d'analyser les conséquences sur la tumeur et les organes à risque.

1.2.2 Principe et étapes du traitement par radiothérapie

L'enjeu de tout traitement par radiothérapie consiste donc à maximiser l'efficacité de l'irradiation sur la tumeur, tout en minimisant la toxicité sur les tissus sains et les organes avoisinants, aussi appelés organes à risque (OAR). En effet, les rayonnements abiment les cellules au niveau de leur ADN afin de les empêcher de se multiplier et donc de proliférer. Il est alors indispensable d'éviter d'irradier des organes sains environnants. La radiobiologie soulignant les capacités régénératives des cellules saines comment étant supérieures à celles des cellules cancéreuses, l'administration du traitement est découpé en fractions afin de laisser le temps à ces cellules de se régénérer et ainsi limiter la destruction des tissus sains environnants. La dose moyenne prescrite varie de 60 à 80 Grays et est délivrée en moyenne en 35 fractions étalées sur sept semaines. La charge humaine apportée par la longueur de traitement tant pour le patient qu'au niveau clinique amène la recherche à essayer de réduire le nombre de fractions : dans certains cas, le traitement est hypo-fractionné en étant réduit à

cinq séances avec une dose délivrée plus élevée allant de 7 à 20 Grays.

Le traitement par radiothérapie peut être découpé en deux étapes : la première consiste à planifier le traitement, la seconde à délivrer la dose prescrite. La phase de planification a pour but de délimiter la zone anatomique à irradier et prévoir les paramètres du traitement : balistique, énergie des faisceaux, forme des collimateurs. Pour ceci, un examen de tomodensitométrie (TDM) est effectué afin d'avoir l'imagerie anatomique du patient en 3D. Cet examen assure un excellent contraste des différents constitutifs du corps (eau, aire, graisse, os). De plus, le niveau de gris (en unité Hounsfield) est corrélé avec la densité électronique nécessaire au calcul de la dose dans le volume d'intérêt. Après segmentation de la tumeur et délimitation des organes à risque, le radiothérapeute délimite le volume cible (Gross Tumor Volume ou GTV). Il y rajoute une marge due aux potentielles diffusions des cellules tumorales autour de la tumeur (Clinical Target Volume ou CTV) puis une autre marge qui tient compte des incertitudes géométriques lors du traitement. On obtient alors le Planning Target Volume ou PTV (Figure 1.6).

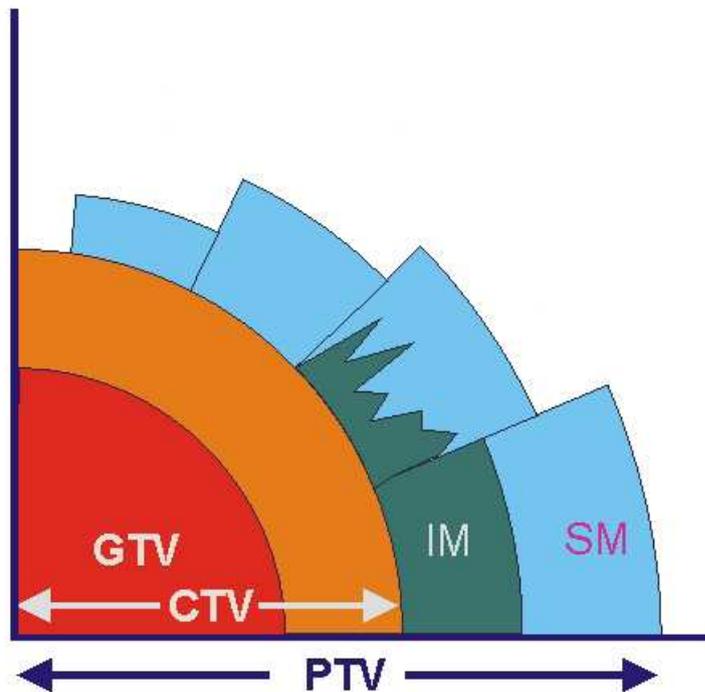


FIGURE 1.6 – Différentes marges sont ajoutées successivement au volume décrivant la tumeur vue par imagerie (GTV). Le CTV prend compte les extensions microscopiques de la tumeur invisibles à l'image. Le PTV ajoute deux autres types de marges au CTV, les marges internes (IM) décrivant le mouvement de la tumeur, et les marges dues au Setup (SM) afin de compenser les incertitudes liées au systèmes.

Le développement technologique a apporté différentes évolutions à l'administration de dose,

1.2. La radiothérapie externe

la plus marquante étant la place prépondérante prise par l'image : on parle alors de radiothérapie guidée par l'image ou IGRT (Image Guided RadioTherapy). La première manière de délivrer la dose était limitée technologiquement par le collimateur qui ne pouvait qu'être de forme rectangulaire, ce qui rendait extrêmement complexe la délivrance complète de la dose à un volume irrégulier en évitant une irradiation de tissus sains. Une des solutions mises en places étaient l'ajout de caches avec des formes spécifiques afin de conformer le faisceau. La mise au point de collimateurs multilames permet ensuite l'application de la radiothérapie conformationnelle 3D lors de laquelle le praticien planifie les angles de tir et la forme du faisceau qui délivre une dose régulière (Figure 1.7). Le perfectionnement des systèmes de planification de traitement ainsi que de l'informatique permettent finalement la Radiothérapie à modulation d'intensité ou IMRT(Intensity Modulated Radiotherapy) qui permet pour chaque position de l'accélérateur de multiples configurations ou faisceaux (de 6 à 12). La planification des angles de tir et des formes de faisceaux devient automatique, le praticien imposant des objectifs de dépôt de dose à atteindre ainsi que des contraintes sur les OARs. L'IMRT connaît enfin une dernière évolution avec l'arc-thérapie à modulation de volume ou VMAT(Volumetric Modulated Arc Therapie) qui remplace les positions fixes de l'accélérateur autour du patient par des arcs de déplacement à 1, 2 ou 3 degrés de liberté du bras au cours desquels l'intensité ainsi que la configuration des lames varie. Cette technique apporte un gain de temps considérable : environ 2 minutes par fraction contre 12 à 20 minutes en IMRT (Figure 1.8).

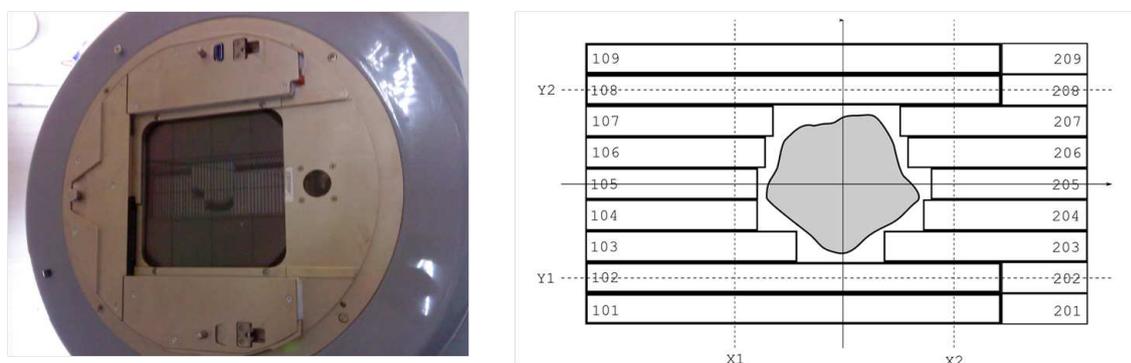


FIGURE 1.7 – Les collimateur multilames permettent d'administrer des faisceaux à géométrie variables afin de cibler celle du volume tumoral.

L'administration de dose étant découpée en fractions, le patient est repositionné à chacune d'entre elles afin de correspondre avec le plan de traitement initial. En routine clinique, des tatouages sur la peau du patient permettent son réalignement le long de lasers calibrés dans la salle de traitement. Ce positionnement est complété par des images de contrôle 2D (imagerie portale, kV) ou 3D (CBCT). Ces images détectent une déviation sur le positionnement au fil des fractions par comparaison de la position des structures osseuses, mais leur dose associée doit être retirée du plan de traitement ce qui diminue la dose réellement déposée dans le PTV.



FIGURE 1.8 – La VMAT change la conformation de l'accélérateur lors de son déplacement en arc de cercle autour du patient

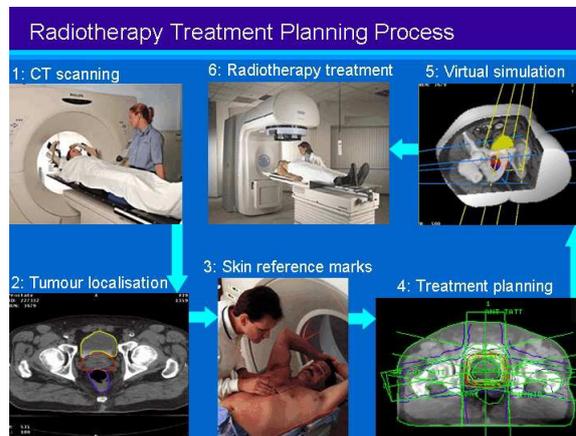


FIGURE 1.9 – Description du processus permettant la planification du traitement avant de pouvoir traiter le patient

1.2.3 Erreur due aux mouvements

On peut distinguer les erreurs de localisation selon trois types :

- l'erreur lors de la définition de l'anatomie. Le rapport AAPM du Task Group 53 relève ce facteur d'imprécision de la définition des volumes anatomiques comme peut-être la plus grande source d'erreur pour l'ensemble du processus de traitement (Fraass et al., 1998). De plus, la variabilité inter et intra physicien lors des segmentations des volumes d'intérêt ajoute une incertitude supplémentaire (Leunens et al., 1993).

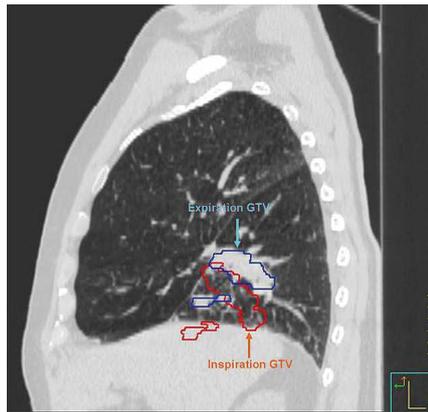


FIGURE 1.10 – Déplacement du GTV associé à une tumeur lors de la respiration (M.Rains (2010))

- les variations anatomiques du patient entre les fractions, qui peuvent être le remplissage d'un organe, la variation du poids du patient ou de la taille de la tumeur. La prise en charge de ce mouvement est faite par l'image, TDM, radiographique ou par ultrasons afin de détecter ces variations physiologiques qui ne sont pas quantifiables à partir d'une observation externe du patient.
- le mouvement intrafraction, qui est principalement le mouvement respiratoire pour les localisations pulmonaires (Seco et al., 2008) et mammaires (Smith et al. (2005), Yue et al. (2007)), mais touche aussi le foie (Wu et al. (2008), Kirilova et al. (2008)) et la prostate (Huang et al. (2002), Aubry et al. (2004)).

L'erreur de définition de la géométrie est limitée en pratique clinique par le suivi d'un ensemble de recommandations émises par des groupes de travail d'experts comme le groupe de travail de l'AAPM 53 (Fraass et al., 1998), mais aussi par le retour d'expérience entre et au sein d'un même service de radiothérapie, l'ensemble des acteurs du traitement étant engagés dans ce processus de qualité.

L'effet du mouvement respiratoire sur la distribution de dose dans le volume tumoral sur l'ensemble des fractions peut être décrit comme un flou sur cette distribution (Leong, 2000). Il a été proposé de modéliser ce flou comme une convolution de la distribution de dose par une fonction de densité de probabilité relative au déplacement du volume par rapport au plan de traitement (Lujan et al. (1999), McKenzie (2000)). Ce modèle n'est pas exact mais donne de bons résultats en pratique afin d'extrapoler les histogrammes dose-volume.

Si l'on s'intéresse à l'effet de ces mouvements sur la délivrance réelle de dose par rapport à la planification, Bortfeld et al. (2002) étudie théoriquement l'effet des mouvements intrafraction en simulant les valeurs de doses dans les organes déplacés. Il montre que l'effet cumulatif sur les fractions amène à ce que la distribution de la dose autour de la valeur prévue forme une gaussienne avec un écart-type égal à 1% de la dose prévue pour un mouvement de

5 mm. George et al. (2003) montre pour les tumeurs du sein que l'impact sur la délivrance de dose est corrélée avec l'amplitude du mouvement respiratoire. Concernant le mouvement du patient durant une fraction pour les localisations tête et cou, Hong et al. (2005) montre que le masque thermoformé ainsi que les méthodes routinières de repositionnement ne suffisent pas à garantir une bonne délivrance de dose.

Une erreur dans la localisation de la zone tumorale a un effet sur la dose administrée mais aussi sur les OARs, surtout si ceux-ci sont proches du volume cible (Peng et al., 2011). Wang et al. (2008) étudie l'impact d'erreurs de repositionnement pour des tumeurs spinales et montre qu'une erreur de 2 mm quelle que soit la direction réduit la couverture de la tumeur de plus de 5% et rajoute plus de 25% de dose aux organes à risque. Concernant les erreurs en rotation, il souligne un effet plus important en cas de cibles multiples ou si la tumeur est éloignée de structures osseuses qui servent au contrôle du repositionnement.

Que ce soit pour des tumeurs du cerveau (Peng et al., 2011) ou du du sein (Jain et al., 2009), toute erreur de localisation induit un traitement moins efficace et plus risqué pour le patient. C'est pourquoi les deux principaux facteurs d'erreur de localisation, la respiration et le repositionnement, sont l'objet de recherches intensives. La prochaine partie présente les solutions cliniques existantes afin d'atténuer leurs effets sur l'administration de la dose.

1.3 Prise en charge du mouvement en radiothérapie externe

1.3.1 Le mouvement respiratoire

Description

La fonction respiratoire a comme objectif d'effectuer des échanges gazeux avec le milieu extérieur. Le lieu où se passe l'échange se situe au niveau des poumons, et le renouvellement de l'air est effectué par une première phase d'expiration qui permet l'évacuation de l'air pauvre en dioxygène (O_2) et riche en dioxyde de carbone (CO_2), suivi d'une seconde phase d'inspiration qui remplit à nouveau les poumons. L'échange gazeux entre le sang (véhicule de l' O_2 et du CO_2 dans le corps) et l'air des poumons s'effectue dans les bronchioles lobulaires par lesquels le dioxygène est capturé et le dioxyde de carbone produit par le cycle de Krebs est rejeté. Un être humain respire en moyenne 23 000 fois par jour, et la mécanique respiratoire se base sur la contraction du muscle diaphragmatique.

Mécanique de l'appareil respiratoire

La mécanique de la respiration est décrite depuis maintenant beaucoup d'années (Gray (1918), Clippe et al. (2003), Mines (1993), Takahashi (1994), Towle (2000)) et peut se décrire

1.3. Prise en charge du mouvement en radiothérapie externe

selon 3 types : ventilation normale, forcée ou assistée.

Lors de la ventilation normale, on considère qu'un volume équivalent à 500 ml est renouvelé à chaque respiration. Lors de l'inspiration, le diaphragme, qui est dans ce cas de figure le muscle inspirateur principal, va se contracter. Lorsque de cette contraction, les dômes diaphragmatiques s'aplatissent en se dirigeant dans la direction caudale ce qui amène automatiquement à une augmentation du volume des poumons (Figure 1.11). La cage thoracique qui protège les poumons est alors soulevée dans la direction antéropostérieure et entraîne avec lui le diaphragme qui lui est rigidement relié.

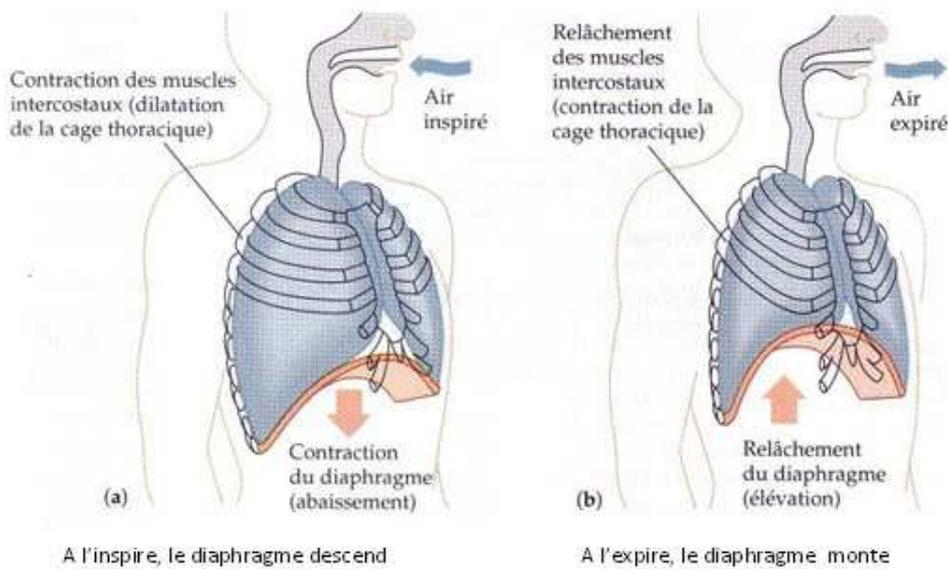


FIGURE 1.11 – Biomécanique respiratoire chez l'homme

Amplitude du mouvement respiratoire

Dans cette partie nous présentons l'amplitude de mouvement due à la respiration en fonction des régions anatomiques. Ce mouvement est présenté selon trois directions : la direction craniale-caudale ou tête-pied (SI), la direction antéro-postérieure (AP) et la direction latérale (LR) comme indiqué sur la figure 1.12. Cette convention sera valable pour l'ensemble du manuscrit.

Le mouvement de tumeurs lors d'un cycle respiratoire a fait l'objet de nombreuses études afin de donner au clinicien un ordre de grandeur lorsqu'il ajoute une marge en fonction de la localisation tumorale lors de la planification. Keall et al. (2006) fait un état des lieux des mouvements respiratoires dans la littérature et les tableaux 1.13 et 1.14 présentent les résultats pour des tumeurs pulmonaires et abdominales. On remarque que l'amplitude de mouvement dans la direction craniale-caudale varie en fonction de la localisation tumorale

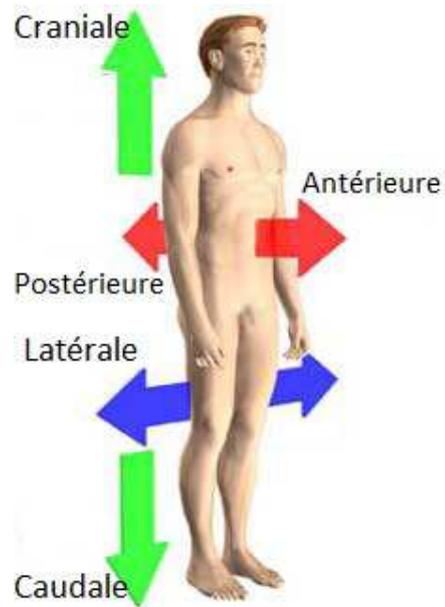


FIGURE 1.12 – Indication des différentes directions anatomiques utilisées dans ce manuscrit

par rapport au diaphragme : dans les poumons, les tumeurs situées dans le lobe supérieur bougent significativement moins dans la direction SI que dans le lobe inférieur. L'observation est aussi valable pour les tumeurs abdominales avec une amplitude de mouvement diminuant lorsque l'on s'éloigne du diaphragme. On remarque que le mouvement dans les directions AP et LR est sensiblement plus faible avec cependant des cas où l'amplitude est forte.

Prise en compte du mouvement respiratoire en radiothérapie

La pratique clinique cherche et développe des méthodes qui prennent en compte le mouvement respiratoire. On peut adapter la planification en rajoutant une marge censée représenter le mouvement de la tumeur, ou tenter de bloquer la respiration du patient ou bien encore n'irradier celui-ci qu'à un seul instant du cycle respiratoire en partant du postulat que la tumeur est à la même localisation. On peut enfin laisser le patient respirer librement et asservir l'irradiation à la respiration, ce qui nécessite de connaître la position de la tumeur en temps réel.

Prise en compte dans la dosimétrie Comme présenté section 1.2.2, lors de la dosimétrie le radiophysicien détermine des marges autour du volume tumoral visible (GTV) afin de tenir compte des erreurs potentielles : cellules voisines cancéreuses, erreur géométrique lors de l'irradiation. Le comité international pour les unités et mesures en radiothérapie (ICRU) définit en plus des marges existantes une marge supplémentaire prenant en compte tout mouvement, dont le mouvement respiratoire, l'ITV (Internal Target Volume) :

" The ITV includes a margin to account for physiological patient movements that are unable

1.3. Prise en charge du mouvement en radiothérapie externe

Référence	Direction		
	Cranio-caudal	Antéro-Postérieur	Latéral
Barnes et al. (2001), BP	18.5 (9-32)	–	–
Barnes et al. (2001), MP, HP	7.5 (2-11)	–	–
Chen et al. (2001)	(0-50)	–	–
Ekberg et al. (1998)	3.9 (0-12)	2.4 (0-5)	2.4 (0-5)
Engelsman et al. (2001), MP,HP	(2-6)	–	–
Engelsman et al. (2001), BP	(2-9)	–	–
Erridge et al. (2003)	12.5 (6-34)	9.4 (5-22)	7.3 (3-12)
Ross et al. (1990), HP	–	1 (0-5)	1 (0-3)
Ross et al. (1990), MP	–	0	9 (0-16)
Ross et al. (1990), BP	–	1 (0-4)	10.5 (0-13)
Grills et al. (2003)	(2-30)	(0-10)	(0-6)
Hanley et al. (1999)	12 (1-20)	5 (0-13)	1 (0-1)
Murphy et al. (2002)	7 (2-15)	–	–
Plathow et al. (2004), HP	4.3 (2.6-7.1)	2.8(1.2-5.1)	3.4 (1.3-5.3)
Plathow et al. (2004), MP	7.2 (4.3-10.2)	4.3 (1.9-7.5)	4.3 (1.5-7.1)
Plathow et al. (2004), BP	9.5 (4.5-16.4)	6.1 (2.5-9.8)	6.0 (2.9-9.8)
Seppenwoolde et al. (2002)	5.8 (0-25)	2.5 (0-8)	1.5 (0-3)
Shimizu et al. (2001)	–	6.4 (2-24)	–
Sixel et al. (2003)	(0-13)	(0-5)	(0-4)
Stevens et al. (2001)	4.5 (0-22)	–	–

FIGURE 1.13 – Tableau tiré du rapport du groupe de travail 76 récapitulant le mouvement (moyenne (min-max)) en millimètre des tumeurs situées dans le poumon lors de la respiration dans différentes études (Keall et al., 2006). BP = bas des poumons. MP = milieu des poumons. HP = haut des poumons.

to be accounted for during treatment. This may include movement of the gut, beating of the heart or respiration. The margin required is known as the internal margin (IM) and may vary in height, breadth and depth based on the location within the body. The ITV is a newer concept that attempts to divide treatment inaccuracies into internal patient factors and external factors. If a method to reduce the effect of internal movements is used (eg. respirator gating) then the ITV can be substantially reduced."

Il n'y a pas de norme concernant la taille de ces marges en fonction de la localisation de la tumeur, seulement un ordre de grandeur associé à ces locations, fourni par les études présentées dans Keall et al. (2006). Il n'est cependant pas rare de voir ces marges atteindre 1 ou 1,5

Localisation	Référence	Type de respiration	
		légère	profonde
Pancréas	Suramo et al. (1984)	20 (10-30)	43 (20-80)
	Bryan et al. (1984)	20 (0-35)	–
Foie	Weiss et al. (1972)	13±5	–
	Harauz and Bronskill (1979)	14	–
Rein	Suramo et al. (1984)	25 (10-40)	55 (30-80)
	Davies et al. (1994)	10 (5-17)	37 (21-57)
	Suramo et al. (1984)	19 (10-40)	40 (2-70)
	Davies et al. (1994)	11 (5-16)	–
Diaphragme	Wade (1954)	17	101
	Korin et al. (2005)	13	39
	Davies et al. (1994)	12 (7-28)	43 (25-57)
	Weiss et al. (1972)	13±5	–
	Giraud et al. (2001)	–	35 (3-95)
	Ford et al. (2002)	20 (13-31)	–

FIGURE 1.14 – Tableau tiré du rapport du groupe de travail 76 récapitulant le mouvement (moyenne (min-max)) en millimètre des organes de l'abdomen durant la respiration (Keall et al., 2006).

cm si la tumeur est située sur les coupes diaphragmatiques par exemple. Le clinicien part généralement d'un ordre de grandeur et adapte ensuite en fonction des OARs environnants.

Les traitements en respiration contrôlée On peut aussi chercher à limiter voire contrôler le mouvement respiratoire afin de réduire le déplacement de la tumeur et donc pouvoir se permettre de ne pas rajouter de marges spécifiques au mouvement respiratoire lors de la planification. On peut tout d'abord demander au patient de bloquer sa respiration en fin d'inspiration : c'est le blocage respiratoire (en anglais DIBH : Deep Inspiration Breath Hold) qui consiste à demander au patient de retenir sa respiration en fin d'inspiration profonde (Figure 1.15, Barnes et al. (2001)). Pour ceci, un apprentissage est effectué au cours duquel son signal respiratoire mesuré par un spiromètre lui sera indiqué avec un seuil qui, une fois franchi, indique le moment d'arrêt de la respiration. Le spiromètre est indiqué dans cette application puisqu'il mesure la variation du volume pulmonaire, la stabilité du signal respiratoire indiquant à priori une position constante des poumons. Cette technique est très adaptée pour les tumeurs pulmonaires mais aussi en cas de cancer des seins et apporte un deuxième avantage médical : le choix d'effectuer le blocage respiratoire à inspiration complète est justifié par

1.3. Prise en charge du mouvement en radiothérapie externe

le fait qu'à cet instant respiratoire, les tissus des poumons sont dilatés, moins denses, ce qui limite l'effet négatif de l'irradiation induite par les rayonnements X.



FIGURE 1.15 – Exemple du système d'asservissement avec blocage respiratoire volontaire avec le feedback vu à l'aide de lunettes.

Une évolution du DIBH est le système ABC (Active breathing coordinator, Elektra) qui bloque la respiration du patient activement par l'intermédiaire d'une valve à travers laquelle le patient respire (Wong et al. (1999), Remouchamps et al. (2003)). Le signal respiratoire aussi fourni par un spiromètre associé au seuil d'amplitude respiratoire désigné comme instant de blocage permet le blocage automatique de la valve pour une durée allant de 15 à 30 secondes. L'irradiation est alors effectuée jusqu'au déblocage de la valve.

En condition clinique, certains patients ne peuvent pas tenir leur respiration sur des durées supérieures à quelques secondes. De plus, l'apprentissage nécessaire allonge le temps de traitement clinique. Dans l'optique de dépasser ces problèmes, on a rapidement cherché à synchroniser l'irradiation avec le signal respiratoire afin de laisser le patient respirer et irradier lorsque l'instant respiratoire requis est atteint : c'est le "Gating" (Kubo and Hill, 1996). La principale contrainte dans ce système est d'arriver à mesurer en temps réel le signal respiratoire avec précision et robustesse. La majorité des méthodes mesurent le signal respiratoire externe à la surface du patient et deux appareils commerciaux sont particulièrement utilisés. Le premier est le système Varian Real-time Position Management (RPM) (Varian Medical Systems, Palo Alto, CA) qui est constitué d'un bloc plastique transparent sur lequel sont fixés deux marqueurs infrarouge passifs (Figure 1.16). Ce bloc est attaché sur l'abdomen du patient à l'aide d'une ceinture. Une caméra CCD émet des infrarouges et suit en temps réel le déplacement des marqueurs dans l'espace. Le deuxième système est la ceinture de pression commercialisée par la société Anzai qui est constituée d'un capteur de pression lui aussi fixé sur l'abdomen du patient avec une ceinture. Lorsque le patient inspire, la ceinture se serre et le capteur détecte une variation de pression, capturant ainsi le signal respiratoire.

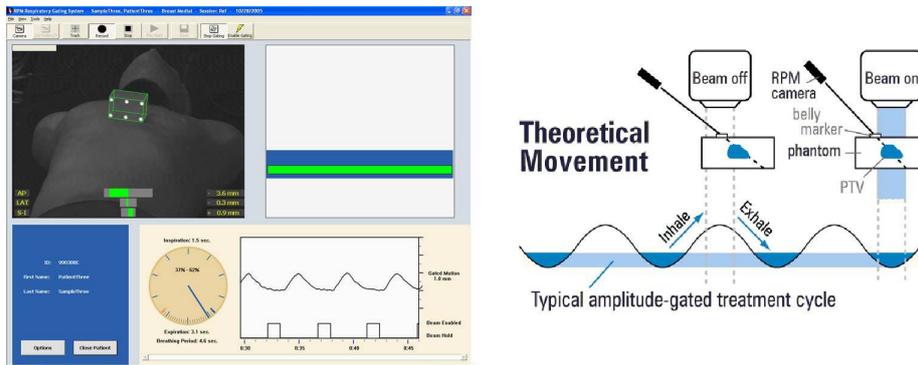


FIGURE 1.16 – A gauche : capture d'écran du logiciel associé au système RPM. A droite : schéma décrivant le principe de fonctionnement du Gating.

Les scanners 3D peuvent aussi fournir le signal respiratoire, et ceci sans contact avec le patient et sur différentes zones anatomiques. Johnson et al. (2004) et Tarte et al. (2006), avec le système commercial VisionRT utilisant la lumière structurée, Xia and Siochi (2012) avec une caméra Kinect ainsi que Penne et al. (2008) et Müller et al. (2009) avec une caméra Temps-de-Vol, montrent qu'il est possible de suivre différents signaux respiratoires et de distinguer respiration abdominale et respiration thoracique.

Le Gating peut enfin être effectué en mesurant le signal respiratoire à l'aide de marqueurs internes préalablement fixés chirurgicalement, en percutané ou à l'aide d'un cathéter bronchoscopique si l'on cherche à atteindre les poumons (Kitamura et al., 2002). Le système est développé conjointement par l'université d'Hokkaido et Mitsubishi et suit des billes d'or de 2 mm qui sont implantées près de la tumeur. Deux imageurs kV stéréotactiques capturent la position des marqueurs en 3D plusieurs fois par secondes. La segmentation automatique des images permet le suivi en temps réel des marqueurs.

Les traitements en respiration libre Il est possible de tenir compte du mouvement respiratoire en radiothérapie sans rajouter de marges lors de la planification et sans limiter l'irradiation à un instant du cycle respiratoire.

La première solution est mécanique et consiste à utiliser un cadre stéréotaxique qui viendra comprimer le diaphragme du patient à l'aide d'une vis de serrage afin de limiter les mouvements internes : on considère cette respiration comme légère et limitant fortement le mouvement des organes internes (Lax et al. (1994), Blomgren et al. (1995)). Par contre, Bouilhol et al. (2012) montre que pour les tumeurs du poumons, l'intérêt est limité sauf si la tumeur est situé dans les lobes inférieurs.

L'autre manière d'effectuer le traitement en respiration libre est de suivre en temps réel la tumeur pour modifier l'orientation des lignes de tir et ainsi coller au plus près de la tumeur.

1.3. Prise en charge du mouvement en radiothérapie externe

Cependant, connaître ou du moins estimer dynamiquement où se situe le volume à irradier sans imagerie irradiante continue est complexe, c'est pourquoi ce domaine est actuellement sujet à énormément de recherches et développements technologiques et sera traité en détail dans le chapitre 2.



FIGURE 1.17 – Le système commercial Cyberknife estime la position d'une tumeur en temps réel par suivi de la position de marqueurs externes fixés sur la veste du patient (Bondiau et al., 2007).

1.3.2 Repositionnement entre les fractions

Le repositionnement entre les différentes fractions du traitement de radiothérapie est une étape cruciale dans le processus assurant une délivrance de la dose optimale. Les techniques mises en place en clinique sont basées sur un mécanisme double, c'est-à-dire un positionnement systématique à chaque fraction basé sur des tatouages alignés dans la salle de traitement sur des lasers fixes, couplé à une imagerie régulière qui détecte toute déviation au fil des jours. Ce schéma est respecté quelque soit l'emplacement de la tumeur à traiter, seul la fréquence à laquelle une imagerie est effectuée change : les localisations sensibles au mouvement ou la position du patient peuvent se voir appliquer une vérification journalière.

Lasers

Le système de positionnement systématique en clinique se repose sur un système de lasers installés dans la salle (Figure 1.18). Ils sont disposés de manière à décrire les trois plans anatomiques du patient de manière à ce que ces trois plans se coupent au niveau de l'isocentre de l'accélérateur. Ce système est aussi installé dans la salle d'examen TDM afin de pouvoir repérer comment les lasers marquent le patient lors de la planification. Des tatouages sont dessinés à l'intersection des lasers et de la peau du patient. Un fois en salle de radiothérapie, les manipulateurs n'ont plus qu'à aligner les tatouages avec les lasers : le volume tumoral est donc repéré dans la même position que lors de la planification.



FIGURE 1.18 – Exemple clinique de ciblage d’une tumeur du sein. On y distingue les trois lasers qui permettent au manipulateur d’aligner le patient à la position de référence

Cette méthode de positionnement est extrêmement pratique d’un point de vue clinique puisqu’elle est rapide et n’ajoute pas de dose additionnelle. Malgré tout, sa précision est limitée par plusieurs facteurs tels que l’épaisseur des tatouages, leur dessin à main levée et le réalignement manuel, ce qui entraîne une incertitude due à la variabilité intra et inter utilisateur. Elle permet malgré tout une première estimation du repositionnement mais se doit d’être confirmée ou corrigée par une méthode intermédiaire pour prévenir tout risque de déviation au fil des fractions : en effet, les repères sont situés sur la peau, et en cas d’amaigrissement ou de changement de physionomie (possible lors de la radiothérapie) ils peuvent ne plus du tout indiquer la bonne localisation interne. La fréquence de ce contrôle dépend de la localisation tumorale et de la propension de la région anatomique à se déformer soit naturellement soit par amaigrissement : une tumeur ORL ne sera quasiment soumise à aucune dérive alors que la zone prostatique est continuellement déformée et nécessite souvent une image de contrôle par jour.

Imageries de contrôle

Imagerie portale Le système d’imagerie portale (EPID ou " Electronic portal imaging device ") est un dispositif situé sous la table de traitement qui capture un faisceau de rayons X provenant de l’accélérateur linéaire. L’image ainsi acquise donne l’information 2D de son positionnement dans le repère de l’accélérateur. L’image résultante est alors généralement comparée visuellement à l’image prise à la fraction 1 (Figure 1.19), même si des études sont menées pour développer des algorithmes de comparaison automatique associant deux images et le volume TDM de planification (Clippe et al., 2003). La technologie, quoique vieille d’une quarantaine d’année, est utilisée commercialement depuis 1988 mais sa principale limitation,

1.3. Prise en charge du mouvement en radiothérapie externe

en dehors de l'irradiation, est la faible qualité de l'image, tant au niveau contraste que niveau détail. De plus, de par l'image 2D fournie, on ne peut que déterminer le déplacement selon deux axes. Afin de positionner le patient correctement, il est alors nécessaire d'effectuer deux images portales, une dans la direction antéropostérieure, l'autre selon la direction latérale.

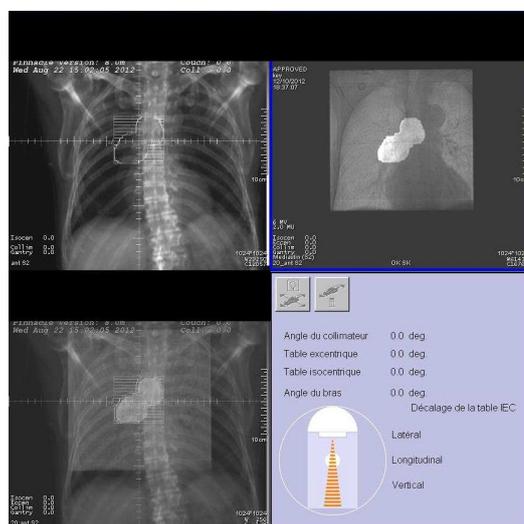


FIGURE 1.19 – Capture d'écran du logiciel permettant la comparaison visuelle entre deux images portales avec surimpression du PTV préalablement prévu

Images kV et MV Les images kV et MV renvoient à la technologie "cone beam" dans laquelle l'acquisition d'une image est faite à partir d'une source de rayons X diffusée en cône. Cette technologie permet la reconstruction 3D de l'objet observé, et peut être vu comme la suite logique des recherches cherchant à déduire un volume 3D de deux images portales collatérales. Il n'est ainsi nécessaire d'effectuer qu'une image pour avoir toute l'information.

1.3.3 Synthèse et enjeux pour la recherche scientifique

Le cancer est à l'heure actuelle un enjeu de santé publique majeur, et son traitement est à l'image de la maladie, varié et complexe. La pratique clinique a recours à trois types d'actions thérapeutiques qui sont la chirurgie, la chimiothérapie et la radiothérapie, et l'action d'ensemble est la plupart du temps constituée d'une combinaison de ces actions. Lorsque l'on se concentre sur la radiothérapie, on se rend compte que les techniques de ciblage évoluent considérablement tout comme l'administration de la dose grâce à l'apport de l'image et les avancées technologiques. Cependant, le domaine de la localisation de la tumeur durant les fractions reste sensible. Le repositionnement inter-fraction influe sur la qualité de l'administration de la dose quelque soit la localisation. De plus, lorsque la tumeur est située dans une région anatomique sensible au mouvement respiratoire, il est indispensable d'en tenir

compte et ceci est pratiqué de différentes manières :

- On peut tout d'abord élargir la zone d'irradiation afin de s'assurer que la tumeur reçoive de la dose à tout instant du cycle respiratoire. Cette solution présente l'avantage de gérer la problématique du mouvement une fois pour toute lors de la planification par l'ajout de marges. Le défaut majeur est clairement d'augmenter la zone d'irradiation et donc de toucher les tissus sains environnants voire les organes à risque, augmentant d'autant les chances de récurrence.
- On peut ensuite décider de n'irradier le patient qu'à un seul instant du cycle respiratoire. Ceci permet de ne pas rajouter de marges supplémentaires. Les méthodes imaginées nécessitent le suivi du mouvement respiratoire afin de déterminer l'instant du cycle où l'irradiation peut s'effectuer. Le patient peut être mis à contribution (respiration tenue), ou alors le système peut décider de lui-même quand irradier ("Gating"). Le principal défaut de ces méthodes est l'allongement de la durée du traitement puisqu'il nécessite la mise en place d'un système complexe, l'apprentissage du patient dans certains cas et n'irradie pas en continu mais seulement sur une fraction du cycle respiratoire.
- On peut chercher à annihiler le mouvement respiratoire à l'aide de moyens de contention, ce qui permet l'irradiation continue en diminuant les marges ajoutées à la planification. Cependant, la mise en place du système est complexe et l'on manque de certitude sur le mouvement interne réel.
- On peut enfin chercher à suivre la tumeur afin de déplacer l'accélérateur linéaire en conséquence et ainsi irradier la tumeur de manière continue. Deux systèmes commerciaux existent pour les localisations pulmonaires et prédisent le déplacement tumoral grâce au suivi du mouvement respiratoire externe connu par des marqueurs accrochés au thorax du patient. Ces systèmes sont une avancée importante pour augmenter la qualité du traitement tout en limitant les irradiations d'organes sains, ils nécessitent cependant de multiples images irradiantes.

Chapitre 2

Contexte scientifique

Sommaire

2.1	Systèmes commerciaux de suivi tumoral	48
2.2	Observation du mouvement surfacique	51
2.2.1	Systèmes acquérant un signal 1D	51
2.2.2	Marqueurs ponctuels	53
2.2.3	Les scanners tridimensionnels	55
2.2.4	Les caméras Temps-de-Vol	58
2.3	Les modèles de respiration	64
2.3.1	Les modèles basés sur un point	65
2.3.2	Les modèles basés sur un champ de déformation	65
2.4	Le repositionnement par l'analyse de la surface du patient	66
2.5	Synthèse	68

Ce chapitre va aborder les dernières innovations technologiques, commercialisées ou non, qui cherchent à éliminer les incertitudes de localisation tumorale dues au mouvement du patient entre et pendant les fractions. Nous présenterons tout d'abord les systèmes de suivi tumoral en temps réel qui sont déjà disponibles dans le commerce. Nous listerons ensuite les technologies permettant de caractériser et/ou quantifier le signal respiratoire, puis présenterons les systèmes permettant le suivi de ce mouvement à la surface du patient. Parmi ceux-ci, les scanners 3D ouvrent la voie à des applications requérant l'analyse du mouvement sur de larges portions du corps, c'est pourquoi nous parlerons ensuite d'une première application qui voit le développement de modèles établissant une corrélation entre le mouvement respiratoire externe et les déformations internes, le but étant de pouvoir prédire le déplacement d'une tumeur par l'analyse de la déformation du thorax et de l'abdomen. La deuxième application concerne l'apport de ces scanners 3D pour l'aide au repositionnement du patient entre et pendant un fraction par analyse de surface.

2.1 Systèmes commerciaux de suivi tumoral

Trois systèmes permettant de suivre la tumeur en temps réel existent déjà et permettent d'atteindre une précision sub-millimétrique. Le système Calypso 4D (Calypso Medical Technologies, Seattle, WA) est utilisé pour la localisation de tumeurs ou de localisation post-opération au niveau de la prostate (Figure 2.1). Le système détecte trois marqueurs électromagnétiques actifs ou passifs de $8 \times 2 \text{ mm}^2$ qui sont insérés dans le lit prostatique par guidance ultrasonore. Une utilisation pour d'autres localisations est en cours de validation. Un émetteur de champ magnétique est fixé au dessus du patient à faible distance avec le détecteur et permet une localisation des marqueurs à une fréquence de 10 Hz. Le calibrage spatial entre l'isocentre de l'accélérateur linéaire et le système de détection est effectué par la présence de marqueurs infrarouges fixés sur le générateur de champ : observés par trois caméras infrarouges, une procédure de calibrage permet de relier les repères ensemble.

Le système agit sur le processus médical par l'arrêt automatique de l'irradiation si la tumeur sort d'un espace préalablement déterminé par l'utilisateur. Au delà de l'acte médical qui permet l'insertion des marqueurs dans la prostate, la principale limitation de ce système est la migration des marqueurs électromagnétiques au fil des séances : le système tient cependant compte de ce phénomène en vérifiant la position des trois marqueurs entre eux afin de détecter toute déviation. Une autre limitation est le faible champ de détection électromagnétique qui est limité à 20 cm, ce qui limite le traitement de patients à forte corpulence.

Comme la prostate est touchée par un mouvement et une déformation complexe et difficilement reproductible à chaque fraction, il semble pour l'instant difficile de se passer de l'imagerie et des marqueurs internes pour s'assurer de la position de celle-ci, malgré les difficultés engendrées par l'insertion des marqueurs et l'ajout de dose par l'imageur. Cependant, cette limitation n'existe pas dans le cas du mouvement respiratoire car l'on peut le caractériser par son mouvement externe, à la surface du patient. Même si ce mouvement surfacique n'est pas directement fiable au mouvement interne de par la physiologie séparant la peau des organes spécifique à chaque patient, une relation semble exister malgré une controverse scientifique quant au degré de corrélation atteignable. C'est d'ailleurs ce qu'effectue le système Cyberknife (Accuray Incorporated) en établissant une corrélation temporaire entre des marqueurs externes infrarouges et la physiologie interne observée par imagerie (Figure 2.2). Cette corrélation est ensuite confirmée ou corrigée par imagerie radiographique régulière durant la fraction, ce qui permet ensuite par suivi des marqueurs infrarouges de déduire la position de points d'intérêt comme une tumeur. Cependant cette solution n'est pas totalement satisfaisante sur au moins trois points :

- Elle nécessite des images à chaque fraction afin d'établir le modèle de corrélation mouvement externe/mouvement interne.

2.1. Systèmes commerciaux de suivi tumoral

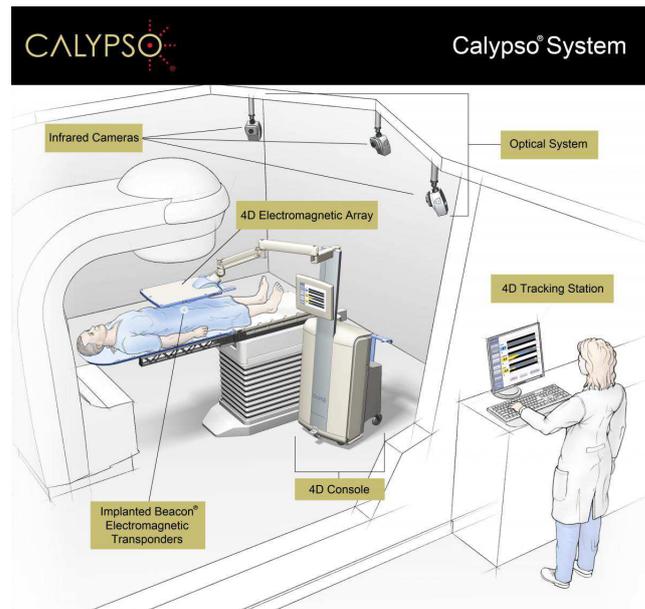


FIGURE 2.1 – Illustration du dispositif Calypso permettant de suivre des tumeurs de la prostate en temps réel (SiteCalypso (2013)).

- Le modèle de corrélation est temporaire et nécessite une mise à jour régulière durant la fraction.
- Le mouvement externe est caractérisé par des marqueurs externes qui nécessitent une interaction avec le patient. De plus, on a une information ponctuelle sur le mouvement de la surface. Enfin, il est enfin nécessaire de contrôler que les marqueurs décrivent bien le mouvement surfacique car ils sont fixés à une veste, et non directement sur la peau du patient.

Un dernier système existant, conçu initialement pour effectuer le repositionnement interfraction, est le système ExacTrac de BrainLAB. Il est lui aussi constitué de deux imageurs à rayons X d'énergie kV positionnés obliquement par rapport au patient ce qui permet, à l'aide des deux images 2D de déterminer la position 3D de localisations internes ou de marqueurs (Figure 2.3). Une fois le repositionnement à effectuer déterminé par fusion des images radiographiques avec un volume CT préalablement acquis, des marqueurs infrarouges réfléchissants fixés au patient sont observés afin de repositionner le patient. La dernière version propose le suivi tumoral par le tracking de marqueurs infrarouges fixés sur le patient de manière similaire au Cyberknife.

En conclusion, les deux systèmes présentés sont assez similaires dans leur approche, prédisant la position tumorale par l'intermédiaire de capteurs 3D ponctuels. Le système Calypso est à part car sa région d'application, la prostate, requiert une insertion directe de ces capteurs car aucun lien ne peut être fait entre un mouvement externe et la position de la prostate.

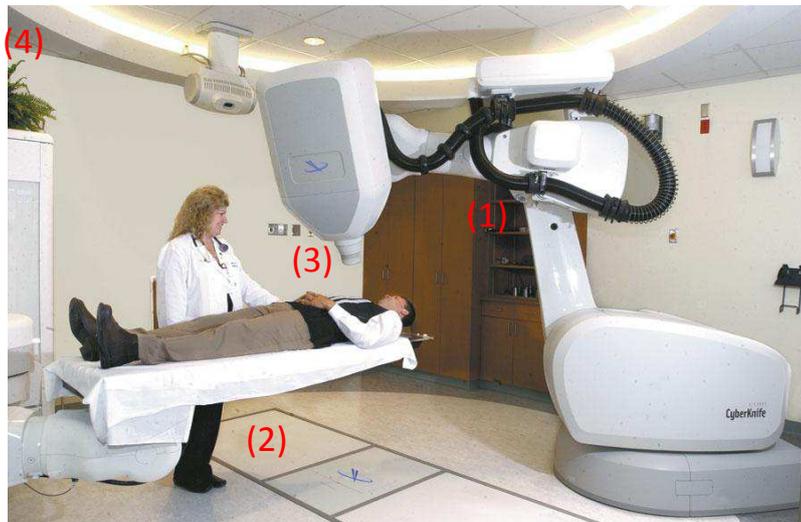


FIGURE 2.2 – Illustration du dispositif Cyberknife qui est constitué de : (1) l'accélérateur linéaire guidé par un bras robotisé ; (2) deux plaques radiographiques scellées au sol ; (3) une veste couverte de marqueurs infrarouges suivies par (4) trois caméras CCD.



FIGURE 2.3 – Le dispositif ExacTrac est constitué de deux imageurs par rayons X fixés obliquement par rapport au patient. Conçu initialement pour le repositionnement, un suivi des marqueurs optiques infrarouge permet le suivi d'une tumeur.

Le choix des marqueurs électro-magnétiques s'explique par leur capacité à être détectés sans ligne de vue avec le détecteur (voir partie 2.2.2). Les deux autres systèmes cherchent à prédire le mouvement respiratoire interne par le suivi du mouvement externe, postulant sur l'existence d'une corrélation entre les deux. Cette relation est temporaire et déterminée à partir d'images irradiantes qui fournissent une vérité terrain sur le mouvement interne, et de capteurs externes ponctuels. Les systèmes fournissent une solution commerciale fiable mais limitée par deux facteurs. Le premier est une prise en compte du mouvement externe limitée par le nombre de capteurs utilisés, alors que seule l'information sur l'ensemble du tronc permet de distinguer les types de respiration (Müller et al., 2009). Le danger est alors d'établir une corrélation moins fiable, c'est pourquoi des images irradiantes sont effectuées régulièrement afin de vérifier la véracité du modèle durant la fraction. La deuxième limitation est la nature temporaire de ce modèle, obligeant à réeffectuer des images à chaque fraction, ajoutant ainsi une dose au patient.

Dans la partie suivante, après la description des technologies fournissant une trace respiratoire 1D en temps réel, nous présenterons les systèmes de suivi du mouvement surfacique en temps réel, en commençant par les différents types de marqueurs. Nous présenterons ensuite les scanners 3D avec un focus particulier sur la caméra Temps-de-Vol.

2.2 Observation du mouvement surfacique

Il existe différentes manières de mesurer le signal respiratoire en temps réel. On peut tout d'abord citer les appareils fournissant un signal 1D qualitatif, celui-ci servant dans des applications tel que le "Gating". La catégorie des marqueurs fournit elle une information quantitative. Ceux-ci sont fixés sur la peau ou par l'intermédiaire d'un support, et fournissent une information 3D quantitative ponctuelle. Il y a enfin les scanners 3D qui fournissent sans contact une cartographie 3D complète du tronc.

2.2.1 Systèmes acquérant un signal 1D

Le spiromètre

Le spiromètre fournit en temps réel le signal respiratoire en mesurant la variation du volume d'air respiré. Ce signal n'est pas concrètement une mesure du mouvement respiratoire, mais permet son analyse qualitative. Cette mesure est robuste puisqu'elle impacte directement sur la variation de volume des poumons, et une corrélation semble exister avec la position de localisations comme le coeur et le diaphragme (Nojonen et al., 2008). Le spiromètre est utilisé comme instrument de diagnostic dans différentes pathologies impliquant des désordres respiratoires (Chowdhary et al. (2011), Jandt et al. (2011)) mais aussi pour

effectuer le Gating respiratoire tant en imagerie qu'en traitement (Kalender et al. (1990), Moroni et al. (2003)). Il est cependant montré que les patients atteints de pathologies respiratoires peuvent avoir des difficultés à utiliser un spiromètre sur une longue durée (Werner et al., 2010).

La ceinture de pression

Un ceinture de pression est constituée d'un capteur de pression fixé à une ceinture attachée au niveau de l'abdomen du patient. La variation de volume lors d'un cycle respiratoire est directement mesurée comme la variation de pression au niveau du capteur. Ce signal analogique est amplifié puis numérisé pour pouvoir faire office de marqueur temporel dans des applications tel que le scanner 4D-TDM ou 4D-TEP (où chaque acquisition est associée à un instant du cycle respiratoire par l'intermédiaire du signal 1D fourni par la ceinture de pression) ou lors d'une irradiation soumise à un "Gating" respiratoire. Pour l'utilisation du modèle commercial le plus utilisé, la ceinture AZ-733V (Anzai Medical Co. Ltd.), deux capteurs sont fournis avec une sensibilité différente, l'un étant plus adapté aux physionomies larges que l'autre. La principale limitation clinique est la perte de signal si la ceinture glisse et se détend, ce qui peut arriver pour les physionomies où l'abdomen est bombé.

Le système infrarouge passif de Varian

Le système RPM (Real-time Position Management) mesure le déplacement du torse ou de l'abdomen du patient dans les trois directions anatomiques. Ce système, non-invasif, est basé sur un système vidéo qui suit en temps réel le déplacement de deux marqueurs infrarouge réfléchissants fixés sur un bloc en plastique transparent, lui-même fixé à l'endroit désiré par une ceinture. Ce système est intégré dans des systèmes d'imagerie et de traitement commercialisés par Varian et General Electrics, et est utilisé pour effectuer le "Gating" du traitement de cancers du poumon (Korreman et al., 2008) et en imagerie TEP et CT.



FIGURE 2.4 – Le système infrarouge passif RPM commercialisé par Varian est composé d'un émetteur de lumière infrarouge (à gauche), d'un bloc plastique transparent sur lequel sont fixés deux réflecteurs infrarouges visible sur l'image infrarouge à droite.

2.2.2 Marqueurs ponctuels

Marqueurs infrarouges

Un marqueur infrarouge est une sphère fixée sur la surface à suivre et qui est dit soit "passif" dans le cas où il ne fait que réfléchir la lumière infrarouge fournie par une source, soit "actif" lorsque c'est lui-même qui émet le signal.



FIGURE 2.5 – En haut, illustration des deux types de marqueurs infrarouges, passifs et actifs. Le système Polaris (NDI Company) est présenté en bas avec des marqueurs réfléchissants associés au système.

Le système de détection est alors constitué d'un émetteur et de deux récepteurs de rayonnement infrarouge, la position du capteur étant connue par triangulation entre celui-ci et les deux récepteurs (Voir figure 2.5). Ce système est extrêmement rapide avec un large champ de détection : le système Polaris de la société NDI détecte à 60 Hz avec une précision de 0.25 mm jusqu'à 3m de distance. Wagman et al. (2003) utilise cette technologie pour effectuer le Gating respiratoire, alors que le Cyberknife en couple quelques dizaines sur une veste afin de caractériser la déformation surfacique d'un patient en temps réel (Section 2.1). On peut quand même citer certains défauts inhérents à cette technologie : la naturelle ponctuelle de l'information qui limite la description totale de la surface et le système de fixation qui implique un contact avec le patient.

Marqueurs optiques

Les marqueurs optiques sont généralement des motifs composés de carrés noirs et blancs afin d'optimiser leur détection par deux caméras classiques qui, une fois calibrées,

forment un système de vision stéréoscopique (Figure 2.6). Sur chaque vue, les coins composant le motif sont détectés qui déterminent des lignes reliant le repère de la caméra à ce point : on peut ainsi en déduire la position 3D grâce au calibrage spatial préalablement effectué. Cette méthode de détection est extrêmement précise et utilisée dans la détection d'instruments chirurgicaux en condition clinique (Nicolau et al., 2005). Le mouvement respiratoire externe est aussi caractérisable par la fixation de multiples marqueurs sur le patient, comme sur la figure 2.7 (Hostettler, 2008).

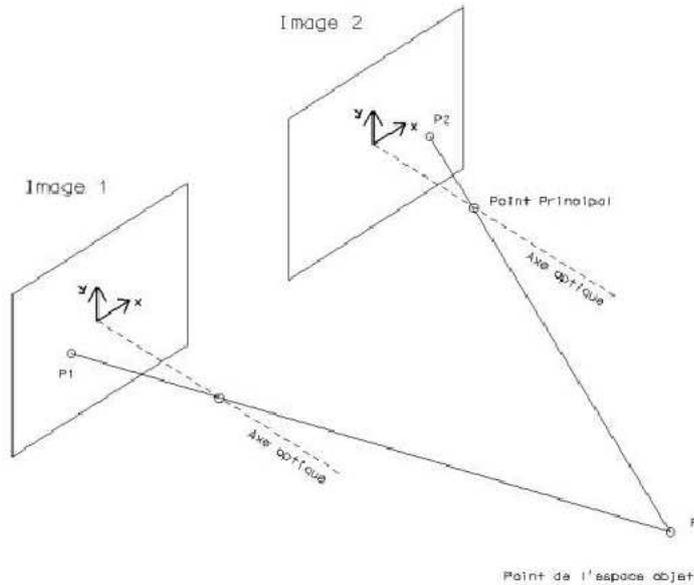


FIGURE 2.6 – Principe de la détection stéréoscopique.



FIGURE 2.7 – Exemple de patient sur lequel sont fixés 31 marqueurs optiques (Hostettler, 2008).

Marqueurs électromagnétiques

Ces marqueurs sont très prisés car peu onéreux, précis et ne demandant pas la conservation de la ligne de vue entre le détecteur et le marqueur, ce qui procure un avantage en terme de confort dans les situations où un praticien doit se déplacer dans l'espace de travail, comme en chirurgie par exemple. Le système détermine l'orientation et la position de la bobine composant le capteur en générant et mesurant les champs électromagnétiques orthogonaux (Raab et al. (1979), Kuipers (1980)). La plus grande limitation de ce système est clairement sa sensibilité aux éléments métalliques qui déforment le champ de détection. Cependant, des études cherchent à compenser cette distorsion afin de pouvoir suivre en temps réel la position et l'orientation d'instruments de chirurgie ou de caméras dans des applications de réalité augmentée (Nakamoto et al., 2008).

En conclusion, les marqueurs surfaciques, qu'ils soient infrarouges, optiques ou magnétiques, possèdent de nombreux intérêts :

- Leur suivi est rapide et précis.
- Leur coût est faible.
- Ils permettent de suivre des points de repère anatomique de manière dynamique.

Ils possèdent aussi certaines limitations :

- Ils requièrent un contact avec le patient.
- Il y a risque de déplacement pour les marqueurs sphériques.
- Les capteurs magnétiques sont extrêmement sensibles à toute modification du champ magnétique.
- Leur nombre est par définition limité. En cas de volonté de suivre la déformation complète du patient, un grand nombre de capteurs devient nécessaire, ce qui allonge considérablement l'ensemble de l'action à mener.

2.2.3 Les scanners tridimensionnels

Un scanner tridimensionnel analyse une scène en caractérisant les surfaces qui y sont présentes. Il fournit généralement un nuage de points 3D pouvant être constitué de milliers de points. Ce type de scanner sert généralement à caractériser des objets en 3D, et trouve des applications dans des domaines variés allant de la numérisation d'objets d'art à la conception d'orthèses et de prothèses en passant par la "motion capture" utilisée au cinéma afin d'humaniser des mouvements de personnages de synthèse.

Le terme de scanner 3D est ici un abus de langage dans le sens où un vrai scanner 3D est

capable de décrire l'ensemble d'une scène en 3D et ceci sans occlusion (la caméra Temps-de-Vol ainsi que le scanner laser peuvent constituer un tel système si différentes vues sont prises autour d'une scène, complétant ainsi la description de la scène en la rendant exhaustive, c'est à dire sans occlusion). Le marqueur magnétique est à ce titre le plus proche d'une vraie mesure 3D dans le sens où la mesure de sa position n'est pas soumise à une ligne de vue. Ceci dit, les trois types de scanners tridimensionnels présentés ici fournissent une information 3D par l'intermédiaire d'une géométrie spatiale permettant la conversion de l'information 2D mesurée en information 3D, sauf pour la caméra Temps-de-Vol où l'information mesurée est directement une information de profondeur. La géométrie du capteur ou de la configuration spatiale entre les différents éléments du système de mesure permet de faire correspondre chaque mesure avec un système de coordonnées sphérique où chaque point est identifié par des coordonnées (r, φ, ϑ) où r représente la distance entre le capteur et le point, φ et ϑ sont les angles formés par la ligne allant de l'origine au point et deux plans passant par l'origine, l'un horizontal, l'autre vertical.

Différentes technologies sont apparues ces dernières années qui permettent d'observer une surface à haute vitesse et dans sa totalité sans poser d'objet sur celle-ci et sans nocivité. Dans cette section, nous présenterons les technologies les plus prometteuses, discuterons leurs avantages et limitations actuelles, et présenterons les différentes applications existantes ou en développement dans le domaine médical. Un focus particulier sera fait sur la caméra Temps-de-Vol puisque c'est l'appareil qui a été principalement utilisé lors de ce travail de thèse.

La lumière structurée



FIGURE 2.8 – Le système commercial AlignRT de VisionRT utilise la lumière structurée pour reconstruire les surfaces en 3D (SiteVisionRT (2013))

Un scanner 3D par lumière structurée projette sur la surface observée un motif connu

2.2. Observation du mouvement surfacique

et observe ses déformations par un système de caméras calibrées préalablement. Le motif le plus répandu et un ensemble de franges observées par deux caméras positionnées latéralement par rapport au projecteur. Grâce à un calibrage préalable, la transformation reliant les deux caméras permet au final d'estimer une position 3D sur la surface par triangulation des points correspondant aux franges déformées. D'autres motifs peuvent être utilisés, comme des points formant un motif pseudo-aléatoire. C'est ce motif qui est utilisé par le système commercial AlignRT (de VisionRT) ainsi que par la caméra Kinect de Microsoft (Voir section 3.1). Cependant, l'algorithme recherchant les correspondances est complexe, ce qui ralentit la reconstruction des données 3D : Peng et al. (2010) relève une fréquence d'acquisition de l'AlignRT allant de 1 Hz à 1.7Hz pour une surface contenant l'ensemble du tronc d'un patient. Le scanner par projection de lumière structurée est une évolution du système stéréoscopique classique dans lequel deux caméras RGB observent une même scène et détectent des localisations identiques sur chaque image, la géométrie reliant les caméras permettant de déduire la position 3D correspondant à l'objet détecté dans les deux vues. La principale limitation de ce système est qu'il est très dépendant de la texture des objets à observer, tout en étant sensible à l'illumination de la scène. Au contraire, la lumière structurée est infrarouge, ce qui permet de diminuer la sensibilité à la texture surfacique tout en limitant les effets des illuminations parasites en lumière visible.

De part l'existence de deux systèmes commerciaux médicaux, de nombreuses publications testent ces appareils dans un large domaine d'application. Le système 3dMD a fait l'étude de recherche concernant l'anthropométrie (Weinberg et al., 2006), la fabrication de prothèses (Eggbeer et al., 2012), l'analyse de visages (Foroud et al., 2012), alors que la société VisionRT propose des systèmes pour le repositionnement ainsi que pour le "Gating" respiratoire lors d'acquisitions 4D-TDM et pour la radiothérapie (Spadea et al. (2010), Kauwelo et al. (2012)). Le signal est alors extrait d'une zone anatomique mesurant 20x20 mm², le système ne pouvant reconstruire les données pour une surface plus grande à très haute fréquence. Plus globalement, la faible vitesse d'acquisition limite l'utilisation de la lumière structurée pour des acquisitions dynamiques sur une large surface tel que le suivi de la respiration. Cependant, des recherches sont effectuées pour reconstruire les données plus rapidement : ainsi Price et al. (2012) arrive à atteindre une vitesse de reconstruction atteignant 23 Hz d'une surface entière, ce qui ouvre le champ à du suivi surfacique en temps réel. Des tests sur fantômes indiquent un accord avec l'imagerie CBCT avec une distance inférieure à 2 mm sur 90% de la surface.

Le scanner 3D par laser

Le scanner par laser par triangulation balaye le champ de vue avec un point ou une ligne laser. Une caméra infrarouge capture la scène, puis, par triangulation avec l'émetteur

laser, déduit la position 3D du point sur la surface. L'émetteur, la caméra et le point sur la surface forment un triangle dont on connaît la distance émetteur-caméra, l'angle entre les deux ainsi que l'angle formé par la caméra et le point de mesure : on connaît donc la position du point laser. Cette méthode est réputée très précise de part la nature très peu bruitée du signal laser. Cependant, elle requiert un calibrage extrêmement précis entre l'émetteur et la caméra, de sa qualité dépendant l'erreur de mesure. De plus, le laser prend un certain temps à balayer une région, et descend difficilement en dessous du Hertz pour observer une surface conséquente. C'est pourquoi la principale application concerne le repositionnement d'un patient par comparaison des surfaces (Petersson et al. (2010), Pallotta et al. (2012a)).

2.2.4 Les caméras Temps-de-Vol

La caméra dite Temps-de-Vol (ToF) est une technologie relativement jeune qui observe une scène en 3D à haute fréquence, en fournissant une information d'intensité et de profondeur pour chaque pixel. Les progrès récents de la micro-électronique et de l'optique permettent le développement de nouveaux modèles de caméras ToF ayant des caractéristiques les rendant intéressantes dans un large spectre d'application. Les modèles les plus récents sont plus compacts et légers et possèdent un meilleur ratio signal-sur-bruit, et des recherches sont en cours afin d'améliorer ces caractéristiques. Les deux derniers modèles, la SwissRanger 4000 (SR4000 de MesaImaging, Suisse) et la CamCube 3.0 de PMD[Vision] montrent de belles capacités et ont poussé de nombreuses recherches (Foix Salmerón et al., 2010).

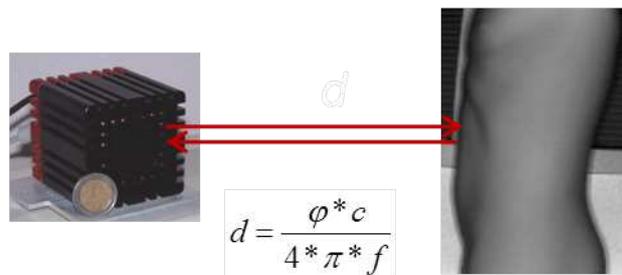


FIGURE 2.9 – La mesure par temps de vol consiste à émettre puis mesurer de la lumière infrarouge réfléctée par les surfaces présentes dans la scène.

Les caméras dites Temps-de-Vol acquièrent les surfaces dans le champ de vue grâce à l'émission de lumière infrarouge suivi de la mesure de la lumière réfléctée (Xu et al., 1998). Cette émission peut être soit pulsée, soit continue à intensité modulée. Pour les capteurs utilisant la lumière pulsée, la mesure du temps de voyage de l'impulsion donne une information directe sur la profondeur. Ceux utilisant de la lumière modulée mesurent la différence de phase entre la lumière émise et celle collectée. De la lumière infrarouge proche (780 à

2.2. Observation du mouvement surfacique

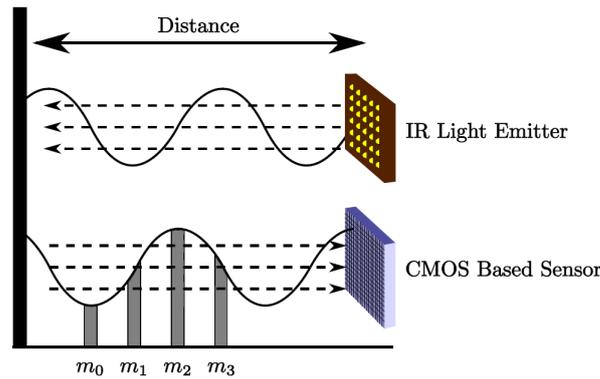


FIGURE 2.10 – Principe de mesure de la profondeur par mesure de l'offset

1400nm) est émise par des diodes LED puis collectée par des capteurs CCD/CMOS. Chaque pixel mesure la quantité de lumière reflétée par la scène quatre fois à intervalle régulier (voir m_0 , m_1 , m_2 et m_3 sur la figure 2.10), ce qui permet le calcul de la phase

$$\varphi = \arctan \frac{m_3 - m_1}{m_0 - m_2} \quad (2.1)$$

Connaissant les propriétés de la lumière émise, la différence des phases permet de calculer directement la distance parcourue par la lumière par la relation

$$d = \frac{\varphi * c}{4 * \pi * f} \quad (2.2)$$

avec d , la distance parcourue par la lumière, φ la différence de phase, c la vitesse de la lumière et f la fréquence de modulation.

La grande majorité des modèles commercialisés est basée sur l'émission de lumière modulée, même si des nouveaux prototypes basés sur la détection synchronisée d'un seul photon émis semblent pouvoir fournir de bonnes performances (Niclass et al. (2009), Stoppa et al. (2007)).

La géométrie de la caméra associe à chaque pixel une ligne de tir dans l'espace, ce qui permet de déduire une position 3D dans le repère cartésien de la caméra grâce à la distance d déduite par la mesure. La rapidité de la constitution de l'information 3D permet des acquisitions pouvant atteindre 50 Hz. Cependant, cette vitesse est limitée par la précision de mesure désirée : en effet, lors la fréquence s'élève, la quantité de lumière collectée diminue, le bruit de mesure en est augmenté. La résolution de ces caméras est à l'heure actuelle de maximum 200x200 pixels, même si on peut citer des modèles en développement ayant une résolution supérieure, comme la PMDTec 100k qui est censée avoir une résolution de 352x288 pixels. La précision de mesure varie en fonction des spécifications d'utilisation et de la gamme d'observation. L'utilisation de lumière modulée induit une profondeur au delà de laquelle la caméra ne peut différencier la phase mesurée de celle mesurée à faible distance : cette limite

de désambiguïté est directement proportionnelle à la modulation, elle est par exemple de 4,8 m pour le modèle SR4000 lorsque la fréquence de modulation est de 31 MHz. Un récapitulatif des caméras disponibles dans le commerce est disponible dans Piatti and Rinaudo (2012). La figure 2.11 donne un exemple des propriétés d'une caméra temps-de-vol.

Physical dimensions	$65 \times 65 \times 68 \text{ mm}^3$
Resolution	176×144
Maximal acquisition rate	54 Hz
Measurement range	0.3 to 5 m
Nominal accuracy at 1 m or less	1% of the distance between the camera and the object
Repeatability at 2 m or less	<5 mm
Angular resolution	0.23°

FIGURE 2.11 – Propriétés de la caméra ToF SR4000 de SwissRanger

La caméra ToF a clairement des avantages directs sur les autres types d'imageurs 3D, les systèmes à lumière structurées et les scanners lasers. Même si un système à lumière structurée ne requiert qu'une seule caméra qui observe un motif lumineux infrarouge, des phénomènes d'occlusion partielles peuvent apparaître. De plus, le système requiert une puissance conséquente ainsi qu'une lumière focalisée tout en nécessitant un environnement lumineux très contrôlé. Enfin le système est volumineux et peut observer à une distance qui dépend de la puissance lumineuse émise alors qu'un calibrage extrinsèque est nécessaire. La résolution est par contre bien meilleure par rapport à une caméra ToF, et les mesures sont réputées peu bruitées. On retrouve finalement les mêmes qualités et défauts pour les scanners lasers, à la grande différence que la vitesse d'acquisition est ici limitée à la vitesse de balayage de la scène par le système laser, là où un système à lumière structurée peut théoriquement atteindre 25Hz. Au final, la caméra ToF a une précision de mesure plus faible, mais une capacité théorique à analyser à haute vitesse un mouvement surfacique bien supérieur, puisqu'elle n'est limitée ni par la vitesse d'acquisition, ni par la reconstruction des données.

Etudes techniques

La littérature suit l'évolution technologique de la caméra ToF et cherche à comprendre et limiter les sources d'erreur de mesure de la profondeur (Karel et al. (2007), Guomundsson et al. (2007), Weyer et al. (2008)). Ces sources d'erreurs peuvent être divisées en deux sous-catégories, les erreurs systématiques et les non-systématiques et Foix et al. (2011) fournit une bonne description de ces erreurs et des moyens afin de les corriger.

Sources d'erreur systématique de mesure Une distorsion sur la mesure de profondeur apparaît car la lumière émise n'est pas exactement sinusoïdale, ajoutant alors un offset à la mesure de profondeur. Il a été aussi observé que différents temps d'intégration pour une même scène renvoyaient des valeurs différentes. Une autre source d'erreur systématique provient du fait que deux pixels peuvent avoir une mesure différente pour une même distance, ceci étant dû à une différence au niveau du matériau ou lors du processus de traitement. Ce phénomène est pourtant négligeable à l'échelle de deux pixels voisins. L'amplitude du signal mesuré peut aussi causer une erreur due à une sous ou une sur-exposition entraînant un effet de saturation. Une faible amplitude apparaît en bord de scène car la puissance lumineuse y est moins forte qu'au centre de la scène. Au contraire, un objet trop proche, trop réfléchissant ou un temps d'intégration trop long entraînent une saturation au centre de la scène. Enfin, la température interne joue aussi un rôle sur la mesure, les valeurs étant sur-estimées avant que la caméra n'atteigne une température stable.

Des études mettent en avant des améliorations significatives de la mesure lorsque l'ensemble de ces erreurs est prise en compte (Fuchs and Hirzinger (2008), Kahlmann et al. (2006)).

Sources d'erreur non-systématique de mesure Les erreurs non-systématiques sont au nombre de quatre. Tout d'abord, le ratio signal-sur-bruit est variable le long de la scène si celle-ci est illuminée de manière non-homogène, ce qui arrive si des profondeurs très variables sont observées. Les zones peu éclairées auront alors tendance à être plus sujettes au bruit. Le phénomène de réflexion multiple engendre aussi de larges erreurs de mesures de profondeurs (voir figure 2.12). Un effet de diffusion est aussi potentiellement existant de part de multiples réflexions entre la lentille et les capteurs : cet effet n'est présent qu'en cas de présence d'un objet dans le champ proche de la caméra. La dernière erreur non-systématique est l'apparition de flou dû au mouvement de l'objet observé (Figure 2.13).

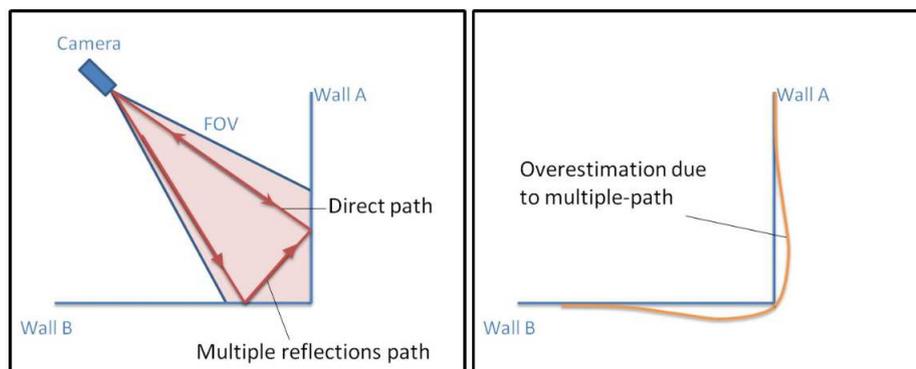


FIGURE 2.12 – Principe de mesure de la profondeur par mesure de l'offset

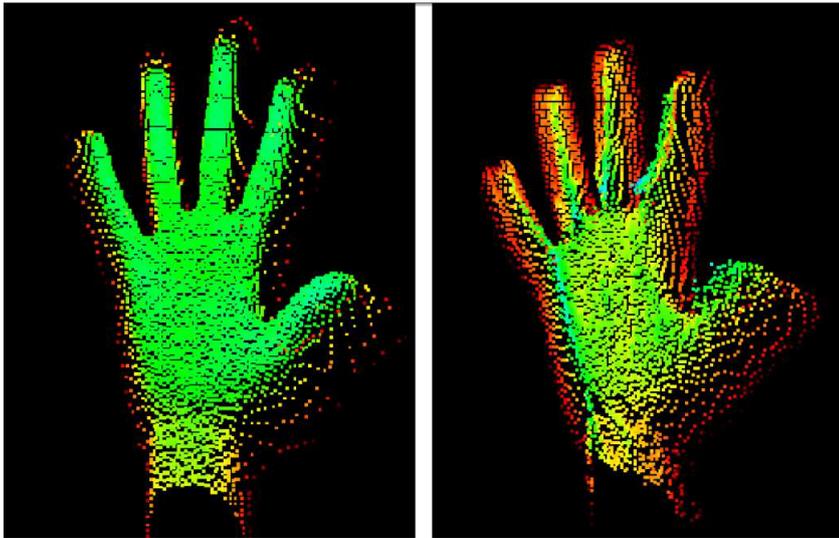


FIGURE 2.13 – Un flou apparait lorsque l'objet observé se déplace pendant l'acquisition (Foix et al., 2011)

Applications

Les caméras ToF ont rapidement trouvé un large champ d'applications dans les domaines de la surveillance et de l'assistance routière de part leur capacité à observer des surfaces courbées, leur grande profondeur d'observation et leur fréquence élevée. Falie and Buzuloiu (2008) a proposé la détection d'intrusions à l'aide de l'analyse de l'image de profondeur offerte par la caméra ToF, en se basant principalement sur l'histogramme des valeurs de profondeur. Ici, la longue portée et la rapidité de mesure sont avantageuses, et la faible précision due à une distance d'observation supérieure à 3.5 m est peu préjudiciable. Gallo et al. (2008) utilise lui la technologie ToF afin d'aider au guidage des voitures lors d'une marche arrière le long d'obstacles ou de rampes, mais aussi afin de détecter rapidement la présence d'humains. Le système est capable de détecter un objet de la taille d'un enfant à une distance de 5.5 m.

Le domaine de la vision par ordinateur utilise les caméras ToF pour leur capacité à observer de larges scènes tant statiques que dynamiques et proposent la fusion de l'information de profondeur "basse résolution" avec une image 2D "haute résolution" (Bartczak et al. (2007), Lindner et al. (2007)) avec la possibilité d'améliorer la résolution des images de profondeur par super-résolution (Yang et al. (2007), Crabb et al. (2011)). Zhu et al. (2011) considère la fusion entre une caméra ToF et un système stéréoscopique comme complémentaire : la mesure de temps de vol est optimale pour des surfaces blanches et lisses et dégradée pour des surfaces texturées, alors que dans le même temps, la vision stéréo permet l'extraction d'information en présence de texture et de contraste.

Dans le cas de scènes statiques, l'estimation de pose (quand l'imageur tourne autour de la

2.2. Observation du mouvement surfacique

scène à observer) ainsi que la cartographie d'un lieu sont des applications où la fusion 2D/3D est extrêmement complémentaire avec extraction de points d'intérêt sur les images 2D qui permettent le recalage des images 3D malgré le déplacement de la source (Prusak et al., 2008).

Dans le cas de scènes dynamiques, la segmentation en profondeur trouve des applications sur les plateaux TV (Langmann et al., 2008) et en réalité augmentée (Thomas (2006), Bartczak et al. (2008)). Il y a enfin le domaine de l'interaction homme-machine avec l'interaction logicielle sans souris (Oggier et al., 2005) et la navigation dans des images médicales (Soutschek et al., 2008), ainsi que la détection de mouvement, du nez (Boehme et al., 2008), de la marche (Jensen et al., 2008).

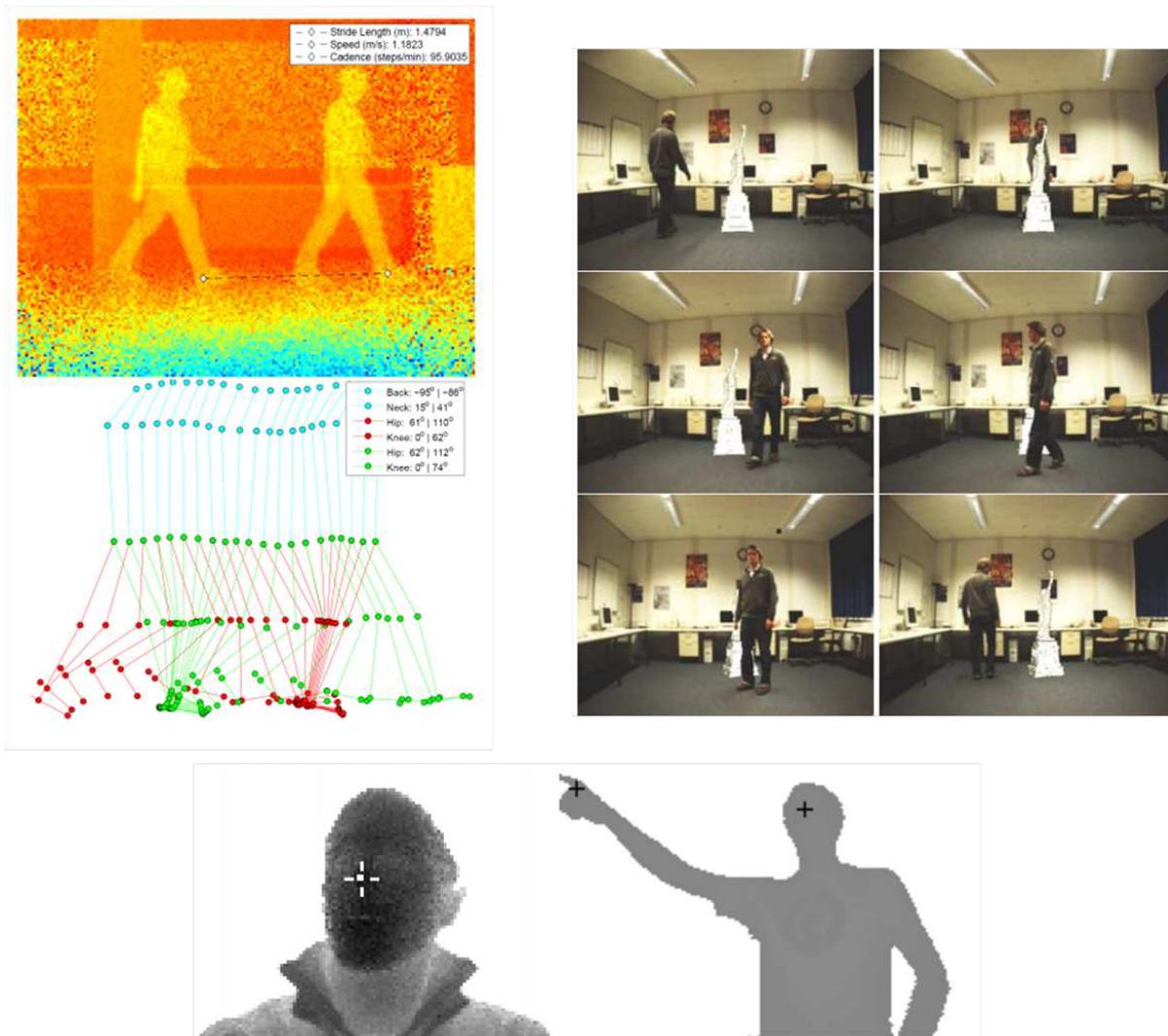


FIGURE 2.14 – Illustrations de différentes applications utilisant une caméra ToF en vision. En haut à gauche, le suivi dynamique de la marche par fitting d'un modèle articulé (Jensen et al., 2008). En haut à droite, ajout d'un objet virtuel dans une scène réelle (Bartczak et al., 2008). En bas, détection en temps réel de points d'intérêt tels que le nez et une main (Boehme et al., 2008).

2.3 Les modèles de respiration

Un modèle de respiration établit une relation entre le mouvement externe respiratoire à la surface du patient et la déformation interne associée. Partant de l'hypothèse qu'il existe bien une corrélation entre ces deux mouvements, de multiples méthodologies ont été présentées, qui peuvent être classifiées en deux catégories en fonction de la considération du mouvement interne, certains cherchant à estimer le mouvement de la tumeur qui est consi-

dérée comme ponctuelle, les autres cherchant le champ de déformation touchant la tumeur ainsi que les tissus environnants. Un tour d'horizon exhaustif de ce sujet est disponible dans la revue de McClelland et al. (2012). Nous présenterons ici l'évolution des modèles proposés en insistant sur la nature de la relation proposée entre mouvement externe et mouvement interne.

2.3.1 Les modèles basés sur un point

Les modèles qui considèrent le mouvement interne comme la trajectoire de la tumeur (vue comme un point) établissent des relations entre le signal externe et une trajectoire, celle de la tumeur le long du cycle respiratoire. L'hypothèse sous-jacente est que le comportement de la tumeur est identique à celui du signal respiratoire externe

Le modèle mathématique établi va du plus simple (fonction sinusoidale périodique avec un signal externe 1D (Chetty et al., 2004)) au plus complexe (Analyse en Composante Principale sur 18 paramètres issus de trois marqueurs externes (Khamene et al., 2004)) afin de décrire au mieux une trajectoire tumorale complexe de part la présence d'une hystérésis. La formulation sinusoidale ne fonctionnant qu'en présence de mouvement régulier et symétrique entr la phase d'inspiration et d'expiration, des formulations plus complexes basées sur une fonction cosinus modifiée sont proposées par Lujan et al. (1999) et validées cliniquement par George et al. (2005). Le cycle respiratoire est découpé en trois phases distinctes (expiration, fin d'expiration et inspiration) par Wu et al. (2004) afin de prendre en compte l'asymétrie de ce cycle tandis que des relation polynomiales sont aussi proposées (Sahih et al. (2005), Seppenwoolde et al. (2007)), tout comme l'ajout de paramètre supplémentaires qui améliorent la corrélation tels que le volume pulmonaire et sa variation (Low et al., 2005). Malgré les développements apportés au fil des études, la réduction de la tumeur à un point sans tenir compte des tissus environnants limite la capacité de cette approche à tenir compte de la nature élastique du mouvement de la tumeur et c'est pourquoi d'autres études cherchent à caractériser le champ de déformation englobant la tumeur et ses tissus proches.

2.3.2 Les modèles basés sur un champ de déformation

Ces modèles obtiennent le champ de déformation interne en recalant élastiquement les différents volumes TDM du patient à chaque instant du cycle respiratoire, fournis par l'imagerie scanner 4D-TDM. L'idée est alors de caractériser ce champ de déformation le long du cycle respiratoire afin de déduire d'un signal externe la déformation interne à l'aide de ces caractéristiques. Les volumes anatomiques du patient peuvent être interpolées d'un TDM à expiration complète et d'un autre à inspiration complète (Sarrut et al. (2006), Schreibmann et al. (2006)) ou à partir de l'imagerie Cone-beam CT (Zeng et al. (2005), Zeng et al. (2007)). Le champ de déformation peut être réduit par PCA pour ne le décrire qu'à l'aide de 4 paramètres (Sohn et al., 2005), on peut aussi le corrélérer avec des paramètres respiratoires (Zhang

et al., 2007). On a enfin cherché à créer un modèle patient-spécifique qui permet de relier la phase du cycle respiratoire avec le champ de déformation associé (McClelland et al., 2006). Cette méthode a comme faiblesse de ne pouvoir fonctionner que si le cycle respiratoire est régulier en période : dans le cas contraire, la phase mesurée du cycle respiratoire ne correspond pas forcément à celle des images qui ont servi à créer le modèle.

Fayad et al. (2010) propose lui une approche plus complète permettant la création d'un modèle global de respiration, modèle qui permettra pour un patient donné de générer la série TDM dynamique à partir de seulement deux examens TDM classique. Ce modèle de respiration global est basé sur la création d'une relation entre les mouvements internes décrits par les vecteurs de déformation qu'on peut l'extraire à partir des données TDM dynamique (4D-TDM), le mouvement externe étant acquis en utilisant un marqueur externe de type 'RPM' (voir partie 1.16). Cela signifie que chaque vecteur de déformation est défini par rapport à l'amplitude et la phase du signal respiratoire. L'hypothèse sous-jacente à l'utilisation de l'amplitude et de la phase du signal respiratoire externe est l'existence d'une forte corrélation entre le mouvement interne des organes et le signal respiratoire unidimensionnel externe.

Un modèle générique est tout d'abord créé avec 6 patients, leurs scanners 4D-TDM ainsi que leur signal externe. Après recalage élastique pour chaque série d'examen TDM par patient puis entre les patients, une Analyse en Composantes principales permet de réduire la quantité d'information et ainsi permettre de relier le champ de déformation interne au signal externe à l'aide d'une matrice de corrélation. Le modèle est appliqué à un nouveau patient par l'incorporation de deux volumes TDM appartenant à celui-ci (un à expiration complète, l'autre à inspiration complète) qui viennent contraindre le modèle générique aux spécificités anatomiques du patient. Le modèle proposé a différents avantages tels que sa robustesse aux variations de respiration et la réduction de dose imposée au patient vu que seuls deux scanners sont nécessaires à sa génération. Le modèle a été validé par comparaison des volumes intermédiaires générés entre inspiration et expiration totale avec la vérité terrain issue d'examens 4D-TDM.

La création d'un modèle patient-spécifique, ne nécessitant pas de modèle global préalable, est aussi proposé (Fayad et al., 2009). Il permet, pour une série de scanners incomplète mais accompagnée de l'information externe complète, de générer les volumes TDM manquants.

2.4 Le repositionnement par l'analyse de la surface du patient

Les scanners 3D capturent l'ensemble de la surface du patient sans ajout de temps au traitement ni de dose additionnelle, ils sont donc une solution intéressante pour effectuer son repositionnement à chaque fraction du traitement. L'opération consiste alors à acquérir cette

2.4. Le repositionnement par l'analyse de la surface du patient

surface au J0 du traitement, puis, lors de chaque fraction, d'analyser la nouvelle position au regard de celle de référence. Cette analyse passe généralement par un recalage entre les deux surfaces.

Les différents types de scanners 3D ont été l'objet d'études pour cette application, le plus connu étant déjà commercialisé sous le nom d'AlignRT, qui est un des produits de la société VisionRT. De nombreuses études ont été publiées à ce sujet et sont consultables sur le site de la compagnie SiteVisionRT (2013). La technologie est basée sur le principe de la lumière structurée (voir section 2.2.3) et permet l'acquisition de la totalité de la surface du patient à une vitesse de 7.5 Hz. Un logiciel de comparaison tridimensionnelle permet de visualiser l'acquisition avec celle de référence puis calcule le déplacement du lit nécessaire pour corriger la position du patient. Les essais sur fantôme valident une précision de 0.4 ± 0.3 mm, ceux sur patients atteignent 1 ± 0.5 mm (Schöffel et al., 2007). La prise en compte du mouvement respiratoire est effectuée en effectuant des acquisitions à un certain instant du cycle respiratoire (Bert et al., 2005). Peng et al. (2010) compare le système AlignRT avec un système d'imagerie CBCT et le système de localisation optique SonArray de Varian sur 5 patients pour des localisations de la tête et obtient une concordance avec le système optique atteignant 2.1 mm et 0.8° au 90^{ième} pourcentile. La localisation mammaire est aussi sujette à différentes études avec des résultats indiquant l'influence de la modification de la surface : en présence d'une chirurgie modifiant la forme du sein, le système ne permet pas de procéder efficacement au repositionnement (Zhao et al., 2011). Cependant les études montrent de très bons résultats pour cette localisation en absence de grande déformation (Schöffel et al. (2007), Miller et al. (2008), Zhao et al. (2011)). La localisation prostatique a aussi fait l'objet d'études : Krengli et al. (2009) compare le système à l'imagerie portale et détecte une erreur de repositionnement supérieure à 5 mm dans la direction x pour 14% des fractions, 2% dans l'axe y et 5% selon z. Bartoncini et al. (2012) détecte les déplacements supérieurs à 5 mm dans 5.3% des cas selon x, 47.4% des cas selon y et 42.1% des cas selon z.

Le système Sentinel (C-Rad AB, Uppsala, Suède) est un scanner laser (Section 2.2.3) qui est aussi commercialisé pour le repositionnement (Brahme et al., 2008). Les tests sur fantôme atteignent une précision de 0.5 mm pour des déplacements inférieurs à 20 mm alors que le repositionnement sur patient avec des tumeurs de la prostate est entaché d'une erreur de 5-6 mm. Une autre étude sur fantôme donne des résultats équivalents tout en montrant que prendre une surface Scanner comme référence diminuait la qualité du repositionnement (Pallotta et al., 2012b). Des études cliniques ont aussi été menées : Moser et al. (2012) étudie 120 patients pour 6 localisations différentes en comparant avec l'imagerie CBCT. Les résultats indiquent l'importance d'utiliser comme image de référence une surface acquise par le scanner laser. Si l'on se concentre sur les localisations thoraciques et pelviennes, l'erreur relative moyenne d'estimation du repositionnement atteint -0.1 mm, 1.3 mm et -1.2 mm (res-

pectivement 1.6 mm, 0.4 mm, -0.9 mm) pour les directions latérales, cranio-caudales et antéro-postérieures pour le thorax (respectivement le pelvis). Les valeurs absolues d'erreurs ne sont pas connues. Pallotta et al. (2013) étudie le repositionnement de 33 patients atteints de tumeurs pulmonaires et au pelvis, avec comme point de comparaison le repositionnement à l'aide de l'imagerie CBCT. La surface de référence est extraite du Scanner TDM de planification et le mouvement respiratoire est pris en compte en moyennant cinq acquisitions successives. Les résultats indiquent une différence absolue de recalage inférieure à 6,4 mm et 3.8° pour 90% des estimations. Un deuxième groupe de patients à qui on demande de retirer leurs vêtements sur une région plus large que d'habitude est aussi étudié et les résultats montrent qu'une région anatomique plus large comme référence améliore le repositionnement. Le système Galaxy de LAP est lui aussi l'objet de quelques études dont celle de Moser et al. (2010) qui montre des erreurs sur fantôme avec une moins bonne précision dans la direction latérale. Un essai sur un patient montre l'erreur minimale dans la direction cranio-caudale (1.7 ± 1.5 mm).

Le repositionnement avec une caméra ToF est testé sur fantôme et volontaires avec une erreur moyenne de 2.9 mm et de 3.4 mm respectivement, un filtre gaussien passe-bas étant appliqué aux données brutes (Schaller et al., 2009). Placht et al. (2010) montre l'amélioration apportée par l'application d'un filtre bilatéral directement sur la carte de distances. L'utilisation d'un filtre de Kalman et d'un calibrage en fonction de la distance est aussi proposée ainsi que l'initialisation de l'ICP par extraction de points d'intérêt sur les surfaces (Placht et al., 2012). L'influence de la déformation de la surface y est aussi mis en évidence.

Toutes ces études se concentrent sur la vérification du setup du patient, après alignement par le système de lasers. Bauer et al. (2011) propose lui d'étudier si l'on peut se passer de ce système en effectuant un premier recalage grossier. En effet, la majorité des systèmes sont conçus pour recaler des surfaces à de faibles distances (un maximum de 50 mm), et ne fonctionnent généralement plus lorsque la translation est de plusieurs dizaines de centimètres. L'utilisation de descripteurs de surfaces permet un tel repositionnement initial.

2.5 Synthèse

La prédiction de la déformation des organes internes due au mouvement respiratoire, et à fortiori le déplacement tumoral induit, est un axe extrêmement dynamique de recherche de part le fort potentiel clinique dans l'amélioration du traitement par radiothérapie externe : l'asservissement en temps réel de l'accélérateur linéaire au mouvement de la tumeur autorise le radiophysicien à ne plus ajouter de marges spécifiques à la respiration lors de la planification. Des systèmes commerciaux voient le jour et font le choix de technologies extrêmement différentes. Elles ont cependant le point commun de nécessiter de nombreuses images afin de suivre un marqueur ou de confirmer un modèle de corrélation temporaire avec des mar-

queurs externes. Le développement de modèles respiratoires globaux ou spécifiques à un patient permet d'espérer l'établissement d'un modèle de corrélation stable entre mouvement interne et mouvement externe et plus encore que les modèles considérant la tumeur comme un point localisé, ceux considérant la déformation sur la tumeur et les tissus environnants semblent plus enclins à fournir une solution robuste. Les derniers modèles proposés ont la particularité d'être robuste aux variations de schémas respiratoires, que se soit en phase ou en amplitude, tout en ne nécessitant que peu d'examen irradiants, effectués en début de traitement pour déterminer le modèle. Il est montré de plus que le modèle gagne en robustesse et en précision si l'on considère le mouvement externe sur de multiples régions localisées le long du tronc plutôt que sur une seule comme dans le cas d'une ceinture de pression ou un système RPM (Fayad et al., 2011). Cette constatation corrobore l'intuition générale qui est que la déformation interne est un phénomène complexe nécessitant une estimation elle aussi complexe du mouvement externe pour permettre une relation stable et robuste.

Plus que les marqueurs fixés sur la peau du patient, l'avenir est clairement aux scanners 3D par leur capacité à observer à très haute vitesse une très grande région anatomique sans interaction avec le patient. Ces scanners font généralement utilisation de lumière infrarouge, que ce soit sous forme laser, de lumière structurée ou pour mesurer un temps de vol : leur utilisation est déjà commercialisée dans des applications statiques tel que le repositionnement, mais le suivi en temps réel de la déformation d'une surface comme le tronc d'un patient avec un scanner 3D est encore non achevé pour deux raisons qui sont liées. Tout d'abord, ces scanners fournissent des nuages de points qui ne sont pas appareillés entre eux, il est donc nécessaire d'extraire la déformation (soit le lien entre deux nuages de points) par des méthodes algorithmiques complexes. Le corrolaire est le temps de calcul associé à ces méthodes qui rendent pour l'instant inaccessible une application en temps réel.

Le repositionnement du patient entre chaque fraction est aussi un domaine dans lequel les scanners 3D ont un avenir prometteur. Le besoin de limiter les images de contrôle irradiantes ainsi que la future prise en compte du mouvement du patient pendant la fraction rend indispensable un mode d'observation en temps réel et précis. La gamme de technologies déjà proposées est large, mais la grande majorité des études concernent des localisations non-touchées par la respiration. Ceci est dû au fait que les systèmes à lumière structurée et lasers sont soit trop lents à acquérir la scène, soit limités par le temps de reconstitution des données. Les solutions proposées sont alors généralement le moyennage des acquisitions sur différents instants temporels ou la prise en compte des surfaces à un instant du signal respiratoire. Même si la solution du moyennage des acquisitions peut apporter satisfaction pour le repositionnement inter-fraction, ce n'est clairement pas une solution envisageable à long terme pour la gestion du mouvement intra-fraction.

Apports méthodologiques et études menées

Chapitre 3

Caméras ToF SR4000, CamCube 3.0 et Kinect : évaluation pour le suivi de mouvement surfacique

Sommaire

3.1 Principe de fonctionnement et caractéristiques	76
3.2 Expérimentations sur fantôme statique	78
3.3 Expérimentations sur fantôme dynamique	80
3.4 Estimation du mouvement respiratoire	82
3.5 Synthèse et conclusion	85

Ce chapitre vise à comparer les deux modèles de caméras ToF qui sont les plus utilisés pour les études ayant trait au domaine médical : le premier modèle est celui utilisé comme fil conducteur de ce travail de thèse, le modèle SR4000 du fabricant suisse MESA Imaging lancé en 2008. Le second est le modèle Camcube 3.0 de la société allemande PMD[Vision] qui est sorti en 2009. L'objectif est de déterminer quels critères d'utilisation privilégient l'un ou l'autre des modèles lorsque l'on veut suivre la déformation du thorax et de l'abdomen d'un patient. Parallèlement à ce travail, la Kinect de Microsoft est aussi testée. Les deux dernières années ont vu le nombre d'études faisant usage de ce nouvel appareil très fortement augmenter, et si ses principales propriétés de mesures commencent à être connues, il manque encore des études lui assignant une place parmi les scanners 3D.

Les trois caméras ont été comparées pour l'estimation d'un mouvement nul d'un fantôme anthropomorphique, ainsi que pour un déplacement linéaire. Une fois mise en évidence l'influence cruciale de la distance d'observation sur la qualité de mesure de la Kinect, nous avons évalué la SR4000 et la Kinect lors de l'extraction d'un signal respiratoire. Ce chapitre se conclut par une synthèse qui replace ces résultats dans l'état de l'art et cherche à proposer



FIGURE 3.1 – Le motif de points infrarouge projeté permet à la Kinect d’estimer une carte de profondeur associant à chaque pixel un point 3D.

les conditions optimales d’utilisation des différentes caméras.

3.1 Principe de fonctionnement et caractéristiques

Le principe de fonctionnement de la Kinect diffère radicalement de celui d’une caméra ToF. Pour rappel, une caméra Temps-de-Vol observe une scène en 3D par la mesure de la lumière infrarouge qu’elle a elle-même émise. La composante ayant rebondi sur les surfaces présentes dans le champ de vue est mesurée par les capteurs CCD de la caméra. La distance les séparant du point de réflexion sur la surface est déduite de la différence de phase entre lumière émise et lumière mesurée. La géométrie associée à la caméra assignant à chaque pixel une ligne de mesure dans l’espace, la distance permet de déduire une position 3D pour chaque pixel. Les propriétés d’un modèle découlent du nombre de capteurs ainsi que de la fréquence de la lumière émise qui influe sur la profondeur maximale d’observation. La nature infrarouge de la lumière rend les mesures sensibles à la lumière naturelle ainsi qu’aux surfaces noires ou réfléchives. Pour plus de précision, les détails sont disponibles section 2.2.4. La Kinect mesure une carte de profondeur d’une scène sur le principe de la lumière structurée en projetant un motif de points infrarouge sur la scène (voir figure 3.1). Une caméra infrarouge forme un système stéréo avec le projecteur et voit les points sur la scène plus ou moins grand en fonction de leur profondeur par rapport à un plan où le calibrage de référence est effectué. Pour chaque pixel de l’image RGB haute résolution (640x480) est alors associé un ou plusieurs points du motif, une fenêtre glissante permet la détermination d’un point 3D qui lui sera relié. Ce système est conçu pour la détection de mouvements dans une optique ludique, c’est pourquoi il produit un nuage de points très dense (640x480 points 3D). La

3.1. Principe de fonctionnement et caractéristiques

	Caméras ToF		
	SR4000	CamCube 3.0	Kinect
Résolution (pixels)	176x144	200x200	640x480
Fréquence d'acquisition maximale (Hz)	50	40	30
Dimensions (mm ³)	65x65x68	180x60x60	280x63x40
Gamme d'observation (m)	0.8-5	0.3-7	0.5-5
Réglage du temps d'exposition	Automatique / Manuel	Manuel	Automatique

FIGURE 3.2 – Tableau répertoriant les principales caractéristiques des trois caméras comparées dans ce chapitre.

visualisation du nuage de points fourni par la Kinect donne l'impression que les données ne sont quasiment pas bruitées. De plus, un effet escalier apparaît lorsque la surface est courbée. Le traitement appliqué aux données par le système embarqué dans la caméra n'est pas connu.

La table 3.2 indique les spécifications principales des trois caméras. Chacune d'entre elle a un ou plusieurs points forts : la Camcube 3.0 a un nombre de pixels significativement plus grand que la SR4000 pour une profondeur d'observation maximale plus grande. Cependant la SR4000 a une vitesse d'acquisition maximale plus élevée, ainsi qu'une taille et une compacité très adaptées aux environnements surchargés dans un milieu médical. L'interface de cette caméra offre de plus une fonction de gestion du temps d'exposition afin de l'optimiser tout en évitant les phénomènes de saturation. La Kinect, elle, se distingue surtout par son nombre de points impressionnants, plus de 300 000, ce qui, associé à un angle d'ouverture assez similaire à ceux des caméras ToF, à une résolution latérale bien meilleure. Les performances de ces caméras ont déjà été l'objet d'études : pour la Kinect, Khoshelham and Elberink (2012) montre que les erreurs de mesure de la Kinect augmentent de manière quadratique avec la distance, de même que sa résolution en profondeur. Concernant les caméras ToF, Falie and Buzoloiu (2007) et Lindner and Kolb (2006) présentent les paramètres dégradant la précision de mesure alors que Chiabrando et al. (2010) montrent que l'angle d'observation d'une surface n'influe pas sur la mesure d'une caméra SR4000, pour un angle allant jusqu'à 50°, ce résultat étant confirmé par les expérimentations qui seront présentées section 4.2.2. Piatti and Rinaudo (2012) compare les modèles SR4000 et Camcube 3.0 et met en valeur un léger avantage de la SR4000 concernant la précision de l'estimation de distance. Il souligne en plus l'effet crucial de l'échauffement de cette caméra sur la qualité de mesure.

Des détails sur les causes de bruit et d'erreurs systématiques sont disponibles section 2.2.4. Pour les expérimentations suivantes, les paramètres d'acquisition de la caméra ToF SR4000 seront les mêmes que ceux détaillés ultérieurement section 4.2.1. Pour le modèle CamCube 3.0, le temps d'exposition est réglé préalablement par un algorithme sélectionnant la durée la plus longue sans pixels saturés. Il est à noter que cette méthode n'est valable qu'en conditions d'observation stables. Enfin, la Kinect n'est pas réglable, et sa fréquence d'acquisition est de 30 Hz.

3.2 Expérimentations sur fantôme statique

La distance d'observation ainsi que l'angle entre la caméra et la surface observée sont deux paramètres qui sont souvent contraints par l'environnement, c'est pourquoi connaître leur influence sur la mesure est capital. Dans cette section, nous présentons le résultat de l'estimation du mouvement nul du fantôme. Pour ceci, nous l'observons 100 fois et considérons pour chaque pixel d'une région d'intérêt (ROI) centrée la distance entre deux estimations successives comme l'erreur de l'estimation du mouvement nul. Ce processus est répété pour une gamme d'observation allant de 0.8 m à 2 m, mais aussi à 1 m pour des angles d'observation allant de 0° à 60°, en utilisant cette fois un plan blanc.

La figure 3.3 indique l'erreur d'estimation du mouvement nul pour les trois caméras en fonction de la distance et montre pour les deux caméras ToF une faible influence de la distance avec une erreur pour la SR4000 allant de 2.2 mm à 0.8 m à 2.5 mm à 2 m avec un maximum de 3.1 mm à 1.2 m. De même, pour la Camcube 3.0, l'erreur est de 2.9 mm à 0.8 m comme à 2 m avec un maximum de 3.5 mm à 1.2 m. La SR4000 surclasse la Camcube 3.0 pour toutes les distances avec une erreur en moyenne 20% plus faible sur toute la gamme de distance. La Kinect présente elle des résultats étonnants avec une erreur quasi nulle à 0.8 m (0.1 mm) qui augmente graduellement pour atteindre 1.8 mm à 2 m.

La figure 3.4 présente cette même erreur associée aux points de mesures appartenant à un plan blanc observé à 1 m. L'observation de ce plan permet de connaître l'angle entre celui-ci et la caméra, grâce à l'algorithme de RANSAC qui sélectionne automatiquement les points 3D de la scène appartenant au plan, puis une minimisation permet la détermination des paramètres du plan (L'algorithme RANSAC est présenté section 6.1.2).

Les résultats indiquent :

- Une incidence quasi-nulle de l'angularité sur la variabilité de la mesure de la Kinect sur la gamme 0 à 50° (de 0.4 ± 0.4 mm à 0.5 ± 0.6 mm). Une légère augmentation est perceptible à 60°, à 0.7 ± 0.7 mm.
- Pour la SR4000, l'effet de l'angularité se fait ressentir à 50°, avec une stabilité pour des angles inférieurs (2.3 ± 0.5 mm à 0° à 2.3 ± 0.5 mm à 40°) puis une augmentation

3.2. Expérimentations sur fantôme statique

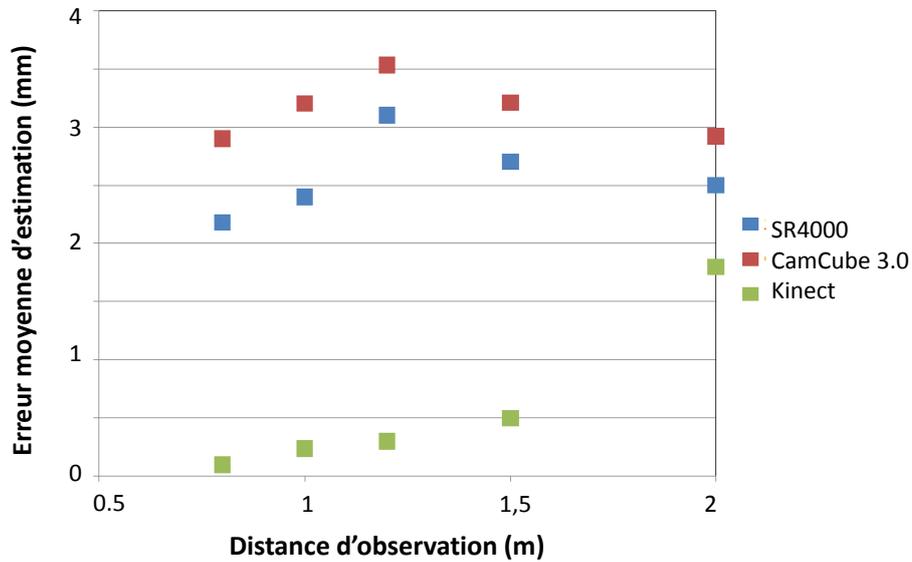


FIGURE 3.3 – Erreur moyenne d'estimation du mouvement surfacique nul du fantôme pour les trois caméras en fonction de la distance.

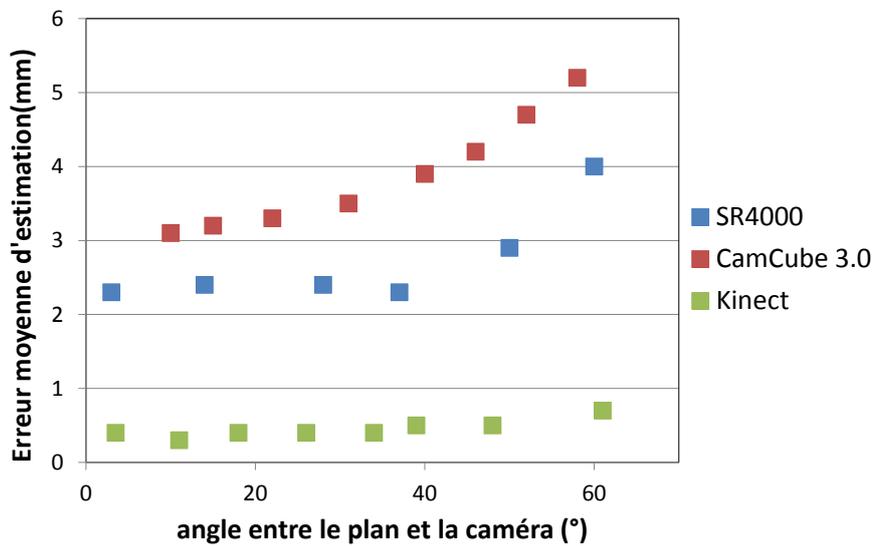


FIGURE 3.4 – Influence de l'angle d'observation sur la mesure de mouvement pour les trois caméras.

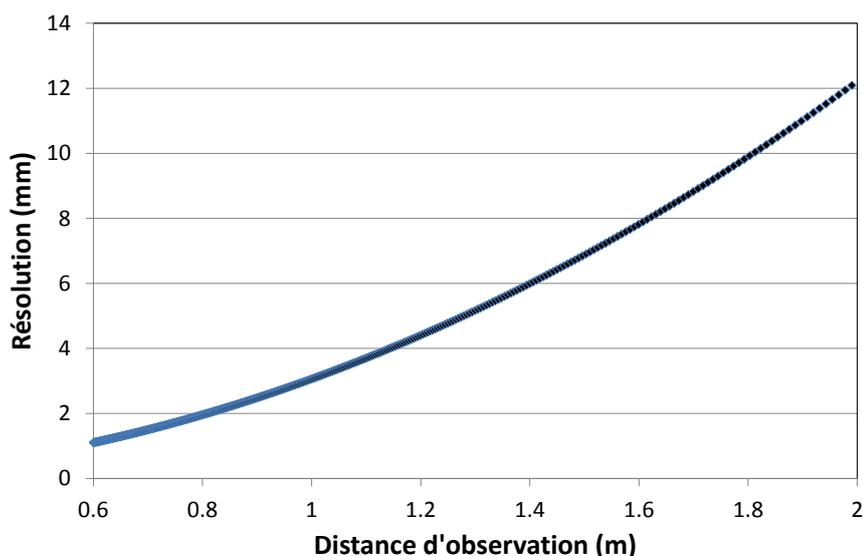


FIGURE 3.5 – Résolution en profondeur de la Kinect sur une gamme allant de 0.6 à 2 mètres.

sensible atteignant 4 ± 1 mm à 60° .

- Concernant la PMD, l'augmentation de la variabilité est tout d'abord lente jusqu'à 30° (de 3.1 ± 0.6 à 0° à 3.5 ± 0.7 mm à 30°) puis augmente sensiblement au-delà, atteignant 5.2 ± 1.1 mm à 60° .

Le niveau de bruit de la mesure de la Kinect est particulièrement faible à courte distance, mais augmente de manière abrupte avec elle. Ce schéma d'évolution confirme que la résolution en profondeur évolue quadratiquement avec la distance (Khoshelham and Elberink, 2012). La mesure de profondeur est codée sur 11 bits dont 1 contenant un flag décrivant la qualité de mesure, elle est donc discrétisée en 1024 gradations. La fonction de conversion (proposée empiriquement par des développeurs) fait ressortir une gamme d'observation allant de 0.3 à 5.3 m. L'évolution n'étant pas linéaire, la résolution spatiale varie avec une augmentation sensible quand la distance augmente, comme illustré figure 3.5 avec une la plage d'observation [0.6 m - 2 m]. La résolution est ainsi de 2 mm à 0.8 m, 3 mm à 1 m, 4.5 mm à 1.2 m, 7 mm à 1.5 m et 12 mm à 2 m.

Les conséquences de cet effet sur l'estimation d'un mouvement non-nul est observé dans la prochaine section, en appliquant un déplacement ou un mouvement connu à l'objet observé.

3.3 Expérimentations sur fantôme dynamique

Dans un premier temps, nous nous sommes intéressés à la détection d'un faible mouvement par ces trois caméras, en déplaçant le fantôme de 2 mm en profondeur à 1.2 m à l'aide d'un actionneur linéaire. Cette distance est choisie car la résolution en profondeur de

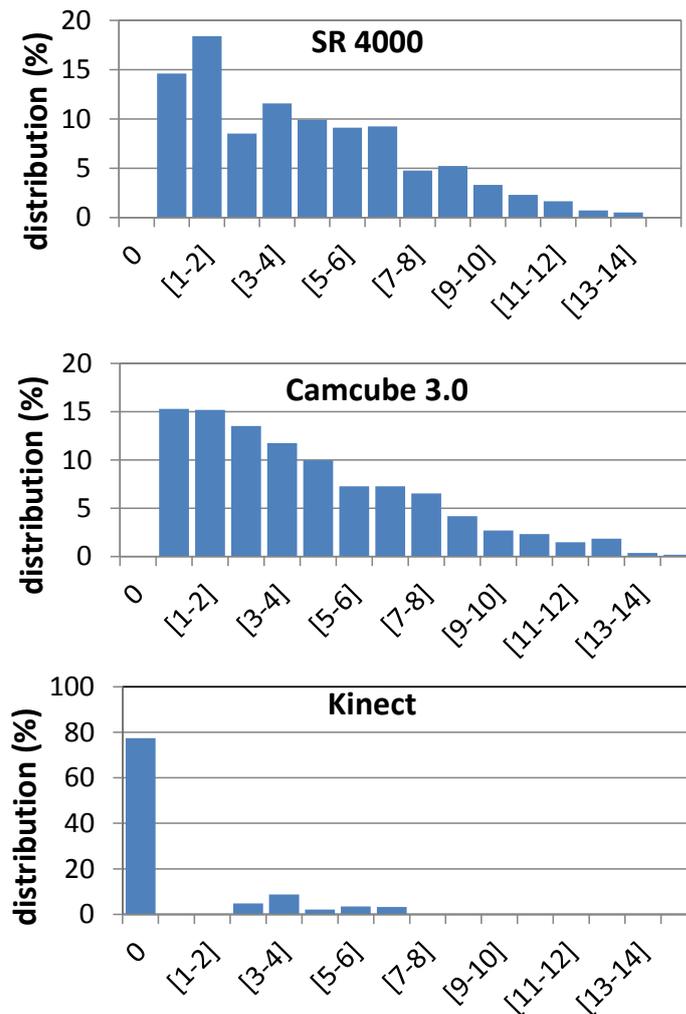


FIGURE 3.6 – Distribution du déplacement estimé pour chaque point d’une ROI pour les trois caméras. Le déplacement réel est de 2 mm

la Kinect y est alors de 4.5 mm, ce qui signifie qu’un point ne pourra prendre une valeur en profondeur que tout les 4.5 mm. Un mouvement de la surface de 2 mm devrait normalement indiquer des résultats confirmant cette hypothèse. Pour chaque point 3D appartenant à une ROI centrée, nous avons mesuré la distance parcourue entre les deux positions par chaque point. La figure 3.6 présente les histogrammes de distribution des valeurs absolues des distances mesurées pour les 3 appareils. Concernant les deux caméras ToF, on assiste à un étalement des estimations extrêmement similaire, avec une erreur moyenne d’estimation du mouvement de 2.7 ± 2.4 mm pour la SR4000 et 2.9 ± 2.5 mm pour la PMD. Par contre, l’histogramme de distribution de l’estimation de mouvement pour la Kinect montre le comportement attendu, c’est-à-dire une immense majorité (76%) des points qui ne sont pas détectés comme ayant bougé. La structure même de l’histogramme, avec 0 points ayant un mouvement entre 0 et 2 mm indique l’influence de la résolution en profondeur, qui interdit

Distance (m)	PMD	SR4000	Kinect
0,8	1,3±1	1,1±1	0,7±0,5
1	1,6±1,4	1,3±1	1,2±0,8
1,2	1,5±1,2	1,8±1,5	1,8±1,2
1,5	1,8±1,3	1,7±1,3	2,8±1,7

FIGURE 3.7 – Erreur moyenne d’estimation du mouvement linéaire du fantôme en suivant le point 3D correspondant au pixel au centre du champ de vue

une telle mesure.

Afin de savoir quelle caméra est la plus adaptée à mesurer un mouvement surfacique donné, un plan est ensuite déplacé linéairement et perpendiculairement à la caméra par l’actionneur linéaire à une vitesse d’un centimètre par seconde. Le mouvement décrit par le pixel au centre du champ de vue est mesuré pour différentes distances d’observations (0.8, 1, 1.2 et 1.5 m). L’erreur d’estimation du mouvement est considérée comme la distance entre la mesure et la droite interpolée depuis l’échantillon complet de points.

Les résultats présentés figure 3.7 indiquent que la Kinect est clairement plus précise pour suivre le mouvement à une distance inférieure à 1 m avec une distance moyenne à la droite d’interpolation de 0.7 ± 0.5 mm à 0.8 m. Au fur et à mesure que la distance augmente, le rapport s’inverse avec une distance moyenne à la droite d’interpolation augmentant puis devenant supérieure à celle des deux modèles de caméra ToF. A 1.5 m, l’effet de quantification devient extrêmement significatif, avec un crénelage maximal (voir figure 3.8).

Ces expériences montrent tout d’abord que le modèle de caméra ToF SR4000 est effectivement plus performant que le modèle CamCube 3.0, que ce soit pour l’estimation d’un mouvement nul ou d’un déplacement de 2 mm. Cependant, le suivi du plan se déplaçant montre que la CamCube 3.0 fait jeu égal voire est plus performante que la SR4000 au delà d’un mètre. Le cas de la Kinect montre tout d’abord une mesure quasiment non-bruitée à faible distance. On voit ensuite que la résolution en profondeur influence profondément l’estimation du mouvement avec des erreurs dues à l’échantillonnage des positions possibles d’un point 3D dans l’espace.

3.4 Estimation du mouvement respiratoire

L’extraction du mouvement respiratoire selon la direction antéro-postérieure est réalisée avec la caméra ToF SR4000 et la Kinect aux distances de 0.9 et 1.5 m, et ceci pour deux

3.4. Estimation du mouvement respiratoire

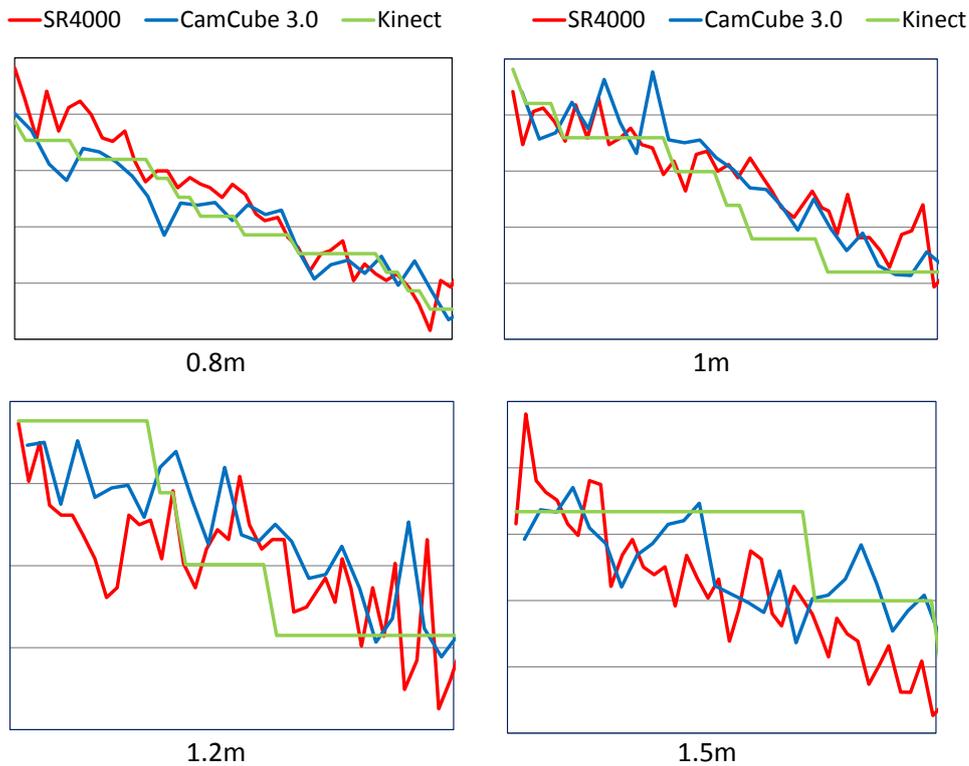


FIGURE 3.8 – Illustration du suivi du mouvement linéaire du fantôme en suivant le point 3D correspondant au pixel au centre du champ de vue. La vitesse est d'un centimètre par seconde.

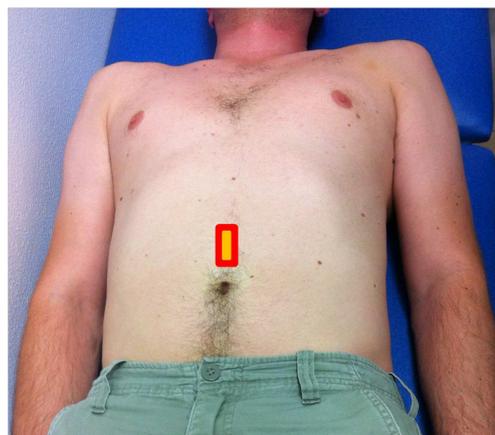


FIGURE 3.9 – Illustration des deux ROIs prises en compte pour l'extraction du mouvement respiratoires. Leur taille est respectivement de $1 \times 5 \text{ cm}^2$ et $5 \times 8 \text{ cm}^2$.

tailles de régions d'intérêt, visibles figure 3.9. Ces distances sont choisies afin d'observer l'effet de la résolution en profondeur sur l'estimation d'un mouvement surfacique dépassant rarement 2 cm. Le mouvement respiratoire est considéré comme le déplacement selon l'axe z (donc la direction antéro-postérieure) du centre de gravité de l'ensemble des points présents dans la ROI. Pour réaliser ces mesures, les deux caméras sont montées sur un même bras articulé, et ceci en les rapprochant au maximum afin d'obtenir un champ de vue, une distance et un angle d'observation similaire. La figure 3.10 indique le nombre de points approxima-

	SR4000		Kinect	
	1x5 cm ²	5x8 cm ²	1x5 cm ²	5x8 cm ²
0.9 m	85	300	150	800
1.5 m	20	75	60	350

FIGURE 3.10 – Récapitulatif du nombre approximatif de points par ROI et par type de caméra.

tifs par ROI et par type de caméra pour chaque distance d'observation. Un volontaire est observé pendant quelques cycles respiratoires d'amplitude variable. La figure 3.11 illustre les traces respiratoires évaluées simultanément par les deux caméras. La taille de la ROI joue fortement sur la qualité du signal avec un niveau de bruit sensiblement diminué sur les traces correspondant à la plus grande zone. De plus, les observations à 0.9 m montrent une très bonne conformité entre la Kinect et la SR4000, même si une sous-estimation de l'amplitude du signal peut être observée quand on considère la grande ROI. Ceci peut être soit dû à une erreur lors de la sélection des régions sur les deux acquisitions, soit à la résolution en profondeur (qui est de 2.5 mm à 0.9 m). Cependant, pour les observations à 1.5 m (où la résolution en profondeur est de 7 mm pour la Kinect), on distingue tout d'abord la sous-estimation du cycle respiratoire de large amplitude (entre 10 et 18 secondes). On voit ensuite que les deux premiers cycles respiratoire d'une amplitude ne dépassant pas 4 mm selon la caméra ToF ne sont pas détectés par la Kinect, ce qui est cohérent étant donné le rapport amplitude du signal/résolution en profondeur.

La figure 3.12 illustre cet effet pour l'ensemble de la surface. Un code couleur est appliqué, indiquant pour chaque point 3D son mouvement en profondeur avec en bleu un mouvement nul et en rouge un mouvement de 5 mm. Les mesures effectuées à 0.9 m sont indiquées dans la partie haute de la figure. Elles montrent une corrélation entre les deux schémas de couleur même si il semble que la Kinect sous-estime les mouvements, ceci étant particulièrement visible dans la partie centrale où le mouvement oscille entre 2 et 3 mm (couleur jaune), mais n'apparaît pas sur la carte de couleur issue des données Kinect. Il est cependant important de souligner que ce résultat est soumis à la sélection des acquisitions sélectionnées, et l'expérience montre que la carte de couleur varie d'une acquisition à une autre. Cet effet pourrait s'expliquer par cette quantification qui implique le basculement d'un grand nombre de points 3D d'une valeur de profondeur à la suivante, et ceci d'une acquisition à une autre.

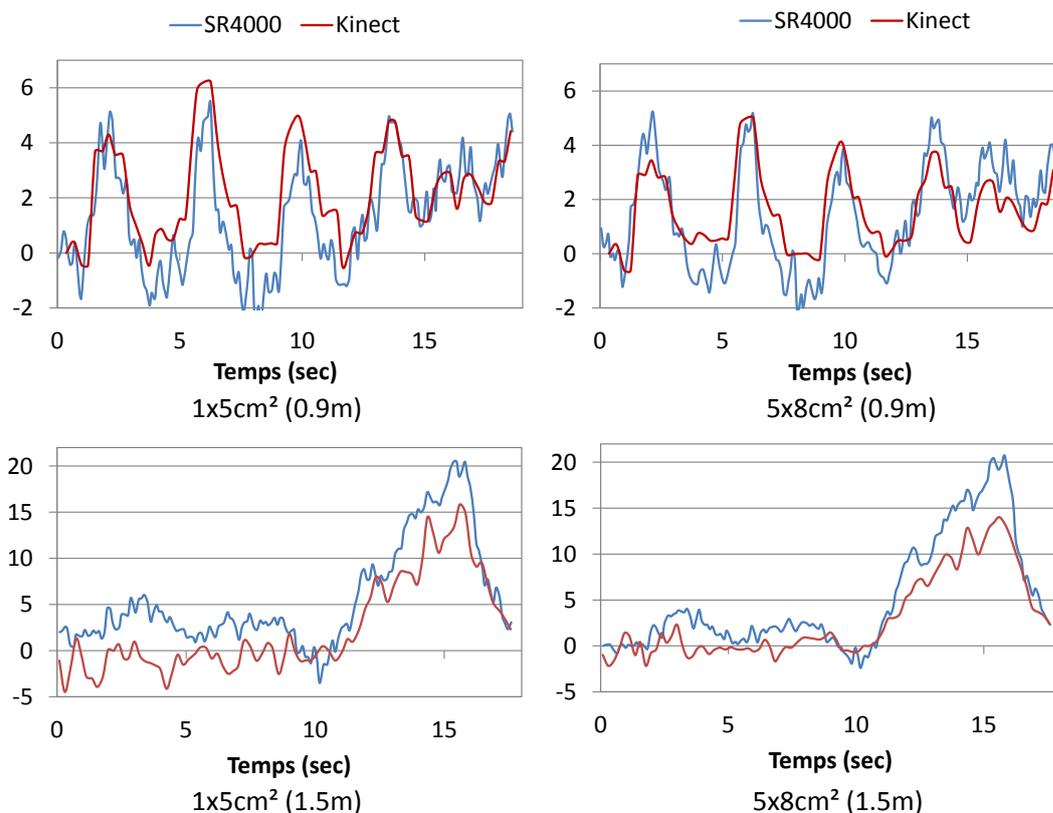


FIGURE 3.11 – Mouvement respiratoire observé par la caméra ToF SR4000 et la Kinect à 0.9 m et 1.5 m avec à gauche une ROI abdominale de 1x5 cm² et à droite une ROI abdominale de 5x8 cm²

On remarque quand même une forte corrélation qualitative entre les deux cartes. Les figures du bas montrent la même analyse pour une observation à 1.5 m. Alors que l'observation par la caméra ToF montre de son côté un important mouvement respiratoire au niveau thoracique, la carte de couleur associée à la Kinect ne permet aucune analyse. Le déplacement mesuré semble être situé aléatoirement le long de la surface par blocs de points se déplaçant conjointement, suivi de larges espaces où les points 3D ne se déplacent pas du tout.

3.5 Synthèse et conclusion

Ce chapitre cherche à caractériser les deux modèles de caméra ToF ainsi que la Kinect pour leur capacité à détecter et estimer le mouvement surfacique. La première série d'expérimentations se focalise sur l'influence de la distance et de l'angle d'observation sur cette estimation afin de prendre en compte ces paramètres lors de l'élaboration de dispositifs d'observations en conditions cliniques, conditions souvent limitantes de part la configura-

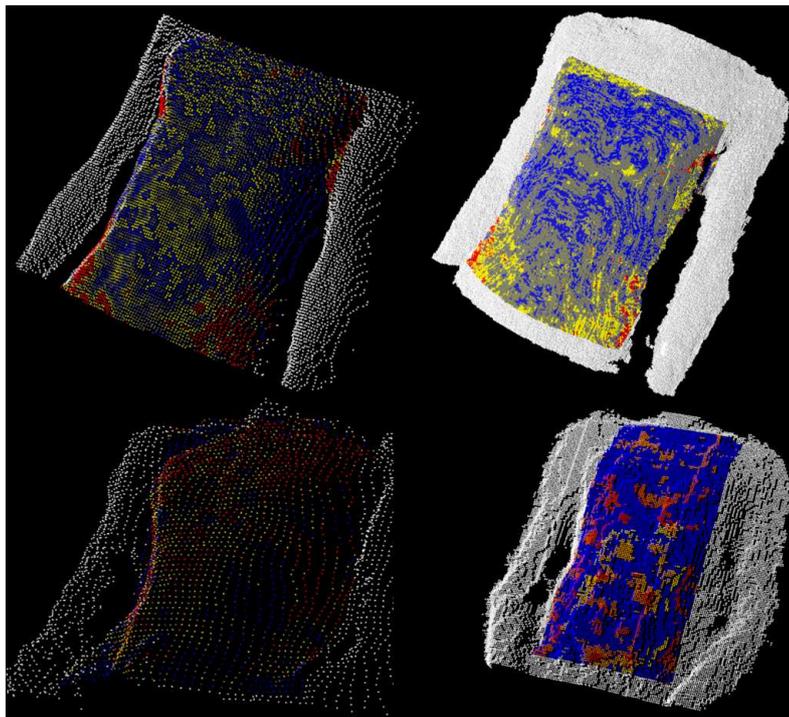


FIGURE 3.12 – Le déplacement en profondeur pour les points d’une large région d’intérêt est montré ici par un code couleur allant de 0 mm (bleu) à 5 mm (rouge) avec à gauche les données de la caméra ToF SR4000 et à droite celles de la Kinect. En haut, observations à 0.9 m et en bas à 1.5 m.

tion de l’environnement de travail. L’angle d’observation a une influence quasi-nulle jusqu’à 50° pour les trois caméras même si la Camcube 3.0 voit son erreur augmenter légèrement à partir de 30°. De même, la distance d’observation n’influe pas sur la qualité de mesure des caméras ToF jusqu’à 2 m, même s’il est à souligner que le nombre de points décrivant une même surface diminue très fortement entre 1 m et 2 m d’observation, limitant les possibilités de traitement des données 3D (filtrage, moyennage, modélisation). La Kinect voit son erreur augmenter quadratiquement avec la distance, ce qui correspond aux résultats obtenus par Khoshelham and Elberink (2012). Cependant il est à souligner que cette erreur est sous-estimée car l’effet de quantification tend à empêcher un point de mesure de passer au niveau de quantification précédent ou suivant, amenant à un déplacement exactement égal à 0. L’erreur augmente cependant à cause des points de mesure étant au bord du basculement vers la valeur suivante, et donc basculant avec un fort déplacement 3D (12 mm pour une surface à 2 m par exemple).

Les résultats montrent que les mesures de la Kinect sont caractérisées par l’effet de quantification mais aussi par un bruit sur les mesures très faible, ce qui pourrait s’expliquer par du post-traitement des données intégré dans la caméra. Finalement, les surfaces ainsi que leurs mouvements sont crénelés, et d’autant plus avec la distance, comme on peut le voir

sur la figure 3.8 où le déplacement d'un seul pixel n'est entaché d'aucun bruit à quelque distance que ce soit. Un bruit de mesure est cependant possible, mais caché par l'amplitude de la quantification en profondeur. La conséquence de cet effet est l'impressionnante qualité d'estimation d'un mouvement au niveau d'un seul point à faible distance : 0.7 ± 0.5 mm à 0.8 m. La Kinect a donc un énorme potentiel de suivi du mouvement surfacique lorsque les conditions permettent d'être près de la surface. L'état de l'art montre des études cherchant à estimer le déplacement de la tête d'un patient (Noonan et al., 2011), ce qui ouvre la perspective de la correction du mouvement en imagerie TEP par exemple. En effet, la tête d'un patient est observable pour une caméra à faible distance, ce qui fait de la Kinect la caméra idéale pour l'estimation du mouvement. Pour des distances d'observation supérieures, tout dépend de la précision que l'on recherche ou l'amplitude du mouvement à détecter. La Kinect garantit une mesure très peu bruitée, mais la résolution en profondeur empêche la détection de mouvements plus faibles qu'elle, tout en sous-estimant les mouvements de plus large amplitude. L'exemple du signal respiratoire observé à 1.5 m, visible à droite sur la figure 3.11 est éloquent à ce sujet : les cycles respiratoires d'environ 4 mm d'amplitude ne sont pas détectés avec la Kinect, les cycles plus amples étant sous-estimés.

Au final, c'est le rapport entre l'effet de quantification en profondeur de la Kinect et le bruit de mesure aléatoire des caméras ToF qui va décider quelle outil est le plus adapté pour estimer un mouvement surfacique. En-deçà d'un mètre, la Kinect est supérieure, le rapport s'inversant au-delà et au fur et à mesure que la distance d'observation augmente.

Concernant les deux modèles de caméras ToF, on peut proposer des conditions qui vont décider laquelle utiliser. Tout d'abord, le modèle SR4000 montre globalement de meilleures estimations du mouvement que son concurrent. Cependant, dans le cas où la résolution du nuage de points est un paramètre important, la Camcube 3.0 fournira un plus grand nombre de points avec une meilleure résolution latérale (40 000 au lieu de 25 000). Dans le cas de scènes où les conditions de luminosité sont variables, la SR4000 présente l'avantage d'avoir une fonction d'auto-illumination qui garantira un temps de pose empêchant l'effet de saturation tout en garantissant la collecte du maximum de signal. Un dernier critère est la compacité de la caméra : la SR4000 est significativement plus petite que la Camcube 3.0, ce qui est un avantage dans les environnements confinés.

Chapitre 4

L'utilisation d'une caméra ToF pour l'extraction en temps réel du mouvement respiratoire externe

Sommaire

4.1	Modélisation surfacique des données	91
4.1.1	Les fonctions B-Splines	92
4.1.2	L'algorithme MBA	93
4.1.3	Extraction du mouvement	95
4.2	Caractérisation de la qualité de mesure	96
4.2.1	Paramètres de la caméra	99
4.2.2	Impact de la modélisation	100
4.3	Comparaison avec un examen 4D-TDM comme vérité terrain	102
4.3.1	L'examen 4D-TDM	103
4.3.2	Conditions d'observation cliniques	104
4.3.3	Mise en correspondance des données	106
4.3.4	Traitement des données	108
4.3.5	Résultats	109
4.4	Prédiction du mouvement interne	115
4.5	Synthèse et conclusion	117

La caméra Temps-de-Vol possède des caractéristiques uniques qui pourraient lui permettre d'observer précisément et en temps réel le mouvement respiratoire externe. Son principe de mesure de profondeur, basé sur l'émission puis la mesure de lumière infrarouge, permet une estimation directe de la distance séparant le pixel de mesure et le point de réflexion de la lumière, sans calibrage préalable et sans reconstruction des données. De plus,

sa compacité par rapport aux autres types de scanners 3D rendent une utilisation clinique plus pratique dans un environnement encombré comme celui d'une salle de traitement de radiothérapie. Ce deuxième chapitre présente la démarche suivie afin d'évaluer l'utilisation de cette caméra pour l'extraction du mouvement de la surface thoraco-abdominale du patient en temps réel.

Nous présenterons tout d'abord la modélisation des données brutes à l'aide de fonction B-splines, évaluerons son apport au niveau la mesure, ainsi que l'influence de deux paramètres essentiels que sont la distance d'observation et l'angle entre la surface observée et la caméra. La validation est effectuée par comparaison avec des données issues d'examens 4D-TDM : dans le cadre de la collaboration entre le LaTIM et le service de radiothérapie du CHU de Brest, des patients qui subissent un tel examen sont observés parallèlement avec une caméra ToF. Nous finirons par utiliser le modèle de respiration patient-spécifique préalablement développé au sein du laboratoire afin de générer des volumes TDM en utilisant le mouvement des surfaces capturées par une caméra ToF.

4.1 Modélisation surfacique des données

La principale raison qui motive le choix de modéliser la surface fournie par une caméra ToF est fournie par la nécessité de mesurer des mouvements ou déformations surfaciques de faible amplitude avec des points de mesures réputés bruités. Si l'on se base sur le travail de Hostettler (2008), dans le cas d'une respiration libre où le patient est allongé sur le dos, les ordres de grandeurs du mouvement de la peau de patient selon l'axe antéro-postérieur varient en fonction de la localisation pour atteindre le centimètre dans la région abdominale (Figure 4.1). Pour la région thoracique, théoriquement moins déformée par la respiration abdominale, il est observé un mouvement atteignant 3 mm.

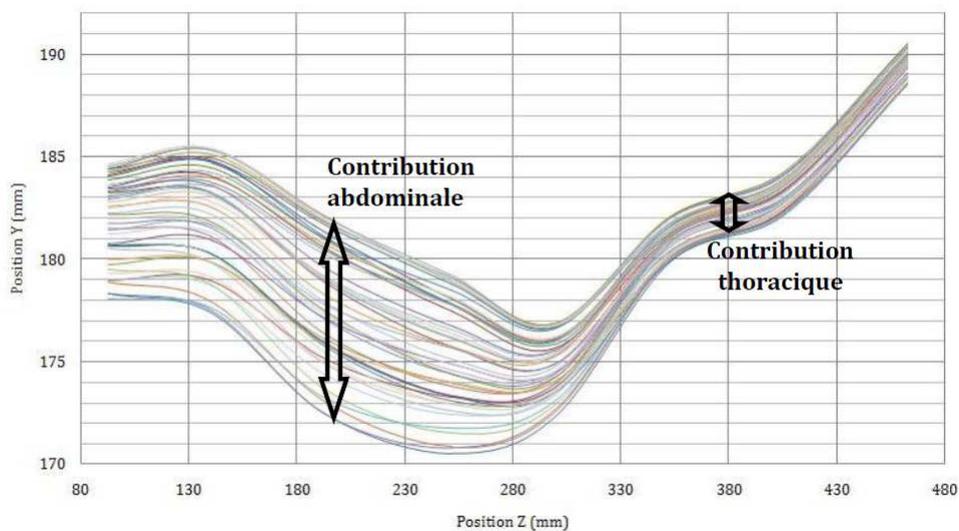


FIGURE 4.1 – Mouvement de la peau selon l'axe antéro-postérieur le long d'un cycle respiratoire (Hostettler, 2008). Le mouvement est mesuré à l'aide de marqueurs optiques, plus précisément des ARTags dont la position est déterminée par un système de caméras stéréoscopiques. L'abscisse représente la direction cranio-caudale, le long de laquelle les marqueurs sont situés, avec à gauche la région abdominale et dans la partie droite la région thoracique.

Dans ces conditions, il semble indispensable de réduire le bruit de mesure au maximum. En effet, Piatti and Rinaudo (2012) montre que la moyenne de l'écart-type de 100 mesures de distances à 1 m est de 3 mm pour le modèle de caméra SR4000. De même, pour la Camcube 3.0, cette moyenne est de 4.8 mm avant une saturation excessive des pixels. Enfin, le chapitre 3 confirme cet ordre de grandeur avec un écart-type de 2.4 mm pour le modèle SR4000 pour une observation à 1 m. Sachant qu'estimer une déformation revient à trouver le déplacement local entre deux acquisitions, il devient alors difficilement envisageable

d'estimer cette déformation avec un tel ordre de grandeur de bruit, à moins de moyenniser le mouvement sur de larges régions d'intérêt, ce qui diminue l'intérêt des scanners 3D qui est de permettre l'analyse surfacique localisée sur de petites régions. Nous avons donc considéré comme nécessaire de nous pencher en premier sur une modélisation qui permettrait de réduire cette variabilité de mesure.

Dans le cas de données éparées, l'interpolation d'un ensemble de points 3D permet d'adapter une fonction à ce jeu de données afin de caractériser la structure sous-jacente, ici une surface. Un lissage est aussi effectué par cette interpolation. De plus, la caractérisation de la surface par une fonction permet de remplacer cet ensemble de données difficilement manipulable par des points de contrôle appartenant à la fonction sous-jacente. Notre choix s'est arrêté sur les fonctions B-Splines pour leur réputation dans le domaine de la modélisation des surfaces mais aussi pour un temps de calcul réduit même en présence de nombreux points de mesure. Enfin, l'utilisateur peut gérer les paramètres de description de la surface par la disposition des points de contrôle en grilles régulières avec un espacement connu. Après un rappel sur la nature des fonctions Splines et B-splines, nous présenterons l'algorithme Multilevel B-Spline Approximation (MBA) proposé par Lee et al. (1997), qui sera utilisé tout au long de ce travail. Cette méthode ménage un bon compromis précision/intensité de lissage de la fonction tout en étant extrêmement rapide d'un point de vue computationnel. Nous présenterons enfin une méthode d'extraction du mouvement surfacique local et selon la direction de la profondeur, ce qui nous permettra dans la section suivante une validation sur fantôme et sur humain de l'estimation en temps réel du mouvement surfacique.

4.1.1 Les fonctions B-Splines

Une Spline est une fonction définie par morceaux par des polynômes. Au contraire de l'interpolation polynomiale qui cherche à faire correspondre à un ensemble de données un seul polynôme passant par chacune de ces valeurs, la méthode des Splines découpe l'espace d'interpolation en sous-ensembles et détermine une fonction polynomiale pour chacun de ces sous-espaces. L'intérêt de cette méthode est d'éviter la recherche de polynômes de haut degré, ce qui requiert un temps de calcul conséquent. De plus, il évite le phénomène de Runge qui montre que rechercher un polynôme de haut degré sur un jeu de données peut engendrer des oscillations entre deux valeurs de mesure afin de faire correspondre la fonction aux valeurs (Runge (1901), 4.2).

Une continuité des différents polynômes locaux est requise afin d'obtenir une fonction finale continue. Une fonction B-Splines est une combinaison linéaire de fonctions splines : si elle est uniforme, la fonction est décrite par un ensemble de noeuds ou points de contrôle équidistants.

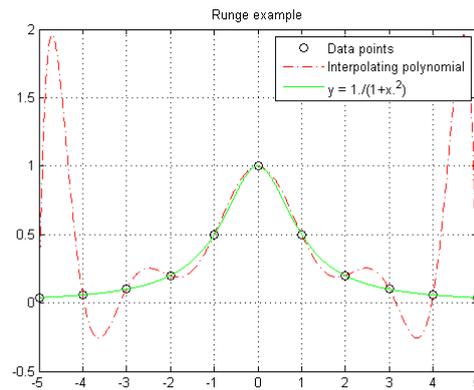


FIGURE 4.2 – Illustration du phénomène de Runge où un jeu de données simple est interpolé par un polynôme de degré 4

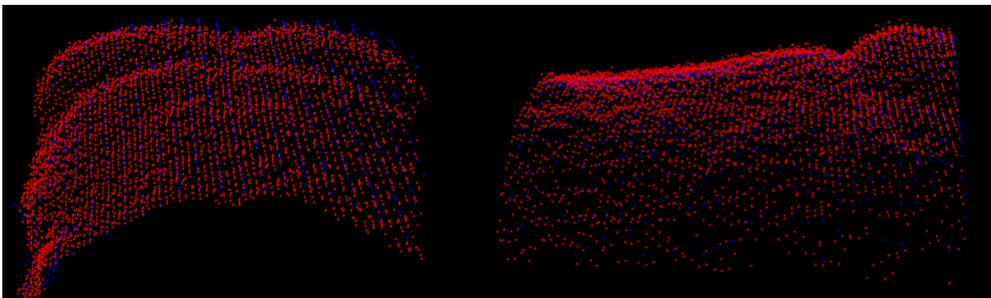


FIGURE 4.3 – Modélisation d'un nuage de points de mesure (en rouge) par une fonction B-Splines (en bleu).

4.1.2 L'algorithme MBA

L'algorithme Multilevel B-Spline Approximation (MBA) présenté par Lee et al. (1997) propose l'approximation de données éparses ou irrégulières avec des fonctions B-Splines en optimisant la vitesse de calcul à l'aide de deux principes, une évolution itérative de la taille des grilles de contrôle afin de partir d'une approximation grossière pour aller vers une approximation fine, et une juxtaposition des grilles de différentes résolutions au fil des itérations.

L'approximation B-Spline part du principe d'apposer à un ensemble de points 3D irréguliers une grille régulière de points de contrôles décrivant la fonction f sous-jacente aux données. La résolution latérale de cette grille est déterminée par l'utilisateur, ce qui permet une approximation plus ou moins fine en fonction de l'espacement des points de contrôle puisque chaque point de mesure influe sur les 16 points de contrôle environnants : la figure 4.4 illustre l'influence de cette résolution sur la forme de la surface modélisée. Réciproquement, chaque point de contrôle est influencé par les points de mesures présents dans son voisinage 4×4 : sur la figure 4.5, les deux points de mesure z_1 et z_2 influencent les points de contrôle verts

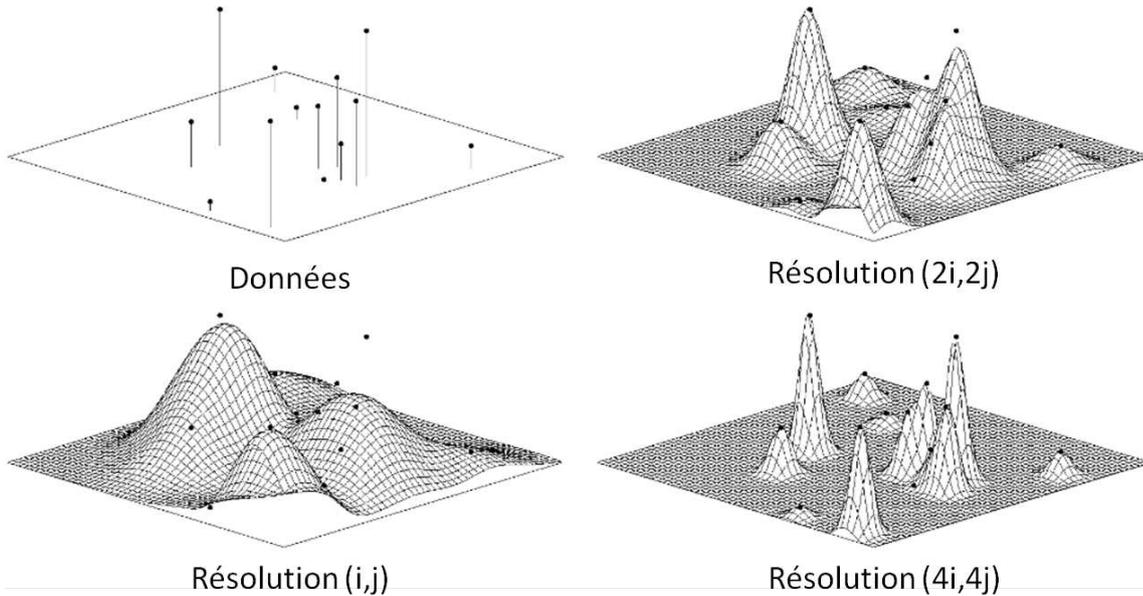


FIGURE 4.4 – Exemple de l’influence de la résolution de la grille de contrôle sur la modélisation des données de mesures et l’intensité de lissage associée. Une grille hautement résolue aura tendance à modéliser parfaitement les points de mesure au détriment du réalisme.

qui reçoivent donc une contribution des deux points proportionnelle à leur éloignement.

Au final, La fonction f approximant les données est définie en chaque localisation (x,y) comme la somme des contributions des 16 points de contrôle voisins, ces contributions étant fournies par des fonctions B-splines cubiques uniformes : elle est donc bien définie par morceaux par des fonctions B-splines.

La difficulté de la méthode de base est la nécessité de trouver un compromis entre l’intensité de lissage de la fonction finale et la précision de l’approximation : une grille très résolue approximera précisément les points de mesure mais conduira à des aberrations lorsqu’il manquera des points dans une région de la surface (trous dans la fonction, pente irréaliste). Afin de résoudre ce dilemme, l’algorithme MBA propose une première optimisation : une approche multi-niveau verra une approximation itérative en commençant par une grille composée de peu de points de contrôle. L’itération suivante consiste à modéliser avec une grille deux fois plus résolue, non plus les points de mesure, mais l’erreur résiduelle d’approximation de la première itération. Le processus se répète avec des grilles deux fois plus résolues à chaque itération jusqu’à qu’un seuil d’erreur d’approximation ou une trop faible distance entre les points de contrôle arrêtent l’algorithme. Finalement, les fonctions d’approximation déterminées à chaque itération sont additionnées afin de constituer la fonction finale. Le compromis intensité de lissage/précision est ici résolu puisque la fonction finale a comme composante de base la fonction issue de la première itération (lorsque la grille est très peu

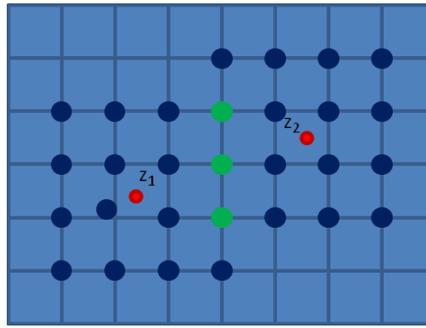


FIGURE 4.5 – Les deux points de mesure z_1 et z_2 influencent chacun leur 16 points de contrôle proches. Réciproquement, les points de contrôle sont influencés par tout points de mesure dans l'espace délimité par les 24 points de contrôle voisins : les points verts reçoivent l'influence de z_1 et de z_2

résolue), alors que dans le même temps la précision est assurée par les grilles aux itérations ultérieures.

Une deuxième amélioration est apportée par rapport à la vitesse de calcul. Comme la structure hiérarchisée présentée additionne des fonctions issues de grilles de résolution différentes, l'addition finale permettant de déterminer la fonction décrivant au mieux le jeu de données est de résolution égale à la dernière grille : il faut donc interpoler chaque grille de résolution inférieure afin d'avoir des tailles de grille similaire en effectuant un rééchantillonnage. Afin de limiter ces opérations, Lee et al. (1997) propose de rééchantillonner les grilles entre chaque itération : l'interpolation ne concerne pour chaque degré qu'un rééchantillonnage d'un facteur 2 dans les directions x et y (Voir figure 4.6).

4.1.3 Extraction du mouvement

Lorsque l'on considère des acquisitions successives de la peau d'un patient avec une caméra fixe, les nuages de points de mesure ne peuvent être reliés entre eux car ils ne décrivent pas une même position anatomique au fil du temps. La seule solution pour estimer la déformation 3D passe par un recalage élastique à partir des nuages de points 3D dépareillés, ce qui empêche à l'heure actuelle toute analyse en temps réel (Le sujet sera abordé section 5.2). Cependant, les modélisations successives par des fonctions B-splines fournissent une occasion de mettre de côté les points de mesure pour les remplacer par la grille de contrôle issue de la modélisation. La position de cette grille est déterminée par l'utilisateur, ce qui ouvre la possibilité de garder à la même position en (x,y) les points de contrôle : la modélisation dynamique permet d'acquérir des grilles successives où les noeuds voient uniquement leur composante z se modifier : leur déplacement 1D décrit donc bien la déformation de la surface à cet endroit selon la direction z de la caméra, soit la profondeur (Voir figure 4.7).

Bien que cette méthode n'estime pas la déformation 3D de la surface (il manque une

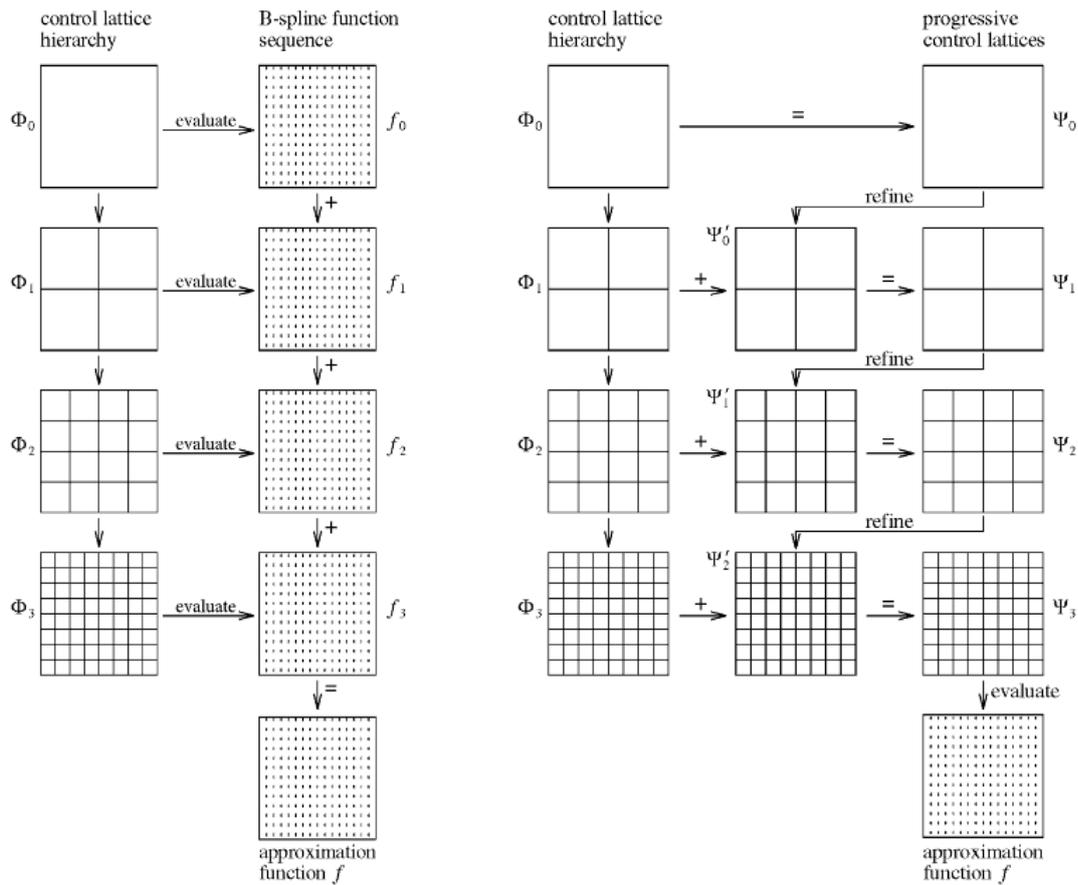


FIGURE 4.6 – Extrait de Lee et al. (1997) qui illustre l’optimisation du temps de calcul induit par le rééchantillonnage des fonctions entre chaque itération plutôt qu’avec la résolution finale.

composante latérale à chaque vecteur de déformation), la déformation générale est tout de même prise en compte, le rapport entre deux vecteurs voisins décrivant une déformation : si l’on considère le volume décrit par la différence de deux surfaces successives (soit la déformation d’un point de vue général), le champ de vecteurs 1D décrit la même variation tout en ajoutant une information locale qui ne permet cependant pas l’estimation du déplacement d’un point de repère précis le long des acquisitions. Par contre, cette manière d’appréhender la déformation surfacique permet une analyse qualitative partielle puisqu’elle permet par exemple de différencier une respiration abdominale d’une respiration thoracique.

4.2 Caractérisation de la qualité de mesure

Les expérimentations effectuées afin d’estimer l’amélioration apportée par l’extraction du mouvement par le biais d’une modélisation B-Splines ont été effectuées avec un fantôme

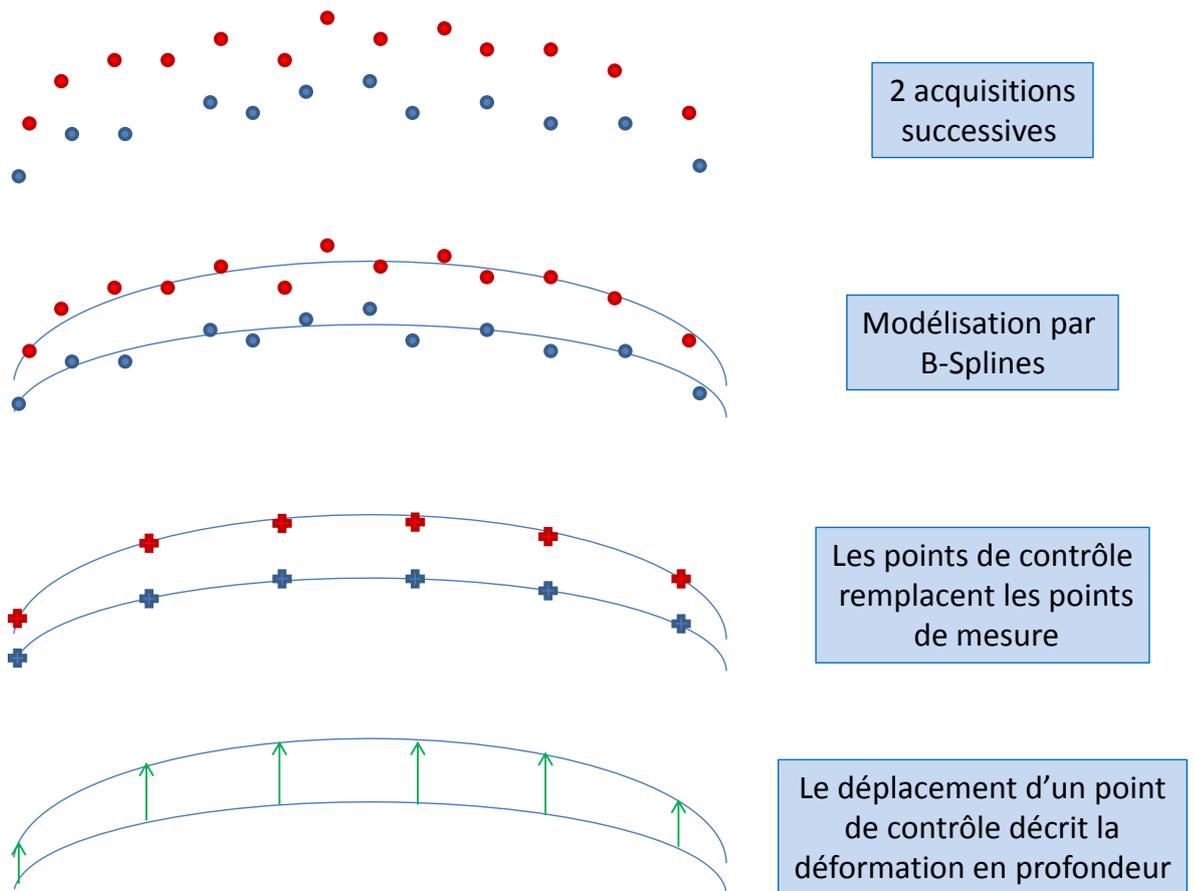


FIGURE 4.7 – La modélisation des surfaces permet de les manipuler aisément et de permettre une extraction extrêmement rapide d'un champ de vecteurs 1D décrivant la déformation locale d'une surface.

anthropomorphique (Voir la figure 4.8). Même si il n'apporte pas un réalisme comparable aux expérimentations sur humain, sa forme, sa couleur permettent un premier degré d'analyse. Les déplacements du fantôme sont fournis par un actionneur linéaire de marque IAI de 80 cms de long avec une précision garantie de 0.02 mm, ce qui nous permet d'avoir une vérité terrain sur les mouvements réels appliqués au fantôme.

Afin d'observer l'impact de la modélisation sur l'estimation du mouvement, nous nous sommes intéressé à l'erreur aléatoire d'estimation du mouvement. Nous avons en effet considéré que l'erreur systématique de mesure du mouvement est négligeable lorsque les deux localisations entre lesquelles le mouvement est mesuré ne sont séparées que de quelques millimètres selon la profondeur : l'erreur systématique sur l'estimation de la première position est compensée par l'erreur systématique à la deuxième position. Nous avons observé

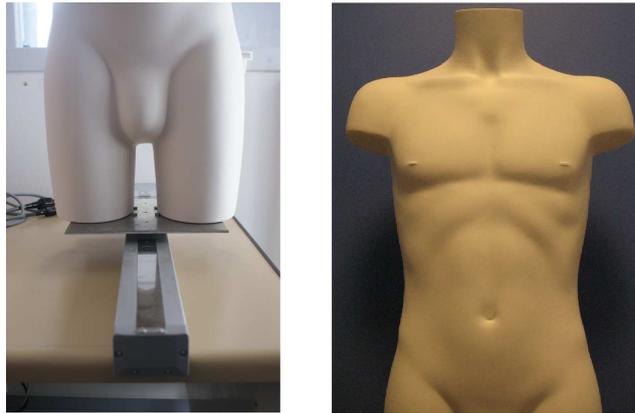


FIGURE 4.8 – Vue du fantôme anthropomorphe utilisé pour simuler le type de relief caractéristique du torse humain, ainsi que de l’actionneur linéaire utilisé pour le déplacer.

le fantôme statique et considéré une région d’intérêt (ROI) proche du centre du champ de vue et contenant 500 points de mesure. L’erreur aléatoire ou variabilité de mesure de mouvement est considérée comme la distance entre deux positions d’un même point de mesure, le mouvement réel étant nul. Nous avons estimé ce mouvement 100 fois. De même, après modélisation de ces 100 acquisitions, la même analyse est effectuée pour chaque point de contrôle.

L’influence de la distance est observée en effectuant cette analyse sur une gamme de distance allant de 0.6 m à 1.5 m. Cette gamme se justifie par les conditions cliniques d’observation tant près d’un imageur type TDM que d’un accélérateur linéaire de traitement en radiothérapie, où l’emcombement impose une telle distance. Les conditions cliniques d’observations seront abordées plus en détail section 4.3.2. De même, tester l’influence de l’angle d’observation est primordial car les conditions cliniques imposent une observation latérale du thorax et de l’abdomen du patient : la caméra ne peut être fixée ni dans le tunnel du Scanner TDM ni au dessus du patient pendant la fraction de traitement. Pour tester cette influence, nous n’observerons plus la surface du fantôme immobile car la forme de celui-ci induit des artefacts de mesure : un point de mesure étant soumis à une variabilité de plusieurs millimètres, la ligne de tir associée à un pixel de la caméra verra la position du point 3D se déplacer le long de cette ligne de tir. Cependant, si la surface est à cet endroit concave en plus d’être observée avec un fort angle, le point pourra être vu alternativement au sommet puis au pied de cette concavité, ce qui augmentera artificiellement le bruit de mesure. Nous avons donc choisi d’observer un mouvement triangulaire de 2cm d’amplitude fourni par l’actionneur linéaire. Les résultats ne sont fournis que pour l’analyse avec les points de contrôle, puisqu’il n’est pas aisément possible de lier les acquisitions brutes successives lorsque la surface se déplace : pour les points de contrôle, le mouvement est analysé comme le déplacement selon l’axe z. Un changement de référentiel spatial permet de faire pivoter les données de mesure afin de les réaligner avec la direction du mouvement.

4.2.1 Paramètres de la caméra

La section 2.2.4 présente les différents types d'erreur associées à la mesure de la profondeur. Ils se divisent en deux types, selon leur nature systématique ou aléatoire. Dans le cadre de cette thèse, les sources d'erreur systématique ne rentrent pas en ligne de compte lors d'une mesure de mouvement, qui est donc une mesure relative. De part les faibles amplitudes de mouvement mesurées, on peut stipuler un très faible impact de ces erreurs, un erreur systématique pour une acquisition à une profondeur donnée étant quasiment identique pour une acquisition à quelques millimètres. Les erreurs non systématiques peuvent au contraire poser problème. Le ratio signal-sur-bruit variable dans une scène en cas de fortes variations d'illumination est limité dans le cas où l'on observe une surface dont les différentes localisations sont situées à une distance similaire de la caméra. De même, les phénomènes de réflexion multiples sont dépendants des conditions d'observation, en fonction des objets ou instruments dans le champ de vue. Ces conditions seront stables pour les acquisitions en parallèle d'un examen 4D-CT (section 4.3). Par contre, lors de l'étude de repositionnement présentée chapitre 6, le déplacement de l'accélérateur linéaire ainsi que des supports d'imagerie portale autour du patient peuvent être une source d'erreur si les surfaces comparées ne sont pas dans le même environnement. L'effet de diffusion n'est sensible qu'en présence d'un objet dans le champ proche de la caméra, ce qui n'est pas le cas de nos mesures. Enfin, l'effet de flou associé au déplacement rapide d'un objet n'apparaît pas dans nos observations car le mouvement latéral de la surface n'est pas assez rapide pour induire un tel effet.

Les mesures sont effectuées avec le modèle de caméra ToF SR4000. Sauf indication contraire, ce sera également le cas dans la suite de ce manuscrit. La caméra SR4000 propose tout d'abord une fonction d'auto-illumination qui adapte le temps de pose afin d'optimiser la mesure tout en évitant la saturation des capteurs : cette fonction sera utilisée systématiquement dans ce travail de thèse. En effet, Piatti and Rinaudo (2012) montre que cette fonction garantit les meilleures mesures tant au niveau du bruit de mesure que de l'erreur systématique. Le constructeur conseille de plus l'utilisation d'un modèle de correction du bruit de mesure conçu en fonction des caractéristiques de la caméra, associé à un filtre adaptatif 5x5. Nous utiliserons pour toutes les mesures ces paramètres d'acquisition. Lors des expériences qui suivent, l'ajout d'un filtre spatial gaussien 5x5 sur les points de contrôle est testé. Il est en effet intéressant d'effectuer un filtrage sur une grille régulière dont la résolution est connue, le filtrage devenant une convolution très peu coûteuse en temps de calcul, au contraire d'un filtrage effectué sur un ensemble de points épars.

4.2.2 Impact de la modélisation sur l'estimation du mouvement surfacique

La figure 4.9 montre l'erreur moyenne d'estimation du mouvement nul du fantôme pour 3 types de données : les données brutes, les points de contrôles des fonctions B-Splines et ces mêmes points de contrôle sur lesquels est appliqué un filtrage spatial gaussien. On distingue tout d'abord l'impact de la modélisation : l'erreur moyenne est réduite en moyenne d'un facteur 3 sur toute la gamme de distance avec une réduction de l'écart type associé, c'est à dire la variabilité de l'erreur le long de la surface. Concernant les données brutes, l'erreur moyenne va de 2.6 mm à 0.6 m à 3.2 mm à 1.5 m avec une pointe à 3.8 mm à 1 m. Dans le même temps, l'erreur moyenne pour les points de contrôle ne dépassent pas 1.2 mm, et l'apport du filtrage gaussien garantit qu'aucun point de contrôle n'a une erreur moyenne supérieure à 2 mm.

Lorsque l'on regarde l'évolution générale, la distance a une influence jusqu'à 1 m avec une augmentation quasi linéaire de l'erreur. Au delà, l'erreur reste stable, voire diminue dans le cas des données brutes. Ce schéma d'influence correspond assez bien avec les données du constructeur qui annonce une précision nominale de 1% de la distance avant 1 m et 1 cm au delà, ce qui correspond à une évolution linéaire suivie d'une stabilisation.

La figure 4.10 montre pour une observation à 1 m la distribution des erreurs moyennes pour chaque jeu de données de même que l'erreur et l'écart-type associés pour un pixel le long des 100 estimations de mouvements. Ce pixel est choisi comme celui ayant l'erreur médiane sur l'ensemble des points. On y observe que la modélisation permet une homogénéisation des mesures le long de la surface avec un écrasement des erreurs moyennes : 95% des erreurs sont comprises dans un intervalle de 5.1 mm pour les mesures brutes, cet intervalle se réduisant à 0.5 mm pour contenir 95% des points de contrôle, 0.3 mm en cas de filtrage additionnel. De même, les résultats à l'échelle d'un pixel confirment cette baisse appuyée de la variabilité de mesure.

La figure 4.11 montre la distribution de l'erreur moyenne d'estimation du mouvement triangulaire du fantôme anthropomorphique pour des angles d'observation allant de 0° à 45°. Cet angle est déterminé par l'angle formé entre le plan moyen formé par le nuage de points de la ROI et le plan x-y de la caméra. Les résultats indiquent une faible influence de l'angle sur la qualité de mesure, même si des outliers apparaissent au delà de 30° lorsqu'aucun filtre n'est appliqué, pouvant atteindre 2.6 mm. Quelque soit l'angle, l'erreur moyenne est limitée à 1.2 mm. Avec filtrage, l'erreur moyenne maximale est de 2 mm.

Ces résultats montrent que le fait d'estimer le mouvement d'une surface est fortement amélioré lorsque l'on considère la surface à travers les points de contrôle plutôt qu'avec les mesures provenant de la caméra ToF l'ordre de grandeur de cette erreur, autour du millimètre, nous permet d'espérer une estimation de qualité d'un mouvement surfacique de quelques

4.2. Caractérisation de la qualité de mesure

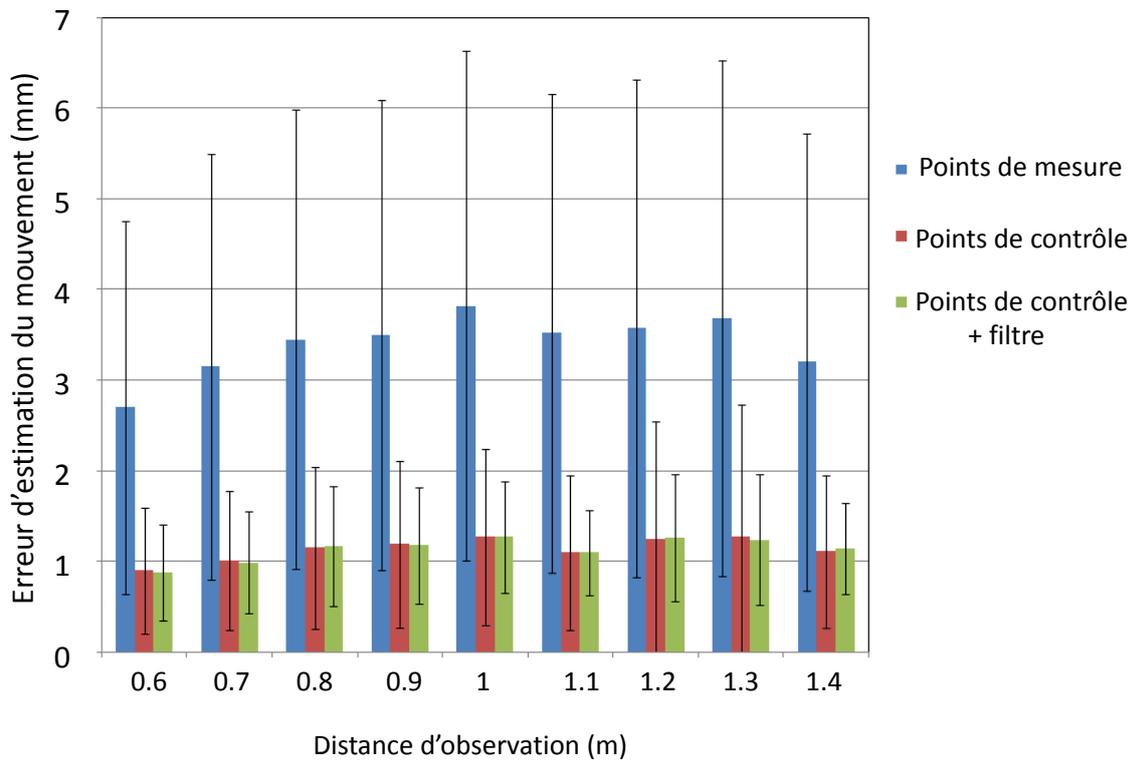


FIGURE 4.9 – Erreur moyenne d'estimation d'un déplacement nul en fonction de la distance d'observation.

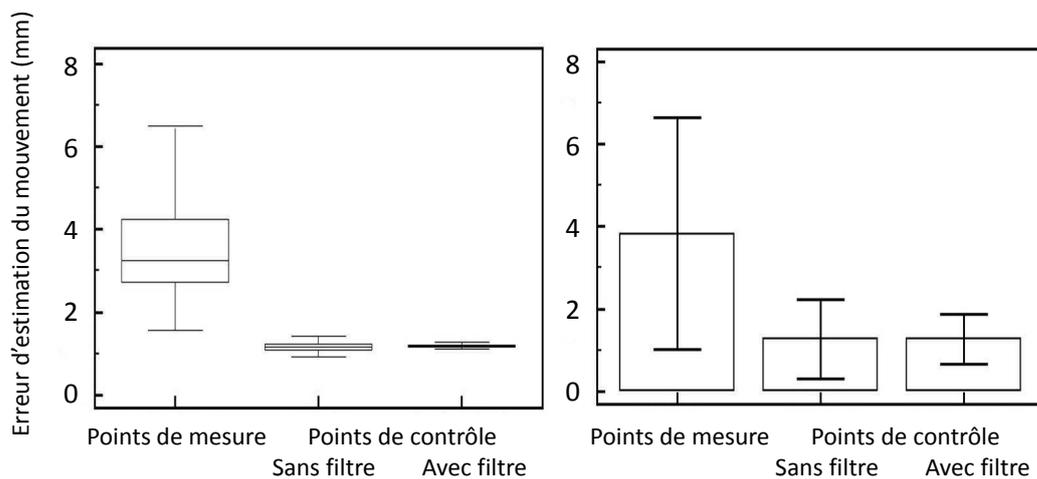


FIGURE 4.10 – A 1m, distribution des erreurs moyennes le long de la surface (gauche), puis à l'échelle d'un pixel ayant l'erreur médiane (droite)

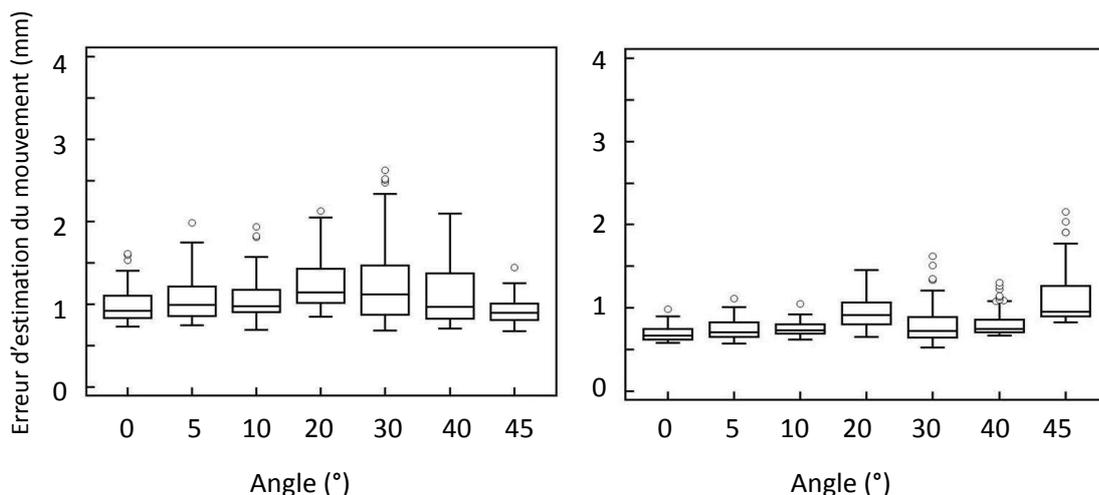


FIGURE 4.11 – Erreur d’estimation d’un mouvement triangulaire de 2cm d’amplitude pour différents angles d’observation. L’estimation est effectuée avec les points de contrôle sans (gauche) et avec (droite) filtrage spatial gaussien.

millimètres d’amplitude, comme dans le cas du mouvement respiratoire. De même, l’influence de l’angle n’est pas significative jusqu’à au moins 45°, ce qui corrobore la littérature (Chiabrando et al., 2010) ainsi que les résultats présentés précédemment dans la partie 3. Ces résultats sont cependant obtenus sur un fantôme anthropomorphe qui ne décrit pas parfaitement les propriétés surfaciques de la peau, surtout au niveau de la réflectivité. Il est donc indispensable d’effectuer des tests sur humains. Nous avons choisi comme validation la comparaison avec un examen 4D-TDM qui fournit les surfaces d’un patient pour chaque instant d’un cycle respiratoire moyen.

4.3 Comparaison avec un examen 4D-TDM comme vérité terrain

Les expérimentations sur fantômes indiquent que le millimètre de précision est atteignable à l’échelle du point de contrôle. Afin de valider plus en avant, des patients passant un examen 4D-TDM sont observés simultanément avec une caméra ToF. L’idée directrice est d’extraire les surfaces du patient des reconstructions TDM pour un cycle respiratoire moyen, de procéder de manière équivalente avec les données de la caméra ToF, pour finalement comparer le mouvement en profondeur de régions d’intérêt. Le travail nécessaire concerne tout d’abord la mise en correspondance spatiale des données scanner et des acquisitions ToF afin de les mettre dans le même repère. Le traitement complet des données dans le but de sélectionner les acquisitions de manière équivalente à un binning est aussi illustré.

4.3.1 L'examen 4D-TDM

L'examen 4D-TDM fournit l'information anatomique 3D d'un patient en prenant en compte la dimension temporelle afin de restituer une reconstruction pour chaque phase d'un cycle respiratoire moyen. De par le champ axial limité des systèmes d'acquisition, il est difficile d'acquérir l'image d'une large région anatomique d'un seul tenant. C'est ainsi que les données sont acquises pour différentes parties de l'anatomie à différents moments par la translation du patient dans le champ d'acquisition. En parallèle de l'acquisition, un système externe permet de capturer le signal respiratoire 1D du patient. Ces systèmes sont quasi systématiquement soit une ceinture de pression, soit un système infrarouge passif type RPM (Section 2.2.1), et le signal est mesuré au niveau de l'abdomen, à l'endroit où l'amplitude de mouvement est maximale (Figure 4.12). Les acquisitions sont ensuite agencées morceaux par morceaux en fonction de l'instant respiratoire qui leur est associé (Figure 4.13).



FIGURE 4.12 – Le système RPM de Varian est constitué d'une caméra fixée au bout du lit de traitement qui observe le déplacement de deux marqueurs réfléchissants infrarouges, visibles sur cette image. Le logiciel intégré fournit alors en temps réel une estimation du déplacement dans les trois directions.

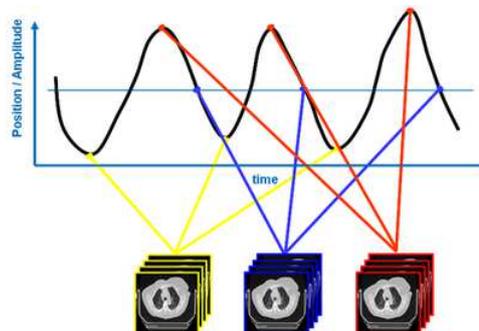


FIGURE 4.13 – Le Binning des données TDM acquises dynamiquement est effectué en sélectionnant les projections pour un même instant respiratoire. Cet instant peut être considéré par une même amplitude du signal respiratoire ou par une même phase à chaque cycle de respiration

Il existe deux méthodes d'acquisitions différentes, l'acquisition hélicoïdale et l'acquisition cinétique. Lors de l'acquisition en mode hélicoïdal, Les projections sont acquises pendant que le lit se déplace linéairement, puis liées au signal respiratoire associé à chacune d'entre elles. Les volumes TDM sont ensuite reconstruits pour une position du cycle respiratoire en associant les projections appropriées pour obtenir une série de volumes 4D-TDM (3D + temps). Le principe de l'acquisition hélicoïdale est à l'origine proposé pour les scanners à une barette (Ford et al. (2003), Vedam et al. (2002)) puis étendu aux scanners multi-barrettes (Keall et al., 2005). L'utilisation d'un scanner 4D-TDM montre une très forte réduction des artefacts dûs au mouvement de la tumeur (Vedam et al., 2002). Lors d'une acquisition en mode cinétique, le patient se déplace par petits décalages dans la direction cranio-caudale. Pour chaque position, l'acquisition dure au minimum un cycle respiratoire et une seconde afin d'acquérir de l'information pour un cycle respiratoire entier. Une fois l'acquisition terminée, le champ se coupe puis le patient est déplacé, et ainsi de suite tant que le champ complet n'est pas balayé. Les volumes 4D-TDM sont ensuite créés par concaténation des tranches issues des projections correspondant à un instant respiratoire (Low et al. (2003), Pan et al. (2004)).

4.3.2 Conditions d'observation cliniques

La configuration de la salle d'examen 4D-TDM limite les emplacements possibles d'observation du patient pendant l'examen. La géométrie du scanner ainsi que la taille du tunnel, combinée à l'ouverture du champ de vue de la caméra ToF SR4000, ne permet que deux emplacements pour la caméra : au dessus de sa tête et au dessus de ses pieds (Figures 4.14 et 4.15). L'angle d'observation est en moyenne de 30° mais peut atteindre 45° pour les patients à forte corpulence qui ont donc leur surface externe située plus haut dans le tunnel. Il est à signaler qu'à cause des conditions d'observations, il a été observé que chez certains patients de forte corpulence, des zones anatomiques étaient quasi invisibles quelque soit la position de la caméra ToF, de part l'ombre faite par un ventre ou un poitrine trop imposantes. Pour nos acquisitions, la caméra est située au dessus des pieds du patient dans l'axe cranio-caudal, ces zones d'ombres sont donc généralement situées au niveau thoracique.

Le Scanner utilisé pour effectuer les examens est un modèle Siemens SOMATOM Sensation qui autorise une acquisition 4D-TDM en mode hélicoïdal avec un maximum de trente centimètres de longueur du champ de vue pour une taille de voxel de $0.97 \times 0.97 \times 3 \text{ mm}^3$. L'idéal pour notre étude serait de centrer ces trente centimètres sur la partie abdominale du patient afin d'observer les régions où le mouvement surfacique est maximal. Or, dans le cadre d'une étude parallèle effectuée sur ces mêmes données, la vue complète des poumons est nécessaire afin de permettre l'étude de la planification effectuée à partir d'un examen 4D-TDM plutôt qu'avec un simple examen TDM. Dans ces conditions, le champ de vue commence à la limite supérieure des poumons puis est étendu au maximum dans la direction caudale. En

4.3. Comparaison avec un examen 4D-TDM comme vérité terrain

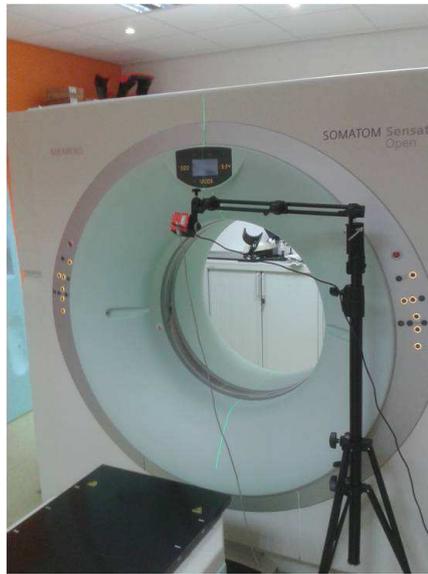


FIGURE 4.14 – Le système de fixation de la caméra ToF est constitué d'un bras girafe prolongé d'une rotule afin de permettre un maximum de degrés de liberté tant en hauteur, profondeur et orientation. La matériel est de la marque Manfrotto (Vitec Company, Italie).



FIGURE 4.15 – Vue du dispositif complet depuis l'arrière du tunnel d'acquisition. Une des difficulté est l'impossibilité de monter plus haut la caméra afin d'avoir un meilleur angle d'observation puisque le tunnel du scanner est arrondi vers l'extérieur.

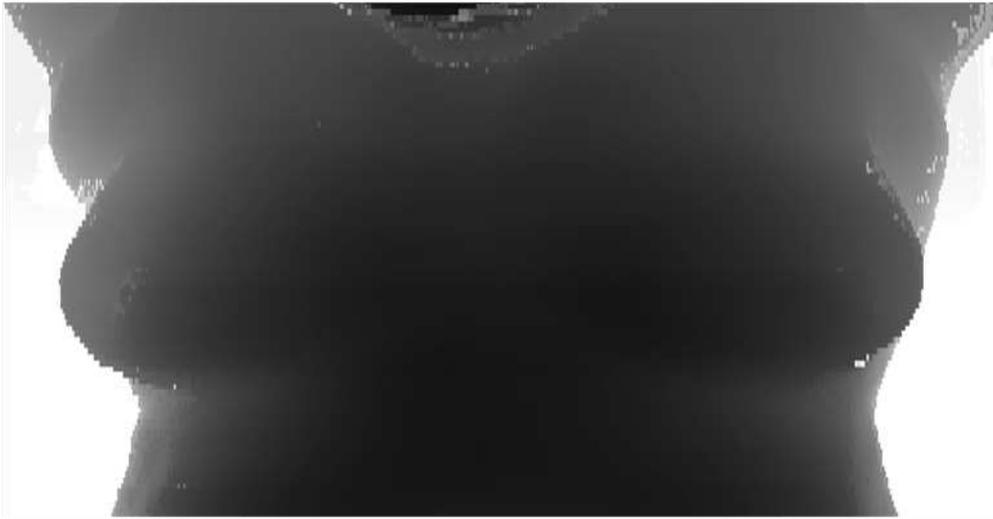


FIGURE 4.16 – Un exemple de la surface d'un patient lors d'un examen 4D-TDM utilisé dans cette étude. On voit que la restriction du champ à trente centimètres dans la direction cranio-caudale ne permet d'observer qu'un bande d'approximativement dix centimètres en dessous des seins.

fonction de la physionomie du patient, cette zone peut s'arrêter avant le diaphragme, ou juste après. Concrètement, ceci implique un faible taux d'examens visualisant la partie abdominale du patient, ce qui limite notre étude aux régions de plus faible mouvement respiratoire externe, les régions thoraciques et intermédiaires.

4.3.3 Mise en correspondance des données

Soit \mathbf{R}_{SR4000} le repère spatial propre à la caméra ToF observant la scène et dans lequel sont contenus les points de mesures, tel que l'axe z corresponde à la profondeur. Soit \mathbf{R}_{AD-TDM} le repère spatial propre au système d'observation scanner, et son axe z correspondant à l'axe antéro-postérieur du patient, pointé de la face antérieure vers la face postérieure. La mise en correspondance des données scanner et ToF nécessite de trouver la transformation reliant ces deux repères (Figure 4.17). Pour déterminer la rotation associée à la transformation rigide, nous observons préalablement la table sur laquelle le patient est allongé, puis déterminons le plan qui lui est associé par minimisation. Cependant, ce vecteur normal ne suffit pas à constituer un repère orthonormé décrivant l'orientation de la table. Il faut alors chercher le vecteur directeur de plus forte pente du plan. Nous utilisons l'analyse en Composantes Principales (ACP) sur le nuage de points décrivant la table : le vecteur propre associé à la valeur propre de plus fort poids indique la direction de la longueur de la table alors que la deuxième valeur propre fournit le vecteur propre en direction de la largeur de la table (Figure 4.18), ce qui nous permet de déterminer le repère complet $\mathbf{R}_{table\ d'examen}$ associé à la table

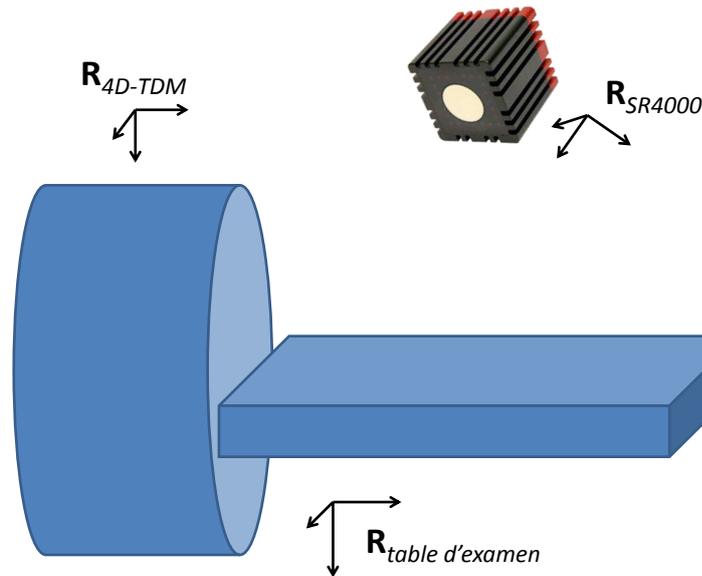


FIGURE 4.17 – La comparaison des données Scanner avec les données ToF nécessitent une mise en correspondance spatiale des données qui est effectuée en déterminant la matrice de passage reliant le repère de la caméra à la table d'examen.

d'examen. La matrice de passage entre les deux repères \mathbf{R}_{SR4000} et $\mathbf{R}_{table\ d'examen}$ permet, une fois appliquée aux données de la caméra ToF, d'aligner ces données à l'axe z du scanner, et ainsi avoir le mouvement en profondeur des deux modalités d'observation dans le même repère. L'alignement final des surfaces nécessite une dernière translation qui est fournie par le recalage de points caractéristiques fixes tels que les épaules du patient. Au final, une région d'intérêt sur un type de données devrait décrire la même région anatomique sur l'autre type.

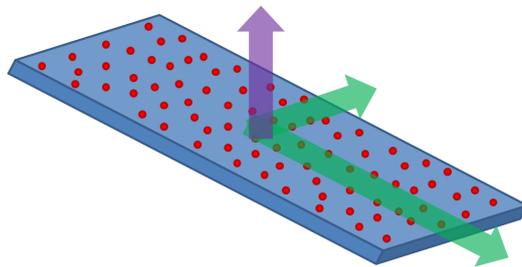


FIGURE 4.18 – Le nuage de points 3D décrivant la table permet la détermination du plan associé (en bleu) qui fournit le vecteur normal associé (violet). Après analyse par ACP, les deux vecteurs propres associés aux deux valeurs propres ayant le plus de poids fournissent un repère complet décrivant la table et son orientation complète, $\mathbf{R}_{table\ d'examen}$

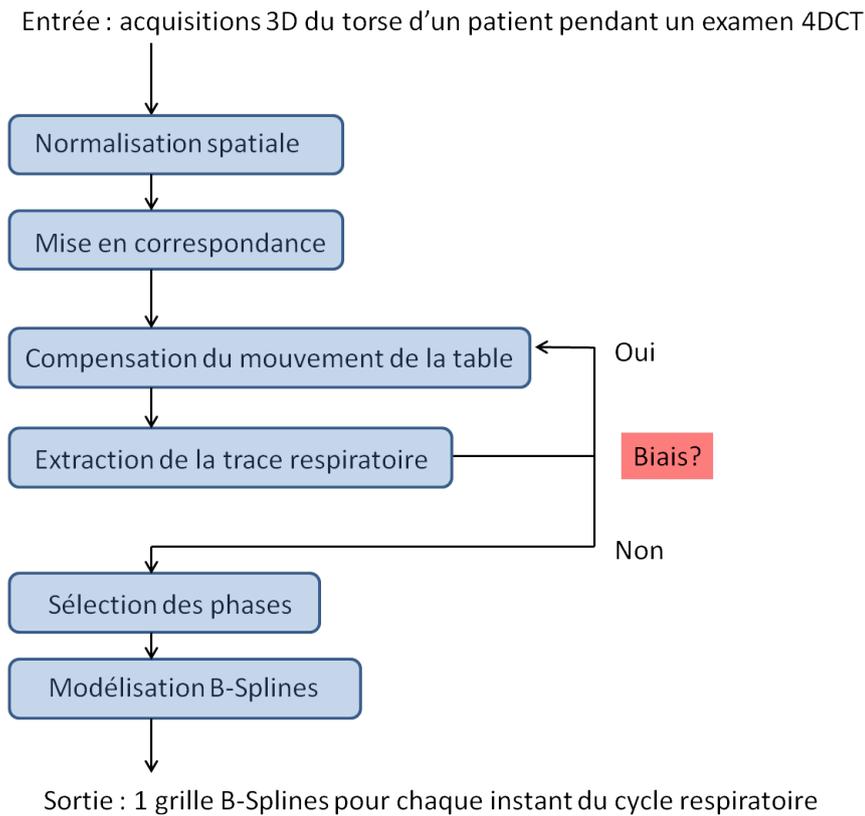


FIGURE 4.19 – Le traitement des données est découpé en trois phases : la transformation des données de la caméra afin de correspondre aux données TDM (Section 4.3.3), la correction du mouvement du patient et l'extraction de la trace respiratoire, puis la sélection et la modélisation des surfaces décrivant un cycle respiratoire moyen.

4.3.4 Traitement des données

L'acquisition des données fournit un film, soit une succession d'acquisitions à 10Hz, correspondant à l'acquisition TDM dynamique qui dure approximativement 99 secondes. Après mise en correspondance spatiale avec le repère de l'appareil d'acquisition (4.3.3), le premier traitement à effectuer est de compenser le mouvement linéaire du patient pendant l'acquisition. Connaissant par les paramètres d'acquisition la translation du patient ainsi que la durée, la vitesse de son déplacement est connue, ainsi que le décalage à appliquer à chaque acquisition en fonction du temps écoulé depuis le début du mouvement. Une fois ceci effectué, un signal respiratoire le moins bruité possible est extrait d'un large région d'intérêt sur la durée totale du film. Cette trace respiratoire nous indique si la compensation du mouvement linéaire durant l'acquisition est bien corrigée : si un offset apparaît, l'erreur de recalage fait que la ROI d'observation ne contient pas la même zone anatomique au fil des 99 secondes. A partir de ce signal, le binning est effectué en sélectionnant un ou plusieurs cycles respiratoires : au contraire des reconstructions 4D-TDM, nous avons accès à l'intégralité de

l'information surfacique à chaque instant, ce qui nous évite d'avoir à réorganiser itérativement les données au fil des cycles respiratoires. Nous sélectionnons un cycle respiratoire moyen, et à partir de ses deux bornes temporelles, les acquisitions équivalentes aux instants de ce cycle. Tout comme dans le cas du Binning des données scanner, la sélection des données peut se faire selon la phase et/ou l'amplitude du signal de référence, ainsi notre choix est de suivre le mode de Binning de l'appareil utilisé pour cette étude, c'est-à-dire de sélectionner les surfaces en fonction de la phase du signal. Il est aussi possible de moyenniser les acquisitions équivalentes sur plusieurs cycles respiratoires. Au final, les 22 acquisitions décrivant le cycle respiratoire moyen sont modélisées par des fonctions B-Splines en étant attentif à ce que les points de contrôles aient la même position relative à la caméra au fil des 22 modélisations.

Pour toute cette étude, la résolution des grilles est de 1 cm selon les directions x et y. Ce choix s'est empiriquement imposé par la nécessité de ne pas appliquer une résolution trop faible afin de ne pas voir apparaître des trous dans les surfaces là où les données manquent (Voir section 4.3.2) tout en assurant une modélisation de qualité.

4.3.5 Résultats

Influence de la taille de la ROI

Dans une première section, une analyse qualitative est effectuée où l'influence de la taille de la ROI considérée sur la qualité de l'information est analysée par l'extraction du signal respiratoire après modélisation pour différentes tailles de régions centrées sur la même position (Figure 4.20). La taille de ces régions est de 1x3, 9x27 et 64x90 mm³. La figure 4.21 montre un exemple de signal respiratoire et illustre le niveau de bruit inhérent à la taille de la surface considérée. On y voit que la taille minimale considérée permet de suivre le mouvement respiratoire mais le signal reste très bruité, ce qui implique, en cas d'utilisation en temps réel, l'usage additionnel de filtres prédictifs de type Kalman. La deuxième taille de région semble plus appropriée au suivi du mouvement car le bruit associé au signal est faible. De plus, l'amplitude des cycles respiratoires est très proche de ceux associés à la région la plus grande.

Concordance spatiale et temporelle

Afin de pouvoir comparer le mouvement sur les données TDM et les données de la caméra ToF, nous avons visualisé pour l'ensemble du champ de vue commun le motif formé par la carte de mouvement issue des deux types de données. Pour ceci, le déplacement selon l'axe z entre les deux phases respiratoires opposées, à expiration et inspiration complète, est déterminé. Pour les données TDM, la différence entre chaque pixel des surfaces segmentées est calculée. De même pour les données ToF, où les grilles B-Splines décrivant les deux sur-

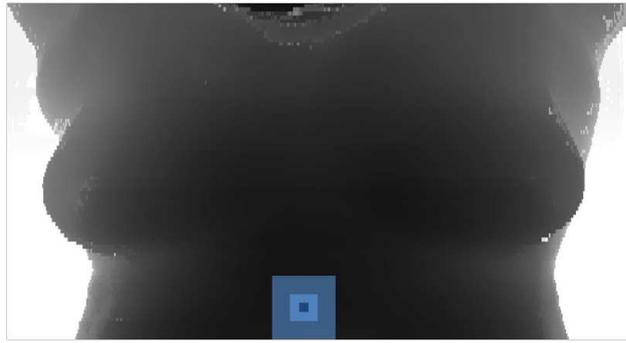


FIGURE 4.20 – Les traces respiratoires sont extraites de régions d'intérêt de taille différentes.

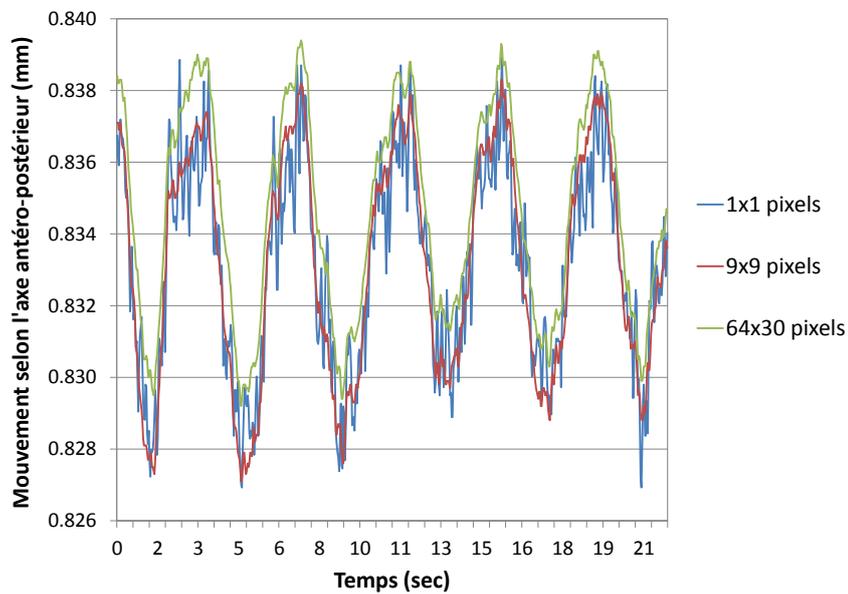


FIGURE 4.21 – Traces associées au ROIS illustrées figure 4.20.

faces fournissent des points de contrôle : leur déplacement selon l'axe z est aussi calculé. Nous présenterons ici trois cas distincts, en premier lieu une patiente avec une corpulence et une poitrine créant une ombre importante, puis une autre patiente dont la morphologie est plus standard, et enfin un homme de morphologie maigre. Bien que ces trois cas ne retranscrivent pas l'ensemble des situations, ils permettent une description de la qualité de l'estimation du mouvement par les données ToF tout en soulignant les limites inhérentes à l'observation.

La figure 4.22 illustre le cas de la patiente dont la poitrine limite l'observation du thorax. En haut à gauche est illustré la carte de couleur illustrant le déplacement selon l'axe z de la surface des volumes TDM entre expiration et inspiration complète pour un patient de sexe féminin. En bleu y est délimité le champ de vue commun avec la caméra ToF, et en

4.3. Comparaison avec un examen 4D-TDM comme vérité terrain

vert la région exploitable. Cette région est plus étriquée que le champ de vue commun car la morphologie de la poitrine de la patiente crée des zones d'ombre sur la partie supérieure des seins et au delà. La zone d'analyse est alors réduite à 17 centimètres dans la direction cranio-caudale, limitée dans sa partie inférieure par l'acquisition scanner et dans sa partie supérieure par le positionnement de la caméra ToF. Pour la région commune exploitable, les deux cartes de couleurs sont montrées dans la partie basse de la figure, avec à gauche la carte issue des données scanner, et à droite celle issue des données de la caméra ToF. On y voit une forte corrélation entre l'information issue des deux types de données dans la partie inférieure de la région analysée, ce qui montre la capacité de la caméra ToF à décrire la réalité de la déformation surfacique, du moins lorsque le nombre de points décrivant la surface est satisfaisant. Cependant, les données scanner détectent un déplacement supérieur à 4 mm sur la face supérieure des seins (peut-être dû à des artefacts, vu la variation locale forte) : ce mouvement n'est pas détecté par la caméra ToF.

Un deuxième cas est illustré figure 4.23 : c'est une patiente ayant une poitrine de taille standard. La surface capturée par la caméra ToF ne possède pas de zones d'ombre dans la région thoracique, ce qui nous permet d'effectuer l'analyse comparative sur l'ensemble du champ de vue commun, soit 25x20 centimètres. Les cartes de couleurs sont là aussi similaires, avec la détection par la caméra ToF de l'asymétrie dans la partie inférieure du champ de vue, de même que dans la partie supérieure droite. Les valeurs de déplacement sont remarquablement proches, avec les régions atteignant 5 mm dans les données scanner étant toutes détectées par les données issues de la caméra. On voit cependant apparaître dans la partie supérieure gauche et centrale une légère différence avec un mouvement externe atteignant 2 à 3 mm pour la caméra là où elle atteint 1 mm pour les données scanner. De plus, on remarque une variation de couleur plus brusque sur les données ToF par rapport à la carte de couleur très lisse fournie par les données scanner.

Le troisième cas est un homme de corpulence maigre (voir figure 4.24). Là aussi, une zone commune atteignant 23x20 centimètres est bien décrite par une ensemble de points 3D homogènes sans zones d'ombre. les données ToF détectent un déplacement surfacique atteignant 15 mm dans la partie inférieure du champ de vue, soit bien plus que dans les données scanner qui détectent un déplacement maximal de 8 mm. L'hypothèse la plus probable est un mauvais alignement des champs de vue, le bas du champ de vue scanner correspondrait à 5cm au dessus de la limite inférieure du champ de vue de la caméra. Cet alignement est rendu difficile par le peu de relief thoracique du patient, compliquant la mise en correspondance des surfaces. Une deuxième explication peut provenir d'une erreur lors de la sélection des données en fonction de la phase respiratoire. On peut en effet imaginer une situation où le cycle respiratoire sélectionné est d'amplitude plus grande que le cycle respiratoire moyen après reconstruction des données TDM (Ce cycle moyen retranscrivant une moyenne sur 99 secondes, soit 15 à 25 cycles respiratoires). Ainsi, l'inspiration complète est plus profonde avec les données de la caméra, entraînant un mouvement plus ample.

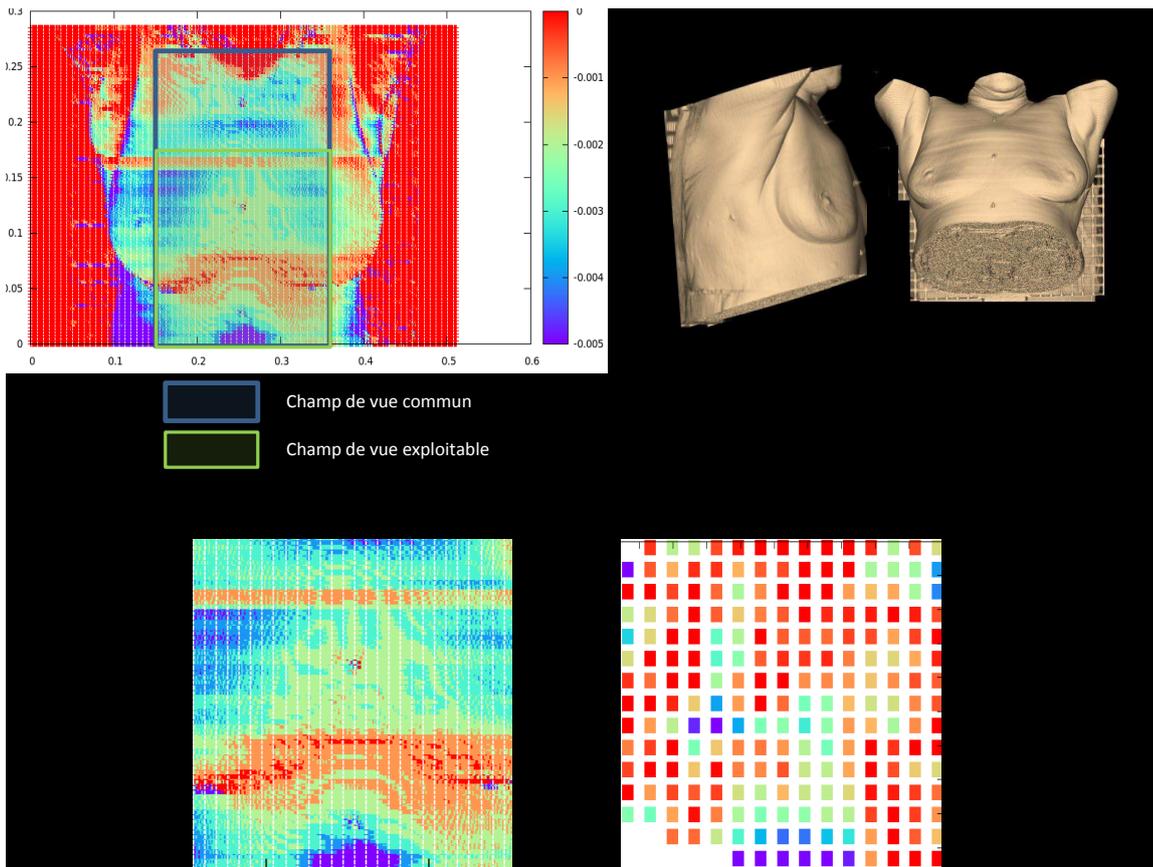


FIGURE 4.22 – Illustration du cas 1 présenté d’une patiente avec une forte poitrine. En haut à droite, la représentation 3D de la surface du volume TDM. En haut à gauche, une carte de couleur représentant la différence en profondeur entre inspiration et expiration complète. Les deux cartes du bas montrent pour la ROI exploitée, les résultats à partir des données TDM (gauche) et de la caméra ToF (droite).

Analyse quantitative

La comparaison quantitative est effectuée en observant le mouvement de régions d’intérêt identiques pour chaque modalité. Comme expliqué section 4.3.2, la région anatomique contenue dans l’acquisition TDM dynamique est limitée à 30 cm dans la direction tête-pieds en partant de la coupole supérieure des poumons. Pratiquement, pour la majorité des patients, seule la partie inférieure de la zone observée est touchée par un mouvement respiratoire, puisque la respiration abdominale est quasi automatique en position allongée. C’est pour cette raison que les ROIs étudiées le sont dans les 5 à 10 derniers centimètres inférieurs du champ de vue, et ceci pour différentes tailles et localisations (Voir les figures 4.22, 4.23 et 4.24 pour la répartition du mouvement selon les localisations anatomiques). Une deuxième raison découlant des observations de la section précédente 4.3.5, est la difficulté à détecter de manière correcte le déplacement dû à la respiration dans la région thoracique de par la

4.3. Comparaison avec un examen 4D-TDM comme vérité terrain

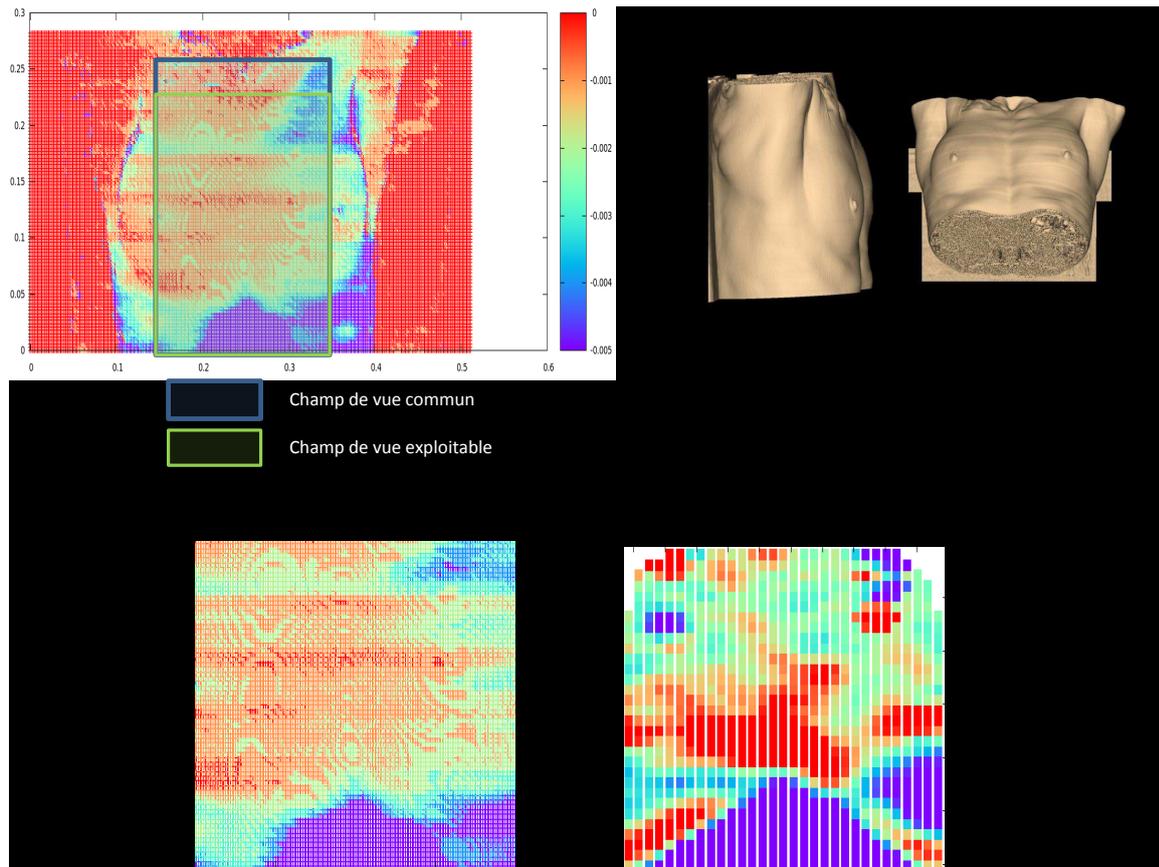


FIGURE 4.23 – Illustration du cas 2 présenté d'une patiente avec une poitrine de taille standard. En haut à droite, la représentation 3D de la surface du volume TDM. En haut à gauche, une carte de couleur représentant la différence en profondeur entre inspiration et expiration complète. Les deux cartes du bas montrent pour la ROI exploitée, les résultats à partir des données TDM (gauche) et de la caméra ToF (droite).

position de la caméra.

Les ROIs considérées sont de taille $4 \times 3 \text{ cm}^2$, $8 \times 3 \text{ cm}^2$ et $20 \times 3 \text{ cm}^2$ dans la partie inférieure du champ de vue (voir figure 4.25). Pour chaque région, la profondeur moyenne des points de mesures ou de contrôle lui appartenant est calculée pour les 22 phases d'un cycle respiratoire. Le mouvement respiratoire est alors composé de 21 éléments, soit la différence entre la position moyenne des 21 dernières phases par rapport à la première, à expiration complète. La figure 4.26 est un exemple du signal respiratoire pour la ROI la plus large et un patient au hasard. Le mouvement extrait des points de contrôle de la ROI après filtrage gaussien montre une erreur moyenne de $0.22 \pm 0.14 \text{ mm}$ sur 21 estimations par rapport à la première acquisition à expiration complète. L'erreur minimale est à 40% de l'inspiration (0.01 mm) et maximale à 80% de l'inspiration (0.42 mm).

La figure 4.27 montre l'erreur moyenne sur les 21 instants respiratoires par rapport à celui

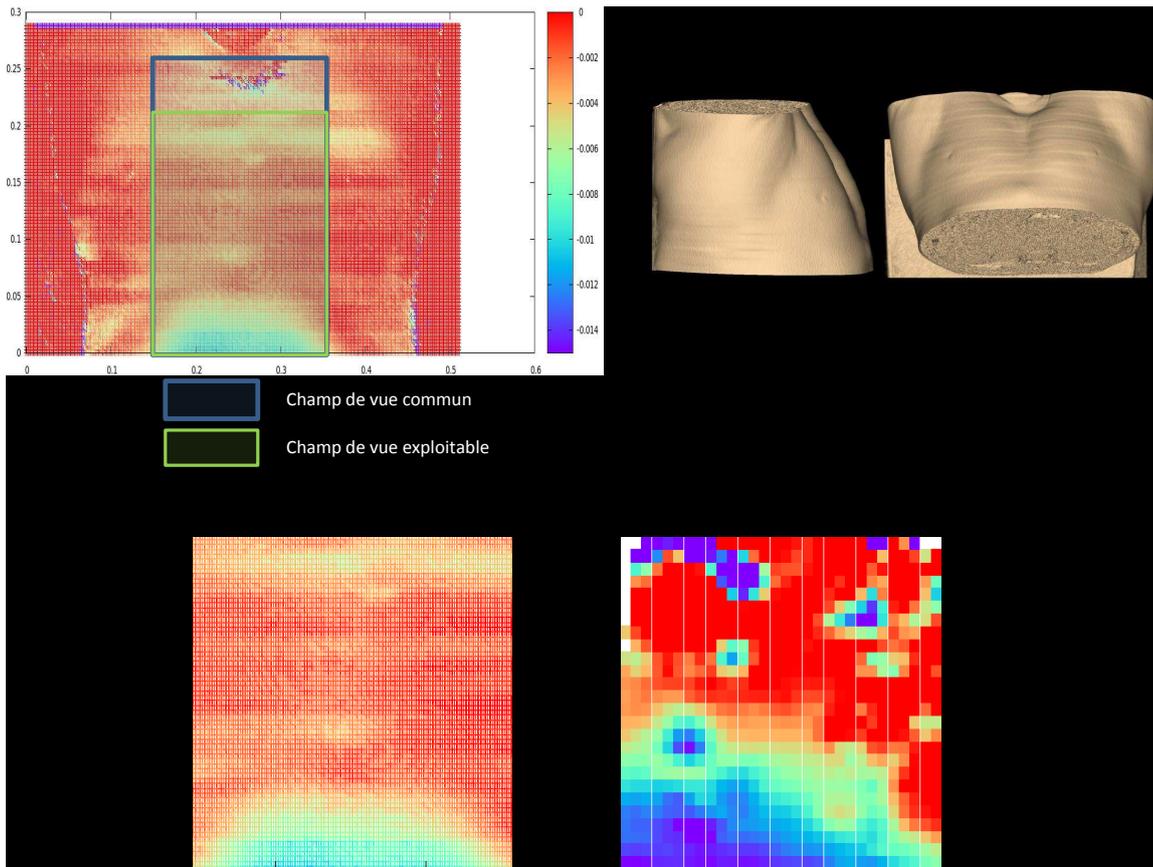


FIGURE 4.24 – Illustration du cas 3 présenté d'un patient avec une morphologie maigre. En haut à droite, la représentation 3D de la surface du volume TDM. En haut à gauche, une carte de couleur représentant la différence en profondeur entre inspiration et expiration complète. Les deux cartes du bas montrent pour la ROI exploitée, les résultats à partir des données TDM (gauche) et de la caméra ToF (droite).

initial, pour 7 patients, et chaque taille de ROIs. D'un point de vue général, l'erreur moyenne est de 0.33 ± 0.26 mm pour la région de taille 4×3 cm², 0.35 ± 0.24 mm pour celle de taille intermédiaire et 0.39 ± 0.28 mm pour la zone la plus grande. Dans le détail, trois types d'erreur ressortent de l'analyse de ces 7 cas. Nous avons tout d'abord les trois premiers patients dont l'erreur moyenne sur l'ensemble du cycle respiratoire ne dépasse pas 0.3 mm quelque soit le patient ou la taille de la région. De plus, l'erreur moyenne augmente légèrement avec la diminution de la taille de la région, de 0.23 mm pour la région 20×3 cm² à 0.26 pour la région 4×3 cm². Lorsque l'on trace les signaux, on voit une très bonne corrélation entre les signaux "TDM" et les signaux "ToF" tant en amplitude qu'en phase, comme on peut le voir sur la figure 4.26. Nous avons ensuite les patients 4 et 6 qui présentent une erreur moyenne plus élevée, quelque soit la taille de la ROI. Le tracé des courbes montre pour l'un une sous estimation du signal due à une mauvaise mise en correspondance spatiale des données, comme

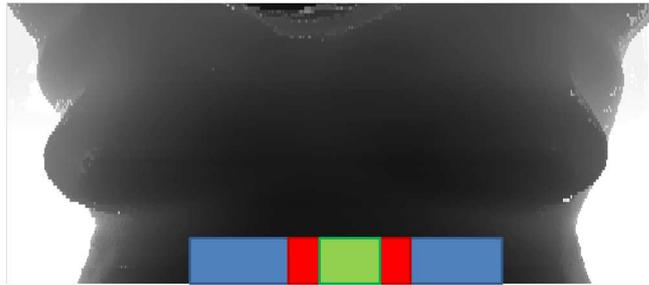


FIGURE 4.25 – Les différentes ROIs utilisées pour estimer le mouvement surfacique ont une taille allant de 4x3 cm² (vert) à 20x3 cm² (bleu).

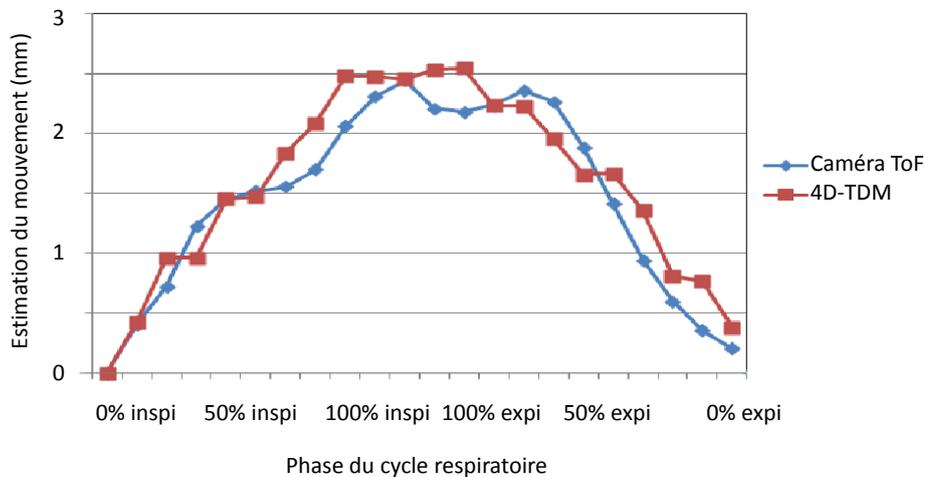


FIGURE 4.26 – Mouvement d’une ROI de 20x3 cm³ lors d’un cycle respiratoire, estimé à partir des données d’une caméra ToF et d’un Scanner 4DCT.

illustré pour le troisième cas d’études dans la section précédente, et pour l’autre un décalage de phase dû à un mauvais binning des données. Les deux derniers patients étudiés (5 et 7) présentent une erreur plus élevée quand la région d’intérêt est la plus large. Ce résultat étonnant est expliqué par des artefacts en bord de ROI qui augmentent fortement l’erreur.

4.4 Prédiction du mouvement interne

Nous avons considéré un patient et cherché à générer les volumes 4D-TDM pour ensuite analyser la distance entre la localisation d’un point caractéristique sur le vrai TDM et celui généré. Pour ceci, un professionnel médical localise 13 points caractéristiques décrivant les poumons, du haut des coupoles au diaphragme. Ces points sont illustrés figure 4.28. L’erreur de génération des volumes TDM est alors la moyenne des distances pour chacun de ces 13 points. Le modèle de respiration patient-spécifique utilisé dans ce travail est celui développé par Fayad et al. (2009). Il nécessite une estimation du mouvement surfacique ex-

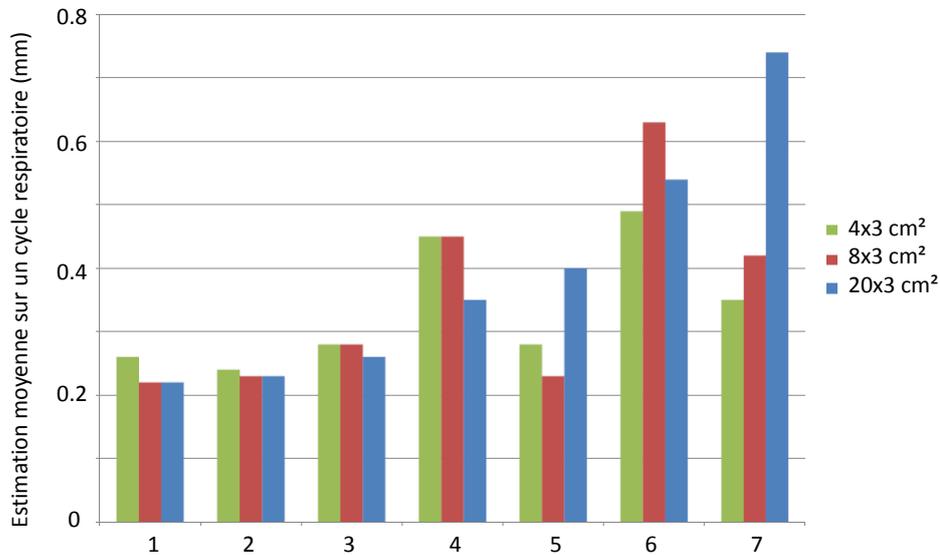


FIGURE 4.27 – Différence moyenne d’estimation du mouvement par une caméra ToF par rapport à celui extrait des données TDM pour un cycle respiratoire complet, pour 7 patients et 3 tailles de ROIs

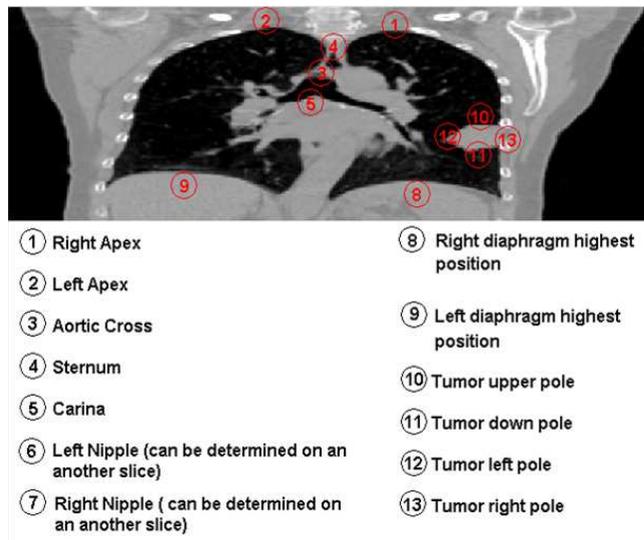


FIGURE 4.28 – Les 13 points d’intérêt sont des points caractéristiques des poumons. Leur trajectoire est déterminée par analyse des 22 volumes TDM par un expert médical.

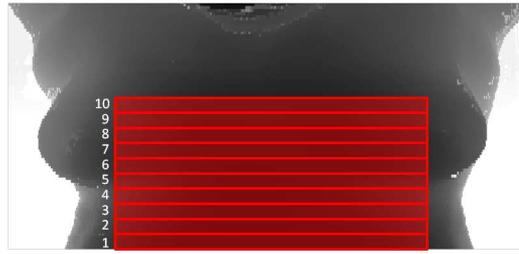


FIGURE 4.29 – Les 10 régions servant à l'extraction des 10 signaux décrivant le mouvement externe surfacique sont des bandes horizontales mesurant $2 \times 20 \text{ cm}^2$.

terne, cette estimation pouvant prendre n'importe quelle forme, allant d'un signal issu d'un spiromètre au déplacement de capteurs 3D fixés sur la peau. Dans un premier temps, nous avons généré ces volumes TDM en créant le modèle patient spécifique selon le principe du "leave one out", soit la génération du modèle à partir des 22 phases moins celle générée. Nous avons pour ceci considéré le mouvement externe comme constitué du signal issu de 10 régions d'intérêt vues chacune comme une bande de $2 \times 20 \text{ cm}^2$ (Figure 4.29). Nous avons ainsi atteint une erreur moyenne de 3.45 mm sur deux phases générées, à 40% en inspiration et 70% en expiration. Afin d'avoir un point de comparaison, nous avons généré simultanément les volumes TDM à partir des surfaces segmentées des données scanner. Dans ce cas là, l'erreur moyenne est de 2.61 mm. Enfin, nous avons généré ces même volumes en appliquant prosaïquement une interpolation linéaire : ainsi, le volume TDM correspondant à 60% inspiration est déduit du volume à 0% inspiration combiné avec 60% du champ de déformation entre les phases 0% inspiration et 100% inspiration. L'erreur moyenne est alors de 6.15 mm.

4.5 Synthèse et conclusion

L'objectif de ce travail était de rendre utilisable une caméra Temps-de-Vol comme instrument de mesure du mouvement respiratoire externe en temps réel. Partant du constat que les données brutes étaient trop bruitées pour caractériser un mouvement souvent inférieur au centimètre, nous avons proposé la modélisation des surfaces avec des fonctions B-splines. Ce choix est justifié par une première raison qui est la réduction du bruit apportée par l'intensité de lissage associée à la fonction décrivant la surface. De plus, ce choix algorithmique nous permet de proposer une méthode simple d'extraction du mouvement surfacique par le déplacement des points de contrôles au fil des acquisitions. Enfin, le temps de calcul de la modélisation est déjà compatible avec une analyse en temps réel grâce au choix de l'algorithme MBA qui autorise la modélisation d'une surface composée de plusieurs milliers de points en moins de 100 millisecondes.

Les résultats sur fantôme confirment et complètent l'état de l'art existant sur l'influence de la distance et de l'angle d'observation sur la qualité de mesure. On peut définir un cadre pour

observer une surface dans un environnement encombré : l'angle d'observation ne doit pas dépasser 50° pour éviter la dégradation de la mesure. De même, cette dégradation est proportionnelle à la distance d'observation jusqu'à 1 m. Au-delà, l'erreur de mesure est stable, jusqu'à 2 m si l'on prend en compte les expérimentations du chapitre 3. Cependant, la distance d'observation influe sur la densité de points décrivant la même surface, limitant ainsi les possibilités de traitement des données.

La méthode d'extraction du mouvement 1D surfacique est basée sur le mouvement des points de contrôle décrivant la surface B-Splines, ce qui permet une extraction du mouvement local à une vitesse élevée, seule la modélisation prenant du temps. Elle est validée sur patient par comparaison du mouvement surfacique avec des surfaces issues de scanners 4D-TDM. Les résultats montrent une forte correspondance qualitative par rapport aux zones anatomiques soumises au mouvement. Les comparaisons quantitatives confirment la capacité de la caméra ToF à suivre un mouvement respiratoire 1D le long de la surface thoraco-abdominale en temps réel. Cependant, des artefacts le long de la surface apparaissent, peut-être causés par le relief de la surface obligeant la caméra à observer des zones de manière quasi tangentielle. Il reste à trouver une solution limitant ces effets, et une de ces solutions serait l'utilisation de deux caméras positionnées de manière à permettre l'observation double des régions d'intérêt : associée à l'estimation de la qualité de la mesure (basée sur le bruit sur un pixel au fil du temps), des régions susceptibles de produire des artefacts ne seront observées que par la caméra ToF la mieux située.

L'utilisation du modèle respiratoire patient-spécifique préalablement développé au laboratoire afin de prédire le mouvement tumoral est effectuée en considérant le mouvement externe comme la surface entière du patient. Les résultats sur les volumes TDM générés montrent une erreur moyenne plus élevée qu'en cas d'utilisation des surfaces externes segmentées depuis les volumes TDM (3.45 mm contre 2.61 mm). Ce résultat n'est pas surprenant dans le sens où les données TDM étant quasi parfaites, observer un mouvement externe de nature similaire ne peut que fournir des résultats moins bons en l'absence d'amélioration au niveau de la description de ce mouvement. De même les résultats sont meilleurs que dans le cas où les volumes TDM sont générés de manière naïve, linéairement (3.45 mm contre 6.15 mm), montrant un apport dû à une bonne prise en compte de la complexité du mouvement surfacique externe par 10 régions d'intérêt.

Ce travail apporte la preuve de la capacité de la caméra TOF à suivre le mouvement respiratoire externe malgré une technologie limitant la précision de mesure par rapport aux autres scanners 3D. Les expérimentations sur humains sont concluantes malgré des conditions d'observation clairement limitées par le milieu clinique, le positionnement des caméras n'étant pas optimal. Enfin, l'impact sur la prédiction du mouvement respiratoire est montré. Cependant, la déformation de la surface n'est décrite que par un champ de vecteur 1D se déplaçant selon l'axe z de profondeur. La déformation est concrètement décrite, mais nous n'avons pas

accès à l'information de déformation locale 3D. C'est pourquoi le prochain chapitre présentera une méthode afin d'estimer la déformation 3D de la surface à haute vitesse en utilisant les propriétés de la caméra ToF.

Chapitre 5

Fusion des informations 2D et 3D pour l'estimation en temps réel du mouvement surfacique 3D

Sommaire

5.1	Le flot optique	122
5.2	Le recalage élastique	124
5.3	Description de la méthodologie	127
5.4	Evaluation de la méthode	128
5.4.1	Tests effectués	128
5.4.2	Respect de la contrainte d'illumination constante	130
5.4.3	Estimation du déplacement 2D	134
5.4.4	Estimation du déplacement 3D	138
5.4.5	Comparaison avec un recalage élastique	139
5.5	Synthèse et conclusion	140

Ce chapitre présente une méthode permettant de déterminer la déformation 3D d'une surface avec un faible temps de calcul. Ce développement est motivé par la complexité d'extraire la déformation d'une surface composée de données éparses : c'est généralement le recalage élastique à partir des nuages de points non-appareillés qui s'impose, avec un temps de calcul inhérent difficilement compatible à l'heure actuelle avec une application en temps réel. Cette méthode fusionne les deux types d'informations fournies par une caméra ToF, une image d'intensité 2D de la scène et un nuage de points 3D, afin de contraindre la déformation 3D à l'aide du flot optique déterminé sur les images d'intensité. Après présentation du flot optique et du recalage élastique de données non-appareillées, nous décrirons la méthodologie, puis présenterons la validation effectuée ainsi que les avantages et limitations actuelles

de la méthode.

5.1 Le flot optique

Le flot optique décrit le déplacement de pixels d'une image à une autre de manière dynamique, et fournit ainsi une estimation du mouvement réel des objets décrits par ces pixels. Les hypothèses sous-jacentes permettant une telle analyse sont :

- La contrainte d'illumination constante : un objet ou point caractéristique garde la même intensité d'une image à l'autre.
- La contrainte de lissage : les objets à identifier forment des ensembles connexes de vitesse homogène, ce qui produit un champ de vecteurs de déplacements lisse.

Rechercher entre deux images le déplacement de pixels peut s'écrire sous la forme :

$$I(x + \omega, t + 1) - I(x, t) = 0 \quad (5.1)$$

où $\omega = [ux, uy]^t$ est le vecteur représentant le déplacement du pixel $I(x, t)$ à l'instant $t + 1$. Lorsque le problème est posé sous forme différentielle, il prend la forme de l'équation :

$$\frac{dI(x(t), y(t), t)}{dt} = \frac{\delta I}{\delta x} \cdot \frac{dx}{dt} + \frac{\delta I}{\delta y} \cdot \frac{dy}{dt} + \frac{\delta I}{\delta t} = 0 \quad (5.2)$$

qui devient l'équation du flot optique :

$$(\nabla I)^t \omega + I_t = 0 \quad (5.3)$$

avec $(\nabla I) = [I_x, I_y]^t$ le gradient spatial de l'intensité lumineuse, I_t la dérivée temporelle et ω la vitesse recherchée. Cependant, il n'est pas possible de déterminer ω sans poser des contraintes sur le système, c'est pourquoi il existe plusieurs approches qui proposent différents types de contraintes.

Les méthodes variationnelles cherchent à minimiser une fonctionnelle regroupant l'équation du flot optique et une contrainte sur la composante de lissage du champ de déplacement. Cette contrainte se traduit généralement par une contrainte de régularisation sur le gradient

ou le laplacien du champ de vitesse. Une résolution globale de cette fonctionnelle est développée par Horn and Schunck (1981) avec la minimisation de la fonctionnelle

$$J = \iint ((\nabla I)^t \omega + I_t)^2 + \alpha((\nabla v_x)^2 + (\nabla v_y)^2) dx dy \quad (5.4)$$

avec le terme de régularisation contraignant les vitesses voisines à être similaires. Malgré d'autres propositions de régularisation (Beauchemin and Barron (1995)), cette approche ne convient qu'à des mouvements très lisses à travers l'image car l'application globale empêche la bonne détection des petits mouvements locaux. Elle est aussi réputée pour le bruit associé au flot optique et son faible temps de calcul.

Une approche variationnelle avec application locale extrêmement utilisée est celle proposée par Lucas et al. (1981) qui considère que sur un voisinage spatial restreint Ω la vitesse locale est constante, ce qui permet de minimiser la fonctionnelle

$$J = \sum_{\Omega} W^2 [\nabla I \cdot \vec{\omega} + I_t]^2 \quad (5.5)$$

où W est une fenêtre locale. Ce type d'approche est réputé moins sensible au bruit et pouvant détecter les mouvements locaux. Cependant le temps de calcul est sensiblement plus long de part l'application locale des contraintes, même si l'algorithme est hautement parallélisable.

Une autre approche très populaire cherche à faire correspondre pour un bloc de taille donnée de l'image, le bloc équivalent sur l'image suivante, la distance séparant les localisations représentant le déplacement (Barjatya, 2004). La recherche du bloc optimal passe par la minimisation d'un critère d'erreur qui décrit en fait une relation de corrélation qui peut être de nature variée : corrélation simple (5.6), somme de carrés des différences (5.7) ou somme des différences absolues (5.8).

$$\int_{\omega} f_1(x + v_x, y + v_y) f_2(x, y) dx dy \quad (5.6)$$

$$SSD = \sum \sum (I(i, j) - J(i + u, j + v))^2 \quad (5.7)$$

$$SAD = \sum \sum |I(i, j) - J(i + u, j + v)| \quad (5.8)$$

Avec I étant les pixels du bloc de référence et J les pixels du bloc comparé. Cette approche ne convient pas à la détection de grands déplacements car la recherche de blocs identiques est souvent limitée à une certaine région voisine afin de ne pas avoir à balayer l'ensemble de blocs de pixels possibles et ainsi réduire le temps de calcul. Cependant, cet algorithme est

lui aussi hautement parallélisable.

Il existe enfin une approche fréquentielle qui cherche à filtrer les transformées de Fourier à l'aide de filtres spatio-temporels qui vont sélectionner des fréquences pouvant être retranscrite en une vitesse dans le domaine temporel (Heeger (1988), Wu et al. (1998)). Ces méthodes présentent le désavantage de perdre de l'information par les opérations de filtrages effectuées. De plus, le paramétrage des filtres rend compliquée leur utilisation.

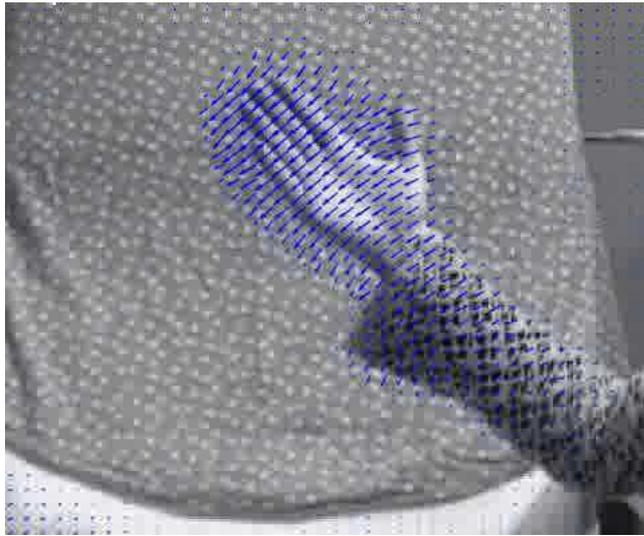


FIGURE 5.1 – Le flot optique fournit un vecteur de déplacement pour chaque pixel de l'image initiale. La contrainte de continuité permet d'éviter des aberrations en partant de l'hypothèse que deux pixels voisins ne peuvent pas avoir des déplacements opposés.

5.2 Le recalage élastique

La déformation de la surface d'un patient pendant sa respiration est de nature élastique, c'est à dire qu'elle n'est pas identique partout, c'est pourquoi son estimation doit être locale. Au-delà du caractère non-rigide du mouvement, cette déformation n'est pas non plus affine (C'est-à-dire une transformation rigide avec une homothétie). Nous avons ici affaire à un mouvement élastique, ce qui implique une estimation locale, assignant à chaque point 3D une nouvelle position après transformation.

Classiquement, ce recalage nécessite la minimisation d'une fonctionnelle qui est composée de deux termes :

- une mesure de similarité entre les deux nuages
- une régularisation qui empêche toute déformation excessive et irréaliste

La résolution est itérative et nécessite généralement une initialisation afin d'éviter les minimaux locaux qui peuvent empêcher la bonne résolution du recalage. Comme on recherche une solution locale pour chaque point, le temps de calcul devient vite très important comparé à un recalage rigide ou affine. C'est pourquoi quelques études cherchent à améliorer le processus afin de réduire les temps de calcul.

Une première approche a été de partir d'un algorithme de recalage rigide, l'ICP (ou Iterative Closest Points, Besl and McKay (1992)) complété d'une déformation locale afin d'aboutir à un recalage élastique. L'ICP est adapté au recalage de vues partielles d'un objet, lorsque les correspondances entre les points des nuages de référence et cible ne sont pas connues, ainsi que dans le cas où les mesures sont nombreuses. Cet ensemble de propriétés en font déjà un algorithme parfaitement adapté aux scanners 3D qui fournissent un nombre élevé de points décrivant une surface, et ceci sans correspondance entre les nuages de points acquis successivement. Feldmar and Ayache (1996) sont les premiers à proposer de partir de l'ICP et déformer localement à partir de celui-ci. Un premier recalage rigide est effectué puis une transformation affine est déterminée pour des sous-ensembles sphériques de la surface source. La régularisation est réalisée en lissant les transformations rigides proches spatialement. B Allen and Popovic (2003) régularise en considérant les sommets de la surface qui sont connectés plutôt que sur un critère de distance puis Amberg et al. (2007) montre qu'il est possible d'appliquer cette méthode suivant une optimisation par les moindres carrés. Schaeffer et al. (2012) ajoute des optimisations algorithmiques et effectue une validation sur 5 patients dont la surface est observée par le système AlignRT qui utilise de la lumière structurée (voir section 2.2.3). La validation est fournie par la comparaison des vecteurs de déformations issus du recalage des surfaces avec le mouvement réel de marqueurs optiques visibles et sélectionnés manuellement sur l'image d'intensité 2D de lumière structurée. Les résultats indiquent une bonne estimation du mouvement surfacique avec une erreur médiane de 1.6 mm. Cependant, des disparités sont montrées avec en particulier une corrélation entre l'erreur d'estimation du mouvement et son amplitude selon les directions S-I et L-R. Le temps de calcul associé est proportionnel linéairement au nombre de points : pour des nuages de points allant de 8 500 à 14 000 éléments, une itération dure de 6.1 à 11.9 sec, et les résultats présentés nécessitent 25 itérations.

Bauer et al. (2012a) propose lui l'observation surfacique avec un système expérimental de triangulation laser actif (AT) qui renvoie un quadrillage de points selon le projecteur. Ce faisant, il propose un modèle variationnel combinant une surface haute résolution fournie par la lumière structurée avec la surface moins résolue fournie par l'AT. Après recalage de la surface haute résolution sur celle basse résolution, il propose par résolution inverse de déterminer pour chaque acquisition basse résolution le mouvement en haute résolution. Les tests

sur fantôme sont très prometteurs mais la validation sur humain est en fait réalisée sur des observations par lumière structurée retravaillées pour ressembler géométriquement aux données AT, et ceci à cause d'interférences lors de la projection simultanée de lumière structurée et de AT. Le temps de calcul est ici estimé à 2.6 ± 0.7 sec.

Globalement, ces méthodes cherchent à optimiser le recalage élastique afin d'approcher des temps de calcul compatibles avec l'application en temps réel. Il est à remarquer que la migration de ces algorithmes sur carte GPU pourrait améliorer significativement le temps de calcul : Nickolls and Dally (2010) passe en revue les différents gains de temps dans de nombreux domaines où le temps de calcul reste une limitation. L'intérêt de cette migration est un gain potentiel au niveau du temps de calcul grâce à la capacité d'une carte graphique à effectuer des calculs en parallèle. Cependant, ce type de calcul rend l'accès à la mémoire plus coûteux qu'un calcul sur CPU : au final, le gain apporté par l'utilisation de cartes GPU dépend du niveau de parallélisation de l'algorithme, mais aussi de facteurs tels que la nécessité pour l'algorithme d'accéder aux données.

Table 3. Representative CUDA application coprocessing speedups.

Application	Field	Speedup
Two-electron repulsion integral ¹²	Quantum chemistry	130×
Gromacs ¹³	Molecular dynamics	137×
Lattice Boltzmann ¹⁴	3D computational fluid dynamics (CFD)	100×
Euler solver ¹⁵	3D CFD	16×
Lattice quantum chromodynamics ¹⁶	Quantum physics	10×
Multigrid finite element method and partial differential equation solver ¹⁷	Finite element analysis	27×
N-body physics ¹⁸	Astrophysics	100×
Protein multiple sequence alignment ¹⁹	Bioinformatics	36×
Image contour detection ²⁰	Computer vision	130×
Portable media converter*	Consumer video	20×
Large vocabulary speech recognition ²¹	Human interaction	9×
Iterative image reconstruction ²²	Computed tomography	130×
Matlab accelerator**	Computational modeling	100×

* Elemental Technologies, Badaboom media converter, 2009; <http://badaboomit.com>.
 ** Accelereyes, Jacket GPU engine for Matlab, 2009; <http://www.accelereyes.com>.

FIGURE 5.2 – Illustration extraite de Nickolls and Dally (2010) qui récapitule les gains en temps de calcul atteignables par transfert sur GPU en fonction du type de traitement à effectuer. Le gain potentiel varie fortement, allant de 9 pour de la reconnaissance vocale à 137 pour une application de simulation moléculaire.

5.3 Description de la méthodologie

La méthode d'estimation du mouvement 3D d'une surface décrite par des points irréguliers propose la fusion des deux types d'information fournies par une caméra ToF, soit une image 2D infrarouge de la scène et un nuage de points 3D. Le point clef est que ces deux jeux de données sont reliés spatialement : chaque capteur CCD de la caméra transforme la lumière infrarouge en une information de profondeur par estimation du temps de vol, mais aussi convertit la quantité de lumière captée en une intensité lumineuse en niveau de gris, comme une caméra CCD classique. Il en résulte que pour un pixel donné de l'image 2D, un point 3D dans l'espace lui est associé. La méthode proposée fusionne ces informations dans le sens où l'analyse de l'image niveau de gris va permettre d'estimer la déformation latérale de la surface, élément manquant de l'estimation du mouvement par les points de contrôle des surfaces B-splines présentées précédemment (voir section 4.1.3).

Soit une surface S observée à deux instants différents. Le scanner 3D observe les configurations S_1 et S_2 avec pour chacune le nuage de points 3D épars P_1 et P_2 . Pour chaque élément de P_1 appartenant à S_1 , nous avons à déterminer sa position sur S_2 . Comme un scanner 3D fournit un ensemble de points 3D, il est hautement improbable de trouver la position de P_1 sur S_2 dans l'ensemble de points P_2 . Nous proposons donc tout d'abord de modéliser S_1 et S_2 avec des fonctions B-Splines afin d'avoir une description continue de ces surfaces. Cette modélisation est à nouveau effectuée par l'algorithme MBA de Lee et al. (1997). De plus, le chapitre précédent a mis en évidence la diminution de l'erreur d'estimation de mouvement d'un facteur 3 lorsqu'elle est effectuée à l'aide des points de contrôle (4.3.5).

Puisque nous avons maintenant une description continue des deux surfaces, une mise en correspondance est envisageable. Cependant, la méthode utilisant les points de contrôle présentée section 4.1.3 permet seulement de caractériser le mouvement selon l'axe z : il manque une estimation du mouvement latéral du point de référence appartenant à S_1 lors de la déformation. Cette estimation est apportée par le flot optique reliant les deux images en niveaux de gris montrant l'intensité lumineuse associée à la scène où S_1 et S_2 existent. Le flot optique va indiquer le déplacement de chaque pixel selon le plan latéral de la caméra : ce mouvement 2D est traduit comme mouvement 3D du point associé au pixel sur les deux images grâce au pont entre 2D et 3D fourni par la géométrie de la caméra qui associe à chaque pixel une ligne de tir 3D le long duquel son point 3D associé sera situé. Après mise en correspondance avec les points 3D associés, nous avons directement le vecteur 3D associé à ce mouvement, et ceci de manière extrêmement rapide. Il y a cependant une limitation due à la résolution latérale : on estime en effet qu'à un mètre, chaque pixel adjacent est séparé en 3D par 3 à 4 mm. Si le flot optique indique un déplacement correspondant à un nombre entier, la nouvelle position du point de référence sur la deuxième image correspondra exactement à un point 3D mesuré. Si ce n'est pas le cas (ce qui arrive souvent), il manque une estimation de cette

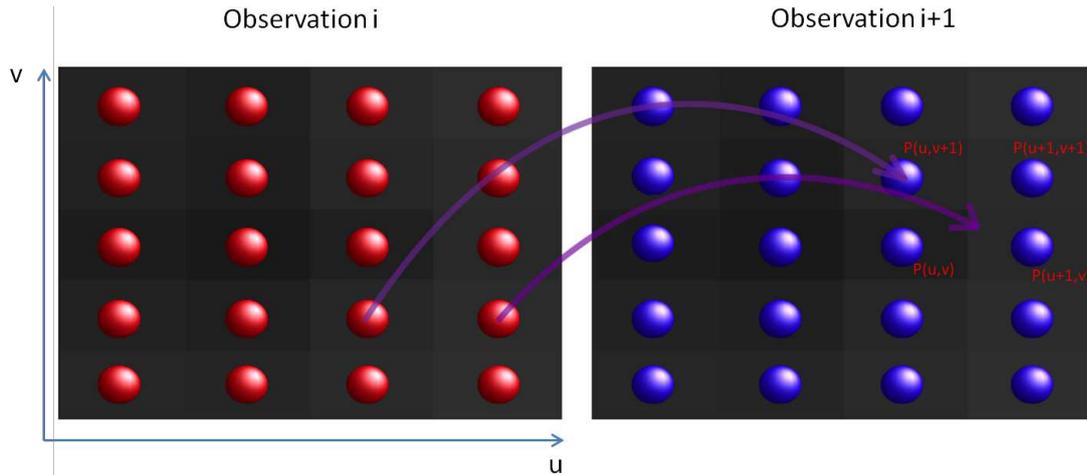


FIGURE 5.3 – Le pixel $P(u, v-1)$ est analysé comme ayant bougé à la localisation $(u, v+1)$ entre les observations i et $i+1$: son mouvement en 3D est le vecteur formé par les points 3D correspondants $P(u, v-1)$ et $P(u, v+1)$. De même, le pixel $P(u-1, v-1)$ s'est déplacé à une position non-définie à l'instant $i+1$: le point 3D correspondant à cette localisation est déterminé à l'aide de la fonction B-splines décrivant la surface, à l'emplacement ayant comme coordonnées x et y interpolé des 4 voisins les plus proches.

position 3D non connue dans les données mesurées (voir figure 5.3) : nous le déterminons à l'aide de la fonction B-Splines décrivant S_2 , en considérant l'emplacement ayant comme coordonnées (x, y, z) avec (x, y) déterminées comme l'interpolation linéaire des i points 3D voisins ayant comme coordonnées (x_i, y_i) .

La nature de la déformation latérale due à la respiration s'adapte a priori assez bien au flot optique puisqu'elle est lisse, avec lors de l'inspiration, un mouvement général dans la direction caudale, et inversement lors de l'expiration. Le respect de la contrainte d'illumination constante est à contrôler, ce qui sera fait ultérieurement.

5.4 Évaluation de la méthode

5.4.1 Tests effectués

La première série de tests effectués cherche à établir l'impact de la méthode d'estimation du flot optique sur, d'une part, la précision et la robustesse, et de l'autre le temps de calcul associé. Nous avons pour ceci utilisé deux méthodes variationnelles parmi les plus utilisées, avec d'une part la méthode proposée par Horn and Schunck (1981), de l'autre celle proposée par Lucas et al. (1981). Nous avons écarté les autres approches pour des raisons de compatibilité avec les données : a priori, les méthodes basées sur le block-matching ne devraient pas être efficaces sur une surfaces se déformant latéralement, et perdant donc leur

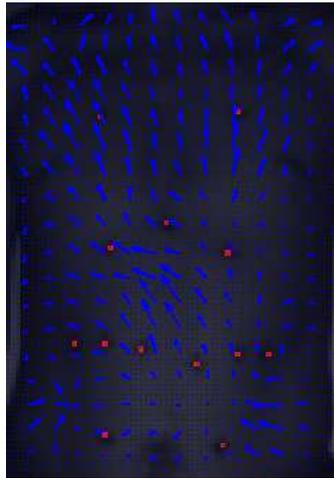


FIGURE 5.4 – Illustration du champ de déplacement estimé par flot optique entre inspiration complète et expiration complète. En rouge, la localisation des marqueurs dessinés sur le patient.

géométrie au fil des acquisitions. De même, les méthodes basées sur le filtrage fréquentiel ont tendance à perdre de l'information à cause des multiples filtrages, ce qui peut poser problème sur des données bruitées et peu résolues comme celles fournies par la caméra ToF. Nous avons donc testé notre méthode en utilisant les deux types de méthodes variationnelles, se différenciant par l'application de la contrainte de régularisation, globale ou locale. La vérité terrain est fournie par le déplacement de croix noires dessinées sur les volontaires dont le déplacement est déterminé manuellement.

Dans un second temps, nous avons étudié la précision atteinte pour l'estimation du déplacement 3D de ces marqueurs optiques en considérant la vérité terrain comme la distance entre les deux points 3D associés aux deux pixels décrivant le déplacement 2D. L'estimation de ce déplacement 3D est donc fournie par la méthode proposée, avec une évaluation pour les deux types d'estimation du flot optique. Nous avons ajouté une troisième méthode, celle présentée en 2012 par Bauer et al. (2012c), qui propose l'utilisation du flot optique pour contraindre l'estimation de la déformation 3D en interpolant la nouvelle position à partir des plus proches voisins. Cette méthode est évaluée sur des données issues de la Kinect, soit un ensemble très dense de points 3D avec un très faible bruit associé, ce qui rend envisageable une simple interpolation. Dans le cas de données ToF, cette méthode pourrait être moins efficace car les données sont moins denses et bien plus bruitées.

Dans un troisième temps, nous avons comparé notre méthode avec un recalage élastique des nuages de points. Le recalage ne fournissant pas de résultats cohérents sur les données issues de la caméra ToF, nous avons effectué cette comparaison sur des données issues de la Kinect.

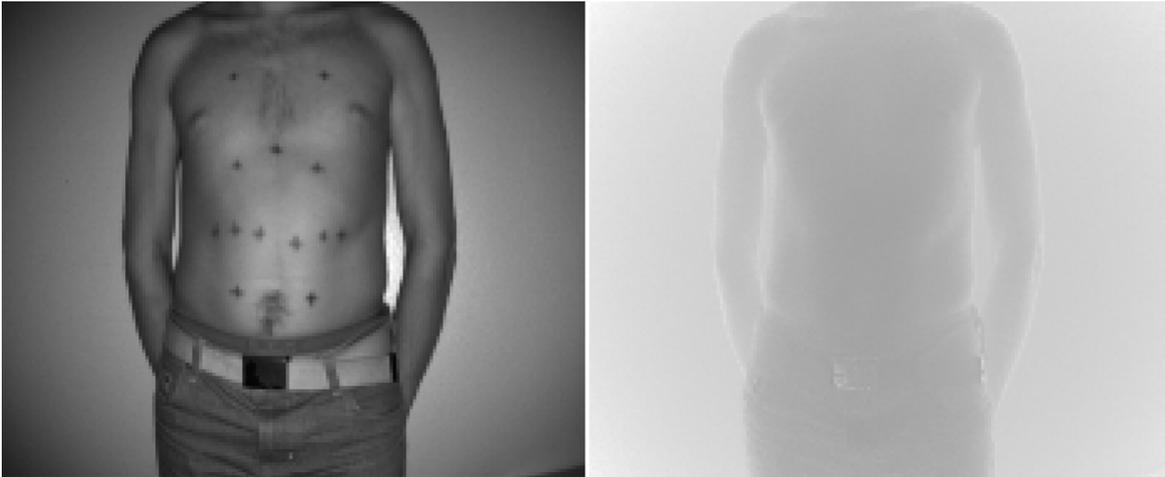


FIGURE 5.5 – Une même scène permet d’avoir deux images : celle de gauche représente l’intensité lumineuse alors que celle de droite représente la carte de profondeur convertie en niveaux de gris.

5.4.2 Respect de la contrainte d’illumination constante

La contrainte d’illumination constante pour la bonne estimation du flot optique stipule que le pixel dont on cherche à connaître le déplacement doit être d’intensité équivalente sur toutes les images. Nous avons à notre disposition deux types d’images en niveaux de gris : l’image de profondeur convertie en niveau de gris et l’image d’intensité lumineuse indiquant la quantité de lumière recue par chaque pixel (Figure 5.5). Nous avons donc cherché à vérifier le respect de la contrainte d’illumination constante en cas d’utilisation des deux types d’image.

La première hypothèse est l’utilisation de la carte de distance convertie en image niveau de gris. Un test synthétique est effectué où un nuage de points 3D en forme de pyramide est créé, puis converti en image de niveau de gris (figure 5.6a). Ce nuage est ensuite translaté de 1cm selon z, ce qui correspond à un mouvement nul selon les axes x et y, puis converti en une image (figure 5.6b). La recherche du flot optique entre les deux images synthétiques montre au contraire un fort mouvement latéral (figure 5.6c) dans le sens de la pente, ce mouvement étant fantôme puisque la surface n’a pas de mouvement latéral. La raison est que la contrainte d’illumination constante est brisée lorsqu’un mouvement selon z est appliqué puisque le niveau de gris évolue justement en fonction de la profondeur associée. Ce que voit en réalité le flot optique, c’est un mouvement fantôme où un pixel est vu décalé puisque c’est son voisin qui va avoir le même niveau de gris (voir figure 5.7).

Pour savoir s’il est envisageable d’utiliser les images d’intensité, il faut identifier les raisons qui pourraient briser la contrainte d’intensité constante. La principale raison est que lorsque la poitrine et l’abdomen sont soulevés, imaginons sans déplacement latéral, le même

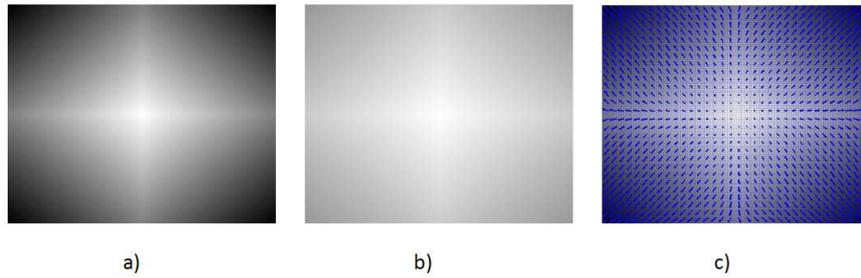


FIGURE 5.6 – Illustration de la pyramide simulée pour test, avec la vue selon la direction z. En c), le flot optique entre les deux images.

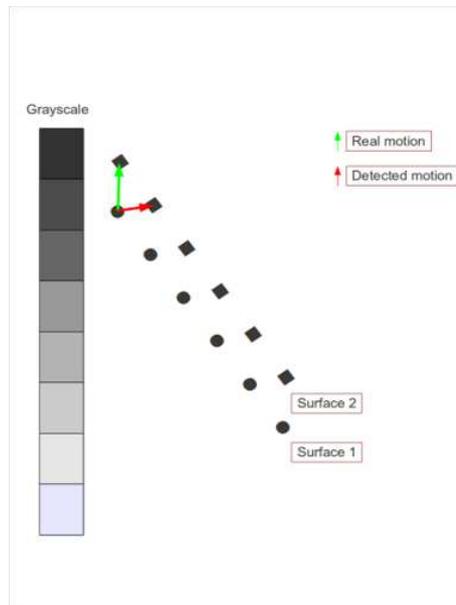


FIGURE 5.7 – La détection d’un mouvement fantôme par le flot optique vient du fait qu’un pixel qui se déplace selon z change de niveau de gris. La flot optique le détecte donc comme ayant bougé latéralement, vers le pixel voisin qui lui, prend la valeur initiale de niveau de gris.

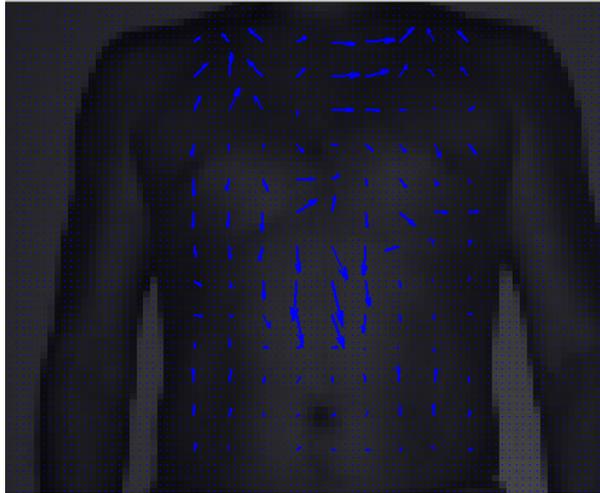


FIGURE 5.8 – Flot optique entre les images de la scène avant et après translation de 4 mm vers la surface. La taille des vecteurs est normalisée.

point anatomique pourrait voir son intensité changer, ce qui briserait la contrainte. On verrait à ce moment là un mouvement fantôme similaire à celui entrevu sur les images de distances, soit un champ de vecteurs décrivant le déplacement d'une pente vers l'extérieur. De plus, un deuxième effet pourrait être géométrique : lorsque la surface s'approche de la caméra, elle est vue comme plus grande en 2D, ce qui créerait un mouvement fantôme vers l'extérieur du champ de vue puisqu'un point surfacique n'ayant qu'un mouvement en profondeur sera détecté par un autre capteur à mesure qu'il se rapproche. Au final, si la contrainte n'est pas respectée, le flot optique devrait être visible au niveau des régions de plus forte pente, soit dans notre cas d'analyse au niveau des seins et des abdominaux. Il peut aussi être dirigé vers l'extérieur du champ de vue.

Afin de tester l'apparition de ces effets, nous avons procédé à une expérimentation lors de laquelle un patient situé à 1m de la caméra ne respire pas, et ceci collé à un mur afin de limiter les mouvements involontaires. La caméra est fixée à l'actionneur linéaire, puis décalée trois fois par pas de 4 mm vers le patient. Le flot optique est appliqué sur l'ensemble du torse et de l'abdomen, soit 3312 pixels. Si l'on observe la figure 5.8 lorsque la caméra est déplacée de 4 mm, on voit que le flot optique ne semble pas dessiner de schéma particulier qui pourrait révéler la présence d'un des deux phénomènes précités. Les statistiques sur les vecteurs de déplacement montrent que l'ensemble est centré (moyenne de 0.1 ± 0.3 pixel en u et 0 ± 0.4 en v) avec respectivement 90% des valeurs entre -0.3 et 0.6 en u et -0.5 et 0.7 en v . La distribution des valeurs entre -1 et 1 pixel est visible figure 5.9.

Les chiffres pour les déplacements de 8 et 12 mm indiquent un bruit du flot optique qui augmente fortement, surtout dans la direction v (Figures 5.10 et 5.11). Cependant, aucun schéma particulier n'apparaît sur les champs de vue, il semble bien que ce qui est détecté soit le bruit associé au flot optique, de plus dans des conditions d'observation qui n'assurent

5.4. Evaluation de la méthode

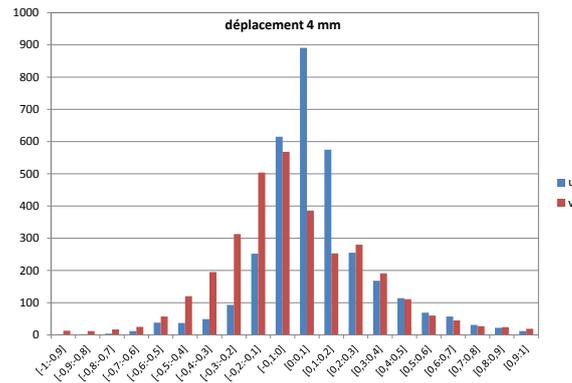


FIGURE 5.9 – Distribution des composantes des vecteurs de déplacement selon les directions u et v entre -1 et 1 pixel. La translation est de 4 mm.

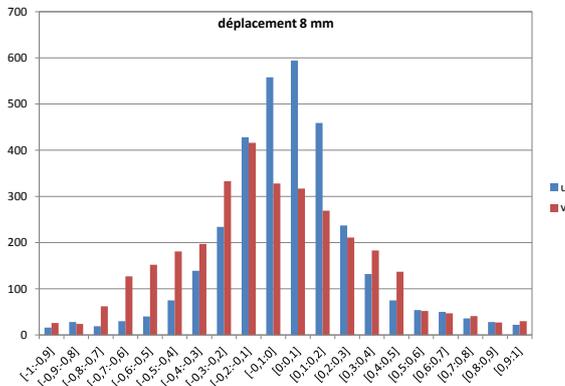


FIGURE 5.10 – Distribution des composantes des vecteurs de déplacement selon les directions u et v entre -1 et 1 pixel. La translation est de 8 mm.

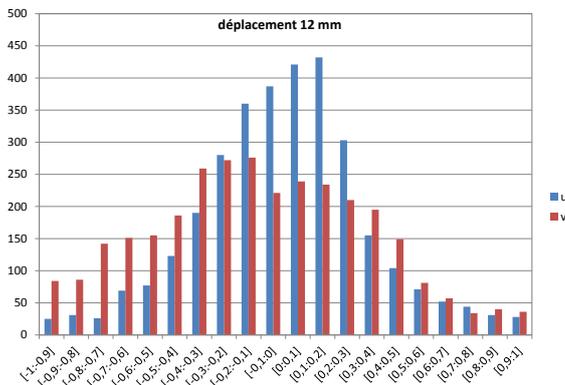


FIGURE 5.11 – Distribution des composantes des vecteurs de déplacement selon les directions u et v entre -1 et 1 pixel. La translation est de 12 mm.

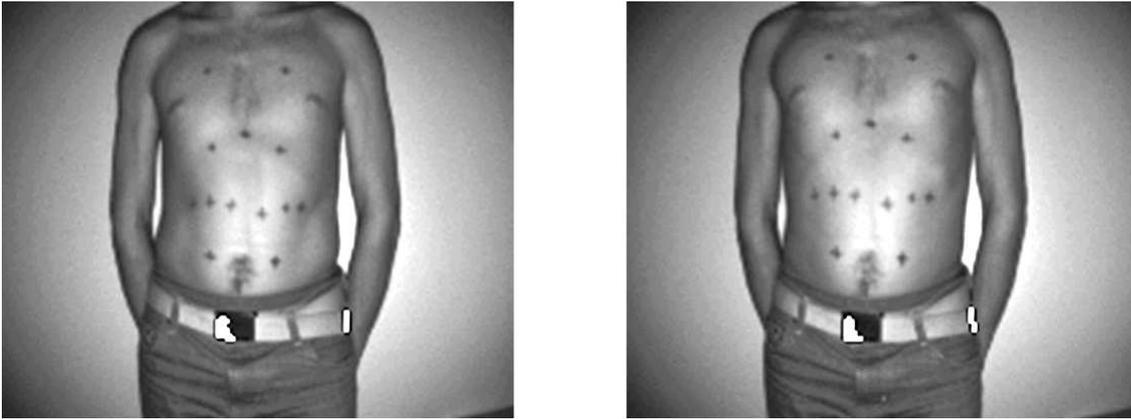


FIGURE 5.12 – Deux images d'un même patient à expiration et inspiration complète.

pas une déformation de la surface réellement nulle. Le déplacement de 8 mm induit un déplacement moyen selon u (respectivement v) de 0 ± 0.3 pixel (0 ± 0.5) avec 90% des valeurs comprises entre -0.5 et 0.6 pixels (-0.7 et 0.9). Pour le déplacement de 12 mm, le mouvement estimé augmente encore avec un déplacement moyen selon u (respectivement v) de 0 ± 0.4 pixel (-0.1 ± 0.6) avec 90% des valeurs comprises entre -0.6 et 0.8 pixel (-0.8 et 1).

Cette expérience nous indique tout d'abord que la contrainte d'illumination constante semble respectée lorsque l'on utilise les images d'intensité pour estimer le flot optique, ou du moins que si elle est rompue, son effet est quasi-indétectable. De même, l'effet géométrique dû à l'avancée de la surface dans le champ de vue n'est pas détectable non plus, en tout cas à 1m d'observation. Cependant, le bruit associé à la mesure d'un flot optique sensé être nul grandit fortement lorsque la translation atteint 8 mm puis 12 mm.

5.4.3 Estimation du déplacement 2D

Afin de valider la capacité du flot optique à estimer le mouvement surfacique avec le type de données fournies (images d'intensité basse résolution), nous avons marqué sur le thorax de volontaires des croix noires qui font office de marqueurs et sont visibles sur l'image d'intensité (voir figure 5.12). Le déplacement de ces marqueurs est estimé manuellement en observant les images. Il est à signaler que de part la faible résolution des pixels (3-4 mm), il y a une incertitude associée à la localisation de ces marqueurs. Mais le fait de valider le flot optique sur ces mêmes données bruitées rend la comparaison homogène et valide. Les 2 images sélectionnées sont situées à inspiration et expiration complète. Au total, 10 cas d'études issus de 3 volontaires sont observés à un mètre avec un nombre de marqueurs visuels allant de 9 à 13. Dans un premier temps, nous avons utilisé la méthode variationnelle de Horn and Schunck (1981) sur les données. Nous avons tout d'abord comparé le déplacement entre une expiration et une inspiration complète lors d'une respiration forcée. Les résultats montrant une sous-estimation systématique du mouvement proportionnelle à son amplitude,

Im(0) / Im(1/3)						
n°	Δu	Δv	Δu (OF)	Δv (OF)	ϵu	ϵv
1	1	2	0,7	1,6	0,3	0,4
2	0	1	0	1,4	0	-0,4
3	0	1	0,7	1	-0,7	0
4	1	1	0,9	1	0,1	0
5	0	1	0,3	0,9	-0,3	0,1
6	1	0	1	0,3	0	-0,3
7	1	1	0,9	0,5	0,1	0,5
8	1	1	0,8	1,1	0,2	-0,1
9	1	1	0,5	0,8	0,5	0,2
10	0	1	0,2	0,9	-0,2	0,1
11	1	1	0,4	1,3	0,6	-0,3
12	0	0	0,7	-0,3	-0,7	0,3
13	1	0	0,5	-0,2	0,5	0,2

moyenne	0,03	0,05
moyenne (abs)	0,32	0,22

FIGURE 5.13 – Erreur d'estimation du flot optique entre Im(0) et Im(1/3).

nous avons donc observé le mouvement à différentes étapes de cette inspiration. En prenant comme référence l'image à expiration complète, nommée Im(0), l'évaluation du mouvement est faite par rapport à différentes images jusqu'à Im(1) correspondant à expiration maximale. Nous avons aussi évalué le mouvement Im(0)/Im(1/3) et Im(0)/Im(2/3).

Les résultats présentés figures 5.13, 5.14 et 5.15 montrent une forte tendance à la dégradation de l'estimation de mouvement lorsque l'amplitude devient forte, qui peut être analysée comme suit (En considérant l'amplitude de déplacement comme une distance de Manhattan, plus adaptée pour appréhender un déplacement ayant deux composantes orthogonales) :

- lorsque l'amplitude est inférieure à 4 pixels, l'estimation est de très bonne qualité, l'erreur est centrée, ce qui indique l'absence de biais (zone verte)
- pour les amplitudes autour de 5 pixels, l'estimation se dégrade, avec la présence d'estimations de bonnes qualité, et des outliers qui amplifient l'erreur moyenne : ces outliers ne sont a priori pas signés (zone orange).
- pour les amplitudes plus fortes, le mouvement est globalement sous-estimé, l'erreur moyenne augmente fortement (zone rouge).

L'explication la plus probable est que le contrecoup de la simplicité algorithmique de la méthode de Horn and Schunck (1981) (et donc sa vitesse de calcul) est qu'il n'arrive pas à détecter des déplacements trop larges. C'est pourquoi dans un deuxième temps, nous avons utilisé l'algorithme de Lucas et al. (1981) sur les mêmes données. Entre Im(0) et Im(1), l'erreur diminue à 1.4 ± 1 pixels selon u et 1.3 ± 1.1 pixels selon v (contre 2.2 ± 1.2 pixels selon u et 2.5 ± 1.6 pixels selon v). De plus, le biais sur les estimations disparaît avec une erreur moyenne de -0.2 pixel selon u et 0.2 pixels selon v. Entre les images Im(0) et Im(1/3), l'erreur absolue moyenne augmente au contraire, atteignant 0.45 ± 0.2 pixel selon u et 0.3 ± 0.2

Im(0) / Im(2/3)						
n°	Δu	Δv	Δu (OF)	Δv (OF)	ϵu	ϵv
1	1	5	0,8	3,7	0,2	1,3
2	-1	4	1,5	2,33	-2,5	1,67
3	0	3	2,4	3,8	-2,4	-0,8
4	0	4	2,8	1,6	-2,8	2,4
5	-1	4	0,5	4,5	-1,5	-0,5
6	1	3	2,4	5,1	-1,4	-2,1
7	0	3	0,6	3,5	-0,6	-0,5
8	0	3	0	3,9	0	-0,9
9	0	4	0,2	5	-0,2	-1
10	-1	4	-4	3,8	3	0,2
11	-1	4	-1,6	6,9	0,6	-2,9
12	-1	1	-0,3	0,8	-0,7	0,2
13	0	1	-0,6	1,2	0,6	-0,2

moyenne	-0,59	-0,24
moyenne (abs)	1,26	1,12

FIGURE 5.14 – Erreur d'estimation du flot optique entre Im(0) et Im(2/3).

Im(0) / Im(1)						
n°	Δu	Δv	Δu (OF)	Δv (OF)	ϵu	ϵv
1	1	8	0,1	5,3	0,9	2,7
2	-1	7	1,5	4,1	-2,5	2,9
3	0	6	3,3	4,5	-3,3	1,5
4	1	7	3,7	2,7	-2,7	4,3
5	-1	7	-0,4	6,1	-0,6	0,9
6	1	5	5,2	1,2	-4,2	3,8
7	0	6	2,8	0,4	-2,8	5,6
8	0	6	1,6	3,3	-1,6	2,7
9	0	6	1,8	4	-1,8	2
10	-1	6	-5	2,4	4	3,6
11	-1	6	-2,4	5,8	1,4	0,2
12	-1	2	0,8	2,5	-1,8	-0,5
13	0	2	-0,5	3,5	0,5	-1,5

moyenne	-1,11	2,16
moyenne (abs)	2,16	2,47

FIGURE 5.15 – Erreur d'estimation du flot optique entre Im(0) et Im(1).

Amplitude max (pixel)	Horn&Schunck	Lukas&Kanade
3	0.27	0.37
3	0.29	0.37
3	0.3	0.4
3	0.34	0.44
4	0.42	0.65
4	0.75	0.75
5	1.2	1.1
6	1.05	0.95
6	1.2	1.2
6	1.25	0.85
8	2.45	1.6
9	2.35	1.95

FIGURE 5.16 – Moyenne des erreurs d’estimation des composantes u et v du flot optique pour 10 cas d’étude et deux méthodes d’estimation, basées sur Horn and Schunck (1981) et Lucas et al. (1981), en fonction du déplacement maximal (au sens de la distance de Manhattan). En vert, le meilleur résultat pour chaque cas.

pixel selon v .

Pour y voir plus clair, nous avons étudié 10 cas distincts (les trois précédents inclus) avec des amplitudes de déplacement maximal allant de trois à neuf pixels, et calculé la moyenne des erreurs pour les deux composantes, u et v (Voir figure 5.16). On y voit surligné en vert le meilleur résultat pour chacun des dix cas. L’influence de l’amplitude de déplacement est visible : en dessous de 4 pixels d’amplitude maximale, la méthode de Horn and Schunck (1981) est la plus performante, la méthode de Lucas et al. (1981) étant plus performante au delà. Cependant, l’erreur moyenne augmente dans les deux cas lorsque l’amplitude augmente. Ces résultats semblent indiquer que l’algorithme de Horn and Schunck (1981) est plus précis que celui de Lucas et al. (1981) sur le type d’images fournies par la caméra ToF lorsque le déplacement est de faible envergure. Cependant, les déplacements plus larges sont mieux estimés avec le deuxième algorithme car il ne les sous-estime pas systématiquement.

La figure 5.17 présente la différence d’estimation du mouvement surfacique 2D pour 5 cas étudiés avec 3 volontaires différents positionnés à un mètre, ce qui représente 53 marqueurs en tout. Nous étudions des couples d’images avec un mouvement maximal de 4 pixels. Lorsque l’on considère les valeurs absolues d’erreur, les résultats indiquent une différence

	Horn & Schunck		Lukas & Kanade	
	ϵ_u (abs)	ϵ_v (abs)	ϵ_u (abs)	ϵ_v (abs)
cas 1	0.3±0.3	0.2±0.2	0.5±0.3	0.3±0.3
cas 2	0.2±0.2	0.4±0.4	0.3±0.3	0.5±0.5
cas 3	0.2±0.3	0.4±0.3	0.3±0.4	0.5±0.4
cas 4	0.3±0.3	0.3±0.3	0.5±0.5	0.4±0.3
cas 5	0.4±0.4	0.5±0.4	0.7±0.5	0.6±0.3

FIGURE 5.17 – Erreur moyenne d’estimation absolue du flot optique pour 5 cas d’étude et deux méthodes d’estimation, basées sur Horn and Schunck (1981) et Lucas et al. (1981), pour les cinq cas où le déplacement maximal est inférieur à 4 pixels.

d’estimation pour la direction u comprise entre $0,2\pm 0,2$ pixel et $0,4\pm 0,4$ pixel (respectivement $0,3\pm 0,3$ pixel et $0,7\pm 0,6$ pixel), et entre $0,2\pm 0,2$ et $0,56\pm 0,3$ pixel (respectivement $0,3\pm 0,3$ pixel et $0,6\pm 0,3$ pixel) pour la direction v . Quelque soit le cas et la direction, la méthode basée sur l’approche de Horn and Schunck (1981) est la plus performante.

5.4.4 Estimation du déplacement 3D

A partir des cinq cas recensés ayant un déplacement maximal de quatre pixels, nous appliquons la méthodologie proposée afin d’estimer le déplacement 3D des marqueurs optiques. Nous avons, de plus, appliqué la méthodologie proposée par Bauer et al. (2012c) sur le même jeu de données. Un filtrage gaussien est appliqué aux données 3D. Il est à noter que le résultat des deux méthodologies est intimement lié puisque la détermination du flot optique est nécessaire dans les deux cas. Nous avons donc pris en compte les mêmes estimations de ce flot pour l’utilisation des deux méthodes. La différence de résultat n’est alors la conséquence que de la détermination de la nouvelle position 3D, véritable différence entre notre approche et celle de Bauer et al. (2012c).

Les résultats sont indiqués dans le tableau 5.18 comme la distance moyenne entre l’estimation de la nouvelle position du point 3D initialement lié à chaque marqueur optique et sa position réelle. La meilleure estimation est dans le cas où la méthodologie présentée ici est utilisée avec le flot optique déterminé par la méthode de Horn and Schunck (1981). Cependant, la différence quantitative avec la méthode de Bauer et al. (2012c) n’est pas significative. On retrouve le même schéma lorsque le flot optique est obtenu par l’algorithme de Lucas et al. (1981), avec cependant des erreurs moyennes bien supérieures, ce qui est logique au vu

des résultats sur l'estimation des déplacements 2D (Figure 5.17).

	Horn&Schunck	Bauer (Horn&Schunck)	Lucas&Kanade	Bauer (Lucas&Kanade)
cas 1	1.5±1.3	1.7±1.4	2.3±2	2.2±2.3
cas 2	1.4±1.1	1.3±1.1	2.6±2.6	2.4±2.6
cas 3	1.1±0.9	1.3±0.9	2.8±3.1	3±3.1
cas 4	1.7±1.7	1.8±1.7	2.7±2.4	2.6±2.6
cas 5	2.8±1.5	2.9±1.7	3.1±2.4	3.2±2.4

FIGURE 5.18 – Erreur moyenne d'estimation en millimètres du mouvement 3D des marqueurs optiques pour les cinq cas dont le déplacement maximal est inférieur à 4 pixels.

Le temps de calcul associé à cette méthode sur un PC Pentium 4500 Bicore est composé des différents éléments suivis :

- Temps de calcul du flot optique : 0.2 sec sur l'image entière 176x144 avec la méthode de Horn and Schunck (1981), 3.5 sec avec celle de Lucas et al. (1981).
- Temps de calcul de la modélisation B-Splines : 0.2sec.
- Pré-filtrage des images 2D (réduction du bruit associé) : 0.1sec.

5.4.5 Comparaison avec un recalage élastique

Les sections précédentes ont cherché à valider la méthodologie présentée avec l'aide d'une vérité terrain fournie par des marqueurs dessinés directement sur la peau. Cependant, comme expliqué précédemment, la méthode classique d'extraction de la déformation tridimensionnelle à partir de nuages de points dépareillés est d'effectuer un recalage élastique "analytique", soit seulement à partir des nuages de point. C'est pourquoi nous avons voulu comparer cette méthodologie avec une telle approche. L'algorithme de recalage élastique de données éparses est ici fourni par Myronenko and Song (2010). L'algorithme "Point Coherent Drift" (CPD) est basé sur une approche probabilistique qui va considérer le lien entre deux jeux de données comme un problème d'estimation de densité de probabilité. Nous avons tout d'abord effectué le recalage élastique sur des acquisitions ToF, mais le résultat semblait incohérent. Cette mauvaise analyse peut s'expliquer par le bruit associé aux données de la caméra (du moins relativement aux autres scanners 3D utilisés dans les utilisations de cet algorithme dans la littérature). Nous avons donc effectué la validation sur des données issues de la Kinect, ce qui assure un niveau de bruit très faible si l'on observe la scène à

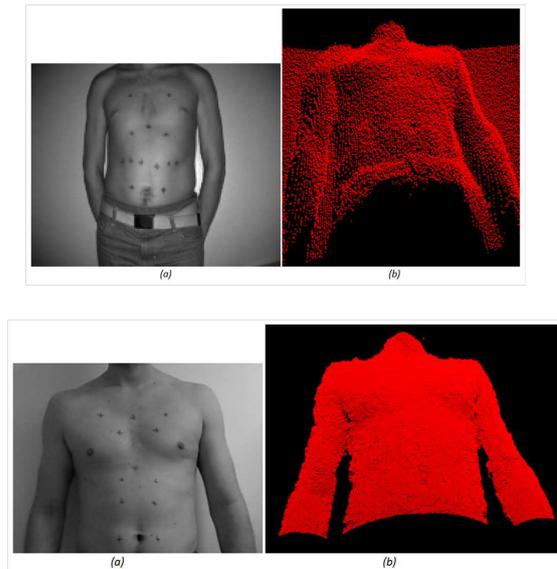


FIGURE 5.19 – Illustration des données fournies par une caméra ToF (en haut) et une Kinect (en bas) avec : a) une image d'intensité et b) un nuage de points 3D. L'image d'intensité de la Kinect est bien plus résolue que celle de la caméra ToF (640*480 contre 176*144) ce qui améliore la précision de la détection du mouvement des marqueurs.

distance raisonnable (cette influence de la distance sur le mouvement d'une surface observée par une caméra Kinect est discutée au chapitre 3).

La figure 5.20 illustre le flot optique sur l'image d'intensité de la Kinect. La figure 5.21 montre les résultats détaillés avec pour chaque marqueur optique, soit au total 12 déplacements, le vrai déplacement ainsi que l'erreur d'estimation par la méthode proposée et par recalage élastique. Le recalage élastique surestime quasi systématiquement le mouvement réel dans la direction tête-pieds, avec une différence atteignant en moyenne 1.6 mm, et un seul marqueur ayant le déplacement cranio-caudal sous-estimé. Dans le même temps, la méthode proposée offre une estimation moins biaisée avec une différence moyenne de 0.3 mm en x, -0.4 mm en y et 0.6 mm en z. La moyenne des valeurs absolues d'erreur montre une estimation du déplacement 3D significativement meilleure pour la méthodologie présentée : erreur moyenne de 0.5 et 0.6 mm selon x et y contre 1.2 et 1.9 mm. Lorsque l'on considère la distance euclidienne du vecteur de déplacement 3D, la moyenne des erreurs absolues est de 0.7 mm. Cette même moyenne atteint 1.3 mm quand la PCD est utilisée.

5.5 Synthèse et conclusion

Ce chapitre avait pour objectif de proposer une méthodologie permettant l'estimation de la déformation 3D d'une surface. Nous avons à l'esprit de tirer profit de la nature complémentaire de l'information 2D et 3D fournie par une caméra ToF afin de résoudre la difficulté

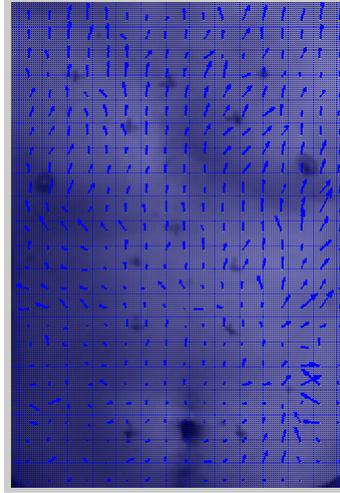


FIGURE 5.20 – Illustration du champ de déplacement estimé par flot optique entre inspiration complète et expiration complète.

	Déplacement des marqueurs (mm)			vs. Flot optique (mm)			vs. Recalage élastique (mm)		
	x	y	z	Δx	Δy	Δz	Δx	Δy	Δz
1	-0.4	-3	-5.8	-0.1	-0.3	-0.1	-1.6	-2.4	-2.3
2	1.4	-3	-5.9	-0.8	0.6	2	-3.4	-1.1	1.2
3	0.5	-3	-5.8	0.5	-0.5	0.2	0.8	-2.5	0.3
4	-0.1	-2.7	-3.7	-0.1	-0.2	-0.4	-2.4	1.3	-0.1
5	0.1	-2.7	-1.9	0.6	-0.4	1.4	-0.3	-0.7	-1.4
6	0	-2.5	-1.8	0.2	-0.2	-1.2	-1.1	-0.3	-1.5
7	-0.1	-0.9	-3.6	-0.1	-1.6	1.8	-0.7	-1.6	0.6
8	0.2	-2.2	-3.5	0.3	-0.2	0.9	-0.1	-2	1
9	-0.1	-2	-3.5	-0.1	0.6	1	0	-2	0.7
10	-1.1	-1	-1.7	1.4	-1.3	0.6	1.5	-3.6	-0.3
11	0	0	0	0	-0.3	0.8	-0.9	-2.4	-1.2
12	-1.1	0.4	-1.7	1.5	-0.5	0.5	1.2	-2.3	0.1
	Erreur moyenne (mm)			0.3	-0.4	0.6	-0.6	-1.6	-0.2
	Erreur absolue moyenne (mm)			0.5	0.6	0.9	1.2	1.9	0.9

FIGURE 5.21 – Détail des résultats pour le patient dont le mouvement surfacique est estimé par la fusion des données 2D/3D et l'algorithme de recalage élastique PCD. Les données sont fournies par une caméra Kinect.

d'estimer la déformation latérale d'une surface. Le flot optique sur les images d'intensité estime finalement le déplacement latéral d'une position anatomique, et c'est exactement l'information qui nous manquait pour la méthodologie présentée au chapitre précédent. L'application de cette information dans le domaine des données 3D fournit une estimation du déplacement d'une localisation anatomique au fil des acquisitions : nous arrivons donc à effectuer un recalage élastique des données. Chaque nouvelle position est déduite de la surface B-Splines associée : en effet, la nature éparse des données rend improbable la nouvelle localisation dans le jeu de données. Il est à noter que Bauer et al. (2012c) propose une fusion similaire sur des données issues de la Kinect, sans toutefois valider la méthodologie en profondeur. Dans cette approche, après estimation du flot optique par une approche variationnelle mêlant application locale et globale (Bruhn et al., 2005), la position 3D associée au déplacement 2D est estimée par interpolation bilinéaire, à priori à partir des points 3D de mesure proche. La validation est qualitative, montrant que cette approche surpasse un recalage géométrique présenté par Bauer et al. (2012b) qui évalue une approche variationnelle débruitant les données d'une caméra ToF simultanément à un recalage sur un surface scanner TDM.

Avant d'estimer l'erreur et le temps de calcul associé à la méthode, il a fallu tout d'abord vérifier que l'on pouvait estimer le flot optique sur les images fournies par la caméra ToF. Il était à craindre que, la surface ayant une déformation latérale mais surtout en profondeur, la luminosité associée à un point anatomique varie au fil des acquisitions. Les images correspondant aux cartes de distances ne pouvant être utilisées (l'intensité des pixels est directement proportionnelle à la profondeur), nous avons étudié les images d'intensité. Les schémas de flots optiques caractéristiques de la brisure de la contrainte d'illumination constante ne sont pas retrouvés entre deux images prises sur un volontaire ne respirant pas, avec un déplacement en profondeur de 4,8 et 12 mm. Ces mesures nous ont quand même permis d'estimer le bruit de flot optique en fonction de ce déplacement, même si, vue l'expérience effectuée, rien n'exclut que les mouvements fantômes mesurés ne soient pas des mouvements involontaires.

Pour la validation, la vérité terrain a été considérée comme étant le déplacement de marqueurs visibles sur les images fournies par la caméra, le déplacement étant connu par sélection manuelle. Ainsi, le déplacement fourni selon deux directions, u et v , est comparé à celui déterminé par le flot optique entre les deux images. Dans un premier temps, nous avons comparé les deux méthodes variationnelles les plus répandues, celles de Horn and Schunck (1981) et Lucas et al. (1981), et les avons évalué pour leur capacité à déterminer le flot optique sur des données de la caméra ToF SR4000. Les résultats indiquent tout d'abord une erreur d'estimation fortement corrélée à l'amplitude de déplacement, avec une sous-estimation des déplacements pour la méthode de Horn and Schunck (1981) au-delà de quatre pixels, l'erreur augmentant moins vite avec la méthode de Lucas et al. (1981). L'analyse détaillée des 5 cas où le déplacement maximal d'un marqueur visuel est inférieur à quatre

pixels montre une meilleure estimation de la méthode de Horn and Schunck (1981) quelque soit le cas considéré. L'erreur d'estimation du déplacement 3D des marqueurs visuels fournit logiquement des résultats comparables avec une erreur moyenne de 1.7 mm contre 2.7 mm. Nous en avons profité pour comparer notre méthodologie avec celle présentée par Bauer et al. (2012c). Les résultats indiquent une performance équivalente des deux méthodologies. Un doute subsiste sur cette évaluation dans le sens où Bauer et al. (2012c) propose une interpolation bilinéaire entre les quatre voisins, soit quatre points de mesure étant touché par le bruit de mesure. De même, notre vérité terrain est elle aussi touchée par le bruit puisque le déplacement est considéré comme la distance euclidienne entre le point 3D de départ et le point 3D d'arrivée, correspondant tous les deux aux localisations des marqueurs visuels. Au contraire, notre méthodologie s'affranchit des points de mesure en modélisant la surface : il est alors possible que notre estimation soit plus proche de la réalité que ce que nous considérons comme vérité terrain. Il manque cependant une meilleur vérité terrain qui pourrait être fournie par des marqueurs optiques ou infrarouges, à la condition qu'ils soient détectables sur les images d'intensité fournies par la caméra.

Si l'on considère d'ores et déjà le cas où l'on observe une surface assez rapidement pour que l'amplitude maximale de déplacement latéral soit inférieure à quatre pixels, l'ensemble des opérations nécessaires à l'estimation de la déformation 3D de l'ensemble de la surface est de l'ordre de 500 millisecondes. Ce résultat est intéressant à mettre en parallèle avec les méthodes de recalage élastiques présentées section 5.2 : plus de 150 secondes pour Schaerer et al. (2012) et 2.6 ± 0.7 secondes pour Bauer et al. (2012a). Il reste cependant une amélioration à apporter au niveau de l'estimation du flot optique : Les images 5.4 et 5.20 montrent des artefacts dans le champ de déplacement, au niveau du centre de l'abdomen pour la première et à droite de la région abdominale pour la seconde. De même, des déplacements supérieurs à 4 pixels sont systématiquement sous-estimés. Normalement, les approches variationnelles locales sont sensées être plus robustes aux images bruitées (ce qui est notre cas) et peu texturées. Cependant, la méthode de Lucas et al. (1981) fournit des résultats moins satisfaisants que l'approche globale. Il est possible qu'un réglage plus fin de l'algorithme utilisé, basé sur une approche pyramidale, soit nécessaire. Les deux paramètres sur lesquels on peut jouer sont la taille de la fenêtre W et le nombre de degrés de la pyramide. Quoi qu'il en soit, avec une erreur moyenne de 1.7 mm sur 53 marqueurs visuels, la précision atteinte est déjà satisfaisante, tant au regard de l'état de l'art (Schaerer et al. (2012) avec une erreur médiane de 1.6 mm avec des observations par lumière structurée, Bauer et al. (2012a) avec une erreur moyenne de 0.22 mm sur des données elles aussi issues d'observations par lumière structurée mais sans validation poussée) que de l'algorithme de recalage élastique utilisé pour comparaison, qui fournit une erreur moyenne deux fois plus élevée. Le temps de calcul est lui aussi en cours d'amélioration avec transposition des méthodes d'estimation du flot optique sur carte GPU, ce qui nous fait espérer des gains de temps substantiels et envisager un processus d'ensemble fonctionnant à une vitesse supérieure à 5 Hz.

Chapitre 6

Le repositionnement inter-fraction avec une caméra ToF

Sommaire

6.1	Création d'un système stéréo d'observation : calibrage de 2 caméras ToF	147
6.1.1	Etat de l'art	147
6.1.2	Proposition de méthode et comparaison avec l'état de l'art	148
6.1.3	Expérimentations	152
6.1.4	Impact sur le repositionnement	157
6.2	Prise en compte du mouvement respiratoire	160
6.3	Etude Clinique	167
6.3.1	Objectifs, moyens et déroulement	167
6.3.2	Dispositif d'observation	168
6.3.3	Système avec 1 seule caméra	170
6.3.4	Système avec 2 caméras	171
6.4	Synthèse et conclusion	174

L'optimisation de la procédure clinique effectuant le repositionnement d'un patient entre chaque fraction du traitement de radiothérapie est essentielle pour garantir en bout de chaîne une délivrance optimale du traitement. A l'heure actuelle, le processus mis en place en routine clinique n'est pas entièrement satisfaisant. Les systèmes lasers permettant l'alignement visuel du patient à l'aide de tatouages atteignent une répétabilité de positionnement satisfaisante, quoique soumise à deux gands facteurs d'erreurs : la variabilité intra et inter-opérateur lors de l'alignement visuel mais aussi la répétabilité du positionnement physique du patient. Un exemple concret concerne la situation où le patient doit avoir ses bras au dessus de la tête afin qu'ils ne soient pas sur les trajectoires de tir vers l'abdomen du patient. La

manière dont les bras sont posés dans les systèmes de contention joue sur l'étirement de la peau le long des flancs du patient et peut induire un large déplacement des tatouages latéraux superficiels : la région visée n'est alors potentiellement plus localisée. De plus, ce système se base sur une information surfacique, ce qui amène toujours à une interrogation sur le bon alignement des structures internes. L'alignement avec lasers est alors complété par une imagerie complémentaire afin d'estimer le repositionnement par visualisation des structures internes. La fréquence de ces imageries dépend de la localisation tumorale et peut atteindre une fréquence journalière dans des cas extrêmes. Dans tous les cas, une dose additionnelle est appliquée au patient, ce qui implique une prise en compte lors de la planification du traitement.

De plus, au-delà du repositionnement inter-fraction, les mouvements intra-fractions ne sont pas du tout pris en compte. Leur gestion nécessiterait une observation continue du patient, ce qui exclut une imagerie irradiante. C'est dans ce contexte que les scanners 3D apportent d'ores et déjà des solutions concrètes par l'observation de l'ensemble de la surface à des vitesses élevées. C'est principalement la lumière structurée qui est l'objet de recherche : sa mesure est connue comme peu bruitée, peu sensible aux conditions d'éclairage avec un champ de vue englobant l'ensemble du thorax et de l'abdomen. Cependant, accéder aux nuages de points 3D nécessite une reconstruction des données qui limitent la capacité à acquérir à haute vitesse une surface. Le scanner laser reste lui limité par sa vitesse d'acquisition d'une telle surface.

Dans ce chapitre, nous cherchons à étudier la caméra ToF comme instrument permettant le repositionnement par l'analyse de la surface. Nous avons pour ceci préalablement déterminé deux points clefs :

- La vitesse d'acquisition permet d'envisager un repositionnement prenant en compte la respiration, ouvrant ainsi la voie à une analyse du mouvement intra-fraction. Cependant, la nature bruitée de l'information rend nécessaire une évaluation de la capacité à déterminer un déplacement d'une région anatomique par recalage des surfaces associées aux positions.
- Le champ de vue de la caméra ToF dans les conditions d'une salle de traitement permet difficilement d'avoir une vue complète du tronc humain.

Afin de remédier au problème du champ de vue, nous avons tout d'abord développé une méthode de calibrage de deux caméras ToF afin d'agrandir le champ de vue et ainsi augmenter l'information surfacique disponible. Le cahier des charges de la méthode est de permettre un calibrage précis et rapide afin de permettre une mise en place simple d'un système d'observation stéréo. Cette méthode a été comparée à l'état de l'art, puis des tests ont été effectués afin d'évaluer l'amélioration apportée lors du repositionnement. Nous avons ensuite étudié

6.1. Création d'un système stéréo d'observation : calibrage de 2 caméras ToF

l'impact du mouvement respiratoire sur le repositionnement et proposé la prise en compte de ce mouvement afin d'améliorer la précision du repositionnement. Dans un second temps, une étude en condition clinique sur 18 mois est actuellement en cours avec le service de radiothérapie du CHRU de Brest. Subdivisée en 3 séquences qui correspondent à nos avancées, elle a pour but de mettre en valeur l'apport de deux caméras puis la prise en compte du mouvement respiratoire pour finalement se situer par rapport aux ordres de grandeur de précision dans l'état de l'art.

6.1 Création d'un système stéréoscopique d'observation : calibrage de 2 caméras ToF

6.1.1 Etat de l'art

Au delà de la caractérisation en trois dimensions d'une scène vue en 2D, les systèmes d'observation stéréo permettent l'augmentation des possibilités d'observation d'une surface, de même que la redondance de l'information pour les régions communes. Le processus de calibrage qui permet une correspondance spatiale entre les données des deux caméras doit permettre de déterminer la géométrie épipolaire décrite par la matrice fondamentale (Longuet-Higgins, 1987). Les systèmes d'observation 2D sont généralement calibrés par l'extraction de points d'intérêt communs sur chaque image : la relation entre les positions permet de déterminer une matrice fondamentale par triangulation. De même certains scanners 3D tels que la Kinect et les systèmes de lumière structurées fournissent une image d'intensité qui permet un tel calibrage. Cependant, une telle opération n'est pas faisable dans le cas d'une caméra ToF car l'image d'intensité n'est pas assez résolue pour permettre une détection fine de points d'intérêts : le modèle actuel le plus résolu est composé de 204x204 pixels. Il est donc nécessaire de procéder différemment afin de permettre de calibrer deux caméras ToF dans le même repère spatial. Mure-Dubois et al. (2008) a montré qu'en utilisant la technique classique de détection des coins d'un quadrillage dans les deux vues, le calibrage résultant n'est pas satisfaisant, sans doute à cause de la faible résolution latérale des images d'intensité. Lindner et al. (2007) propose bien de combiner une caméra ToF avec une caméra RGB haute résolution, mais la combinaison de deux caméras de type différent induit la création d'un nouveau système d'observation (Un peu comme dans le cas de la Kinect de Microsoft, où la caméra RGB doit sans doute servir à interpoler les mesures 3D fournies par la lumière structurée).

Cependant, il est possible de déterminer indirectement des points d'intérêt à l'aide des nuages de points 3D fournis par les caméras : par exemple, l'observation d'une sphère par deux caméras permet de déterminer la coordonnée du centre de cette sphère dans les deux vues par minimisation d'une fonction de coût décrivant la géométrie de la sphère. Peu de travaux ont

été publiés sur le sujet, on peut cependant s'arrêter sur deux méthodes qui diffèrent par l'objet détecté. Böller (2010) propose de déplacer une sphère dans le champ de vue commun et de déterminer le centre de celle-ci par les deux caméras : trois estimations suffisent à déterminer la transformation rigide reliant le repère des deux caméras. Mure-Dubois et al. (2008) propose l'observation d'un plan afin de modéliser le nuage de points correspondant et en extraire le vecteur normal associé à celui-ci, la translation étant elle déterminée par sélection manuelle de points d'intérêt dans les deux scènes. Ces deux méthodes ont en commun de tirer partie du point fort de la caméra ToF qui est la grande quantité de points décrivant la sphère ou le plan observé. Cependant elles ne sont pas automatiques : la première nécessite le déplacement de la sphère dans l'espace alors que la deuxième nécessite une intervention de l'utilisateur. C'est donc avec un cahier des charges requérant une méthode automatique, précise et rapide de calibrage que nous avons proposé un calibrage à l'aide d'un cube qui permet le calibrage complet par estimation de la rotation à l'aide des trois faces visibles et de la translation par localisation du coin du cube.

6.1.2 Proposition de méthode et comparaison avec l'état de l'art

Le but du calibrage spatial entre deux caméras, 2D ou 3D, est de détecter dans les deux prises de vue des points d'intérêt avant de déterminer la relation les liant. Nous allons tout d'abord décrire les deux méthodes déjà existantes puis présenter celle que nous avons développé. Après avoir analysé la qualité d'estimation de paramètres clefs que sont les points d'intérêts, nous avons procédé à l'évaluation des trois méthodes en considérant deux espaces, celui dont les points d'intérêt sont issus et l'espace commun élargi.

Méthode	Détection de la région d'intérêt	Points caractéristiques	Estimation de la rotation	Estimation de la translation
Calibrage avec une sphère	Transformée de Hough	Centre de la sphère	ICP	
Calibrage avec un plan	RANSAC	Normale au plan + points d'intérêt sélectionnés manuellement	Alignement des normales sur l'axe z	Minimisation
Calibrage avec un cube	RANSAC	Normale au plan + coins du cube	Méthode d'orientation absolue par quaternions	Translation entre les estimations d'un coin

FIGURE 6.1 – Description des points clefs des méthodes de calibrage

Calibrage avec un plan

Mure-Dubois et al. (2008) propose l'observation d'un plan afin de déterminer la rotation reliant les repères des caméras, la translation étant elle estimée grâce à l'utilisateur qui va sélectionner des points d'intérêt commun sur les vues 2D de la scène. Les points 3D appartenant au plan sont détectés par l'utilisation de l'algorithme de RANSAC (ou Random Sample Consensus algorithm, Fischler and Bolles (1981)). La détection de plan est un problème classique pour lesquelles différentes méthodes ont été proposées, dont la transformation de Hough ou celle de Radon (Bauer and Polthier, 2008). Le très célèbre algorithme de RANSAC estime itérativement les paramètres d'un plan présent dans la scène en sélectionnant aléatoirement un groupe de points : leur plan associé est modélisé, puis classé selon sa pertinence. La grande force de cette méthode est d'être quasi insensible à la présence de nombreux outliers, généralement très présents dans les jeux de données fournies par la caméra ToF. Cette méthode est non-déterministe, ce qui rend incertain la convergence des solutions fournies lors d'acquisitions multiples.

La détermination de la rotation est faite grâce à la méthode d'alignement sur l'axe z des vecteurs normaux des deux estimations d'un même plan. Soit un plan observé depuis deux points de vue V et V' , N_1 et N'_1 les vecteurs normaux associés aux observations π_1 et π'_1 . Ces vecteurs sont alignés avec l'axe z dans le repère monde. Si R est la matrice de rotation reliant N_1 et z , et R' celle entre N'_1 et z , la matrice de rotation R_z qui aligne V et V' est définie par

$$R_z = R' R^{-1} \quad (6.1)$$

La composante rotationnelle finale est obtenue en prenant en compte le vecteur T_n qui décrit la normalisation de l'ensemble de points P' décrivant la scène V' . Au final, les différentes opérations appliquées à P' sont :

- translation par $-T_n$
- rotation par R_z
- translation par T_n
- translation par T , la translation entre P et P'

La transformation finale prend alors la forme :

$$P = MP' = R_z(P' - T_n) + T_n + T \quad (6.2)$$

La composante translation T est elle connue par sélection de 6 à 10 points aisément détectables sur les deux vues 2D de la scène, ce qui revient à utiliser les images 2D comme dans le cas du calibrage classique stéréoscopique.

Calibrage avec une sphère

Böller (2010) propose d'observer une sphère dont on connaît le rayon pour trois positions différentes au centre du champ de vue. La détection des points appartenant à la sphère est faite sur l'image d'intensité grâce à la transformée de Hough qui détecte automatiquement les géométries circulaires et sphériques sur les images 2D. Les points 3D correspondant aux pixels détectés constituent l'ensemble de données à partir desquelles la minimisation sera effectuée.

L'estimation du centre de la sphère à partir d'un nuage de points décrivant une partie de la surface de la sphère réalise la minimisation d'une fonction de coût définie par

$$F = \sum_{i=1}^n (||p_i - C|| - r)^2 \quad (6.3)$$

où $p_i = \{x_i, y_i, z_i\}^t$ est le $i^{\text{ème}}$ point de la sphère dans \mathbb{R}^3 , $C = \{c_x, c_y, c_z\}^t$ le centre de la sphère, r le rayon de la sphère et n le nombre de points. La minimisation est réalisée avec la méthode de Levenberg-Marquardt. Le processus est répété itérativement en enlevant 5% des points considérés comme des outliers et tant que le centre de la sphère se déplace de plus de 0.1 mm. Une fois trois positions effectuées, la correspondance est déterminée par application de l'ICP entre les groupes de points (Besl and McKay, 1992). On estime donc alors simultanément la rotation et la translation.

Calibrage avec un cube

Tirant le bilan des deux approches déjà existantes, nous nous sommes penchés sur l'utilisation d'un cube comme outil de calibrage. Nous voulions tout d'abord avoir une méthode qui détermine séparément la rotation et la translation, et ceci de manière automatique. Le cube nous a semblé parfaitement adapté puisqu'il fournit 3 plans visibles par chaque caméra (soit une information sur la rotation bien plus complète que pour l'utilisation d'un seul plan). De plus, le coin du cube est un point de repère efficace pour déterminer la composante translation, et très simple à déterminer à partir des équations des trois faces.

La première étape consiste à détecter automatiquement les trois plans dans la scène. Comme pour Mure-Dubois et al. (2008), l'algorithme de RANSAC est utilisé trois fois afin de détecter les nuages de points décrivant chaque plan. Cependant, nous ne déterminons pas le vecteur normal à partir du plan associé pour ne pas avoir de solution non-déterministe. Pour ceci, nous comptons appliquer une minimisation au sens des moindres carrés (Least Squares, LS) afin d'avoir une solution déterministe (Becerro, 2008). La présence d'outliers dans les jeux de données de caméras ToF rend cette méthode potentiellement moins robuste : nous avons donc appliqué la méthode des moindres médians carrés (Least Median Square, LMS) (voir figure 6.2).

6.1. Création d'un système stéréo d'observation : calibrage de 2 caméras ToF

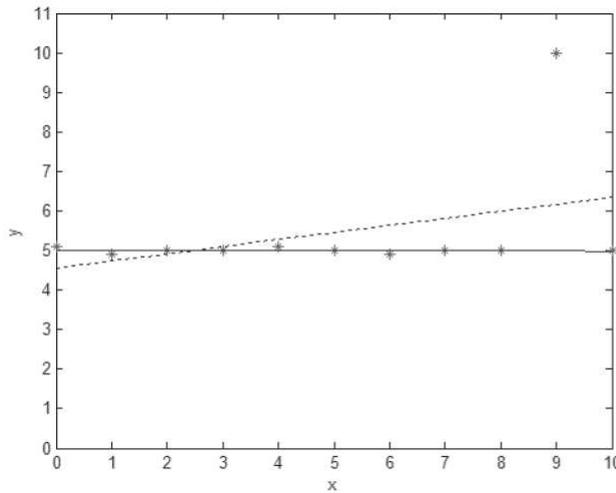


FIGURE 6.2 – Régression linéaire d'un jeu de données (croix) à l'aide de la méthode des moindres carrés (pointillés) et de la méthode des moindres médians carrés (ligne pleine)

Cette méthode ajoute une étape à la méthode LS. Un triplet de points est sélectionné, et le plan qu'il décrit est classé en fonction de sa capacité à décrire le jeu de données. Ce processus est répété un grand nombre de fois, ce qui fournit un classement des plans les plus justes. Au final, le plan associé à l'erreur médiane est utilisé pour enlever de potentiels outliers grâce à un seuil, ici de 5 mm. La méthode LS est finalement appliquée sur le nouveau jeu de données.

Une fois déterminés les vecteurs normaux aux 3 plans dans chaque vue, nous avons des repères quasi-orthonormés, seule l'erreur d'estimation et la géométrie physique du cube empêchant les vecteurs d'être orthonormaux. Afin de permettre l'estimation de la rotation, nous avons utilisé la méthode analytique déterminant l'orientation absolue par l'utilisation des quaternions unitaires pour décrire la rotation (Horn, 1987).

Le quaternion est défini comme un nombre complexe composé d'un nombre réel et de trois composantes imaginaires :

$$q = q_0 + iq_x + jq_y + kq_z \quad (6.4)$$

où le quaternion q est composé du nombre réel q_0 et de trois parties imaginaires q_x, q_y, q_z .

La rotation d'un vecteur p dans l'espace 3D est :

$$p' = Rp \quad (6.5)$$

Lorsque la rotation est décrite par un quaternion, elle peut être décrite comme le produit composite du quaternion purement imaginaire q :

$$p' = qpq^* \quad (6.6)$$

avec q^* le conjugué de q .

La méthode requiert au minimum 3 points afin de contraindre le système d'analyse. soit p_1, p_2 et p_3 les trois points dérivés des vecteurs normaux des plans observés dans la vue V , de même que p'_1, p'_2 et p'_3 les trois points dérivés des vecteurs normaux des plans observés dans la vue V' . La rotation optimale décrite par q minimise la fonction F

$$F = \sum_{i=1}^3 (qp_i p_i^*) p'_i \quad (6.7)$$

La minimisation est effectuée à partir d'une matrice 4x4 construite à partir des sommes des produits des paires des points appareillés. La résolution de cette matrice par Décomposition en Valeur Singulière fournit le quaternion q qui décrit la meilleure rotation comme un vecteur unitaire ayant la même direction que le vecteur propre de plus grand poids.

Enfin, la composante translation est décrite comme le vecteur reliant les estimations du coin du cube. En fonction de la configuration des caméras, la méthode doit être adaptée car chaque caméra ne voit par forcément les mêmes plans ni le même sommet : une translation supplémentaire égale aux dimensions du cube résoudra la difficulté. Le coin du cube est en fait l'intersection des 3 plans adjacents, ou dit autrement l'intersection entre chaque plan et l'arête résultant de l'intersection des deux autres plans. Soient trois plans π_1, π_2 et π_3 , chacuns décrit par leurs vecteurs normaux N_1, N_2 et N_3 et un point p_1, p_2 et p_3 . Nous définissons la ligne l comme l'intersection des plans π_2 et π_3 . Le point p_v est le coin correspondant aux 3 plans si

$$l = (N_2 \wedge N_3) \quad (6.8)$$

et

$$p_v = \pi_1 \cap l \quad (6.9)$$

6.1.3 Expérimentations

Pour l'étude des trois méthodes, la sphère est en polystyrène avec un rayon de 15 cm, alors que le cube est en bois de dimension $30 \times 30 \times 30 \text{ cm}^3$. Le plan est considéré comme une face du cube. Comme on peut le voir figure 6.3, les objets sont blancs afin de limiter l'absorption de la lumière infrarouge par leur surface.

Une première partie d'expérimentations a cherché à évaluer la précision atteinte pour l'extraction des points ou les figures d'intérêt des différentes méthodes : le centre d'une sphère et le vecteur normal à un plan. Une deuxième partie a comparé les méthodes. Pour ceci, nous avons proposé une estimation de l'erreur de recalage après calibrage des différents points d'intérêt, et ceci selon deux sous-espaces, un sous-espace au centre du champ de vue commun qui contient les localisations où sont prises les mesures de calibrage (sous-espace

6.1. Création d'un système stéréo d'observation : calibrage de 2 caméras ToF

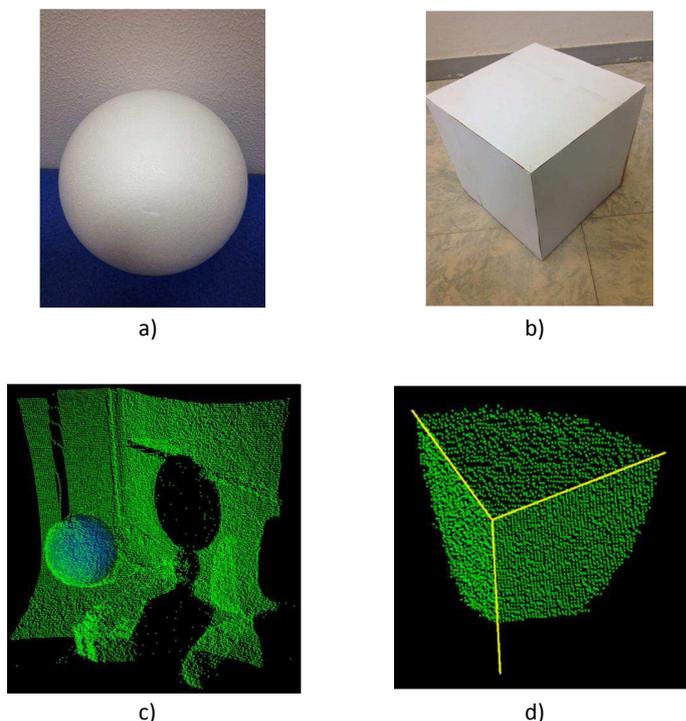


FIGURE 6.3 – Les deux objets physiques utilisés sont (a) une sphère en polystyrène de 15 cm de diamètre et (b) un cube en bois peint en blanc de diamètre 30x30x30 cm qui sert aussi de plan à détecter. En c) et d), les scènes 3D des objets en question.

Calibrage), et un sous-espace éloigné, appelé espace Monde (Figure 6.4). L'intérêt de cette approche est de proposer une analyse homogène aux différentes méthodes tout en détectant quelle composante est la plus touchée par les erreurs : une erreur faible dans le sous-espace Calibrage couplée à une erreur forte dans le sous-espace Monde indiquera une bonne estimation de la composante translation couplée à mauvaise une mauvaise rotation.

L'erreur de calibrage est considérée comme la distance euclidienne séparant l'estimation du point d'intérêt caractéristique d'une méthode dans les deux vues. Pour la méthode avec les sphères, ce sera le centre de la sphère. Pour la méthode avec un plan, il n'y a pas de point caractéristique fixe : nous décidons d'apposer un repère orthonormé sur le plan à l'aide du vecteur normal, la base étant complétée par la méthode de Gram-Schmidt (Golub and Van Loan, 1996). Des points virtuels sont déterminés le long des vecteurs constituant le repère. Nous procédons de même pour la méthode du cube, sauf que nous n'avons pas une base exacte mais trois vecteurs quasi normaux.

Choisir la distance Euclidienne séparant deux estimations d'un même point d'intérêt semble convenir puisqu'un calibrage parfait rendrait cette distance nulle. Soit p_f l'estimation de ce point dans la vue V , p'_f dans la vue V' et M la transformation rigide reliant les deux vues.

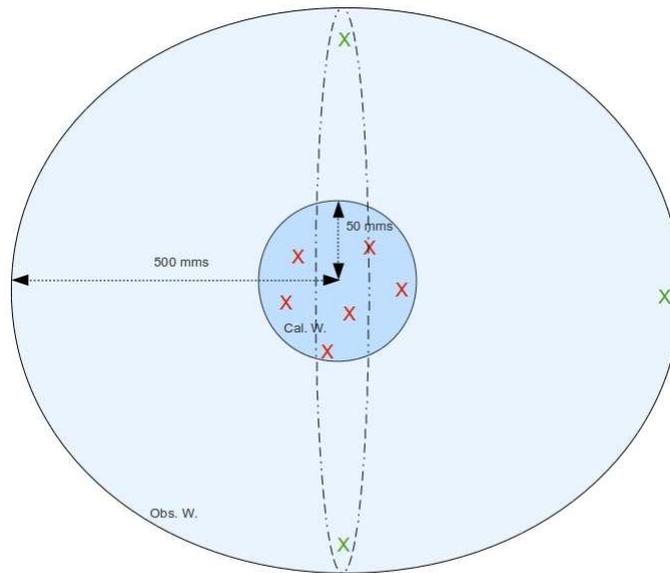


FIGURE 6.4 – Représentation du champ de vue commun aux deux caméras qui est divisé en deux sous-espaces : Le sous-espace Calibrage et le sous-espace Monde. Les croix rouges indiquent les positions ayant servi au calibrage et/ou à la validation alors que les croix vertes désignent des positions ayant seulement servi pour validation.

L'erreur moyenne de calibrage ϵ est obtenu par la formule

$$\epsilon = \frac{1}{n} \sum_n \|p_f - M \cdot p'_f\| \quad (6.10)$$

avec n le nombre d'estimations.

Nous avons tout d'abord analysé la qualité d'estimation de deux étapes clefs des méthodes présentées : l'estimation du centre d'une sphère et celle du vecteur normal à un plan. Pour la sphère, la répétabilité de l'estimation est déterminée en observant et modélisant 100 fois une sphère statique, et ceci pour deux distances, 1 m et 1.5 m. Cependant, la répétabilité ne permet pas de distinguer le bruit sur l'estimation et la présence d'une erreur systématique. Nous avons donc déplacé la sphère avec l'actionneur linéaire pour des translations de 10, 30 et 70 mm selon les directions x et z . Moyennant 20 estimations à chaque position, le biais est défini comme la différence entre les centres de sphère et la vérité terrain.

Pour connaître la répétabilité de l'estimation d'un vecteur normal à un plan, nous avons évalué l'écart-type sur les angles d'Euler pour 20 observations, et ceci pour les deux méthodes préalablement présentées, l'algorithme de RANSAC et RANSAC complété avec la méthode LMS.

6.1. Création d'un système stéréo d'observation : calibrage de 2 caméras ToF

La figure 6.5 montre que la variabilité sur l'estimation du centre de la sphère est significativement plus grande selon l'axe z que selon les autres axes, à 1 m (de 0.2 mm à 0.6 mm) comme à 1.5 m (de 0.5 à 1.1 mm). On voit tout de suite la forte influence de la distance, qui est bien expliqué par le fait qu'à 1.5 m, la même surface est décrite par un nombre de points moins élevé (voir figure 3.10). Le fait que la composante du centre en z soit soumise à une plus grande variabilité confirme les résultats présentés dans Böller (2010), tant au niveau de l'interprétation que de l'ordre de grandeur des erreurs.

Concernant l'estimation de la translation de la sphère, la figure 6.6 montre des biais extrêmement variables allant de 0.1 mm (translation de 10 mm selon l'axe x) à 2.6 mm (70 mm dans la direction z). Ces résultats sont difficiles à interpréter même si on peut distinguer l'influence de l'amplitude de la translation, ce qui pourrait être interprété par le fait qu'une grande translation rend les deux nuages de points décrivant les positions moins similaires.

Ecart-type (mm)		
	Caméra 1	Caméra 2
A 1 m		
X	0.1	0.1
Y	0.1	0.2
Z	0.6	0.6
A 1.5 m		
X	0.4	0.5
Y	0.4	0.4
Z	1.2	1.1

FIGURE 6.5 – Variabilité de l'estimation du centre de la sphère fixe.

La variabilité d'estimation des angles d'Euler (figure 6.7) est réduite significativement grâce à l'apport de la méthode LMS avec une réduction pouvant atteindre un facteur 9. La variabilité sur les angles lorsque RANSAC est utilisé confirme le problème de la détermination non-déterministe. La résolution de cette limite passe par l'ajout d'une méthode déterministe, la méthode LMS permettant en plus d'éliminer les outliers, vecteurs d'erreur lors d'une minimisation.

L'erreur de calibrage est présentée figure 6.8 pour chaque méthode et chaque région spatiale. La méthode de la sphère produit une erreur allant de 1.4 mm à 12.7 mm en s'éloignant du centre du champ de vue. A contrario, notre méthode montre l'erreur la plus forte au centre (5.1 mm) mais la plus faible en dehors (7.1 mm). La méthode basée sur un plan n'obtient pas de résultats satisfaisants avec une erreur forte dans les deux régions. La méthode utilisant une sphère ayant une erreur bien plus grande à la périphérie du champ de

Erreur moyenne d'estimation (mm)		
	Caméra 1	Caméra 2
Translation selon X (mm)		
10	0.4	0.1
30	1.1	0.2
70	2.6	1.7
Translation selon Z (mm)		
10	0.6	1.2
30	1.3	2.5
70	1.7	0.4

FIGURE 6.6 – Estimation de la translation de la sphère.

	RANSAC	RANSAC + LMS
alpha (°)	0.8	0.1
beta (°)	0.7	0.2
phi (°)	0.9	0.1

FIGURE 6.7 – Variabilité sur l'estimation des angles d'Euler associés à un plan.

vue qu'au centre, on peut interpréter que sa translation est bien estimée au contraire de sa composante rotation. L'utilisation de l'ICP pour recalculer les positions des sphères peut expliquer les difficultés à bien estimer la rotation lorsqu'il y a peu de points et que ceux-ci sont bruités selon l'axe z. Le nombre de positions, ici trois, influe peut-être aussi ce résultat, une expérimentation supplémentaire avec plus de points serait à effectuer. Cependant, le résultat obtenu confirme celui présenté dans Böller (2010), soit une erreur médiane de 1.5 mm avec un maximum de 3 mm. A contrario, l'erreur de calibrage lorsque le cube est utilisé montre le schéma opposé avec une erreur plus forte au centre du champ de vue, erreur qui n'augmente que très peu à la périphérie. L'interprétation qu'on peut en faire est une très bonne estimation de la rotation associée à une médiocre estimation de la translation. En effet, la variabilité sur l'estimation du vecteur normal à un plan est excellente, ce qui permet à la méthode du quaternion absolue de déterminer analytiquement une solution optimale pour la composante rotation. A contrario, l'estimation du coin du cube manque de précision, l'explication reste encore à trouver par des expérimentations supplémentaires. Enfin la méthode basée sur un plan semble fournir une estimation médiocre de la rotation et de la translation : les résultats montrent que la détermination de l'orientation d'un plan par RANSAC est limitée, ainsi que la détermination de la translation qui est extrêmement utilisateur-dépendant, en plus d'utiliser des images de faible résolution pour la détection de points d'intérêt.

6.1. Création d'un système stéréo d'observation : calibrage de 2 caméras ToF

	Sphère	Plan	Cube
Sous-espace Calibrage	1.4 mm	4.2 mm	5.1 mm
Sous-espace Monde	12.7 mm	11.4 mm	7.1 mm

FIGURE 6.8 – Erreur de calibrage des différentes méthodes au centre et à la périphérie du champ de vue.

6.1.4 Impact sur le repositionnement

Nous avons évalué l'impact direct de l'utilisation de deux caméras à la place d'une seule, avec dans un premier temps un fantôme anthropomorphique déplacé par un actionneur linéaire, ce qui permet un déplacement extrêmement précis. L'analyse est cependant limitée à deux axes, x et z, le fantôme étant trop large pour l'application d'un mouvement dans la direction cranio-caudale. Dans un deuxième temps, un volontaire est déplacé par la table d'examen de l'accélérateur de traitement ONCOR de Siemens qui a une précision d'un millimètre selon les trois directions. Afin de considérer des surfaces identiques du point de vue de leur déformation, la sélection se fait à expiration complète. Dans les deux cas, la caméra est située à 1.2 m de face par rapport à l'objet ou l'humain observé. Pour les conditions d'observation, la configuration des caméras est une position symétrique par rapport au plan cranio-caudal afin d'avoir une description homogène des flancs du patient (caméras A et B sur la figure 6.9). Le système mono est constitué d'une des deux caméras A ou B. Pour l'observation d'un humain, le système mono est une caméra C fixée au dessus du patient de manière à être aligné avec le plan de la table d'examen (caméra C sur la figure 6.9).

Le repositionnement est fourni par le recalage de régions d'intérêt identiques à chaque position par ICP. La surface du fantôme est équivalente à la zone représentée figure 6.10 à gauche. La surface de l'humain est découpée en deux régions d'intérêt, une région large contenant une large partie du thorax, de l'abdomen plus les bras, et une autre plus exigüe sans les bras (ces régions sont représentées figure 6.10). Le recalage est effectué par un algorithme ICP.

Le champ de vue après fusion des vues (voir figure 6.11) permet une estimation complète des flancs du patient tout en augmentant la quantité de points décrivant une large bande centrale dans la direction tête-pieds, dans une zone touchée par l'effet de l'angularité avec un nombre de points comparablement plus faible que sur les flancs.

Les figures 6.12 et 6.13 montrent l'erreur d'estimation de déplacement du fantôme pour des amplitudes de mouvement allant de 5 à 90 mm. On y voit l'apport du système d'observation stéréo avec une influence plus importante lorsque le déplacement est ample. Les résultats peuvent être divisés en deux : lorsque le déplacement est inférieur au centimètre, le système stéréo n'apporte pas d'amélioration, ni lors d'un mouvement latéral (0.7 ± 0.2 mm contre 0.6 ± 0.3 mm et 0.7 ± 0.1 mm) ni lors d'un déplacement en profondeur (0.7 ± 0.4

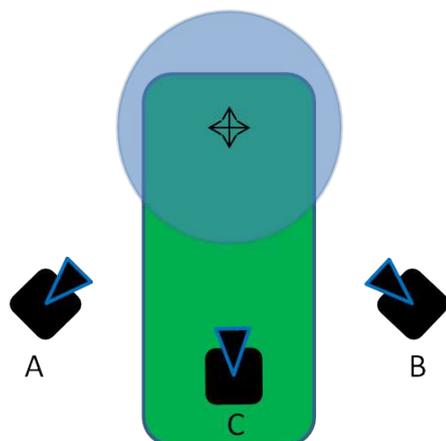


FIGURE 6.9 – Configuration du dispositif d’observation par rapport à la table d’examen. Les caméras sont centrées sur l’isocentre de traitement.

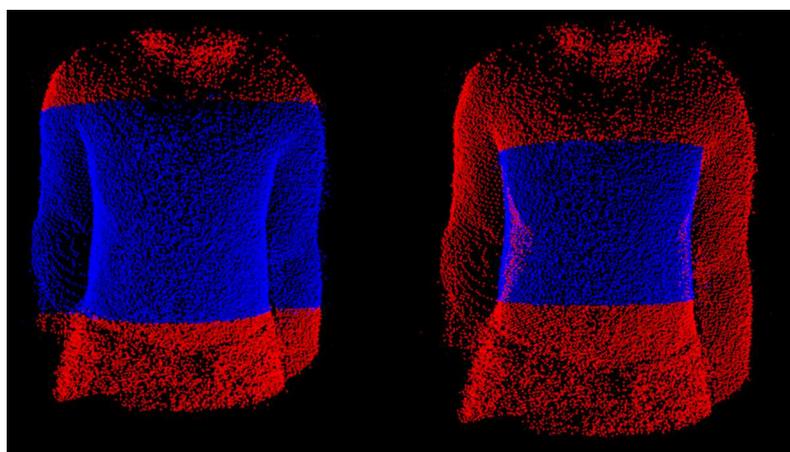


FIGURE 6.10 – Régions d’intérêt pour l’évaluation du repositionnement sur volontaire (région 1 à gauche, région 2 à droite).

mm contre 0.5 ± 0.3 mm et 1 ± 0.7 mm). Au delà du centimètre de déplacement, l’erreur augmente fortement lorsqu’une seule caméra est utilisée, au contraire du système stéréo qui conserve une erreur faible avec une erreur moyenne de 1.3 ± 0.7 mm (contre 3.8 ± 3.1 mm et 3.4 ± 2.8 mm) et 1 ± 0.4 mm (contre 3.1 ± 1.4 mm et 2.8 ± 2.1 mm) selon la direction du mouvement. L’amélioration moyenne est d’un facteur 3, étant plus forte lorsque l’amplitude de déplacement augmente.

Les figures 6.14, 6.15 et la table 6.16 indiquent la différence entre l’estimation de déplacement d’un humain et le mouvement de la table de l’accélérateur de traitement ONCOR. D’un point de vue général, l’amélioration moyenne apportée par le système stéréo va de 10% à 69%. Quand la région d’intérêt est la plus large, l’amélioration est de 18%, 39% and 10% au regard d’un déplacement latéral, en profondeur et longitudinal. De même pour la région 2 : 69%, 40% and 37%. On voit de même que la région 2 permet une meilleure estimation

6.1. Création d'un système stéréo d'observation : calibrage de 2 caméras ToF

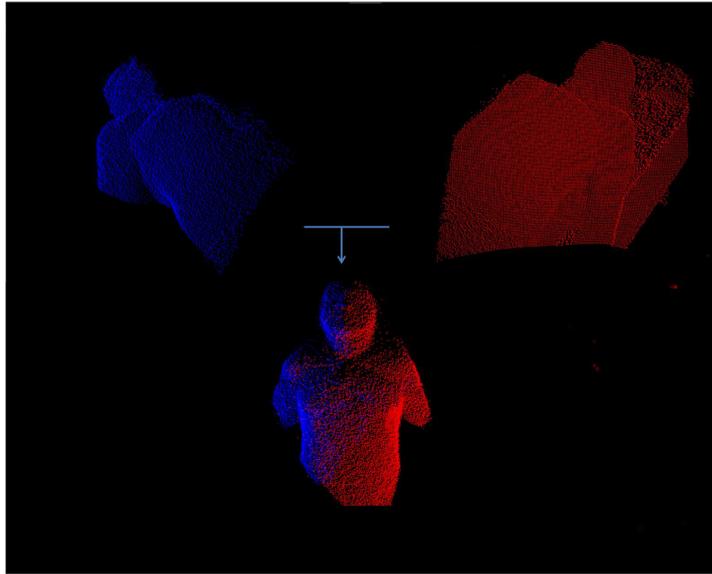


FIGURE 6.11 – Fusion des observations latérales du volontaire avec en bleu les données de la caméra A, et en rouge celles de la caméra B.

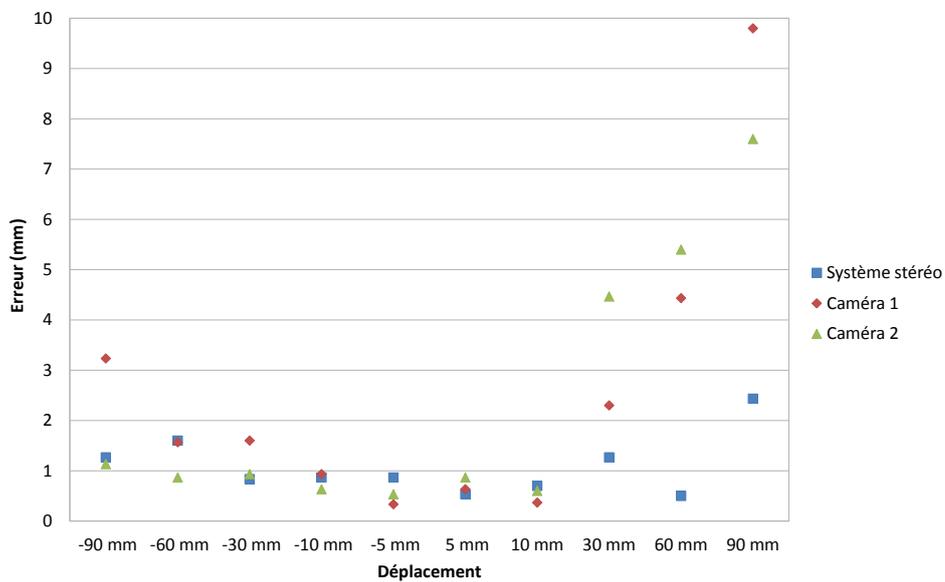


FIGURE 6.12 – Erreur d'estimation du mouvement latéral du fantôme quand l'observation est effectuée par la caméra A, la B ou les deux.

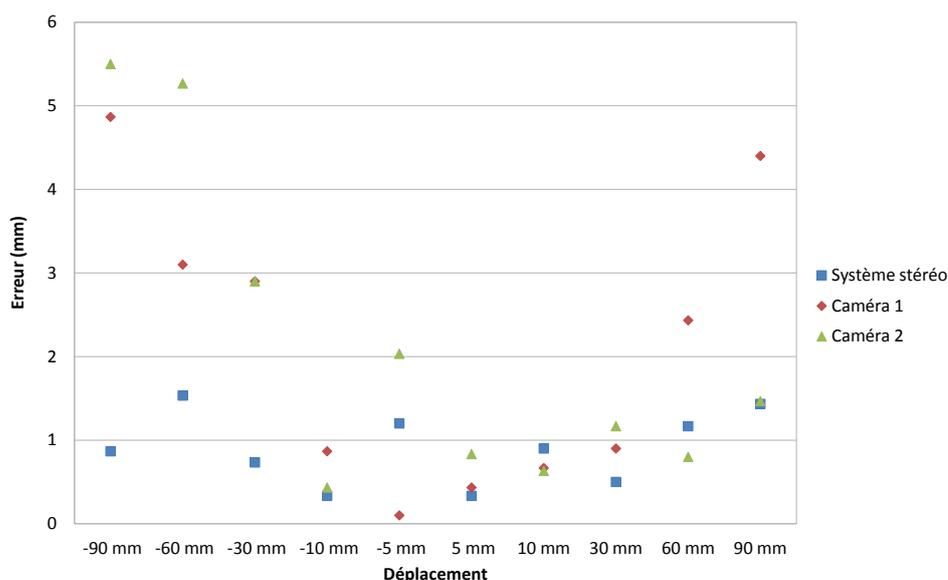


FIGURE 6.13 – Erreur d’estimation du mouvement en profondeur du fantôme quand l’observation est effectuée par la caméra A, la B ou les deux.

de déplacement pour toutes les directions par rapport à la région 1 : 1.4 mm contre 1 mm (latéral), 1.1 mm contre 0.9 mm (profondeur) et 8.9 mm contre 3.4 mm (longitudinal).

6.2 Prise en compte du mouvement respiratoire

Les résultats collectés lors des tests précédents mais aussi lors de la phase 1 de l’étude en condition clinique indiquent une très bonne capacité à estimer un déplacement latéral, mais les directions cranio-caudales et antérieures-postérieures posent clairement un souci de précision. La principale variable qui n’est pas encore prise en compte est le mouvement respiratoire qui déforme la surface du patient, ce qui implique qu’il y a de fortes chances pour que les surfaces prises pour comparaisons ne décrivent finalement pas la même réalité, impactant sur la capacité de recalage. De plus, autant la respiration déforme latéralement le patient, mais de manière symétrique tout au long du cycle respiratoire, autant la déformation n’est symétrique ni en profondeur (la surface est globalement plus élevée en inspiration complète) ni selon la direction tête-pieds (le thorax et l’abdomen ne se soulèvent pas avec la même amplitude. Enfin, l’ordre de grandeur de cette déformation (pouvant atteindre 1 cm) est un obstacle majeur à un recalage d’une précision de quelques millimètres.

Nous avons cherché à tenir compte de ce mouvement respiratoire afin de sélectionner les surfaces à recalculer en fonction de cette information. La caméra ToF nous fournit un instrument précieux puisqu’elle peut acquérir à haute fréquence (jusqu’à 50 Hz), nous permettant

6.2. Prise en compte du mouvement respiratoire

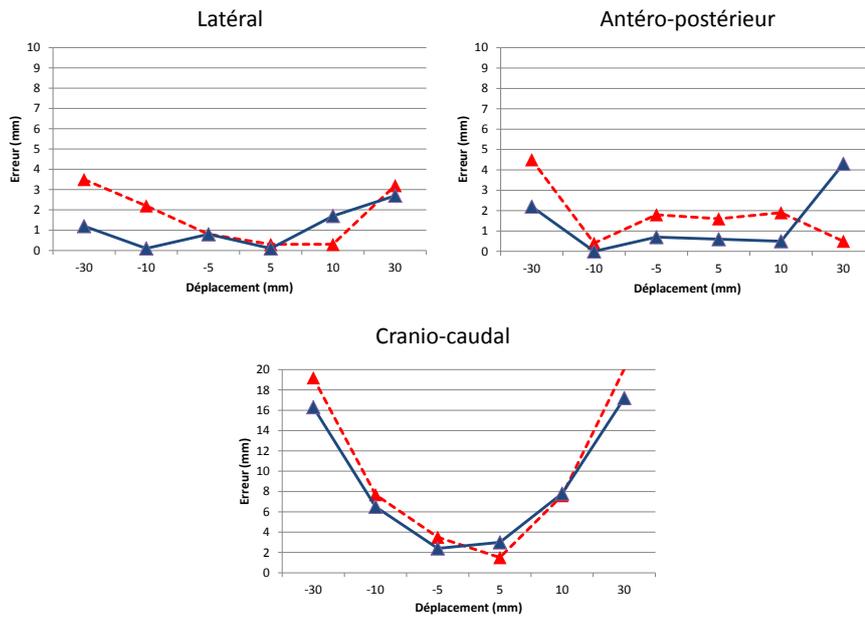


FIGURE 6.14 – Erreur d’estimation de la translation d’un humain quand l’observation est effectuée par la caméra C (en rouge) ou le système stéréo (bleu). La région d’intérêt est la première de la figure 6.10.

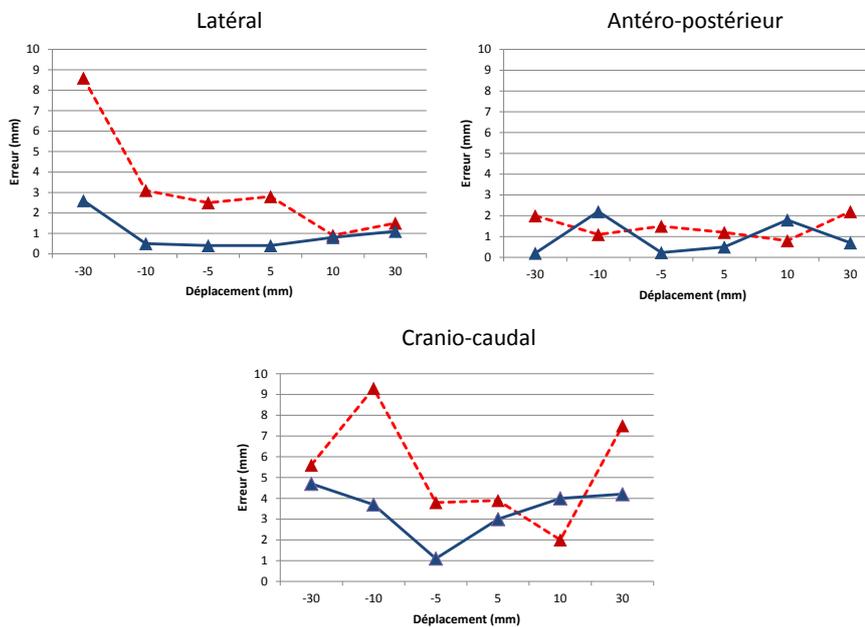


FIGURE 6.15 – Erreur d’estimation de la translation d’un humain quand l’observation est effectuée par la caméra C (en rouge) ou le système stéréo (bleu). La région d’intérêt est la deuxième de la figure 6.10.

	Région 1		Région 2	
	1 caméra	2 caméras	1 caméra	2 caméras
Lateral	1.7	1.4	3.2	1
Antéro-postérieur	1.8	1.1	1.5	0.9
Cranio-caudal	9.9	8.9	5.4	3.4

FIGURE 6.16 – Erreur moyenne d’estimation du mouvement en millimètre pour toute la gamme de translations.

ainsi d’extraire le signal respiratoire avec une bonne résolution temporelle. En prenant en compte trois régions, une thoracique, une abdominale et une dernière intermédiaire, nous pouvons distinguer les types de respiration et ainsi gagner en caractérisation. La figure 6.17 montre une acquisition et les trois ROIs desquelles sont tirés les signaux respiratoires.

Un test est effectué avec un volontaire auquel est demandé de respirer de manière variée. Les traces extraites des trois zones (voir figure 6.18) permettent de sélectionner les surfaces à recalcr afin d’avoir les traces à des instants similaires. La figure 6.19 illustre deux cas. Le premier cas est celui où les surfaces sont sélectionnées en prenant compte le signal abdominal, ce qui peut amener à sélectionner les images 1 et 3. Malheureusement, ces instants ne correspondent pas du tout si l’on tient compte des traces correspondant aux régions thoraciques et abdominales. Le recalage des surfaces considérées comme la région bleue avec l’ICP estime un déplacement de -2.6 mm en latéral, 3.8 mm en tête-pied et 11.1 mm en profondeur, alors que le mouvement réel est nul. Dans un second cas, on cherche deux images pour lesquelles les trois signaux sont dans une configuration identiques, ce qui amène à sélectionner les images 1 et 2, même si on distingue un léger décalage de l’ordre de 2 mm entre les deux configurations. Le recalage renvoie un déplacement de -0.6 mm en latéral, 0.1 mm en tête-pied et 2.6 mm en z.

Un exemple plus détaillé est effectué à partir de données cliniques. Un patient est observé durant la première fraction de son traitement. Les différentes imageries de contrôles lui font appliquer dans un premier temps un décalage de 8 mm vers la tête, suivi d’un déplacement de 8 mm dans la direction antérieure. Nous avons tout d’abord considéré le premier décalage, en sélectionnant deux cycles respiratoires avant et après décalage. La surface de référence choisie correspond à l’expiration complète du patient. Nous avons ensuite recalé cette surface avec les surfaces équivalentes anatomiquement pour chaque instant du cycle

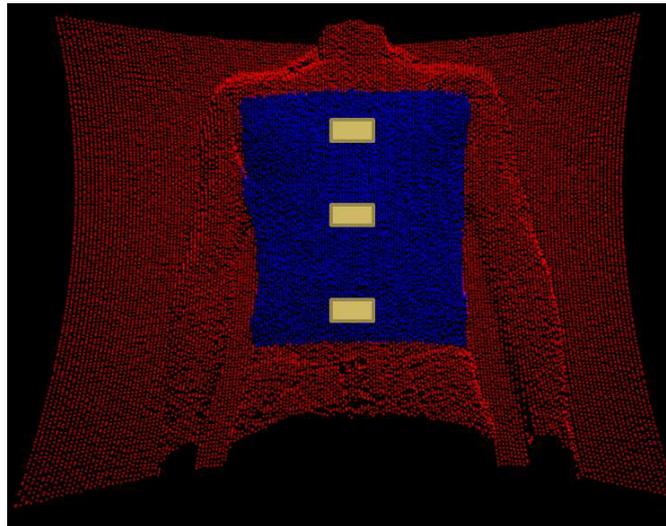


FIGURE 6.17 – Trois régions d'intérêt sont délimitées afin de caractériser l'instant respiratoire associé à la surface acquise.

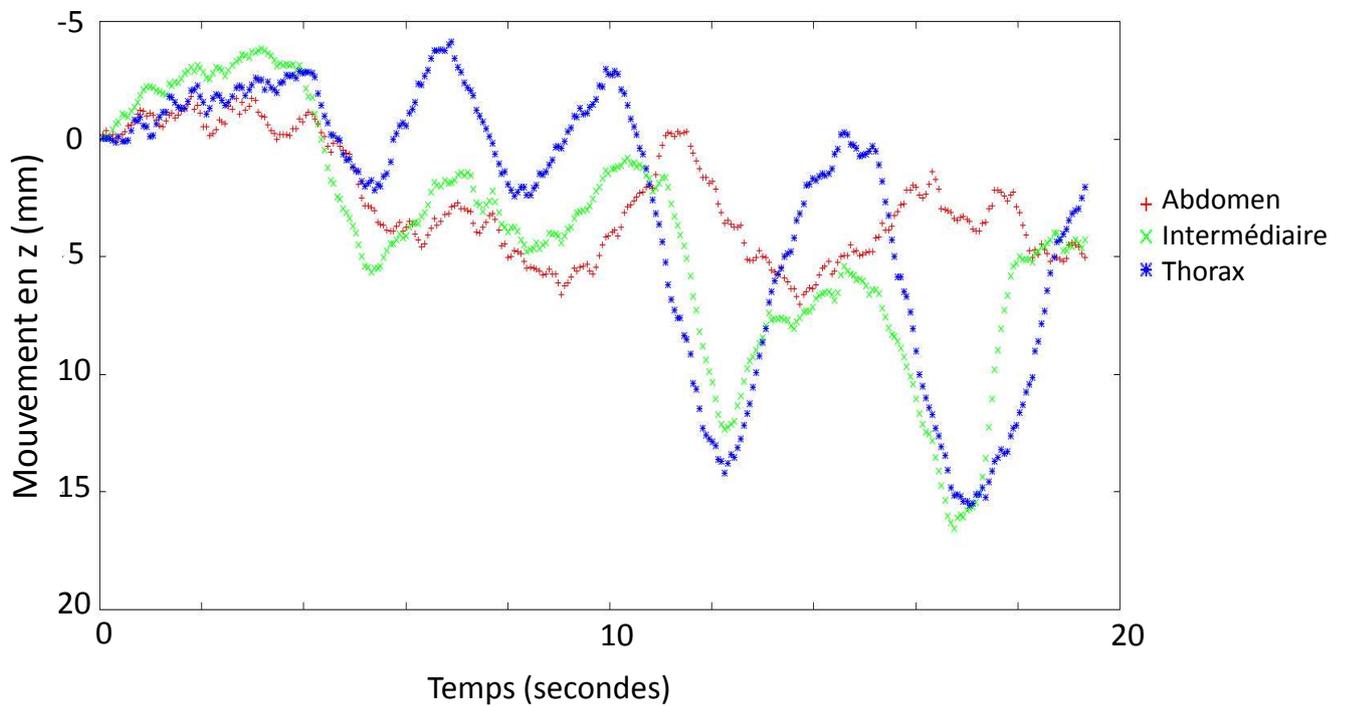


FIGURE 6.18 – Signaux respiratoires correspondant aux trois régions d'intérêt illustrées figure 6.17.

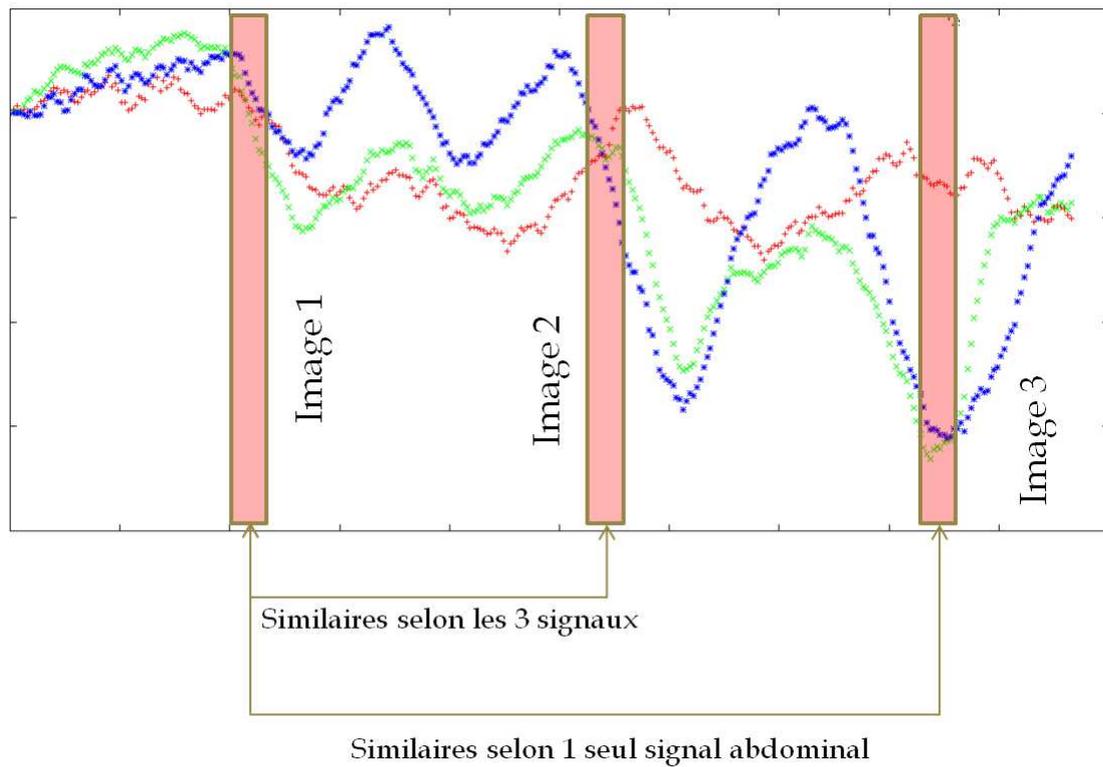


FIGURE 6.19 – Trois acquisitions sont sélectionnées, la deuxième pour sa concordance avec la première selon les trois signaux, la troisième pour sa concordance d’après le signal abdominal uniquement.

respiratoire, en commençant par celle correspondant à l’expiration totale, soit en tout 26 acquisitions. La figure 6.20 montre pour les 26 acquisitions le déplacement estimé par le recalage des surfaces avec l’ICP pour les trois directions. La meilleure estimation du déplacement réel (0 mm, 8 mm, 0 mm) est fournie lors du recalage de la surface à expiration totale (0.7 mm, 8.3 mm, 0.5 mm), soit une erreur totale de 0.9 mm. De même, la dernière acquisition, elle aussi à expiration totale, produit la deuxième meilleure estimation du déplacement, (1.5 mm, 7.2 mm, -0.1 mm) soit une erreur de 1.7 mm. A contrario, l’estimation du déplacement lorsque les surfaces sélectionnées sont situées à un instant respiratoire différent est entachée d’une erreur étant croissante avec l’amplitude du mouvement respiratoire qui atteint son maximum lors de la treizième acquisition. L’erreur est principalement concentrée sur la direction antéro-postérieure avec un déplacement estimé atteignant 7 mm à inspiration complète, avec une corrélation importante avec le mouvement respiratoire. Ce résultat était attendu puisque les surfaces recalées sont situées à des hauteurs différentes à cause de la respiration. Cependant, la similarité médiocre des surfaces se ressent aussi par un mouvement latéral fantôme qui, bien que pas complètement corrélé avec le signal respiratoire, atteint 8 mm avec une forte variabilité.

La même analyse est effectuée par rapport au deuxième décalage. Une démarche si-

6.2. Prise en compte du mouvement respiratoire

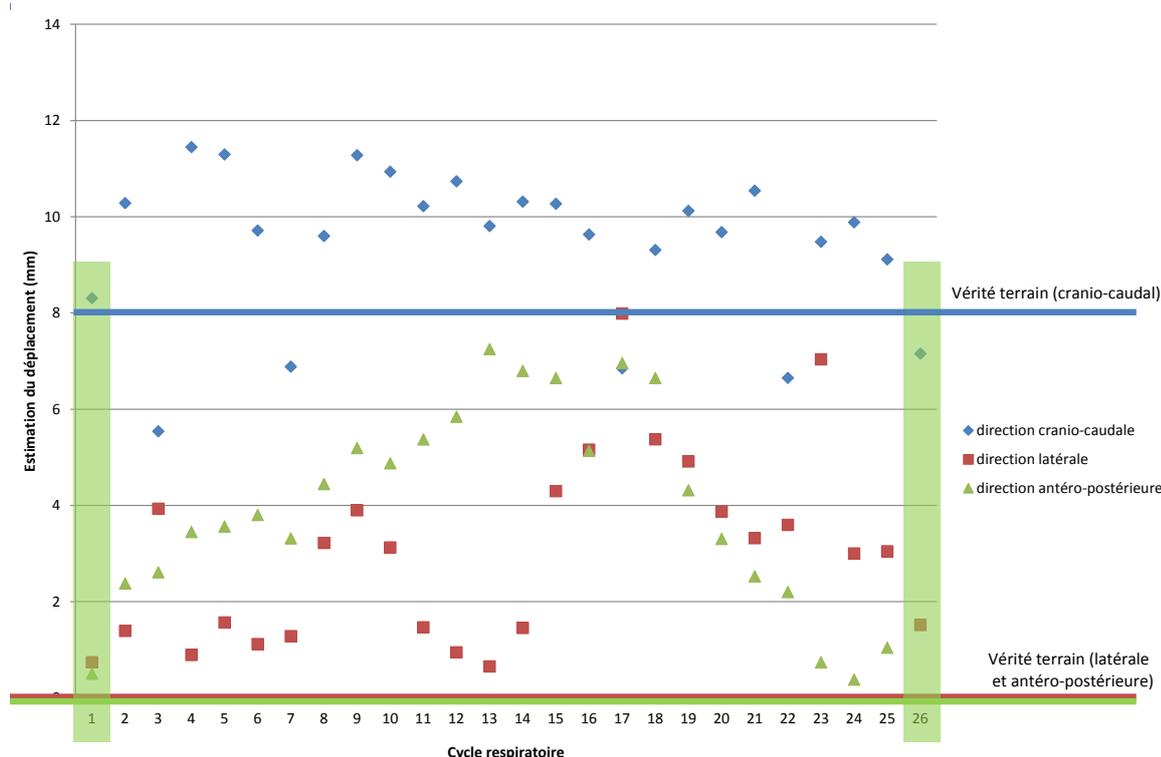


FIGURE 6.20 – Résultats du recalage surfacique d'un patient ayant subi une translation de 8 mm dans la direction craniale-caudale. La surface pré-déplacement est sélectionnée à expiration complète. Elle est ensuite recalée avec chaque acquisition d'un cycle respiratoire équivalent après le déplacement. Les meilleurs recalages sont avec la première et la vingt-sixième images (en vert).

miltaire est suivie, avec cette fois ci un cycle respiratoire après déplacement du patient qui est constitué de 28 acquisitions. Les résultats sont présentés figure 6.21. On y voit encore une fois une estimation de mouvement selon l'axe antéro-postérieur extrêmement corrélé au signal respiratoire. Cependant, les deux meilleures estimation du déplacement sont aux deuxièmes et vingt-troisièmes acquisitions, soit pas exactement à expiration complète. Les erreurs sont alors respectivement de 1.9 mm et 1 mm. Une interprétation possible est que les cycles respiratoires sélectionnés n'ont pas la même amplitude. Le second cycle respiratoire est sans doute plus profond que le premier, ce qui explique que la meilleure similarité des surfaces soient finalement entre l'expiration profonde du premier cycle et approximativement 20% de l'inspiration pour le second.

En conclusion, on peut relever la difficulté à sélectionner des surfaces correspondant au même

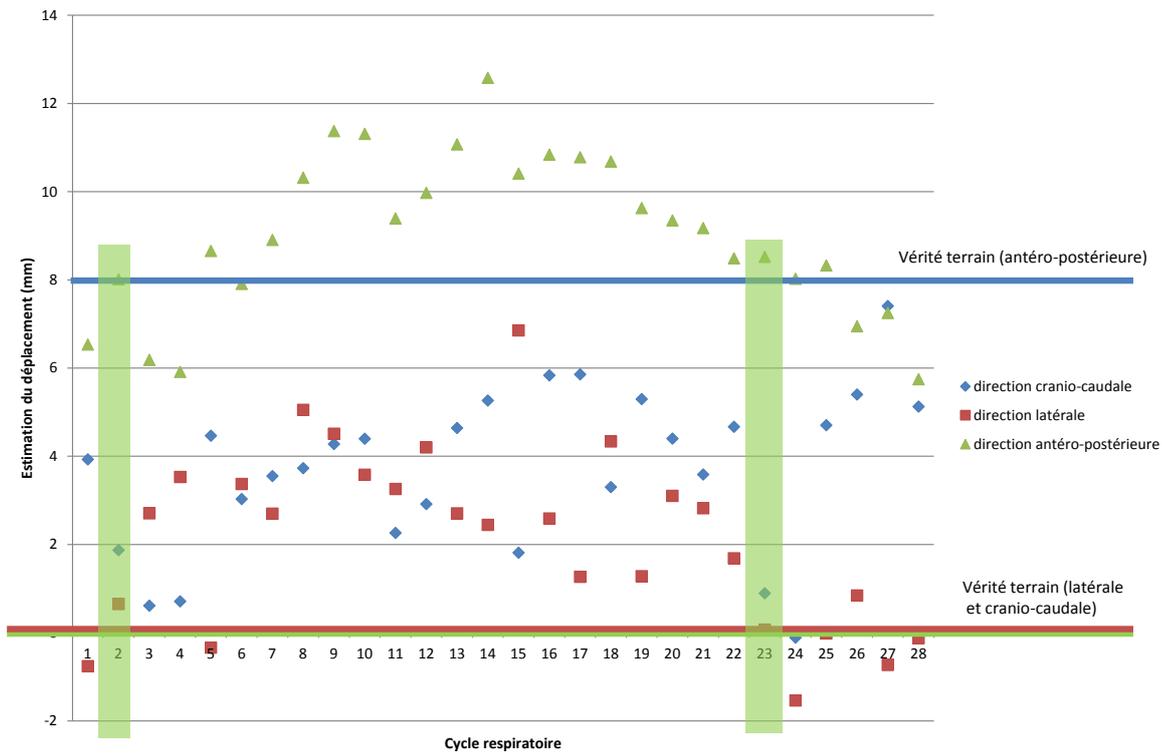


FIGURE 6.21 – Résultats du recalage surfacique d'un patient ayant subi une translation de 8 mm dans la direction antérieure. La surface pré-déplacement est sélectionnée à expiration complète. Elle est ensuite recalée avec chaque acquisition d'un cycle respiratoire équivalent après le déplacement. Les meilleurs recalages sont avec la deuxième et la vingt-troisième images (en vert).

instant respiratoire. La variabilité des cycles respiratoires rend cette tâche compliquée, c'est pourquoi pour l'étude clinique, il a été décidé de sélectionner des surfaces à expiration complète. La première raison est qu'il est facile de repérer sur les signaux respiratoires cette phase. La deuxième raison est que l'expiration est à priori la position la plus reproductible naturellement au fil des cycles respiratoires.

6.3 Etude Clinique

6.3.1 Objectifs, moyens et déroulement

Les différentes approches présentées ont permis de développer des solutions afin d'optimiser le repositionnement d'un patient sans extraire de points caractéristiques, comme souvent dans la littérature. Nous avons donc mis en place une étude clinique avec le service de radiothérapie du CHRU de Brest afin d'évaluer l'apport hypothétique d'une caméra Temps-de-Vol au processus global de repositionnement d'un patient. L'étude est composée d'un objectif primaire, suivi de deux objectifs secondaires. Le principal objectif est de situer la caméra ToF dans la littérature associée à l'utilisation des scanners 3D comme aide au repositionnement d'un patient. La section 2.4 a déjà présenté les systèmes existants, tous basés sur la lumière structurée ou sur le balayage laser. Dans les deux cas, la surface du patient est décrite comme un ensemble de points épars. Les performances sont assez floues en ce qui concerne la localisation thoracique, qui est celle où la surface observée se déforme le plus sous l'effet de la respiration. Or la comparaison de deux acquisitions d'une même surface ayant subi une déformation dégrade fortement l'estimation de son déplacement, comme montré dans la section 6.2 et par Zhao et al. (2011) dans le cas de seins ayant subi un acte chirurgical. L'objectif principal de cette étude est tout d'abord d'évaluer la précision de l'estimation de déplacement d'un patient lorsqu'il est translaté par la table de traitement sur laquelle il est allongé. Il est ensuite de prédire le déplacement à appliquer au patient par rapport à sa position lors du J0 de son traitement.

Les deux objectifs secondaires sont liés au mouvement intrafraction du patient. Le premier est d'évaluer ce mouvement pendant une fraction, surtout dans le cas où celle-ci peut durer quinze à vingt minutes. Le second est d'évaluer le dispositif d'observation afin de suivre en temps réel le mouvement respiratoire externe. L'idée sous-jacente à cette tâche est l'utilisation du modèle de respiration global développé au laboratoire et présenté dans ce manuscrit section 2.3.2 comme outil de prédiction du déplacement des organes internes lors de la respiration. Ce point sera discuté en conclusion de ce chapitre. Toutefois, l'avancement actuel de l'étude ne nous permet pas de présenter de résultats concernant ces objectifs secondaires.

Cette étude est mise en place en liaison avec le service de radiothérapie du CHRU de Brest. Deux médecins internes ont été et sont associés à cette étude, leur tâche étant le recrutement de patients lors de la consultation pré-traitement afin d'obtenir leur accord signé. De même, ils participent à l'acquisition des données ainsi que le traitement de celles-ci par le biais de logiciels spécialement développés à cet effet. Nous avons eu accès à l'accélérateur de type ONCOR de Siemens sur lequel sont traités la totalité des localisations du haut du corps.

Suivant l'ordre chronologique des développements, l'ensemble de l'étude se divise en trois

phases, la dernière étant toujours en cours. Les trois mesures sont : la détection du déplacement, sa prédiction et sa conformité. La détection du mouvement est la capacité à retrouver le déplacement du patient dans l'espace, généralement par le déplacement de la table d'examen. C'est la mesure qui permet d'évaluer la performance du recalage effectué, le déplacement de la table étant précis en dessous du millimètre. La prédiction du mouvement est effectuée en comparant la position du patient avant la fraction avec la position de référence à J0. Le déplacement de comparaison est alors fourni par l'imagerie portale qui détermine ou non un décalage par rapport à J0. Enfin, la conformité est l'estimation du déplacement entre la position du patient après repositionnement par imagerie et sa position à J0.

On peut dire que la précision de la détection du mouvement est la validation de la méthode de repositionnement par surface puisqu'elle est comparée à une vérité terrain et que le laps de temps entre les acquisitions est de l'ordre de quelques secondes, au maximum de quelques minutes. Les prédictions et mesures de conformité permettent d'évaluer la concordance entre le système par scanner 3D et l'imagerie portale de référence. Nous nous sommes enfin intéressés à la détection des déplacements supérieurs à 5 mm. En effet, en pratique clinique, les déplacements inférieurs à une certaine valeur et détectés par imagerie portale ne sont pas corrigés, ce seuil étant variable en fonction des services et des médecins. Un déplacement réel supérieur à 5 mm est considéré comme détecté si son estimation est elle supérieure à 4 mm, ce qui indique une estimation qui n'est à priori pas due à une erreur de recalage.

La première phase n'utilise qu'une caméra et évalue uniquement la détection du déplacement sur des fractions uniques de 39 patients différents. La seconde phase utilise deux caméras calibrées entre elles et suit 3 patients pendant cinq fractions de leur traitement afin d'évaluer la détection, la prédiction et la conformité du déplacement. Conformément aux études, le mouvement respiratoire est alors pris en compte en moyennant 10 acquisitions successives sur une seconde. La dernière phase reprend le protocole précédent mais tient compte du mouvement respiratoire lors de la sélection des surfaces à recalculer. Nous avons pour l'instant les résultats pour deux patients.

6.3.2 Dispositif d'observation

L'accès à la salle de traitement de radiothérapie nous a permis de réfléchir précisément au meilleur dispositif permettant :

- un champ de vue optimal centré sur l'isocentre de traitement.
- un positionnement relatif des deux caméras permettant de visualiser au final les flancs du patient tout en ayant une zone d'observation commune.
- une distance d'observation minimale.
- une gêne minimale pour les manipulateurs, afin d'assurer une insertion harmonieuse dans le contexte médical de traitement.

- un système qui permet de remettre les caméras à la même position, afin de permettre des comparaisons entre différentes fractions.

Les figures 6.22 et 6.23 montrent la configuration expérimentale, avec les deux caméras fixées au dessus de la table de traitement. Leur position est limitée par le champ de déplacement de l'accélérateur linéaire, elle sont donc approximativement à 70 cm de l'isocentre en longueur. Leur hauteur est optimisée pour observer l'isocentre tout en étant gêné le moins possible par l'accélérateur. Au final, la distance à l'isocentre est de 1.1 m.



FIGURE 6.22 – Vue d'ensemble du dispositif d'observation composé de deux caméras ToF fixés au plafond à l'aide de systèmes amovibles. Leur position est limitée par le champ de déplacement de l'accélérateur ainsi que par le système laser de repositionnement. Elle est optimisée avec comme critère la plus faible distance d'observation par rapport à l'isocentre et un champ de vue maximal afin d'englober l'ensemble du thorax et de l'abdomen.



FIGURE 6.23 – Le système de fixation de la caméra située à gauche par rapport à l’accélérateur est amovible, avec une barre métallique qui s’enfile dans un support visé au plafond. La caméra de droite est elle tout le temps fixée à un système mécanique permettant une position de repos non-gênante dans le faux plafond grâce à une fixation magnétique. La position baissée est assurée par une butée.

6.3.3 Système avec 1 seule caméra

Le protocole mis en place tient compte de la pratique clinique et cherche à intervenir le moins possible dans le travail du service. Il est demandé aux manipulateurs de déplacer le patient par rapport à l’isocentre, puis de le recentrer. Des acquisitions sont effectuées à ces deux instants. 39 fractions ont pu être observés ainsi. Deux localisations anatomiques font partie de la cohorte, le poumon et le bassin, qui ont été analysé de manière conjointe. La consigne donnée aux manipulateurs était d’appliquer un déplacement soit faible (inférieur à 2 centimètres) soit large.

Nous avons testé trois recalages des surfaces :

- L’ICP seul.
- L’ICP suivi de la sélection de l’information mutuelle et d’un nouvel ICP afin de compenser les différences des surfaces observées lors de larges déplacements,
- La deuxième approche précédée d’un filtrage gaussien 3D des données.

La figure 6.24 montre l’estimation de déplacement des 39 patients selon l’amplitude de déplacement. Pour le groupe ayant des mouvements supérieurs à 20 mm, la sélection des surfaces communes diminue fortement l’erreur, ce qui est logiquement moins le cas pour des déplacements d’amplitude plus faible. Concernant ce groupe, celui qui nous intéresse par-

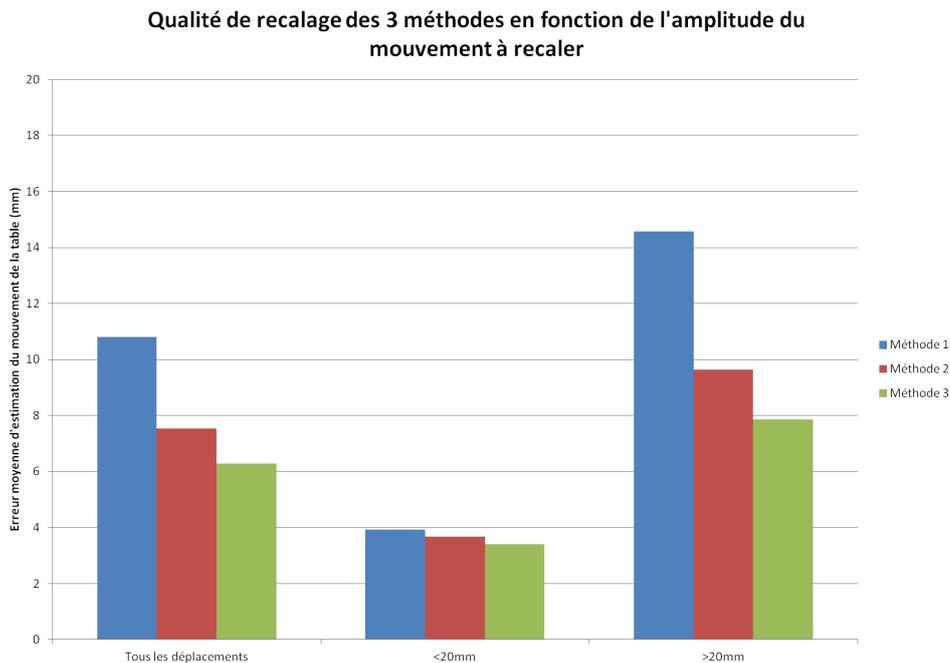


FIGURE 6.24 – Erreur moyenne d’estimation du déplacement d’un patient en fonction de l’amplitude de déplacement

ticulièrement vu que le repositionnement clinique dépasse rarement le centimètre, l’erreur moyenne est de 3.3 mm. Les résultats sont cependant très variables avec des erreurs pouvant atteindre 1 cm, ce qui peut être expliqué par le manque d’information de part le champ de vue, mais aussi à l’effet respiratoire pour les localisations pulmonaires. Ces résultats montrent de plus que pour des faibles déplacements, le filtrage gaussien n’apporte rien, ce qui peut s’expliquer par la perte de pente qui limite la capacité de l’ICP à converger vers la solution optimale. Cette perte est à peine compensée par la diminution du bruit apportée par le filtrage.

Ces premiers résultats indiquent de fortes disparités lors du recalage, ce qui nous a amené à installer une deuxième caméra afin d’améliorer la caractérisation de l’ensemble de la surface du patient.

6.3.4 Système avec 2 caméras

Nous avons pu suivre trois patients avec localisation pulmonaire au cours de leur traitement avec acquisition lors du J0, puis une semaine entière. La vérité terrain est ici le mouvement appliqué au patient lorsque l’imagerie portale détecte un mouvement par rapport à la position de référence à J0. Nous avons effectué trois types d’analyse :

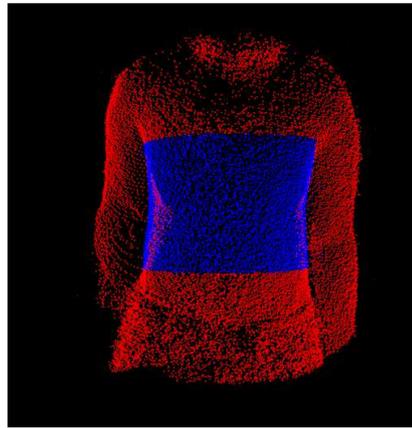


FIGURE 6.25 – La surface fournie par les deux caméras permet une description complète du tronc du patient malgré une plus faible densité de points en haut du thorax. La zone bleue correspond à la région prise pour estimer le déplacement

- la détection du mouvement appliqué au patient, soit la comparaison entre avant et après le déplacement.
- la conformité, soit le recalage entre la surface du patient après repositionnement et lors du J0.
- la prédiction, soit le recalage entre la surface du patient avant repositionnement et lors du J0.

Nous avons ainsi pu suivre trois patients lors de leurs cinq premières fractions de traitement, après consentement. Les résultats sont présentés figure 6.26 pour chaque catégorie d'estimation. Nous avons indiqué à droite les résultats de l'étude clinique de Moser et al. (2012) qui utilise le système Galaxy de LAP pour le repositionnement avec comparaison avec l'imagerie MVCT. L'estimation du déplacement est effectuée pour 12 fractions, et montre une erreur moyenne absolue de 0.8 mm, 2.8 mm et 2.4 mm selon les trois directions. Seule la direction latérale ne donne pas de résultats avec une forte variabilité. La direction cranio-caudale est entachée d'une forte variabilité sur les estimations avec des cas où le déplacement n'est quasiment pas détecté : pour exemple, lors d'une des fractions, un mouvement d'1 cm en direction de la tête est détecté comme un déplacement de 1,5 mm. L'estimation de déplacement selon la direction antéro-postérieure est fortement biaisée avec une erreur médiane de 2,3 mm. Les erreurs moyennes relatives sont de -0.1 mm, -0.9 mm et 2.3 mm selon les trois directions (respectivement LAT, LNG et VRT, cette convention étant appliquée pour la suite de ce chapitre), ce qui illustre une détection d'un mouvement antéro-postérieure lorsqu'il n'y en a pas, ce qui est le cas dans les 12 fractions en question. La détection des déplacements réels supérieurs à 5 mm est de 3/3 pour la direction latérale et 1/6 pour la direction cranio-caudale. Aucun déplacement vertical n'est observé, ce qui empêche de conclure pour cette direction anatomique.

Détection du mouvement : n= 12

	LAT (mm)	LNG (mm)	VRT (mm)	n=138		
Moyenne (rel)	-0.1	-0.9	2.3	0	0.7	-0.8
Moyenne (abs)	0.8	2.8	2.4			
extrema	-1.8/1.7	-8.5/5.4	-0.6/4.2	-2/3.1	-6.8/6.6	-4/1.3

Conformité du mouvement : n= 10

	LAT (mm)	LNG (mm)	VRT (mm)
Moyenne (rel)	0.6	0.8	-2.5
Moyenne (abs)	2.1	1.5	4
extrema	-2.2/7.7	-1.5/3.1	-8.9/3.6

Prédiction du mouvement : n= 9

	LAT (mm)	LNG (mm)	VRT (mm)	n=153		
Moyenne (rel)	-0.7	0.7	5.5	-0.1	1.3	-1.2
Moyenne (abs)	2	3.2	5.7			
extrema	-7.5/2.3	-5/3.1	-0.8/10.9	-3.3/4.6	-5.2/9.3	-4.9/4

FIGURE 6.26 – Résultats de la phase 2 de l'étude, sur trois patients, avec les erreurs de détection et l'estimation de la conformité ainsi que la prédiction du déplacement. A droite, les résultats de Moser et al. (2012) pour les localisations pulmonaires

La prédiction du déplacement est effectuée pour 10 fractions et montrent une erreur moyenne relative de -0.7 mm, 0.7 mm et 5.5 mm et une erreur moyenne absolue de 2 mm, 3.2 mm et 5.7 mm. La conformité entre les surfaces après repositionnement avec la surface de référence est de 0.6 mm, 0.8 mm et -2.5 mm en valeur relative et 2.1 mm, 1.5 mm et 4 mm en valeur absolue. Il est cependant difficile de tirer des conclusions de ces valeurs à cause de la médiocre détection du mouvement (erreur moyenne absolue de 0.8 mm, 2.8 mm et 2.4 mm) et au faible nombre de données. Toutefois, on peut observer une similitude des résultats avec ceux de Moser et al. (2012) pour la détection du déplacement : l'estimation de déplacement dans la direction latérale est plus précise que dans les deux autres directions, moins biaisée aussi. Ces résultats semblent indiquer un effet encore sensible du mouvement respiratoire malgré le moyennage de 10 acquisitions successives durant une seconde. En effet, les directions antéro-postérieures sont les plus touchées par ce mouvement.

La troisième partie de l'étude nous a amené à suivre des patients atteints de localisation thoraciques durant les 5 premières fractions, puis une fois par semaine ultérieurement.

Les objectifs sont multiples :

- Voir si la détection du déplacement est améliorée par la prise en compte du mouvement respiratoire
- Si oui, comparer la prédiction avec ce qui est réellement vu par l'imagerie portale
- Détecter la présence éventuelle d'une déviation au fil des semaines par changement d'anatomie du patient

Pour l'instant, les résultats sont disponibles pour deux patients, et sont présentés figure 6.27. L'erreur relative (respectivement absolue) de détection d'un déplacement est de -0.3 ± 0.8 mm, -0.3 ± 2.6 mm et -0.2 ± 0.9 mm (0.5 ± 0.6 mm, 1.8 ± 1.7 mm et 0.5 ± 0.7 mm) selon les directions latérales, cranio-caudales et antéro-postérieures. Cette détection est fortement améliorée par rapport aux résultats précédents, ce qui démontre l'apport de la prise en compte complète du mouvement respiratoire lors du choix des acquisitions à recalculer. L'amélioration est la plus sensible dans la direction antéro-postérieure (de 2.4 mm à 0.5 mm) même si elle est aussi significative dans les autres directions (de 0.8 mm à 0.5 mm en latéral et de 2.8 mm à 1.8 mm en tête-pieds). La précision atteinte pour estimer le déplacement nous permet d'interpréter le résultat des prédictions de mouvements et de conformité. Les différences moyennes de prédiction avec le système d'imagerie portale sont de 1.8 ± 1.3 mm, 2.1 ± 1.3 mm et 2.3 ± 2.6 mm. La détection des seuils de déplacement à 5 mm est de 2/2 selon la direction latérale et 3/6 dans la direction tête-pieds.

6.4 Synthèse et conclusion

Le travail effectué dans ce chapitre est centré autour de la mise en place et de la validation de la caméra ToF comme scanner 3D pouvant apporter un plus à la problématique du repositionnement en radiothérapie. Les premiers résultats présentés ici situent la caméra TOF dans l'état de l'art comme un instrument performant au niveau des autres imageurs de même type pour les localisations thoraciques, dont certains sont commercialisés pour de telles applications, ce qui amène à une plus large validation qui est en cours d'achèvement. Ces résultats ont été atteints grâce à deux apports méthodologiques : tout d'abord, l'utilisation de deux caméras ToF formant un système d'observation stéréo a permis une erreur d'estimation de repositionnement non biaisées par rapport à l'utilisation d'une seule caméra, et ceci grâce à l'extension du champ de vue permettant l'observation des deux flancs du patient ainsi que la redondance de l'information 3D dans la partie d'observation commune aux deux caméras. Nous avons proposé et développé une méthode de calibrage spatial de ces caméras : cette méthode est automatique et robuste, principalement pour l'estimation de la rotation entre les caméras. Des développements sont cependant à envisager afin d'améliorer

6.4. Synthèse et conclusion

Détection du mouvement : n= 8

	LAT (mm)	LNG (mm)	VRT (mm)	n=138		
Moyenne (rel)	-0.3±0.8	-0.3±2.6	-0.2±0.9	0	0.7	-0.8
Moyenne (abs)	0.5±0.6	1.8±1.7	0.5±0.7			
extrema	-1.9/0.4	-3.9/4.6	-1.9/0.8	-2/3.1	-6.8/6.6	-4/1.3

Conformité du mouvement : n= 6

	LAT (mm)	LNG (mm)	VRT (mm)
Moyenne (rel)			
Moyenne (abs)			
extrema			

Prédiction du mouvement : n= 6

	LAT (mm)	LNG (mm)	VRT (mm)	n=153		
Moyenne (rel)	-1.3±2	0.2±2.7	1.6±3.2	-0.1	1.3	-1.2
Moyenne (abs)	1.8±1.3	2.1±1.3	2.3±2.6			
extrema	-3.3/3.5	-4/1.7	-0.2/6.8	-3.3/4.6	-5.2/9.3	-4.9/4

FIGURE 6.27 – Résultats de la phase 3 de l'étude, sur trois patients, avec les erreurs de détection et l'estimation de la conformité ainsi que la prédiction du déplacement. A droite, les résultats de Moser et al. (2012) pour les localisations pulmonaires

l'estimation de la translation.

Le deuxième apport méthodologique concerne une première tentative de prise en compte de ce mouvement respiratoire en sélectionnant les surfaces en fonction de trois signaux de respiration décrivant le type et l'amplitude du mouvement respiratoire. Les tests sur patients mettent en relief son importance cruciale en réduisant significativement l'erreur (D'un facteur 10 entre le cas le plus favorable et celui le moins favorable). Cette méthode n'est qu'une étape puisqu'elle contient une possibilité d'erreur par la prise en compte de cycles respiratoires d'amplitude et d'offset différents. Il n'est cependant pas aisé d'extraire des signaux respiratoires correspondant aux mêmes localisations anatomiques à deux positions du patient différentes. Une des perspectives serait le développement de paramètres décrivant la forme de la surface, indépendamment de sa localisation : les surfaces à recalculer seraient ainsi sélectionnées selon leur similarité surfacique.

L'étude clinique a tout d'abord mis en relief l'influence de l'amplitude du déplacement à estimer sur la qualité de cette estimation : l'ensemble des déplacements supérieurs à 20 mm sont 3,5 fois moins bien estimés que ceux inférieurs. Dans un second temps, l'observation

avec deux caméras associée une gestion basique du mouvement respiratoire (moyennage sur 10 acquisitions) est testée. Les résultats indiquent une très bonne estimation du déplacement latéral (3/3 déplacements supérieurs à 5 mm détectés). Cependant, les déplacements dans la direction cranio-caudale et antéro-postérieure sont sous-estimés voire non-détectés : seulement 1/6 déplacement supérieur à 5 mm détecté. En conséquence, la prédiction du déplacement par rapport à la position à J0 est de faible qualité. Dans un troisième temps, la prise en compte du mouvement respiratoire est complexifiée. Les résultats sont améliorés, avec toujours une très bonne détection du mouvement dans la direction latérale. La détection dans la direction cranio-caudale est améliorée avec 3/6 détections. Plus intéressant, parmi les trois cas non-détectés, seul un n'est pas réellement détecté, les deux autres ayant des déplacements n'atteignant pas 4 mm mais supérieurs à 3 mm.

La place de ces résultats dans la littérature est d'ors et déjà digne d'intérêt. Le principal problème des systèmes commercialisés est la prise en compte des mouvements respiratoires, c'est pourquoi peu d'études se penchent sur ces localisations, et celles qui existent prennent en compte ce mouvement en moyennant des acquisitions successives. Le scanner laser a une vitesse de balayage d'une large région anatomique trop lente pour espérer sélectionner les acquisitions en fonction du moment respiratoire qui lui est associé. Les systèmes basés sur la lumière structurée rencontrent encore à l'heure actuelle des difficultés à reconstruire les points 3D sur de larges régions à des vitesses permettant la sélection des surfaces dans un délai raisonnable, et ceci alors que la vitesse d'acquisition est suffisante pour caractériser le mouvement respiratoire. La caméra ToF allie une haute fréquence d'acquisition avec une reconstruction des données 3D rapide, ce qui nous permet d'envisager au delà de l'aide au repositionnement entre les fractions un suivi dynamique des mouvements du patient pendant la fraction. Cette perspective est sans doute l'apport futur le plus intéressant d'un point de vue clinique. En effet, les scanners 3D ne visualisant pas les structures osseuses, un contrôle par imagerie ionisante restera sans doute indispensable afin de détecter des déviations dues à l'amaigrissement, par exemple. Par contre, le suivi en temps réel du mouvement du patient pourra dans un premier temps permettre l'arrêt de l'irradiation instantanément en cas de déplacement de la zone anatomique ciblée au delà d'une tolérance, afin de ne pas irradier à côté de la tumeur et ainsi garantir la délivrance complète de dose sans effet nocif supplémentaire. Dans un second temps, il sera imaginable d'asservir l'irradiation au mouvement respiratoire afin de diminuer les marges prises en compte lors de la planification du traitement.

Synthèse, conclusion générale et perspectives

La radiothérapie externe est aujourd'hui un maillon indispensable dans le traitement du cancer. Son efficacité dépend de la planification du traitement, de la qualité de l'irradiation et de la localisation journalière de la tumeur. Ce dernier point est pourtant soumis à deux facteurs d'erreurs qui sont, pour toutes les localisations, le repositionnement du patient à une même position spatiale par rapport à l'isocentre, et pour les tumeurs pulmonaires et du sein, le mouvement respiratoire provoquant un mouvement de la tumeur pendant la délivrance de dose. C'est dans l'optique d'évaluer la caméra Temps-de-Vol comme apport technologique pour l'amélioration de ces deux facteurs d'erreur que se situent ces travaux de thèse.

Bilan des travaux

La localisation de la tumeur en radiothérapie est un domaine où des évolutions sont nécessaires, la preuve en est la commercialisation de systèmes innovants proposant une alternative aux systèmes classiques de repositionnement d'un patient entre chaque fraction du traitement. De même, d'autres systèmes cherchent à suivre la localisation de la tumeur en temps réel afin de couper l'irradiation lorsque celle-ci s'éloigne trop de sa position de référence. D'autres asservissent l'accélérateur linéaire au mouvement estimé en temps réel. C'est dans ce cadre en pleine évolution que la caméra Temps-de-Vol pourrait apporter des solutions car elle fait partie de cette nouvelle génération de systèmes d'observation 3D en temps réel, au même titre que les scanners à lumière structurée ou à balayage laser. Nous pouvons même y ajouter maintenant la Kinect, même si son principe de mesure repose sur le principe de la lumière structurée. Ce type de caméras 3D apporte la possibilité de caractériser directement la surface complète du patient, à des vitesses ouvrant la porte à de l'analyse du mouvement en temps réel. Les études déjà menées approfondissent d'une part les applications déjà existantes dans lesquelles elles remplacent les systèmes de localisation surfacique tels que les marqueurs (suivi de la trace respiratoire, détection du mouvement de la tête lors d'une acquisition TEP...). Elles permettent aussi la mise en place de nouvelles méthodes telles que le repositionnement d'un patient à partir de l'analyse de sa surface. La lumière structurée comme le scanner à balayage laser font d'ailleurs déjà l'objet de commercialisation, les deux dans une optique d'aide au repositionnement.

C'est dans le cadre de cette évolution technologique que ce travail de thèse se situe. La caméra ToF connaît depuis une quinzaine d'années des évolutions techniques qui diminuent le bruit de ses mesures tout en augmentant sa résolution. Les domaines d'application sont multiples, la caméra apportant des propriétés nouvelles tel que l'accès direct à la mesure de profondeur et un encombrement minimal comparé aux autres scanners 3D. La dernière génération est évaluée comme ayant un bruit de mesure de l'ordre de 3-4 mm, ce qui ouvre de nouveaux domaines tels que les applications médicales.

Le premier travail effectué a consisté à évaluer les deux derniers modèles de caméras

ToF pour estimer le déplacement d'une surface en profondeur. Nous avons de même évalué la Kinect. Le but de cette première étape était de déterminer les ordres de grandeurs de la précision associée à chaque caméra. De plus, le positionnement d'une caméra par rapport à la surface d'intérêt étant souvent contraint par l'environnement, caractériser l'influence de la distance à la surface et l'angularité a permis la mise en place de dispositifs d'observation optimaux au regard des contraintes. Les expérimentations montrent que la Kinect est hautement performante à de distances inférieures à 0.8 m avec une erreur de mesure inférieure au millimètre. De plus, elle est capable de détecter un mouvement de très faible amplitude. A mesure que la distance d'observation augmente, la Kinect perd cette capacité à cause de la résolution en profondeur des mesures qui impose un pas de valeurs possibles des points 3D augmentant quadratiquement. Les résultats indiquent qu'au delà d'1.2 m, estimer le déplacement surfacique par une caméra ToF est plus précis qu'avec la Kinect. Plus précisément, le modèle SR4000 est globalement plus performant que le modèle Camcube 3.0, même si la différence est faible au-delà d'un mètre. Finalement, on peut établir un ensemble de conditions permettant le choix de l'une ou l'autre caméra par élimination :

- Si les conditions permettent une observation à moins d'un mètre, la Kinect est plus performante que les caméras ToF. Cette assertion est d'autant plus vraie que cette distance se raccourcit.
- Lorsque la caméra est située à plus d'1.2 m, les caméras ToF sont plus performantes. Dans ce cas, si le nombre de points décrivant la surface est un paramètre plus important que la précision atteignable au niveau du point, ou si les conditions d'observations assurent une luminosité stable, la CamCube 3.0 est un bon choix, malgré une précision moindre. Dans tout les autres cas, la SR4000 est plus adaptée.

Cette ensemble de conditions justifie pour la suite des travaux l'utilisation du modèle SR4000 : en effet, que ce soit l'observation d'un patient pendant un examen scanner 4D-TDM ou subissant une fraction de radiothérapie, les conditions d'observation empêchent d'observer une large fraction du tronc à moins d'1.1, 1.2 m. De même, dans les deux situations, un arrière plan éloigné et réfléchissant rend les conditions d'illuminations fragiles. De plus, dans le cas de la salle de traitement, l'accélérateur linéaire se déplace, modifiant ainsi les conditions. C'est pour cet ensemble de raisons que la caméra ToF SR4000 est utilisée, garantissant ainsi une précision stable grâce à l'utilisation de la fonction d'auto-illumination.

La deuxième partie de cette thèse a consisté à développer et valider une méthodologie permettant l'estimation du mouvement surfacique humain en profondeur, et ceci en temps réel et sur l'ensemble du torse et de l'abdomen. Nous avons pour ceci résolu la question du bruit de mesure de la surface (les mouvements mesurés étant inférieurs au centimètre, voire majoritairement inférieurs à 5 mm, une erreur de l'ordre de 3-4 millimètres rend délicat une quantification précise de ces mouvements) ainsi que de l'extraction de la déformation

en temps réel en modélisant les ensembles de points par des surfaces B-Splines. L'extraction de l'information locale décrivant la déformation surfacique est directement effectuée à partir des points de contrôle décrivant cette surface, l'ensemble du processus requérant moins de 200 ms sans optimisation par transfert sur GPU. Ceci nous a tout d'abord permis de diminuer substantiellement l'erreur d'estimation de mouvement d'un facteur 3 d'après des tests sur fantôme anthropomorphe, atteignant ainsi le millimètre d'erreur au niveau d'un seul point de contrôle. Après avoir caractérisé l'influence de la distance et de l'angle d'observation sur l'estimation du mouvement, la validation sur patient a été effectuée par comparaison avec des données scanner par l'observation de patients durant un examen 4D-TDM. Le champ d'observation du scanner étant limité à 30 centimètres, seul le thorax et une légère partie de l'abdomen est visible sur les données. Nous avons donc validé l'estimation de mouvement sur des régions anatomiques dont l'amplitude de déplacement dans la direction antéro-postérieure ne dépasse généralement pas 4 mm. Malgré cela, la concordance est très bonne pour des régions d'intérêt pouvant atteindre la taille de $4 \times 3 \text{ cm}^2$, et constitué de 9 à 18 points de contrôle : à cette taille là, l'erreur moyenne pour les 4 patients dont la concordance spatiale et temporelle est parfaite est inférieure à 0.4 mm. La concordance spatiale entre les données scanner et les données issues de la caméra ToF est validée au moins pour la région abdominale haute, la région thoracique étant sujette à des erreurs à cause des conditions d'observations empêchant une bonne observation de cette région. D'un point de vue quantitatif, l'estimation du mouvement d'une région anatomique est potentiellement entachée par deux "erreurs" par rapport aux données scanner : le binning des données ToF, acquises à 10 Hz, rend sensible la sélection des acquisitions, un offset en phase pouvant apparaître. De même, une erreur en amplitude peut apparaître si l'on considère un cycle respiratoire d'amplitude supérieure à la respiration moyenne durant l'acquisition. Les résultats sur les 4 patients non-touchés par ces sources d'erreurs sont une erreur moyenne de 0.26 mm ($4 \times 3 \text{ cm}^2$), 0.24 mm ($8 \times 3 \text{ cm}^2$) et 0.28 mm ($20 \times 3 \text{ cm}^2$).

Forts de ces résultats démontrant la capacité à extraire le mouvement respiratoire en profondeur pour des régions anatomiques de taille réduite, nous avons utilisé le modèle de respiration patient-spécifique préalablement développé au laboratoire afin de générer des volumes 4D-TDM à partir de volumes pré-existants et de la déformation surfacique du patient, le mouvement respiratoire externe étant considéré comme le déplacement de 10 régions d'intérêt. Nous avons généré deux volumes TDM, à 40% de la phase d'inspiration et 70% de celle d'expiration. L'erreur moyenne de génération sur 13 points anatomiques décrivant le poumon est alors de 3.45 mm. Nous avons généré de même ces volumes en considérant le mouvement externe depuis les surfaces des volumes réels TDM segmentés, composé des mêmes régions le long de la surface thoraco-abdominale. Cette génération optimale (les surfaces des TDM étant normalement très peu bruitées, tout en étant fortement corrélées aux données TDM décrivant le volume interne) conduit à une erreur moyenne de 2.61 mm. Enfin, la génération des mêmes volumes selon une interpolation naïve linéaire montre une erreur

de 6.15 mm. Même si ces résultats ne concernent qu'un seul cas, ils sont encourageants, montrant que remplacer l'information externe issues des volumes TDM par une observation avec une caméra ToF ne diminue pas excessivement la génération des volumes anatomiques. D'autres cas traités sont en cours, reproduisant le même protocole et testant différentes prises en compte de l'information externe.

Nous avons ensuite développé une méthode permettant l'estimation de la déformation 3D de l'ensemble de la surface thoraco-abdominale. Lorsque la surface est décrite par des 3D points qui ne sont pas reliés à la même localisation anatomique au fil des acquisitions, le recalage élastique demeure la principale solution pour estimer le déplacement local au niveau de la surface. De nombreux développements cherchent à réduire le temps de calcul qui reste une réelle limite à toute estimation en temps réel. Nous avons proposé la fusion des informations 2D et 3D comme voie alternative de l'estimation de la déformation en déterminant le mouvement 2D latéral par recherche du flot optique sur les images d'intensité, puis en contraignant la déformation 3D par cette information. Concrètement, chaque point 3D étant lié physiquement à un pixel dans les images 2D, le déplacement de ce pixel entre deux images décrit parallèlement le déplacement du point 3D associé. Plus précisément, une localisation anatomique peut alors être suivie au fil des acquisitions grâce aux images d'intensité. Cependant, la description échantillonnée des surfaces conduit à chercher une nouvelle localisation n'existant probablement pas dans le nuage de point décrivant la nouvelle surface. Nous avons proposé la modélisation des surfaces par des surfaces B-Splines afin d'avoir une description continue de celles-ci, et déterminer la position 3D pointée par le flot optique dans l'espace des images 2D.

Après vérification du respect de la contrainte d'illumination constante sur les images d'intensité fournies par la caméra ToF, nous avons comparé la qualité d'estimation du flot optique sur les images d'intensité fournies par la caméra ToF SR4000, et ceci avec les deux méthodes les plus usitées, les méthodes variationnelles à application de la contrainte globale (Horn and Schunck, 1981) et locale (Lucas et al., 1981). Les résultats indiquent l'influence de l'amplitude du déplacement à estimer : au-delà de quatre pixels, l'erreur augmente fortement, les méthodes ayant tendance à sous-estimer le déplacement, surtout celle de Horn and Schunck (1981). En deçà, Lucas et al. (1981) est significativement moins performante. Nous avons ensuite validé l'extraction de la déformation 3D par comparaison avec le mouvement de marqueurs visuels dessinés sur des volontaires. La différence entre mouvement 3D réel des marqueurs et l'estimation par la méthodologie proposée avec extraction du flot optique par Horn and Schunck (1981) est de 1.7 mm pour 5 volontaires testés avec pour chacun 9 à 13 marqueurs. De même, nous avons cherché à comparer notre méthodologie avec un algorithme de recalage élastique, l'algorithme CPD développé par Myronenko and Song (2010), avec toujours cette analyse du mouvement de marqueurs optiques. Cette validation n'a pas été réalisable sur des données issues des caméras ToF, le recalage élastique conduisant à des résultats incohérents. Nous avons donc travaillé sur des données issues de la Kinect car elles

sont très peu bruitées et donc plus adaptées à un recalage élastique avec cet algorithme. Notre méthodologie y présente les meilleurs résultats avec une erreur moyenne de 0.7 mm contre 1.3 mm.

Nous avons aussi comparé nos résultats avec ceux fournis par la méthode de Bauer et al. (2012c) qui propose d'utiliser directement le flot optique pour interpoler une nouvelle position sur la nouvelle surface. Les résultats ne montrent pas une différence significative, même si un autre type de validation serait nécessaire afin de départager les méthodes, avec une vérité terrain issues de capteurs, optiques ou magnétiques. Quoi qu'il en soit, la validation montre une précision intéressante, que ce soit pour des données ToF ou Kinect. De plus, la méthode est plus performante que la méthode de recalage élastique PCD. L'erreur moyenne est plus faible sur les données Kinect (0.7 mm contre 1.7 mm). Enfin, le temps de calcul associé, de l'ordre de 500 ms, ouvre d'ors et déjà la porte à une estimation en temps réel.

L'utilisation de scanners 3D comme complément aux imageurs de contrôle afin de détecter la déviation du positionnement par alignement de lasers permettrait de limiter la dose additionnelle appliquée au patient. Plus loin encore, il est envisageable de compléter voire remplacer le positionnement par lasers par l'alignement des surfaces externes du patient au regard de celle de référence. La quantité bien plus grande d'information justifierait du moins une telle opportunité. Des scanners à lumière structurée ainsi qu'à balayage laser sont déjà commercialisés dans cette optique. Nous avons cherché à valider de la même manière la caméra ToF. Pour ceci, nous avons développé un système stéréo composé de deux caméras ToF afin d'agrandir le champ de vue du patient qui contient ainsi l'intégralité des deux flancs et une information double sur la partie centrale du tronc. La méthode de calibrage repose sur l'observation d'un cube dans l'espace de travail afin de déterminer la rotation et la translation entre les deux caméras ToF, et ceci de manière automatique. L'expérience montre une très bonne estimation de la composante rotation, alors que la translation n'est pas assez précise. L'impact sur le repositionnement est testé sur fantôme et volontaire, montrant un impact décisif avec une erreur moyenne de 1 mm selon la direction latérale et 0.9 mm dans la direction antéro-postérieure. Seule la direction cranio-caudale est sujette à de larges erreurs atteignant 3.4 mm. L'impact de la prise en compte du mouvement respiratoire est illustrée par expérimentations sur volontaire et un vrai patient, montrant que pour une translation de 8 mm dans la direction cranio-caudale, considérer les surfaces à recalcr à un même instant respiratoire selon 3 traces permettait de réduire l'erreur d'estimation à 0.9 mm au lieu de 10 mm dans le cas le plus défavorable lorsque les surfaces comparées sont à instant respiratoire opposé.

L'étude clinique menée est divisée en trois étapes, dont la troisième est en cours. Dans un premier temps, l'observation d'un patient sur une fraction avec une seule caméra en limitant le mouvement respiratoire par moyennage sur 10 acquisitions successives a montré des limites avec une sous-estimation du déplacement dans les trois directions, associées à un écart type de plusieurs millimètres. Dans un second temps, l'installation d'une deuxième caméra

et l'étude de patients pendant 5 fractions successives a tout d'abord montré la diminution significative de la sous-estimation du mouvement, phénomène dû à la meilleure contrainte sur le recalage des surface apportée par l'augmentation du champ de vue. Cependant, les mouvements dans les directions antéro-postérieures et cranio-caudales sont encore mal estimée avec une erreur moyenne absolue de 2.8 mm et 2.4 mm respectivement. La troisième phase prend en compte le mouvement respiratoire en sélectionnant les surfaces par similarité des trois instants respiratoires associés à trois régions d'intérêt, abdominale, intermédiaire et thoracique. Les premiers résultats montrent une réduction significative de la sous-estimation dans la direction antéro-postérieure (de 2.3 mm à -0.3 mm). De même, l'erreur moyenne absolue diminue pour les trois directions : de 0.8 à 0.5 mm selon x, de 2.8 mm à 1.8 mm selon y et de 2.4 à 0.5 mm selon z. Il manque des données supplémentaires pour valider définitivement cet ordre de grandeur qui est extrêmement intéressant par comparaison avec les autres systèmes proposés. La détection de déplacements supérieurs à 5 mm est toujours aussi performante pour la direction latérale. La prise en compte du mouvement respiratoire améliore fortement les résultats dans la direction cranio-caudale, avec un taux de détection passant de 1/6 à 3/6, avec un seul cas où le déplacement détecté ne dépasse pas la barre des 3 mm.

Perspectives

Il y a certains points méthodologiques qui requièrent une amélioration, par choix d'une méthode plus performante ou par validation supplémentaire. Une première validation complémentaire concerne le suivi du mouvement respiratoire externe. Afin de confirmer les résultats obtenus, mais cette fois sur la région abdominale, le même type d'expérimentation serait à effectuer avec un examen 4D-TDM ayant un plus grand champ d'observation afin de contenir cette région qui est touchée prioritairement par le mouvement respiratoire. Afin d'observer au mieux l'ensemble du tronc tout en limitant les effets de bord comme c'est le cas dans la région thoracique pour les expérimentations du chapitre 4. Une possibilité serait d'observer le patient par les deux côtés du tunnel du scanner, en mettant en place le système d'observation stéréoscopique présenté section 6.1. Chaque caméra garantirait alors l'observation optimale des sections hautes et basses du tronc du patient. L'impact de la vitesse d'acquisition (et donc du temps d'intégration des données) sera aussi un point à étudier. Dans ce travail de thèse, la vitesse d'acquisition est fixée à 10 Hz, vitesse suffisante pour permettre un binning. Cependant, une phase d'inspiration durant entre une et deux secondes, le nombre d'acquisitions la décrivant est au mieux de 20 acquisitions. Dans le cas de respirations plus courtes, 10 acquisitions seulement décrivent l'inspiration et l'expiration. C'est dans ce cas où la moindre erreur lors de la sélection des acquisitions induit un décalage de phase entre les signaux. L'augmentation de la fréquence d'acquisition, si elle n'induit pas de perte significative de précision, limitera cet effet grâce une meilleure résolution temporelle.

Le premier point méthodologique est l'amélioration de la mise en correspondance spatiale des données TDM et de la caméra ToF. Une des pistes à suivre serait d'extraire des points d'intérêt anatomiques fixes, comme par exemple le sommet des côtes ou de la poitrine, afin de réduire l'erreur de localisation. La deuxième amélioration potentielle est de tester la sélection des surfaces sur plusieurs cycles respiratoires afin d'analyser l'impact sur la robustesse de l'estimation du déplacement. Il sera alors nécessaire de tenir compte de la variabilité des cycles respiratoires, tant en amplitude qu'en phase, afin de ne pas considérer par erreur deux acquisitions comme étant au même instant respiratoire alors qu'elles décrivent deux phases différentes. La question de la variation en amplitude des cycles respiratoires rend difficile la comparaison avec des données scanner qui décrivent au final un cycle respiratoire moyen au long de l'acquisition.

L'estimation du flot optique sur des données ToF est marquée par une sous-estimation des déplacements supérieurs à quatre pixels et par des artefacts surfaciques. Les approches variationnelles classiques fournissent des résultats satisfaisants, mais tester une méthode cherchant à lier les deux types d'approches, limitant ainsi le temps de calcul tout en améliorant la robustesse par rapport à une méthode globale pourrait diminuer les artefacts (Bruhn et al., 2005). Dans tous les cas, le transfert sur cartes GPU du traitement des données devra permettre d'atteindre une vitesse d'analyse en temps réel. La modélisation B-Splines ainsi que le calcul du flot optique ont à priori des structures algorithmiques adaptées à un tel transfert. Pour finir, des validations plus poussées sont nécessaires, la méthode étant comparées à des marqueurs optiques dont la détection est limitée de part la faible résolution des caméras ToF. Cependant, l'application de la méthode sur les données issues de la Kinect fournit de bons résultats, alors que la résolution de cette caméra autorise une détection des marqueurs optiques bien plus précise. Le choix d'autres méthodes de recalage élastique permettra une validation plus poussées, et à cet égard les dernières méthodes développées (Schaerer et al. (2012), Bauer et al. (2012a)) peuvent fournir de bons points de comparaison. Une validation avec estimation du déplacement de marqueurs optiques fixés sur la peau, et donc visibles sur les images d'intensité, est envisageable si l'émission de lumière infrarouge par le système n'interfère pas avec la caméra ToF.

La validation de l'utilisation de caméras ToF pour le repositionnement n'est pas complète, mais les résultats obtenus suggèrent des performances au moins similaires aux systèmes existants. Cependant, les points d'améliorations concernent tout d'abord la méthode de calibrage présentée. Elle répond au cahier des charges exigeant une méthode rapide afin de pouvoir créer rapidement un système stéréoscopique. La détermination de la composante translationnelle de la transformation reliant les caméras n'est cependant pas assez bonne, ce qui en pratique impacte peu sur le recalage des surfaces car la translation recherchée est alors effectuée sur des surfaces ne décrivant pas la véritable anatomie, mais ne se déformant pas après translation, ce qui serait le cas si la rotation était mal estimée. Il est cependant important de compléter la méthode avec une estimation de la translation plus performante afin qu'elle

soit utilisable dans d'autres applications où la surface observée doit correspondre à la réalité, comme dans le cas cité ci-dessus où deux caméras observeraient le mouvement respiratoire externe d'un patient pendant une acquisition scanner. Un autre point d'amélioration concerne le traitement des données brutes. Dans le cas de l'étude, un filtrage gaussien est appliqué aux données. Ce traitement n'est pas optimal car il ne garantit pas la conservation des courbes et reliefs, ce qui peut induire une erreur de recalage principalement dans la direction cranio-caudale qui est la direction dans laquelle la surface anatomique possède le moins de relief. Des méthodes basées sur la diffusion anisotrope sont sans doute plus adaptées au traitement des données de la caméra. Le filtrage bilatéral pourrait alors apporter une solution : basé sur le même principe que le filtrage gaussien, le poids des voisins sur le filtre appliqué dépend aussi de la distance euclidienne au point concerné, mais aussi de la variation locale de couleur ou d'intensité, ce qui rend local le filtre appliqué aux données. Des tests préliminaires sont en cours pour évaluer l'apport d'un tel filtre.

Les travaux qui découlent directement de ces travaux sont l'amélioration de la caractérisation du mouvement respiratoire externe lors de la création des modèles de respiration. Mais avant tout, après avoir testé la génération de volumes TDM par modèle de respiration patient-spécifique, nous envisageons de créer le modèle global, qui permettra pour un patient de générer l'ensemble des volumes TDM à partir d'un examen classique à expiration et inspiration complète. Il semble logique de confirmer l'apport de la caractérisation du mouvement surfacique par 10 régions d'intérêt. L'évolution logique est d'apporter l'information 3D de la déformation surfacique pour caractériser cette déformation. Dans un premier temps, remplacer le déplacement 1D des 10 régions d'intérêt par leur mouvement 3D pourrait améliorer la corrélation mouvement externe/déformation interne. Il faudra pour ceci améliorer l'estimation du flot optique en appliquant une approche plus robuste tout en n'augmentant pas significativement le temps de calcul, quitte à transférer le calcul sur carte graphique. Dans un second temps, remplacer la notion de région d'intérêt par une paramétrisation de la déformation permettra de simplifier le nombre de paramètres du modèle, permettant son application en temps réel afin de prédire le déplacement tumoral.

L'application à long terme sera alors la prédiction du déplacement tumoral, afin dans un premier temps d'asservir l'irradiation à un espace duquel la sortie de la tumeur entrainera un arrêt du traitement, puis à long terme d'asservir, ou l'accélérateur, ou la table de traitement afin de déplacer la tumeur de manière à compenser son mouvement. L'utilisation d'une caméra temps-de-vol pour générer les volumes 4D-TDM à partir de deux Scanners normaux en utilisant le modèle de respiration global permettra une première réduction de dose pour le patient lors de la phase de la planification du traitement. Durant le traitement, l'application du modèle étant robuste au fil des fractions, idéalement aucune image additionnelle ne sera nécessaire pour confirmer la corrélation pré-établie, au contraire du système Cyberknife par

exemple. De même, la caméra ToF comme instrument d'observation en temps réel n'implique aucune interaction avec le patient, comme pour le système Calypso où une insertion de marqueurs est requise. Les développements présentés dans ce manuscrit montrent la possibilité d'estimer en temps réel la déformation surfacique du patient, en 1D comme en 3D. Il reste à confirmer la création des modèles de respirations robustes et précis, ainsi que la validité de la relation de corrélation pour un patient au fil des fractions de traitement.

La suite de l'étude suivant des patients pour des localisations pulmonaires confirmera la capacité du système d'observation en terme de précision atteignable. La prochaine étape concernera l'étude du mouvement du patient pendant la fraction, ce mouvement pouvant provenir de réflexes ou d'une détente musculaire au fil des minutes. Utilisé en combinaison avec le modèle de respiration, il est envisageable d'estimer le déplacement de la tumeur lors d'une séance, ainsi que son impact en terme de délivrance de dose. C'est clairement un domaine dans lequel la caméra ToF peut apporter ses qualités d'acquisition et de construction des données 3D qui sont effectuées dans nos expériences à une fréquence de 10 Hz. Couplée aux perspectives présentées précédemment visant à prédire la position d'une tumeur en temps réel, elle permettrait là aussi un contrôle de l'irradiation en cas de déplacement significatif au fil des minutes. Au final, la caméra ToF pourra analyser en temps réel le déplacement tumoral causé par, d'un côté le mouvement respiratoire, de l'autre les mouvements involontaires du patient.



Bibliographie

- Amberg, B., Romdhani, S., and Vetter, T. (2007). Optimal step nonrigid icp algorithms for surface registration. *Proc. IEEE Conf. Comp. Vis. Patt. Recog. (Minneapolis, MN)*, pages 1–8.
- Aubry, J.-F., Beaulieu, L., Girouard, L.-M., Aubin, S., Tremblay, D., Laverdière, J., and Vigneault, E. (2004). Measurements of intrafraction motion and interfraction and intrafraction rotation of prostate by three-dimensional analysis of daily portal imaging with radiopaque markers. *International Journal of Radiation Oncology* Biology* Physics*, 60(1) :30–39.
- B Allen, B. C. and Popovic, Z. (2003). The space of human body shapes : reconstruction and parameterization from range scans. *ACM Trans. Graph.*, 22 :587–94.
- Barjatya, A. (2004). Block matching algorithms for motion estimation. *IEEE Transactions Evolution Computation*, 8(3) :225–239.
- Barnes, E. A., Murray, B. R., Robinson, D. M., Underwood, L. J., Hanson, J., and Roa, W. H. (2001). Dosimetric evaluation of lung tumor immobilization using breath hold at deep inspiration. *International Journal of radiation oncology, biology, physics*, 50(4) :1091–1098.
- Bartczak, B., Koeser, K., Woelk, F., and Koch, R. (2007). Extraction of 3d freeform surfaces as visual landmarks for real-time tracking. *Journal of Real-time Image Processing*, 2(2-3) :81–101.
- Bartczak, B., Schiller, I., Beder, C., and Koch, R. (2008). Integration of a time-of-flight camera into a mixed reality system for handling dynamic scenes, moving viewpoints and occlusions in realtime. *Int. Symp. on 3D Data Processing, Visualization and Transmission (3DPVT)*.
- Bartoncini, S., Fiandra, C., Ruo Redda, M., Allis, S., Munoz, F., and Ricardi, U. (2012). Target registration errors with surface imaging system in conformal radiotherapy for prostate cancer : study on 19 patients. *La Radiologia Medica*, pages 1–10.

-
- Bauer, S., Berkels, B., Ettl, S., Arold, O., Hornegger, J., and Rumpf, M. (2012a). Markerless reconstruction of dense 4-d surface motion fields using active laser triangulation for respiratory motion management. *Med Image Comput Assist Interv*, 15(1) :412–421.
- Bauer, S., Berkels, B., Hornegger, J., and Rumpf, M. (2012b). Joint tof image denoising and registration with a ct surface in radiation therapy. *Scale Space and Variational Methods in Computer Vision*, pages 98–109.
- Bauer, S., Wasza, J., Haase, S., Marosi, N., and Hornegger, J. (2011). Multi-modal surface registration for markerless initial patient setup in radiation therapy using microsoft’s kinect sensor. In *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, pages 1175–1181. IEEE.
- Bauer, S., Wasza, J., and Hornegger, J. (2012c). Photometric estimation of 3d surface motion fields for respiration management. *Bildverarbeitung für die Medizin 2012*, pages 105–110.
- Bauer, U. and Polthier, K. (2008). Detection of planar regions in volume data for topology optimization. *Advances in Geometric Modeling and Processing*, pages 119–126.
- Beauchemin, S. S. and Barron, J. L. (1995). The computation of optical flow. *ACM Computing Surveys (CSUR)*, 27(3) :433–466.
- Becerro, F. G. (2008). External-self-calibration of a 3d time-of-flight camera in real environments. *system*, 4(3) :33.
- Bert, C., Metheany, K. G., Doppke, K., and Chen, G. T. (2005). A phantom evaluation of a stereo-vision surface imaging system for radiotherapy patient setup. *Medical physics*, 32 :2753.
- Besl, P. J. and McKay, N. D. (1992). A method for registration of 3-d shapes. *IEEE Transactions on pattern analysis and machine intelligence*, 14(2) :239–256.
- Blomgren, H., Lax, I., Naslund, I., and Svantrom, R. (1995). Stereotactif high dose fraction radiation therapy of extracranial tumors using an accelerator. clinical experience of the first thirty-one patients. *Acta Oncol*, 34(6) :861–870.
- Boehme, M., Haker, M., Martinetz, T., and Barth, E. (2008). A facial feature tracker for human-computer interaction based on 3d tof cameras. *Int. J. on Intell. Systems Techn. and App., Issue on Dynamic 3D Imaging*, 5(3-4) :264–273.
- Böller, A. (2010). Evaluation of 3d camera technology for patient positioning within the field of radiation therapy.

-
- Bondiau, P., Bénézery, K., Beckendorf, V., Peiffert, D., Gérard, J.-P., Mirabel, X., Noel, A., Marchesi, V., Lacornerie, T., Dubus, F., Sarrazin, T., Hérault, J., Marcié, S., Angellier, G., and Lartigau, E. (2007). Radiothérapie stéréotaxique robotisée par cyberknife : aspects techniques et indications. *Cancer et Radiothérapie*, 11(6-7) :338–344.
- Bortfeld, T., Jokivarsi, K., Goitein, M., Kung, J., and Jiang, S. B. (2002). Effects of intra-fraction motion on imrt dose delivery : statistical analysis and simulation. *Physics in medicine and biology*, 47(13) :2203.
- Bouilhol, G., Ayadi, M., Rit, S., Thengumpallil, S., Schaerer, J., Vandemeulebroucke, J., Claude, L., and Sarrut, D. (2012). Is abdominal compression useful in lung stereotactic body radiation therapy ? a 4dct and dosimetric lobe-dependent study. *Physica Medica*.
- Brahme, A., Nyman, P., and Skatt, B. (2008). 4d laser camera for accurate patient positioning, collision avoidance, image fusion and adaptive approaches during diagnostic and therapeutic procedures. *Medical physics*, 35 :1670.
- Bruhn, A., Weickert, J., and Schnörr, C. (2005). Lucas/kanade meets horn/schunck : Combining local and global optic flow methods. *International Journal of Computer Vision*, 61(3) :211–231.
- Bryan, P., Custar, S., Haaga, J., and Balsara, V. (1984). Respiratory movement of the pancreas : an ultrasonic study. *Journal of Ultrasound in Medicine*, 3(7) :317–320.
- Chen, Q.-S., Weinhaus, M. S., Deibel, F. C., Ciezki, J. P., and Macklis, R. M. (2001). Fluoroscopic study of tumor motion due to breathing : facilitating precise radiation therapy for lung cancer patients. *Medical physics*, 28 :1850.
- Chetty, I., Rosu, M., McShan, D., Fraass, B., Balter, J., and Tenhaken, R. (2004). Accounting for center-of-mass target motion using convolution methods in monte carlo-based dose calculations of the lung. *Medical Physics*, 31(4) :925–932.
- Chiabrando, F., Piatti, D., and Rinaudo, F. (2010). Sr-4000 tof camera : further experimental tests and first applications to metric surveys. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*, 38(5) :149–154.
- Chowdhary, S., Cloward, T., Shilling, K., Berjaoui, W., Farney, R., Nuttall, T., and Walker, J. (2011). Correlation of spirometry, sleep disordered breathing, and nocturnal hypoxemia in severely obese subjects. *CHEST Journal*, 140(4_MeetingAbstracts) :943A–943A.
- Clippe, S., Sarrut, D., Malet, C., Miguet, S., Ginestet, C., and Carrie, C. (2003). Patient setup error measurement using 3d intensity-based image registration technique. *International Journal of radiation oncology, biology, physics*, 56(1) :259–265.

-
- Crabb, R., Tracey, C., Puranik, A., and Davis, J. (2011). Comparison of depth super-resolution methods for 2d/3d images. *IEEE Conf. on Computer Vision and Pattern Recogn. ; Workshop on ToF-Camera based Computer Vision*, 3 :635–645.
- Davies, S., Hill, A., Holmes, R., Halliwell, M., and Jackson, P. (1994). Ultrasound quantitation of respiratory organ motion in the upper abdomen. *British journal of radiology*, 67(803) :1096–1102.
- Eggbeer, D., Bibb, R., Evans, P., and Ji, L. (2012). Evaluation of direct and indirect additive manufacture of maxillofacial prostheses. *Proceedings of the Institution of Mechanical Engineers. Part H, Journal of engineering in medicine*, 226(9) :718–728.
- Ekberg, L., Holmberg, O., Wittgren, L., Bjelkengren, G., and Landberg, T. (1998). What margins should be added to the clinical target volume in radiotherapy treatment planning for lung cancer ? *Radiotherapy and oncology*, 48(1) :71–77.
- Engelsman, M., Damen, E. M., De Jaeger, K., van Ingen, K. M., and Mijnheer, B. J. (2001). The effect of breathing and set-up errors on the cumulative dose to a lung tumor. *Radiotherapy and Oncology*, 60(1) :95–105.
- Erridge, S. C., Seppenwoolde, Y., Muller, S. H., van Herk, M., De Jaeger, K., Belderbos, J., Boersma, L. J., Lebesque, J. V., et al. (2003). Portal imaging to assess set-up errors, tumor motion and tumor shrinkage during conformal radiotherapy of non-small cell lung cancer. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, 66(1) :75.
- Falie, D. and Buzuloiu, V. (2007). Noise characteristics of 3d time-of-flight cameras. In *Signals, Circuits and Systems, 2007. ISSCS 2007. International Symposium on*, volume 1, pages 1–4. IEEE.
- Falie, D. and Buzuloiu, V. (2008). Wide range time of flight camera for outdoor surveillance. In *Microwaves, Radar and Remote Sensing Symposium, 2008. MRRS 2008*, pages 79–82. IEEE.
- Fayad, H., Pan, T., Clement, J., and Visvikis, D. (2011). Technical note : Correlation of respiratory motion between external patient surface and internal anatomical landmarks. *Medical physics*, 38(6) :3157–3164.
- Fayad, H., Pan, T., Roux, C., Rest, C. C.-L., Pradier, O., and Visvikis, D. (2009). 2d-spline patient specific model for use in radiation therapy. *IEEE International Symposium on Biomedical Imaging : From Nano to Macro*, pages 590–593.

-
- Fayad, H., Pan, T., Roux, C., and Visvikis, D. (2010). A generic respiratory motion model for motion correction in pet/ct. In *Nuclear Science Symposium Conference Record (NSS/MIC), 2010 IEEE*, pages 2455–2458. IEEE.
- Feldmar, J. and Ayache, N. (1996). Rigid, affine and locally affine registration of free-form surfaces. *Int. J. Comput. Vis.*, 18 :99–119.
- Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus : a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6) :381–395.
- Foix, S., Alenya, G., and Torras, C. (2011). Lock-in time-of-flight (tof) cameras : a survey. *Sensors Journal, IEEE*, 11(9) :1917–1926.
- Foix Salmerón, S., Alenyà Ribas, G., Torras Genís, C., et al. (2010). Exploitation of time-of-flight (tof) cameras.
- Ford, E., Mageras, G., Yorke, E., and Ling, C. (2003). Respiration-correlated spiral ct : a method of measuring respiratory-induced anatomic motion for radiation treatment planning. *Medical physics*, 30 :88.
- Ford, E., Mageras, G., Yorke, E., Rosenzweig, K., Wagman, R., and Ling, C. (2002). Evaluation of respiratory movement during gated radiotherapy using film and electronic portal imaging. *International Journal of Radiation Oncology* Biology* Physics*, 52(2) :522–531.
- Foroud, T., Wetherill, L., Vinci-Booher, S., Moore, E., Ward, R., Hoyme, E., Robinson, L., Rogers, J., Meintjes, E. M., Molteno, C., Jacobson, J., and Jacobson, S. (2012). Relation over time between facial measurements and cognitive outcomes in fetal alcohol-exposed children. *Alcoholism : Clinical and Experimental Research*, 36(9) :1634–1646.
- Fraass, B., Doppke, K., Hunt, M., Kutcher, G., Starkschall, G., Stern, R., and Van Dyke, J. (1998). American association of physicists in medicine radiation therapy committee task group 53 : quality assurance for clinical radiotherapy treatment planning. *Medical physics*, 25 :1773.
- Fuchs, S. and Hirzinger, G. (2008). Extrinsic and depth calibration of tof-cameras. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–6. IEEE.
- Gallo, O., Manduchi, R., and Rafii, A. (2008). Robust curb and ramp detection for safe parking using the canesta tof camera. volume 1-3, pages 1558–1565.

-
- George, R., Keall, P., Kini, V., Vedam, S., Siebers, J., Wu, Q., Lauterbach, M., Arthur, D., and Mohan, R. (2003). Quantifying the effect of intrafraction motion during breast imrt planning and dose delivery. *Medical physics*, 30 :552.
- George, R., Vedam, S., Chung, T., Ramakrishnan, V., and Keall, P. (2005). The application of the sinusoidal model to lung cancer patient respiratory motion. *Medical Physics*, 32(9) :2850–2861.
- Giraud, P., De Rycke, Y., Dubray, B., Helfre, S., Voican, D., Guo, L., Rosenwald, J.-C., Keraudy, K., Housset, M., Touboul, E., et al. (2001). Conformal radiotherapy (crt) planning for lung cancer : analysis of intrathoracic organ motion during extreme phases of breathing. *International journal of radiation oncology, biology, physics*, 51(4) :1081.
- Golub, G. H. and Van Loan, C. F. (1996). *Matrix computations*, volume 3. Johns Hopkins University Press.
- Gray, H. (1918). *Anatomy of the human body, 20th ed.* Philadelphia : Lea and Febiger.
- Grills, I. S., Yan, D., Martinez, A. A., Vicini, F. A., Wong, J. W., and Kestin, L. L. (2003). Potential for reduced toxicity and dose escalation in the treatment of inoperable non-small-cell lung cancer : A comparison of intensity-modulated radiation therapy (imrt), 3d conformal radiation, and elective nodal irradiation. *International Journal of Radiation Oncology* Biology* Physics*, 57(3) :875–890.
- Guomundsson, S. A., Aanæs, H., and Larsen, R. (2007). Environmental effects on measurement uncertainties of time-of-flight cameras. In *Signals, Circuits and Systems, 2007. ISSCS 2007. International Symposium on*, volume 1, pages 1–4. IEEE.
- Hanley, J., Debois, M. M., Mah, D., Mageras, G. S., Raben, A., Rosenzweig, K., Mychalczak, B., Schwartz, L. H., Gloeggler, P. J., Lutz, W., et al. (1999). Deep inspiration breath-hold technique for lung tumors : the potential value of target immobilization and reduced lung density in dose escalation. *International Journal of Radiation Oncology* Biology* Physics*, 45(3) :603–611.
- Harauz, G. and Bronskill, M. (1979). Comparison of the liver’s respiratory motion in the supine and upright positions : concise communication. *Journal of Nuclear Medicine*, 20(7) :733–735.
- Heeger, D. J. (1988). Optical flow using spatiotemporal filters. *International Journal of Computer Vision*, 1(4) :279–302.

-
- Hong, T. S., Tomé, W. A., Chappell, R. J., Chinnaiyan, P., Mehta, M. P., and Harari, P. M. (2005). The impact of daily setup variations on head-and-neck intensity-modulated radiation therapy. *International Journal of Radiation Oncology* Biology* Physics*, 61(3) :779–788.
- Horn, B. K. (1987). Closed-form solution of absolute orientation using unit quaternions. *JOSA A*, 4(4) :629–642.
- Horn, B. K. and Schunck, B. G. (1981). Determining optical flow. *Artificial intelligence*, 17(1) :185–203.
- Hostettler, A. (2008). Modélisation et simulation patient-dépendante, préopératoire, prédictive, et temps-réel du mouvement des organes de l'abdomen induit par la respiration libre. *Thesis*.
- Huang, E., Dong, L., Chandra, A., Kuban, D. A., Rosen, I. I., Evans, A., and Pollack, A. (2002). Intrafraction prostate motion during imrt for prostate cancer. *International Journal of Radiation Oncology* Biology* Physics*, 53(2) :261–268.
- Jain, P., Marchant, I., Green, M., Watkins, G., Davies, J., McCarthy, C., Loncaster, J., Stewart, A., Magee, B., Moore, C., and Price, P. (2009). Inter-fraction motion and dosimetric consequences during breast intensity-modulated radiotherapy (imrt). *Radiotherapy and Oncology*, 90(1) :93–98.
- Jandt, S. R., da Sil Caballero, R. M., Junior, L. A. F., and Dias, A. S. (2011). Correlation between trunk control, respiratory muscle strength and spirometry in patients with stroke : An observational study. *Physiotherapy Research International*, 16(4) :218–224.
- Jensen, R., Paulsen, R., and Larsen, R. (2008). Analyzing gait using a time-of-flight camera. *Proceedings of the Scandinavian Conference on Image Analysis*.
- Johnson, U., Landau, D., Lindgren-Turner, J., Smith, N., Meir, I., Howe, R., Rodgers, H., Davit, S., and Deehan, C. (2004). Real time 3d surface imaging for the analysis of respiratory motion during radiotherapy. *International Journal of Radiation Oncology * Biology * Physics*, 60(1) :603–604.
- Kahlmann, T., Remondino, F., and Ingensand, H. (2006). Calibration for increased accuracy of the range imaging camera swissrangertm. *Proc. of IEVM*.
- Kalender, W. A., Rienmüller, R., Seissler, W., Behr, J., Welke, M., and Fichte, H. (1990). Measurement of pulmonary parenchymal attenuation : use of spirometric gating with quantitative ct. *Radiology*, 175(1) :265–268.

-
- Karel, W., Dorninger, P., and Pfeifer, N. (2007). In situ determination of range camera quality parameters by segmentation. In *Proc. 8th Int. Conf. on Opt. 3D Meas. Tech*, pages 109–116.
- Kauwelo, K., Ruan, D., Park, J., Sandhu, A., Kim, G., Pawlicki, T., Watkins, W., Song, B., and Song, W. (2012). Gatectm surface tracking system for respiratory signal reconstruction in 4dct imaging. *Medical Physics*, 39(1) :492.
- Keall, P., Mageras, G., and Yorke, E. (2006). The management of respiratory motion in radiation oncology report of aapm task group 76. *Medical physics*, 33(10) :3874–3900.
- Keall, P. J., Joshi, S., Vedam, S. S., Siebers, J. V., Kini, V. R., and Mohan, R. (2005). Four-dimensional radiotherapy planning for dmlc-based respiratory motion tracking. *Medical physics*, 32 :942.
- Khamene, A., Warzelhan, J. K., Vogt, S., Elgort, D., Chefd’Hotel, C., Duerk, J. L., Lewin, J. S., Wacker, F. K., and Sauer, F. (2004). Characterization of internal organ motion using skin marker positions. In Barillot, C., Haynor, D. R., and Hellier, P., editors, *MICCAI (2)*, volume 3217 of *Lecture Notes in Computer Science*, pages 526–533. Springer.
- Khoshelham, K. and Elberink, S. O. (2012). Accuracy and resolution of kinect depth data for indoor mapping applications. *Sensors*, 12(2) :1437–1454.
- Kirilova, A., Lockwood, G., Choi, P., Bana, N., Haider, M. A., Brock, K. K., Eccles, C., and Dawson, L. A. (2008). Three-dimensional motion of liver tumors using cine-magnetic resonance imaging. *International journal of radiation oncology, biology, physics*, 71(4) :1189–1195.
- Kitamura, K., Shirato, H., Seppenwoolde, Y., Onimaru, R., Oda, M., Fujita, K., Shimizu, S., Shinohara, N., Harabayashi, T., and Miyasaka, K. (2002). three-dimensional intrafractional movement of prostate measured during real-time tumor-tracking radiotherapy in supine and prone treatment positions. *International Journal of radiation oncology, biology, physics*, 53(5) :1117–1123.
- Korin, H. W., Ehman, R. L., Riederer, S. J., Felmlee, J. P., and Grimm, R. C. (2005). Respiratory kinematics of the upper abdominal organs : a quantitative study. *Magnetic resonance in medicine*, 23(1) :172–178.
- Korreman, S. S., Juhler-Nøttrup, T., and Boyer, A. L. (2008). Respiratory gated beam delivery cannot facilitate margin reduction, unless combined with respiratory correlated image guidance. *Radiotherapy and Oncology*, 86(1) :61–68.

-
- Krengli, M., Gaiano, S., Mones, E., Ballarè, A., Beldi, D., Bolchini, C., Loi, G., et al. (2009). Reproducibility of patient setup by surface image registration system in conformal radiotherapy of prostate cancer. *Radiat Oncol*, 4(9).
- Kubo, H. and Hill, B. (1996). Respiration gated radiotherapy treatment : a technical study. *Physics in Medecine and Biology*, 41 :83–91.
- Kuipers, J. (1980). Spasyn :an electromagnetic relative position and orientation tracking system. *IEEE transactions Instrumentation and measurement*, 29(4) :462–466.
- Langmann, B., Hartmann, K., and Loffeld, O. (2008). Real-time foreground segmentation via range and color imaging. *International Journal of Computer Information Systems and Industrial Management Applications*, pages 1–5.
- Lax, I., Blomgren, H., Naslund, I., and Svantrom, R. (1994). Stereotactic radiotherapy of malignancies in the abdomen. methodological aspects. *Acta Oncol*, 33(6) :677–683.
- Lee, S., Wolberg, G., and Shin, S. Y. (1997). Scattered data interpolation with multilevel b-splines. *IEEE Transactions on Visualization and Computer Graphics*, 3 :228–244.
- Leong, J. (2000). Implementation of random positioning error in computerised radiation treatment planning systems as a result of fractionation. *Physics in medicine and biology*, 32(3) :327.
- Leunens, G., Menten, J., Weltens, C., Verstraete, J., and Van der Schueren, E. (1993). Quality assessment of medical decision making in radiation oncology : variability in target volume delineation for brain tumours. *Radiotherapy and Oncology*, 29(2) :169–175.
- Lindner, M. and Kolb, A. (2006). Lateral and depth calibration of pmd-distance sensors. *Advances in Visual Computing*, pages 524–533.
- Lindner, M., Kolb, A., and Hartmann, K. (2007). Data-fusion of pmd-based distance-information and high-resolution rgb-images. *Proceedings of the International Symposium on Signals, Circuits and Systems (ISSCS), Session on Algorithms for 3D TOF-cameras*, 1 :121–124.
- Longuet-Higgins, H. C. (1987). A computer algorithm for reconstructing a scene from two projections. pages 61–62.
- Low, D., Parikh, P., Lu, W., Dempsey, J., Wahab, S., Hubenschmidt, J., Nystrom, M., Handoko, M., and Bradley, J. (2005). Novel breathing motion model for radiotherapy. *International Journal of radiation oncology, biology, physics*, 63 :921–929.

-
- Low, D. A., Nystrom, M., Kalinin, E., Parikh, P., Dempsey, J. F., Bradley, J. D., Mutic, S., Wahab, S. H., Islam, T., Christensen, G., et al. (2003). A method for the reconstruction of four-dimensional synchronized ct scans acquired during free breathing. *Medical physics*, 30 :1254.
- Lucas, B. D., Kanade, T., et al. (1981). An iterative image registration technique with an application to stereo vision. In *Proceedings of the 7th international joint conference on Artificial intelligence*.
- Lujan, A. E., Larsen, E. W., Balter, J. M., and Ten Haken, R. K. (1999). A method for incorporating organ motion due to breathing into 3d dose calculations. *Medical physics*, 26 :715.
- McClelland, J., Blackall, J., Tarte, S., Chandler, A., Hughes, S., Ahmad, S., Landau, D., and Hawkes, D. (2006). A continuous 4d motion model from multiple respiratory cycles for use in lung radiotherapy. *Medical physics*, 33 :3348–3358.
- McClelland, J., Hawkes, D., Schaeffter, T., and King, A. (2012). Respiratory motion models : a review. *Medical image analysis*.
- McKenzie, A. (2000). How should breathing motion be combined with other errors when drawing margins around clinical target volumes? *British journal of radiology*, 73(873) :973–977.
- Miller, D., Klein, E., Zoberi, I., Taylor, M., and Powell, S. (2008). Inter-fraction and intra-fraction breast motion localized using alignrt for early breast cancer. *International Journal of Radiation Oncology* Biology* Physics*, 72(1) :S189–S190.
- Mines, A. (1993). *Respiratory Physiology*. Raven Press.
- Moroni, C., Mascalchi, M., Camiciottoli, G., Bartolucci, M., Falaschi, F., Battolla, L., Orlandi, I., Pistolesi, M., and Villari, N. (2003). Comparison of spirometric-gated and-ungated hrct in copd. *Journal of computer assisted tomography*, 27(3) :375–379.
- Moser, T., Fleischhacker, S., Schubert, K., Sroka-Perez, G., and Karger, C. P. (2010). Technical performance of a commercial laser surface scanning system for patient setup correction in radiotherapy. *Physica Medica*.
- Moser, T., Habl, G., Uhl, M., Schubert, K., Sroka-Perez, G., Debus, J., Herfarth, K., and Karger, C. P. (2012). Clinical evaluation of a laser surface scanning system in 120 patients for improving daily setup accuracy in fractionated radiation therapy. *International Journal of Radiation Oncology* Biology* Physics*.

-
- M.Rains (2010). Tumour in full motion ; inspiration gating versus free breathing. *Radiographer physics*, 57(2).
- Müller, K., Schaller, C., Penne, J., and Hornegger, J. (2009). Surface-based Respiratory Motion Classification and Verification. In Meinzer, H.-P., Deserno, T. M., Handels, H., and Tolxdorff, T., editors, *Bildverarbeitung für die Medizin 2009*, pages 257–261, Berlin.
- Mure-Dubois, J., Hügli, H., et al. (2008). Fusion of time of flight camera point clouds. In *Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications-M2SFA2 2008*.
- Murphy, M. J., Martin, D., Whyte, R., Hai, J., Ozhasoglu, C., and Le, Q.-T. (2002). The effectiveness of breath-holding to stabilize lung and pancreas tumors during radiosurgery. *International Journal of Radiation Oncology* Biology* Physics*, 53(2) :475–482.
- Myronenko, A. and Song, X. (2010). Point set registration : Coherent point drift. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(12) :2262–2275.
- Nakamoto, M., Nakada, K., Sato, Y., Konishi, K., Hashizume, M., and Tamura, S. (2008). Intraoperative magnetic tracker calibration using a magneto-optic hybrid tracker for 3-d ultrasound-based navigation in laparoscopic surgery. *IEEE transactions on Medical Imaging*, 27(2) :255–270.
- Nickolls, J. and Dally, W. (2010). The gpu computing era. *IEEE Micro*, 30 :56–69.
- Niclass, C., Favi, C., Kluter, T., Monnier, F., and Charbon, E. (2009). Single-photon synchronous detection. *Solid-State Circuits, IEEE Journal of*, 44(7) :1977–1989.
- Nicolau, S., Goffin, L., and Soler, L. (2005). A low cost and accurate guidance system for laparoscopic surgery : validation on an abdominal phantom. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, pages 124–133.
- Noonan, P., Cootes, T., Hallett, W., and Hinz, R. (2011). The design and initial calibration of an optical tracking system using the microsoft kinect. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011 IEEE*, pages 3614–3617. IEEE.
- Noponen, T., Kokki, T., Lepomaki, V., Kajander, S., Durand-Schaefer, N., Teras, M., and Knuuti, J. (2008). Spirometry based respiratory gating method for cardiac pet and mri imaging. In *Nuclear Science Symposium Conference Record, 2008. NSS'08. IEEE*, pages 4832–4834. IEEE.
- Oggier, T., Büttgen, B., Lustenberger, F., Becker, G., Rüegg, B., and Hodac, A. (2005). Swissranger sr3000 and first experiences based on miniaturized 3dtof cameras. *n Proc. of the First Range Imaging Research Day at ETH Zurich*.

-
- Pallotta, S., Marrazzo, L., Ceroti, M., Silli, P., and Bucciolini, M. (2012a). A phantom evaluation of sentinel, a commercial laser/camera surface imaging system for patient setup verification in radiotherapy. *Medical physics*, 39(2) :706–712.
- Pallotta, S., Marrazzo, L., Ceroti, M., Silli, P., and Bucciolini, M. (2012b). A phantom evaluation of sentinel, a commercial laser/camera surface imaging system for patient setup verification in radiotherapy. *Medical physics*, 39 :706.
- Pallotta, S., Simontacchi, G., Marrazzo, L., Ceroti, M., Paiar, F., Biti, G., and Bucciolini, M. (2013). Accuracy of a 3d laser/camera surface imaging system for setup verification of the pelvic and thoracic regions in radiotherapy treatments. *Medical physics*, 40 :011710.
- Pan, T., Lee, T.-Y., Rietzel, E., and Chen, G. T. (2004). 4d-ct imaging of a volume influenced by respiratory motion on multi-slice ct. *Medical physics*, 31 :333.
- Peng, J., Lui, C., Chen, Y., Amdur, R., Vanek, K., and Li, J. (2011). Dosimetric consequences of rotational setup errors with direct simulation in a treatment planning system of fractionated stereotactic radiotherapy. *J Appl Clin Med Phys*, 12(3).
- Peng, J. L., Kahler, D., Li, J. G., Samant, S., Yan, G., Amdur, R., and Liu, C. (2010). Characterization of a real-time surface image-guided stereotactic positioning system. *Medical physics*, 37 :5421.
- Penne, J., Schaller, C., Hornegger, J., and Kuwert, T. (2008). Robust Real-Time 3D Respiratory Motion Detection Using Time-of-Flight Cameras. *Computer Assisted Radiology and Surgery 2008*, 3(5) :427–431.
- Petersson, K., Enmark, M., and Ambolt, P. (2010). Improving patient throughput with preserved accuracy using a body surface scanning system for patient set-up during tomotherapy. *TomoTherapy EMEA User Symposium*.
- Piatti, D. and Rinaudo, F. (2012). Sr-4000 and camcube3.0 time of flight (tof) cameras : Tests and comparison. *Remote Sensing*, 4(4) :1069–1089.
- Placht, S., Schaller, C., Adelt, A., et al. (2010). Improvement and evaluation of a time-of-flight-based patient positioning system. *BVM2010*.
- Placht, S., Stancanello, J., Schaller, C., Balda, M., and Angelopoulou, E. (2012). Fast time-of-flight camera based surface registration for radiotherapy patient positioning. *Medical Physics*, 39(1) :4.
- Plathow, C., Ley, S., Fink, C., Puderbach, M., Hosch, W., Schmähl, A., Debus, J., and Kauczor, H.-U. (2004). Analysis of intrathoracic tumor mobility during whole breathing

-
- cycle by dynamic mri. *International Journal of Radiation Oncology* Biology* Physics*, 59(4) :952–959.
- Price, G., Parkhurst, J., Sharrock, P., and Moore, C. (2012). Real-time optical measurement of the dynamic body surface for use in guided radiotherapy. *Physics in Medicine and Biology*, 57 :415–436.
- Prusak, A., Melnychuk, O., Schiller, I., Roth, H., and Koch, R. (2008). Pose estimation and map building with a pmd-camera for robot navigation. *Int. J. on Intell. Systems Techn. and App., Issue on Dynamic 3D Imaging*, 5(3-4) :355–364.
- Raab, F., Blood, E., Steiner, T., and Jones, R. (1979). Magnetic position and orientation tracking system. *IEEE transactions Aerospace and Electronic systems*, 15(5) :709–718.
- Remouchamps, V. M., Letts, N., Vicini, F. A., Sharpe, M. B., Kestin, L. L., Chen, P. Y., Martinez, A. A., and Wong, J. W. (2003). Initial clinical experience with moderate deep-inspiration breath hold using an active breathing control device in the treatment of patients with left-sided breast cancer using external beam radiation therapy. *International Journal of radiation oncology, biology, physics*, 56(3) :704–715.
- Ross, C. S., Hussey, D. H., Pennington, E. C., Stanford, W., and Fred Doornbos, J. (1990). Analysis of movement of intrathoracic neoplasms using ultrafast computerized tomography. *International Journal of Radiation Oncology* Biology* Physics*, 18(3) :671–677.
- Runge, C. (1901). Über empirische funktionen und die interpolation zwischen äquidistanten ordinaten. *Zeitschrift für Mathematik und Physik*, 46(224-243) :20.
- Sahih, A., Haas, O., Burnham, K., and Mills, J. (2005). Organ motion modelling and prediction for adaptive radiotherapy. *Proc. IAR-ACD*, 2 :211–216.
- Sarrut, D., Boldea, V., Miguet, S., and Ginestet, C. (2006). Simulation of fourdimensional ct images from deformable registration between inhale and exhale breath-hold ct scans. *Medical physics*, 33 :605–617.
- Schaerer, J., Fassi, A., Riboldi, M., and G Baroni, P. C., and Sarrut, D. (2012). Multi-dimensional respiratory motion tracking from markerless optical surface imaging based on deformable mesh registration. *Med Image Comput Assist Interv*, 15(1) :412–421.
- Schaller, C., Adelt, A., Penne, J., and Hornegger, J. (2009). Time-of-flight sensor for patient positioning. In *SPIE Medical Imaging*, pages 726110–726110. International Society for Optics and Photonics.

-
- Schöffel, P. J., Harms, W., Sroka-Perez, G., Schlegel, W., and Karger, C. P. (2007). Accuracy of a commercial optical 3d surface imaging system for realignment of patients for radiotherapy of the thorax. *Physics in medicine and biology*, 52(13) :3949.
- Schreibmann, E., Chen, G., and Xing, L. (2006). Image interpolation in 4d ct using a b-spline deformable registration model. *International Journal of radiation oncology, biology, physics*, 64 :1537–1550.
- Seco, J., Sharp, G. C., Wu, Z., Gierga, D., Buettner, F., and Paganetti, H. (2008). Dosimetric impact of motion in free-breathing and gated lung radiotherapy : A 4d monte carlo study of intrafraction and interfraction effects. *Medical physics*, 35(1) :356.
- Seppenwoolde, Y., Berbeco, R., Nishioka, S., Shirato, H., and Heijmen, B. (2007). Accuracy of tumour motion compensation algorithm from a robotic respiratory tracking system : A simulation study. *Medical Physics*, 34 :2774–2784.
- Seppenwoolde, Y., Shirato, H., Kitamura, K., Shimizu, S., van Herk, M., Lebesque, J. V., and Miyasaka, K. (2002). Precise and real-time measurement of 3d tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *International Journal of Radiation Oncology* Biology* Physics*, 53(4) :822–834.
- Shimizu, S., Shirato, H., Ogura, S., Akita-Dosaka, H., Kitamura, K., Nishioka, T., Kagei, K., Nishimura, M., and Miyasaka, K. (2001). Detection of lung tumor movement in real-time tumor-tracking radiotherapy. *International Journal of Radiation Oncology* Biology* Physics*, 51(2) :304–310.
- SiteCalypso (2013). Site internet de la société calypso.
- SitePlanCancer (2013). Site internet présentant le plan cancer sur la période 2009-2013.
- SiteVisionRT (2013). Site internet de la société visionrt.
- Sixel, K. E., Ruschin, M., Tirona, R., and Cheung, P. C. (2003). Digital fluoroscopy to quantify lung tumor motion : potential for patient-specific planning target volumes. *International Journal of Radiation Oncology* Biology* Physics*, 57(3) :717–723.
- Smith, R. P., Bloch, P., Harris, E. E., McDonough, J., Sarkar, A., Kassaei, A., Avery, S., and Solin, L. J. (2005). Analysis of interfraction and intrafraction variation during tangential breast irradiation with an electronic portal imaging device. *International Journal of Radiation Oncology* Biology* Physics*, 62(2) :373–378.
- Sohn, M., Birkner, M., Yan, D., and Alber, M. (2005). Modelling individual geometric variation based on dominant eigenmodes of organ deformation : implementation and evaluation. *Physics in Medicine and Biology*, 50 :5893–5908.

-
- Soutschek, S., Penne, J., and Hornegger, J. (2008). 3d gesture-based scene navigation in medical imaging applications using time-of-flight cameras. *IEEE Conf. on Computer Vision and Pattern Recogn. ; Workshop on ToF-Camera based Computer Vision*.
- Spadea, M., Baroni, G., Gierga, D., Turcotte, J., Chen, G., and Sharp, G. (2010). Evaluation and commissioning of surface based system for respiratory sensing in 4d ct. *Journal of Applied Clinical Medical Physics*, 12(1) :3288.
- Stevens, C. W., Munden, R. F., Forster, K. M., Kelly, J. F., Liao, Z., Starkschall, G., Tucker, S., and Komaki, R. (2001). Respiratory-driven lung tumor motion is independent of tumor size, tumor location, and pulmonary function. *International Journal of Radiation Oncology* Biology* Physics*, 51(1) :62–68.
- Stoppa, D., Pancheri, L., Scandiuzzo, M., Gonzo, L., Dalla Betta, G.-F., and Simoni, A. (2007). A cmos 3-d imager based on single photon avalanche diode. *Circuits and Systems I : Regular Papers, IEEE Transactions on*, 54(1) :4–12.
- Suramo, I., Päivänsalo, M., Myllylä, V., et al. (1984). Cranio-caudal movements of the liver, pancreas and kidneys in respiration. *Acta radiologica : diagnosis*, 25(2) :129.
- Takahashi, T. (1994). *Atlas of the Human Body*. Harper Collins.
- Tarte, S., McClelland, J., Hughes, S., Blackall, J., Landau, D., and Hawkes, D. (2006). A non-contact method for the acquisition of breathing signals that enable distinction between abdominal and thoracic breathing.
- Thomas, G. (2006). Mixed reality techniques for tv and their application for on-set and pre-visualization in ?lm production. *Int. Workshop on Mixed Reality Technology for Film-making*.
- Towle, A. (2000). *Modern biology*. Holt, Rinehart and Winston.
- Vedam, S., Keall, P., Kini, V., Mostafavi, H., Shukla, H., and Mohan, R. (2002). Acquiring a four-dimensional computed tomography dataset using an external respiratory signal. *Physics in medicine and biology*, 48(1) :45.
- Wade, O. (1954). Movements of the thoracic cage and diaphragm in respiration. *The Journal of physiology*, 124(2) :193–212.
- Wagman, R., Yorke, E., Ford, E., Giraud, P., Mageras, G., Minsky, B., and Rosenzweig, K. (2003). Respiratory gating for liver tumors : use in dose escalation. *International Journal of radiation oncology, biology, physics*, 55 :659–669.

-
- Wang, H., Shiu, A., Congjun, W., O'Daniel, J., Mahajan, A., Woo, S., Liengsawangwong, P., Mohan, R., and Chang, E. (2008). Dosimetric effect of translational and rotational errors for patients undergoing image-guided stereotactic body radiotherapy for spinal metastases. *International Journal of radiation oncology, biology, physics*, 71(2) :1261–1271.
- Weinberg, S., Naidoo, S., Govier, D., Martin, R., Kane, A., and Marazita, M. (2006). Anthropometric precision and accuracy of digital three-dimensional photogrammetry : Comparing the genex and 3dmd imaging systems with one another and with direct anthropometry. *Journal of Craniofacial Surgery*, 17(3) :477–483.
- Weiss, P. H., Baker, J. M., and Potchen, E. J. (1972). Assessment of hepatic respiratory excursion. *Journal of Nuclear Medicine*, 13(10) :758–759.
- Werner, R., White, B., Handels, H., Lu, W., and Low, D. A. (2010). Technical note : Development of a tidal volume surrogate that replaces spirometry for physiological breathing monitoring in 4d ct. *Medical physics*, 37(2) :615.
- Weyer, C. A., Bae, K.-H., Lim, K., and Lichti, D. (2008). Extensive metric performance evaluation of a 3d range camera. *Int. Soc. Photogramm. Remote Sens*, 37 :939–944.
- Wong, J. W., Shapre, M. B., JAffray, D. A., Kini, V. R., Robertson, J. M., Stromberg, J. S., and Martinez, A. A. (1999). The use of active breathing control (abc) to reduce margin for breathing motion. *International Journal of radiation oncology, biology, physics*, 44(4) :911–919.
- Wu, H., Sharp, G., Salzberg, B., Kaeli, D., Shirato, H., and Jiang, S. (2004). A finite state model for respiratory motion analysis in image guided radiationtherapy. *Medical Physics*, 31(8) :22–74.
- Wu, Q. J., Thongphiew, D., Wang, Z., Chankong, V., and Yin, F.-F. (2008). The impact of respiratory motion and treatment technique on stereotactic body radiation therapy for liver cancer. *Medical physics*, 35 :1440.
- Wu, Y.-T., Kanade, T., Cohn, J., and Li, C.-C. (1998). Optical flow estimation using wavelet motion model. In *Computer Vision, 1998. Sixth International Conference on*, pages 992–998. IEEE.
- Xia, J. and Siochi, R. A. (2012). A real-time respiratory motion monitoring system using kinect : Proof of concept. *Medical Physics*, 39 :2682.
- Xu, Z., Schwarte, R., Heinol, H., Buxbaum, B., and Ringbeck, T. (1998). Smart pixel-photonic mixer device (pmd). In *Proc. Int. Conf. on Mechatron. & Machine Vision*, pages 259–264.

-
- Yang, Q., Yang, R., Davis, J., and Nistér, D. (2007). Spatial-depth super resolution for range images. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE.
- Yue, N. J., Li, X., Beriwal, S., Heron, D. E., Sontag, M. R., and Huq, M. S. (2007). The intrafraction motion induced dosimetric impacts in breast 3d radiation treatment : A 4dct based study. *Medical physics*, 34 :2789.
- Zeng, R., Fessler, J., and Balter, J. (2005). Respiratory motion estimation from slowly rotating x-ray projections : theory and simulation. *Medical physics*, 32(4) :984–991.
- Zeng, R., Fessler, J., and Balter, J. (2007). Estimating 3-d respiratory motion from orbiting views by tomographic image registration. *IEEE Transactions on Medical Imaging*, 26(2) :153–163.
- Zhang, Q., Pevsner, A., Hertanto, A., Hu, Y., Rosenzweig, K., Ling, C., and Mageras, G. (2007). A patient-specific respiratory model of anatomical motion for radiation treatment planning. *Medical physics*, 34 :4772–4781.
- Zhao, H., Wang, B., Rassiah-Szegedi, P., Huang, Y., Sarkar, V., Szegedi, M., Sadinski, M., Kokeny, K., Poppe, M., and Salter, B. (2011). Su-et-527 : Comparison of ct-on-rails and a 3d surface imaging system for image guided partial breast radiation therapy. *Medical Physics*, 38 :3610.
- Zhu, J., Wang, L., Yang, R., Davis, J. E., and Pan, Z. (2011). Reliability fusion of time-of-flight depth and stereo geometry for high quality depth maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(7) :1400–1414.