



HAL
open science

Perception des objets en mouvement Composition bayésienne du flux optique et du mouvement de l'observateur

Francis Colas

► **To cite this version:**

Francis Colas. Perception des objets en mouvement
Composition bayésienne du flux optique et du
mouvement de l'observateur. Autre [cs.OH]. Institut National Polytechnique de Grenoble - INPG,
2006. Français. NNT: . tel-00080229

HAL Id: tel-00080229

<https://theses.hal.science/tel-00080229v1>

Submitted on 15 Jun 2006

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.

INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

N° attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--

THÈSE

pour obtenir le grade de

DOCTEUR DE L'INPG

Spécialité : *Informatique*

préparée au laboratoire GRAVIR et l'INRIA Rhône-Alpes, dans le cadre de l'École doctorale *Mathématiques, Sciences et Technologies de l'Information, Informatique*

présentée et soutenue publiquement par

Francis COLAS

le 27 janvier 2006

Titre :

Perception des objets en mouvement
Composition bayésienne du flux optique et du mouvement de l'observateur

Co-directeurs de thèse :

Pierre BESSIÈRE et Jacques DROULEZ

Composition du jury :

M.	Jeanny HÉRAULT	Président
M.	Pascal MAMASSIAN	Rapporteur
M.	Thierry VIÉVILLE	Rapporteur
Mme.	Valérie CORNILLEAU-PÉRÈS	Examinatrice
M.	Pierre BESSIÈRE	Directeur de thèse
M.	Jacques DROULEZ	Co-directeur de thèse

Remerciements

Je souhaite tout d'abord remercier mes directeurs de thèse, Pierre Bessière et Jacques Droulez pour leur patience, leurs conseils et leur amitié.

Je remercie aussi les membres du Jury pour leur patience, l'intérêt qu'ils ont portés à mon travail et leurs précieux commentaires sur le document.

Je n'oublie pas tous les membres de l'équipe Sharp, renommée ensuite E-Motion, qui m'ont accueilli durant ce travail et avec qui j'ai pu avoir des discussions enrichissantes.

Enfin, j'adresse un grand merci à ma famille et mes amis qui m'ont soutenu durant tout ce temps.

Merci à tous.

Résumé

La perception est le regroupement et recouplement d'un ensemble d'informations en vue de comprendre son environnement. À cette fin, l'homme utilise plusieurs modalités sensorielles telles que la vision, le toucher, l'ouïe ou même la proprioception. Ces sens eux-mêmes peuvent convoier différents types d'information. La vision par exemple, propose des indices de texture, de couleurs, de formes ou de mouvement. Le flux optique est ainsi le déplacement de l'image sur la rétine. Avec lui seul, un homme est capable d'appréhender une partie des formes observées.

Or, extraire la forme à partir du mouvement est un problème très complexe. D'une part, la perception en général est un problème *inverse*. La géométrie permet de déterminer le flux optique à partir de la forme et du mouvement des objets. Ensuite, plusieurs configurations peuvent engendrer un même flux optique. Dans ce cas, on dit que le problème est *mal-posé*. Enfin, il est rarement possible d'être certain d'une perception à une précision arbitraire. Toute perception est donc entachée d'*incertitude*.

Pour résoudre ces difficultés, nous utilisons le formalisme de la *programmation bayésienne*. Celui-ci permet de raisonner sur l'incertitude à l'aide des probabilités. Les multiples solutions d'un problème mal-posé sont prises en compte par la multi-modalité des distributions de probabilité. Enfin, la symétrie de la règle de Bayes permet de faire l'inférence de la même manière que ce soit sur les problèmes inverses et directs.

Nous proposons un modèle bayésien de la perception de plans par le flux optique, basé sur quelques hypothèses explicites. Les principales hypothèses mises en œuvre dans ce modèle sont la *rigidité* et la *stationnarité*. La rigidité porte sur le mouvement des points et exprime que le flux optique est plus probablement dû à un objet rigide qu'à une collection de points de mouvements indépendants. La stationnarité porte sur le mouvement de l'objet en indiquant qu'il est plus probablement petit.

Nous validons notre modèle en nous basant sur cinq expériences issues de la littérature. Pour chacune d'elles, nous expliquons les détails du fonctionnement du modèle et les hypothèses qui conditionnent la bonne reproduction des résultats. Enfin, nous montrons d'une part comment intégrer notre modèle au sein de systèmes plus complexes et, d'autre part, l'applicabilité du modèle bayésien à des expériences de perception du flux optique différentes de part leurs protocoles, mais similaires dans leurs ambiguïtés.

Table des matières

1	Introduction	1
1.1	Perception	1
1.2	Perception de la forme par le mouvement	2
1.2.1	Difficultés	3
1.2.2	Objectifs	4
1.3	Modèle proposé	4
1.3.1	Modélisation bayésienne	4
1.3.2	Contribution	5
1.4	Plan du manuscrit	6
2	État de l’art	7
2.1	Vision par ordinateur	7
2.1.1	Extraction du flux optique	7
2.1.2	Forme par le mouvement	8
2.2	Psychophysique	9
2.2.1	Flux optique	9
2.2.2	Forme par le mouvement	13
2.3	Modélisation bayésienne	14
2.3.1	Outil de modélisation	14
2.3.2	Modèles bayésiens de la perception	18
3	Exemple introductif	21
3.1	Hypothèse et protocole	21
3.1.1	Hypothèse testée	21
3.1.2	Protocole expérimental	21
3.2	Résultats et hypothèses	23
3.2.1	Résultats expérimentaux	23
3.2.2	Hypothèses explicatives	24
3.3	Modèle	25
3.3.1	Variables	25
3.3.2	Décomposition	26
3.3.3	Distributions	27
3.3.4	Question	28

3.4	Résultats	29
3.4.1	Comparaison	29
3.4.2	Validité des comparaisons	31
3.5	Résumé	32
4	Modèle complet	33
4.1	Principe	33
4.2	Variables	34
4.2.1	Positionnement du plan	34
4.2.2	Mouvement propre	36
4.2.3	Mouvement du plan	37
4.2.4	Flux optique	38
4.2.5	Champ de vision	38
4.2.6	Récapitulatif	38
4.3	Décomposition	39
4.4	Distributions	41
4.4.1	Stationnarité	41
4.4.2	Rigidité	45
4.4.3	Position	46
4.4.4	Distributions inconnues	47
4.5	Question	48
4.5.1	Formulation de la question	48
4.5.2	Inférence	48
4.5.3	Paramètres	49
4.6	Test de régression	50
4.6.1	Rappel de l'expérience	50
4.6.2	Obtention de la réponse	50
4.6.3	Comparaison	52
4.7	Résumé	54
5	Résultats	57
5.1	Importance du flux optique	57
5.1.1	Hypothèse et protocole	57
5.1.2	Résultats de l'expérience	58
5.1.3	Résultats du modèle	59
5.1.4	Explication	61
5.2	Influence de la taille du champ de vision	67
5.2.1	Hypothèse et protocole	67
5.2.2	Résultats de l'expérience	67
5.2.3	Résultats du modèle	68
5.2.4	Explication	70
5.3	Influence du mouvement propre	71
5.3.1	Hypothèse et protocole	71

5.3.2	Résultats de l'expérience	72
5.3.3	Résultats du modèle	73
5.3.4	Explication	73
5.4	Influence relative de la stationnarité et de la rigidité	75
5.4.1	Hypothèse et protocole	75
5.4.2	Résultats de l'expérience	77
5.4.3	Résultats du modèle	77
5.4.4	Explication	78
5.5	Influence du <i>shear</i>	82
5.5.1	Hypothèse et protocole	82
5.5.2	Résultats de l'expérience	82
5.5.3	Résultats du modèle	84
5.5.4	Explication	85
5.6	Résumé	89
6	Utilisations avancées	91
6.1	Expérience approchée	91
6.1.1	Hypothèse et protocole	91
6.1.2	Transposition du modèle	93
6.1.3	Résultats de l'expérience	95
6.1.4	Sous-modèle initial	96
6.1.5	Amélioration du modèle <i>top-far</i>	103
6.1.6	Sous-modèle final	107
6.2	Modèle de flux optique	110
6.2.1	Principe	110
6.2.2	Modèle naïf	111
6.2.3	Modèle découplé	113
6.3	Interfaces du modèle de perception	120
6.3.1	Rappel du modèle	120
6.3.2	Entrées	122
6.3.3	Sorties	124
6.4	Résumé	125
7	Conclusion	127
7.1	Hypothèses	127
7.2	Résultats	128
7.3	Utilisations avancées	129
7.4	Perspectives	129
A	Équations du flux optique d'un plan	141

Table des figures

1.1	Animation d'un masque en rotation	2
2.1	Paramètres (X, Υ) d'un plan	11
3.1	Position relative du plan et de l'observateur	22
3.2	Exemple d'affichage présenté aux participants.	23
3.3	Répartition expérimentale de l'erreur d'estimation de l'orientation.	24
3.4	Orientation d'un plan	26
3.5	Résumé du modèle bayésien.	29
3.6	Comparaison du modèle et de l'expérience	30
4.1	Inclinaison et orientation d'un plan	35
4.2	Paramètres (X, Υ) d'un plan	36
4.3	Mouvement propre	37
4.4	Changement de repère : translation sagittale	42
4.5	Changement de repère : translation transversale	43
4.6	Changement de repère : rotation verticale	43
4.7	Résultats expérimentaux	50
4.8	Distribution dans l'espace (X, Υ)	51
4.9	Distribution dans l'espace (Σ, T)	52
4.10	Résultats du modèle	52
4.11	Comparaison du modèle et de l'expérience	53
4.12	Résumé du modèle bayésien de perception.	55
5.1	Conditions de l'expérience 5.1	58
5.2	Résultats de l'expérience 5.1	59
5.3	Évaluation comparée du modèle pour l'expérience 5.1	60
5.4	Flux optiques associés aux conditions de l'expérience 5.1.	62
5.5	Distribution sur (Ξ, Ω) en condition petite inclinaison-grande vitesse	64
5.6	Distribution <i>a priori</i> sur (Ξ, Ω)	65
5.7	Vraisemblance de $\vec{\phi}$ sachant (Ξ, Ω) et (v_x, v_z, r_y, λ) en condition petite inclinaison-grande vitesse	65
5.8	Vraisemblance de $\vec{\phi}$ sachant (Ξ, Ω) et (v_x, v_z, r_y, λ) en condition grande inclinaison-grande vitesse	66

5.9	Distribution sur (Ξ, Ω) en condition grande inclinaison-grande vitesse.	66
5.10	Résultats de l'expérience 5.2	68
5.11	Évaluation comparée du modèle pour l'expérience 5.2	69
5.12	Taux d'inversion en fonction de la covariance sur le flux	70
5.13	Conditions de l'expérience 5.3	72
5.14	Résultats de l'expérience 5.3	73
5.15	Résultats du modèle pour l'expérience 5.3	74
5.16	Percepts inverses en conditions immobile et actif	74
5.17	Conditions de l'expérience 5.4	76
5.18	Résultats de l'expérience 5.4	77
5.19	Évaluation comparée du modèle pour l'expérience 5.4	78
5.20	Flux optiques des différentes conditions de l'expérience 5.4.	79
5.21	Configurations rigides et non rigides en condition A	80
5.22	Configurations rigides et non rigides en condition B	81
5.23	Angle de <i>shear</i> d'un plan en rotation	83
5.24	Résultats de l'expérience 5.5	83
5.25	Résultats du modèle pour l'expérience 5.5	85
5.26	Flux optique en fonction du <i>shear</i>	86
5.27	Percepts extrêmes	87
6.1	Objet ondulé	92
6.2	Amplitude du stimulus	92
6.3	Conditions de l'expérience 6.1.3	93
6.4	Transposition du stimulus	95
6.5	Résultats de l'expérience 6.1.3	96
6.6	Résultats du modèle pour l'expérience 6.1.3	99
6.7	Rappel des résultats de l'expérience 6.1.3	99
6.8	Configurations en conditions A et C	101
6.9	Évaluation de l'amélioration de la question <i>top-far</i>	105
6.10	Évaluation du modèle final de question <i>top-far</i>	108
6.11	Vraisemblance d'une réponse <i>top-far</i>	109
6.12	Distribution sur ${}_x\Phi_x^1$ et ${}_y\Phi_x^1$	118
6.13	Comparaison de la perception entre le modèle de reparamétrisation et le flux optique théorique	121
6.14	Rappel du modèle bayésien de perception.	122

Chapitre 1

Introduction

Agir et percevoir sont des activités cognitives essentielles pour chacun d'entre nous. Les neurosciences modernes ont découvert de fortes interactions entre perception et action : non seulement la perception influence l'action, mais on sait désormais que l'action peut modifier profondément nos percepts.

Dans ce document, nous nous intéressons principalement à l'étude de la perception. Nous verrons, pourtant, que les actions que nous faisons modifient notre perception. Plus spécifiquement, nous étudions la perception visuelle de la forme des objets et nous modélisons la modification de l'appréciation de cette forme par le mouvement.

1.1 Perception

Pour nous, la perception pose essentiellement la question : que peut-on déduire de ce que l'on observe ?

Les observations passent par des sens tels que la vue, l'ouïe, le toucher... La perception opère le regroupement et le recoupement de ces informations en vue de comprendre ou d'agir dans l'environnement.

La perception humaine ne permet pas toujours de reconstruire une représentation objective et précise de l'environnement. Par exemple, la figure 1.1 présente une série d'images extraites d'une animation d'un masque en rotation. Au milieu de cette vidéo, de nombreuses personnes perçoivent une inversion du sens de rotation et du relief du masque de manière à ce qu'il reste convexe. Cette inversion peut s'expliquer avec l'hypothèse, très souvent vérifiée en pratique, qu'un visage est convexe. De nombreuses illusions d'optique sont conçues pour mettre en avant des perceptions erronées d'une image. La ventriloquie est un autre exemple, où la rencontre des informations auditives : « quelqu'un parle », et visuelles : « la marionnette remue les lèvres » induit la perception que la marionnette parle.

Ces exemples montrent surtout qu'une représentation objective et précise de l'environnement n'est en général pas indispensable pour l'action au quotidien.

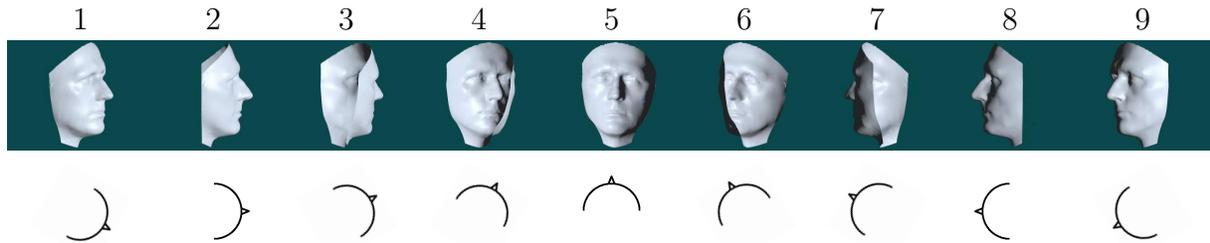


FIG. 1.1 – Animation d’un masque en rotation (<http://www.kyb.mpg.de/bu/demo/index.html>). La ligne du haut présente des images capturées d’une animation d’un masque en rotation de gauche à droite. La ligne du bas montre la position du masque vu de haut. Sur les images 4, 5 et 6 le masque est vu de l’intérieur mais il est la plupart du temps perçu convexe et tournant de droite à gauche.

1.2 Perception de la forme par le mouvement

La perception de la forme par le mouvement consiste à se baser sur le flux optique pour recouvrer des paramètres géométriques de la scène observée.

Le flux optique est le mouvement de l’image sur la rétine. Plus précisément, il s’agit de l’ensemble des vitesses rétiniennes. De ce point de vue, il n’est pas une entrée sensorielle, mais il peut être déduit du flux d’information visuelles (voir 2.1.1). Il dépend à la fois des formes et positions des objets dans l’image, et du mouvement relatif entre chacun des objets de la scène et l’observateur. Le flux optique seul renseigne donc sur la forme, la position et le mouvement des objets observés.

Son utilité première est bien sûr la perception du mouvement. La variation de la taille d’une voiture dans le rétroviseur est une indication de la vitesse à laquelle elle se rapproche et surtout du temps avant qu’elle n’arrive. De la même manière, le défilement du paysage nous renseigne sur notre direction et notre vitesse mais aussi sur les positions relatives entre les arbres ou les collines.

À l’image de nombreux travaux de la littérature, nous nous intéressons plus spécifiquement à la perception d’un plan par le flux optique. Il s’agit d’une instance du problème général de la forme par le mouvement qui présente, de manière simple, plusieurs ambiguïtés génériques (orientation, sens de rotation, inclinaison, etc.).

Les interactions avec d’autres modalités sensorielles, telles que la texture ou l’illumination, peuvent aider à résoudre certaines ambiguïtés. Dans les expériences de psychophysique, les stimuli sont donc simplifiés (grilles, ensembles de points, ...) pour contenir le moins possible d’indices autres que le flux optique.

Dans cette thèse, nous proposons une vue unifiée sur cinq expériences de psychophysique de perception de plan par le flux optique.

1.2.1 Difficultés

L'extraction de formes par le mouvement présente trois difficultés communes aux problèmes de perception : l'inversion, l'ambiguïté et l'incomplétude.

Problème inverse Un problème est qualifié d'inverse s'il s'attache à calculer l'antécédent d'un élément par une fonction donnée. Pour la perception, la physique permet de déterminer la fonction directe, qui donne l'observation possible en fonction de l'état du monde.

De manière générale, la perception est un problème inverse (Poggio 1984 ; Pizlo 2001). Il n'est pas toujours possible d'inverser analytiquement les fonctions directes d'observation. Il faut donc trouver un autre moyen pour résoudre ce type de problème.

Dans le cas de la perception de la forme par le flux optique, l'observation est le déplacement de l'image sur la rétine. Ce déplacement est la projection du mouvement de l'objet. La connaissance de la forme et du mouvement d'un objet suffisent à en donner la projection *via* les règles de géométrie projective.

Problème mal posé Un problème mal posé est un problème pour lequel il n'existe pas une unique solution. Pour la perception, la dimensionnalité de l'environnement est plus grande que celle des observations que l'on peut en faire. En pratique, plusieurs configurations de l'environnement correspondent à une même observation. À partir d'une telle observation, la perception ne peut donc pas discriminer les différentes possibilités. La perception est donc généralement un problème mal posé.

Dans le cas de la perception de la forme par le flux optique, il s'agit d'une projection d'un mouvement sur la rétine. Or une projection est une application non injective, c'est-à-dire que plusieurs formes et mouvements ont exactement la même projection. Comme la fonction d'observation, directe, n'est pas injective, le problème inverse de perception de forme par le flux optique est mal posé.

Incomplétude Les modèles ne sont toujours que des approximations de la réalité. Ces approximations permettent de comprendre et de raisonner sur un phénomène pour en capturer l'essentiel. Un modèle néglige sciemment certains éléments jugés non pertinents vis-à-vis de ses objectifs. Ces éléments, appelés parfois variables cachées, sont la source d'une variabilité des résultats dans des conditions indiscernables des variables retenues dans le modèle. La modélisation de la perception ne fait pas exception et est confrontée à l'incomplétude.

Dans le cadre de la perception de formes par le flux optique, de nombreux éléments peuvent rentrer en compte pour modifier, légèrement ou non, l'observation. Il peut s'agir, par exemple, de défauts oculaire tels que l'astigmatisme¹ ou d'irrégularités dans la propagation de la lumière dont des cas extrêmes sont les mirages.

¹Anomalie de la vision souvent due à des inégalités de courbure de la cornée. En conséquence, l'image peut apparaître floue et déformée.

1.2.2 Objectifs

Nous souhaitons proposer un modèle *psychophysique simple* de la perception de la forme par le mouvement.

Psychophysique La psychophysique est l'étude des corrélations entre les stimuli et les sensations sans s'intéresser à la physiologie précise de ces relations. Cette étude se base sur des expériences mettant en jeu des participants humains. Ces observateurs, généralement naïfs vis-à-vis du but de l'expérience, sont placés dans des conditions de perception précises et doivent répondre à des questions relatives à leurs sensations.

Le but est de formuler et vérifier des hypothèses sur la perception. On construit donc des modèles autour de ces hypothèses pour prédire des réponses. Ces réponses sont alors comparées à celles des participants pour décider de la validité des hypothèses.

En particulier, il est nécessaire d'éliminer le plus possible les interférences dues à d'autres modalités sensorielles. À cette fin, l'expérimentateur conçoit et contrôle l'environnement. De plus les instructions données au sujet, ainsi que les contraintes imposées aux modalités de réponses (choix forcé, temps limité...) sont soigneusement définies en fonction des buts de l'expérience, pour éviter autant que possible l'interférence avec les capacités conceptuelles des sujets.

L'objectif d'un modèle psychophysique est de reproduire les sensations des observateurs et non pas de reconstruire la scène présentée. Un bon modèle doit se tromper de la même manière que les participants à l'expérience.

Simplicité Notre ambition est de restituer la complexité de la perception sous une forme accessible à la compréhension.

La simplicité d'un modèle se présente sous plusieurs aspects. Le premier est le nombre et la complexité des hypothèses. À partir de ces hypothèses, il y a la facilité de spécification du modèle. Cela inclut le nombre et la sémantique des paramètres. Enfin, il y a l'interprétabilité du fonctionnement et des résultats du modèle.

Nous souhaitons construire un modèle avec le moins possible d'hypothèses, facile à spécifier avec un nombre réduit de paramètres réifiables.

1.3 Modèle proposé

1.3.1 Modélisation bayésienne

La modélisation bayésienne s'attache à manipuler des distributions de probabilité à l'aide, notamment, de la règle de Bayes.

Probabilités Nous désirons modéliser les connaissances d'un sujet confronté à une expérience partielle du monde. Pour ce faire, nous utilisons les probabilités pour indiquer ce à

quoi il est raisonnable de s'attendre² (Cox 1946). Il est important de noter que nous nous plaçons plutôt dans une approche dite subjectiviste des probabilités. En particulier, les modèles que nous proposons n'ont pas pour objectif de décrire le monde, mais le raisonnement que peut faire un observateur (un sujet) confronté à des informations incomplètes et incertaines sur son environnement. Les probabilités sont donc un outil pour manipuler ces informations et pas une caractéristique intrinsèque de l'environnement. Ainsi, les hypothèses que nous sommes amenées à exprimer sont des simplifications dans le raisonnement de l'observateur et non pas des règles de fonctionnement du monde.

Les variables cachées, qui rajoutent de la variabilité dans les résultats du modèle sont prises en compte *via* l'ajout d'incertitude sur les distributions de probabilité sur les variables étudiées.

Les distributions de probabilité peuvent être multimodales. Cela signifie en particulier que la distribution peut exprimer que plusieurs possibilités sont plus probables que les autres. Ces possibilités peuvent être les différentes solutions d'un problème mal posé. Une distribution de probabilité peut aussi représenter les probabilités relatives de ces différentes solutions.

L'utilisation des probabilités permet de résoudre des problèmes mal posés et de tenir compte de l'incomplétude en la transformant en incertitude.

Règle de Bayes Les distributions de probabilité, qui représentent les connaissances d'un modèle bayésien, sont manipulées par deux règles d'inférence : la règle de marginalisation et la règle de Bayes. La règle de marginalisation permet de passer de la connaissance conjointe sur deux variables à la connaissance sur l'une d'elles.

La règle de Bayes relie la connaissance sur deux variables à celle sur l'une de ces variables en supposant être certain de la valeur de l'autre. Cette règle est symétrique et l'on peut passer de la connaissance d'une variable conditionnée par une seconde à celle de la seconde conditionnée par la première, *via* la distribution conjointe.

Cette règle permet d'inverser une spécification : à partir de connaissances probabilistes sur l'observation sachant la configuration observée, on peut inférer des connaissances sur la configuration observée sachant l'observation. Un problème inverse peut donc se résoudre à l'aide de l'application de la règle de Bayes à des distributions de probabilité.

1.3.2 Contribution

Nous proposons un modèle simple de la perception du plan par le flux optique.

Formalisme bayésien Ce modèle est exprimé dans le formalisme de la programmation bayésienne. Ce formalisme permet de traiter les caractères inverse et mal posé du problème de perception et de prendre en compte l'incomplétude du modèle. Nous verrons qu'il s'agit ici d'un véritable mécanisme de programmation bayésienne, comme cela sera illustré au chapitre 3 et illustré ensuite.

²*reasonable expectation*

Expression des hypothèses Dans ce formalisme, nous construisons un modèle basé sur des hypothèses issues de la littérature. Les principales sont la rigidité qui suppose l'objet observé non déformable et la stationnarité qui fait l'hypothèse d'un petit mouvement.

Ces hypothèses se retrouvent dans la structure du modèle et dans les choix des distributions de probabilités. Les paramètres nécessaires à ces distributions, par exemple l'amplitude moyenne du mouvement attendu, sont réifiables.

Compréhension des résultats Nous nous attachons à reproduire les résultats de 5 expériences de psychophysique dans lesquelles les participants observent le flux optique de plans en mouvement.

Par exemple, nous montrons que l'on préfère parfois percevoir un objet qui se déforme mais qui ne bouge pas, plutôt qu'un objet qui bouge beaucoup mais qui ne se déforme pas. Notre modèle explique aussi que l'on arrive mieux à évaluer l'inclinaison d'un plan dont on voit une grande partie qu'une petite partie. Enfin nous exhibons une relation sur l'amplitude des rotations et des translations attendues par les observateurs humains.

1.4 Plan du manuscrit

Nous commençons ce manuscrit par une étude bibliographique organisée en trois parties : le flux optique en vision par ordinateur, la psychophysique de la perception de plan par le mouvement et l'outil de modélisation bayésienne. De cette étude, nous dégageons le formalisme dans lequel nous construisons notre modèle, les hypothèses sur lesquelles il se base ainsi que les expériences pour le valider.

Dans une deuxième partie, nous présentons plus en détails une expérience de perception d'un plan. Nous proposons un modèle bayésien simple comme exemple didactique de programmation bayésienne.

Le chapitre 4 est consacré à l'exposé de notre modèle bayésien de perception d'un plan par le flux optique.

Nous confrontons ensuite les résultats de ce modèle à cinq expériences de psychophysique issues de la littérature. À chaque fois, nous expliquons en détails le fonctionnement du modèle et les hypothèses principales induisant ces résultats.

Dans le chapitre 6, nous nous intéressons à l'intégration de notre modèle au sein de systèmes perceptifs plus importants. Nous abordons la problématique d'extraction du flux optique et étudions l'adaptation de notre modèle à des expériences de perception du flux optique ne portant pas sur des plans. Pour cela nous introduisons la notion d'interface à partir de laquelle envisager les combinaisons de modèles.

Enfin nous concluons en récapitulant nos contributions et présentons des perspectives de recherche ouvertes par ce travail.

Chapitre 2

État de l'art

Dans ce chapitre nous résumons les travaux existants sur la perception du flux optique. Nous distinguons le point de vue de la vision par ordinateur et l'approche psychophysique de ce problème. Enfin nous nous intéressons à l'outil de modélisation bayésienne que nous employons dans ce document.

2.1 Vision par ordinateur

La vision artificielle regroupe un large spectre d'applications. Nous examinons les travaux effectués sur l'extraction du flux optique et de forme par le mouvement.

2.1.1 Extraction du flux optique

De nombreux travaux portent sur l'extraction de flux optique (revues par Barron *et al.* (1994a) ; Beauchemin et Barron (1996)). Ces algorithmes se basent sur une suite de deux ou plusieurs images. Une image est en général un champ de *pixels*, avec une information d'intensité lumineuse attachée à chacun.

Le flux optique de sortie de ces algorithmes est un champ de vecteurs bidimensionnels. Chaque vecteur indique le déplacement en ce point de l'image. En général, on souhaite obtenir un champ de vecteurs dense, c'est-à-dire une valeur du flux en chaque point.

Dans la littérature de vision par ordinateur, on distingue habituellement trois techniques d'extraction du flux optique : les méthodes différentielles, fréquentielles ou par corrélation.

Une hypothèse de base commune à beaucoup de modèles est que l'intensité d'un point en mouvement est constante. En notant $I(x, y, t)$ l'intensité d'un point en position (x, y) de l'image à l'instant t et (u_x, u_y) la vitesse de déplacement de ce point en cet instant, on écrit que la dérivée par rapport au temps du champ d'intensité est nulle, ce qui amène à l'équation fondamentale (Horn et Schunck 1981) :

$$\frac{\partial I}{\partial x} \cdot u_x + \frac{\partial I}{\partial y} \cdot u_y + \frac{\partial I}{\partial t} = 0 \quad (2.1)$$

Le système d'équations ainsi obtenu est sous-contraint et biaisé. Pour le résoudre, les méthodes différentielles utilisent les dérivées spatio-temporelles en rajoutant des contraintes notamment de continuité.¹

Les méthodes fréquentielles se basent sur du filtrage de l'image. On applique un filtre sur l'image d'intensité lumineuse de telle sorte que les valeurs obtenues se trouvent sur un hyperplan caractérisant le flux optique.²

Les méthodes par corrélation consistent à mettre en correspondance un certain nombre de vignettes entre les différentes images d'entrée au cours du temps. Une première approche consiste en la caractérisation de la vignette de départ puis en la recherche de cette vignette dans les images suivantes.³ Une seconde approche part du principe que l'on peut calculer la vignette suivante à partir de la vignette courante et de son mouvement. La comparaison de la vignette prévue avec celle observée permet d'évaluer la plausibilité du mouvement.⁴ À la limite des méthodes par corrélations, existent aussi des méthodes d'appariement d'amers entre plusieurs images.

Ces différentes méthodes sont plus ou moins adaptées à des conditions différentes d'image, de mouvement et de bruit (revue par Barron *et al.* (1994b)). Par exemple, les méthodes par corrélation ne peuvent pas mesurer de mouvement en dessous d'un pixel mais sont plus efficaces sur des images faiblement échantillonnées. D'un autre côté, les méthodes différentielles sont assez sensibles à la qualité des images et leurs performances peuvent parfois être améliorées par des prétraitements.

En résumé, l'extraction du flux optique à partir d'une série d'image n'est pas un problème totalement résolu. Les résultats de tels algorithmes sont entachés d'incertitude qu'il est important de prendre en compte dans leur interprétation.

2.1.2 Forme par le mouvement

En général, une scène réelle comporte plusieurs objets et plusieurs mouvements associés. Le flux optique n'est qu'un indice parmi d'autres. Même si les problèmes de segmentation et de fusion ne sont pas explicitement traités ici, ils constituent des perspectives de développements futurs importants. Il existe de nombreux travaux en vision par ordinateur sur ces questions (revue par Faugeras (1993) ; Shapiro et Stockman (2001) ; Forsyth et

¹Horn et Schunck (1981) font l'hypothèse supplémentaire que le flux est continu. Lucas et Kanade (1981) appliquent une méthode des moindres carrés autour d'un voisinage dans lequel le flux est supposé constant. Nagel (1987) utilise les dérivées secondes du champ d'intensité pour minimiser la variation du flux perpendiculairement au gradient d'intensité.

²Adelson et Bergen (1985) utilisent par exemple la transformée de Fourier. Les coefficients non nuls se trouvent sur un plan dont l'équation donne la vitesse de translation. Selon le même principe, Heeger (1988) utilise des familles de filtres de Gabor et une méthode des moindres carrés pour obtenir l'équation du plan. Fleet et Jepson (1990) utilisent des filtres passe-bande dont ils extraient la phase qui donne la vitesse de déplacement.

³Par exemple, Barnard et Thompson (1980) sélectionnent des vignettes de taille 5×5 en fonction des différences d'intensité avoisinantes et calculent itérativement des disparités à l'aide de probabilités.

⁴Anandan (1989) utilise une méthode hiérarchique dans laquelle une première corrélation à grande échelle permet de mieux appairer les échelles plus fines à l'aide d'une descente de gradient.

Ponce (2002)) et nous allons nous concentrer ici sur la vision d'un seul objet à l'aide du seul flux optique.

L'extraction de la forme par le mouvement est le processus de recouvrer des informations sur la scène observée à partir du flux optique⁵ ou directement à partir d'une série d'images (revue par Huang et Netravali (1994)).

Les approches utilisant une série d'images cherchent le plus souvent à suivre le mouvement de quelques points d'intérêt entre les images. Cette approche est similaire aux méthodes par corrélation pour l'estimation du flux optique, en se limitant à quelques points au lieu de chercher à obtenir un champ de mouvement dense. À partir du mouvement de ces points, ou du flux optique, vient une étape d'estimation du mouvement et de la structure. Le problème se pose en général en termes de résolution d'un ensemble d'équations qui dépendent du modèle de caméra choisi.⁶

Le mouvement extrait peut être celui d'un objet, mais aussi le mouvement propre de la caméra. À partir d'informations sur les mouvements des objets, voire en les supposant fixes, l'extraction de la structure et du mouvement à partir d'une série d'images devient également un indice sur le mouvement en trois dimensions de la caméra (Fermüller et Aloimonos 2000). De la même manière, ces algorithmes sont parfois utilisés avec deux ou plusieurs caméras comme information sur leurs positions relatives. On s'intéresse alors à la stéréovision sans calibration de la position relative des caméras.⁷

2.2 Psychophysique

À côté de ces travaux en vision par ordinateur, on trouve des études de la perception du flux optique d'un point de vue psychophysique. Nous nous intéressons à l'utilisation du flux optique avant d'entrer dans les détails de la perception de formes par le flux optique.

2.2.1 Flux optique

La capacité d'extraction d'informations sur la forme à partir du flux optique est l'objet d'étude en psychophysique depuis longtemps (von Helmholtz. 1867 ; Gibson 1950 ; Wallach et O'Connell 1953).

⁵Adiv (1985, 1989) propose de segmenter le flux optique par rapport au mouvement rigide d'une surface plane puis de regrouper ces segments en fonction de leur mouvement 3D pour correspondre à un objet rigide

⁶Koenderink et van Doorn (1991) utilisent une projection affine. La géométrie projective a été utilisée en particulier par Faugeras (1992) et Triggs (1996). Kahl et Triggs (1999) utilisent aussi la géométrie euclidienne. Pour résoudre ces équations, McLauchlan et Murray (1995) utilisent des processus gaussiens, Oliensis (2005) propose simplement une méthode des moindres carrés. Viéville *et al.* (1996) utilisent les colinéations, transformations respectant l'alignement, pour caractériser et segmenter les plans en mouvement dans une série d'images.

⁷Faugeras (1992) montre que les informations obtenues à partir de caméras non calibrées sont relatives à des points de référence. Hartley *et al.* (1992) proposent une méthode non itérative de calcul du placement de 3 points en se basant sur la matrice essentielle.

Une première utilisation du flux optique est l'estimation du mouvement propre (Warren 1995). La majorité de notre environnement est plutôt statique (sol, murs, arbres...). Le flux optique que l'on perçoit est donc principalement dû à notre propre déplacement dans l'environnement. Le flux optique peut induire la perception de mouvement : il s'agit du phénomène devection.⁸ Le flux optique renseigne sur la direction du déplacement : c'est ce que l'on appelle le *heading* (Lappe *et al.* 1999). Enfin, le flux optique donne des indications de temps avant impact.⁹ Dans ce document, nous nous restreignons à l'utilisation du flux optique pour la perception de formes.

Ce phénomène regroupe en premier lieu la parallaxe de mouvement (*motion parallax*). Il s'agit de l'effet d'une translation relative entre l'observateur et la scène. La vitesse angulaire est inversement proportionnelle à la distance entre l'observateur et l'objet. En conséquence, la variation de vitesse angulaire, mesurée sur la rétine, informe sur l'écart de profondeur.¹⁰

Un second phénomène est l'effet cinétique de profondeur (*kinetic depth effect*) qui s'intéresse à la perception de mouvement et de formes en rotation. On s'est d'abord intéressé à la projection d'un cylindre et l'effet de profondeur induit (Wallach et O'Connell 1953). On a ensuite utilisé des stimuli générés par ordinateur et en particulier des *kinematograms* et stéréogrammes à base de points aléatoires (Green 1961 ; Braunstein 1962). L'utilisation de points pilotés par ordinateur permet d'avoir une plus grande liberté sur le stimulus et même de présenter des mouvements qui ne correspondent pas à un objet réel.¹¹

Le flux optique est généré principalement par le mouvement relatif. Si on présente un objet immobile à un observateur en mouvement ou un objet en mouvement à un observateur immobile, le flux optique est le même si le mouvement relatif est le même. Dans les deux cas, un participant est capable d'extraire des informations sur la forme et le mouvement de l'objet (Wallach *et al.* 1974 ; Rogers et Graham 1979). Les premiers travaux se sont donc intéressés au flux optique seul.

Des études se sont focalisées sur les aspects physiques et calculatoires du flux optique. Ullman (1979) montre qu'avec trois vues de quatre points d'un objet rigide, il est possible d'en reconstruire la structure. C'est ce qu'il appelle *structure from motion*. Des études suivantes (Longuet-Higgins et Pradny 1980 ; Longuet-Higgins 1984 ; Ullman 1986)

⁸Un exemple devection peut être expérimenté dans un train en gare. Lorsque le train d'à côté démarre, on a parfois l'impression de bouger soi-même.

⁹Lee (1981) montre que les fous de Bassan utilisent le flux optique, et en particulier le taux d'expansion, pour décider du moment de replier leurs ailes au moment de la rentrée dans l'eau à l'issue d'un piqué.

¹⁰Gibson *et al.* (1959) font des expériences avec deux vitesses de translation relative, ce qui amène à la perception de deux profondeurs, et d'un gradient de vitesse, perçu comme une surface inclinée. À l'aide d'un oscilloscope, Rogers et Graham (1979) simulent des surfaces de différents profils (carré, sinusoïdal, ...) et montrent les perceptions majoritairement correctes de ces profils. Cornilleau-Pérès et Droulez (1989) étudient la sensibilité de la perception de courbures cylindriques.

¹¹Todd (1982) s'intéresse d'abord à la perception de la rigidité à partir de trajectoire prédéfinies puis, avec Norman, à la discrimination de surfaces sinusoïdales (Todd et Norman 1991) et la perception de transformations affines (Norman et Todd 1993). Sperling *et al.* (1989) définissent un catalogue d'objets 3D et montrent qu'ils sont reconnaissables par le flux induit par une rotation. Mulligan (1992) examine aux biais de perception induits par les courbes de Lissajou possédant plusieurs interprétations. Mamassian et Bühlhoff (1996) s'intéressent à la perception de bords arrondis ou aigus d'un objet en rotation et montrent que le mouvement propre permet de mieux percevoir les bords aigus.

s'intéressent à la perception en tant que processus calculant les paramètres décrivant la forme observée, sa position et son mouvement à partir du flux optique.

La forme observée est une surface définie dans un repère (x, y, z) (avec z l'axe sagittal) par l'équation $z = f(x, y)$. Les paramètres de la forme sont les valeurs à l'origine des dérivées d'ordre 1 et 2 de cette fonction selon les deux axes. Dans le cas d'un objet planaire, f est linéaire et les trois dérivées d'ordre 2 sont nulles. Dans ce document, nous notons χ et v (*chi* et *upsilon* minuscule) les dérivées respectives selon les axes x et y (voir figure 2.1).¹² L'équation du plan dans le repère (x, y, z) est donc :

$$\begin{aligned}(\tilde{x}, \tilde{y}, \tilde{z}) \cdot (\chi, v, -1) &= 0 \\ \tilde{x}\chi + \tilde{y}v - \tilde{z} &= 0\end{aligned}$$

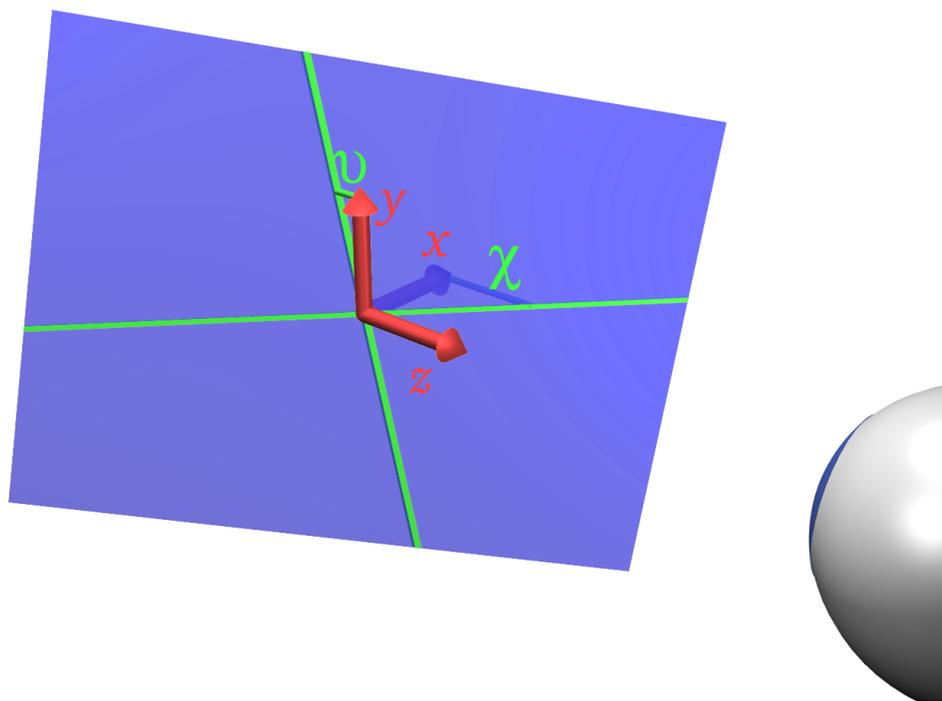


FIG. 2.1 – Illustration des paramètres (X, Y) caractérisant un plan. L'objet est le plan bleu et l'observateur est placé en $(0, 0, d > 0)$. En vert sont figurées les intersections du plan avec les plans horizontal et sagittal dont les pentes sont les paramètres X et Y respectivement.

¹²Nous supposons que le plan passe par le point de fixation. Ainsi deux paramètres suffisent pour en décrire la position. Cette hypothèse nous permet d'éviter le problème du facteur d'échelle indéterminé en vision monoculaire. Notons que les tâches auxquelles sont confrontés les observateurs ne sont pas affectées par cette indétermination.

Les observations sur lesquelles se basent les algorithmes proposés par Longuet-Higgins (1984) et Ullman (1986) sont des combinaisons linéaires des dérivées spatiales du champ de vitesse.

Nous notons le champ de vitesse $\vec{\Phi}$; la vitesse en un point (x, y) de l'image est $\vec{\Phi}(x, y) = (u_x, u_y)$ avec u_x et u_y les vitesses dans l'image selon les axes x et y respectivement. Nous appelons $(\phi_x^0, \phi_y^0) = \vec{\Phi}(0, 0)$ la vitesse dans l'image du point de fixation. Il s'agit de la dérivée d'ordre 0 du flux optique en l'origine. De la même manière, nous notons ${}_y\phi_x^1$ la valeur en $(0, 0)$ de la composante selon l'axe x de la dérivée selon l'axe y du champ de vitesse. Les quatre composantes ${}_x\phi_x^1$, ${}_y\phi_x^1$, ${}_x\phi_y^1$ et ${}_y\phi_y^1$ forment la matrice jacobienne du flux optique $\vec{\Phi}$ en $(0, 0)$. Dans le cas d'un plan, seules deux des composantes des dérivées secondes du flux sont non nulles. Nous les appelons ${}_x\phi^2$ et ${}_y\phi^2$. Considérant le mouvement d'un objet rigide planaire et un modèle de caméra perspective,¹³ le flux optique s'écrit alors (voir annexe A pour une démonstration) :

$$\begin{aligned} \vec{\Phi}(x, y) &= \vec{\phi}^0 + \vec{\phi}^1 \cdot {}^t(x, y) + {}^t(x, y) \cdot {}^t\vec{\phi}^2 \cdot {}^t(x, y) \\ \begin{pmatrix} u_x \\ u_y \end{pmatrix} &= \begin{pmatrix} \phi_x^0 + {}_x\phi_x^1 \cdot x + {}_y\phi_x^1 \cdot y + {}_x\phi^2 \cdot x^2 + {}_y\phi^2 \cdot xy \\ \phi_y^0 + {}_x\phi_y^1 \cdot x + {}_y\phi_y^1 \cdot y + {}_x\phi^2 \cdot xy + {}_y\phi^2 \cdot y^2 \end{pmatrix} \end{aligned}$$

où tA est la transposée d'une matrice A .

À partir de ces 8 composantes, il s'agit de recouvrer les deux paramètres χ et v nécessaires pour définir un plan ainsi que les 6 paramètres de son mouvement : 3 pour la rotation (que nous notons $\vec{\omega} = (\omega_x, \omega_y, \omega_z)$) et 3 pour la translation (notés $\vec{t} = (t_x, t_y, t_z)$).

Avec ces notations, les équations à résoudre sont :

$$\begin{cases} \phi_x^0 &= t_x \\ \phi_y^0 &= t_y \\ {}_x\phi_x^1 &= t_z + \chi(t_x + \omega_y) \\ {}_y\phi_x^1 &= -\omega_z + v(t_x + \omega_y) \\ {}_x\phi_y^1 &= \omega_z + \chi(t_y - \omega_x) \\ {}_y\phi_y^1 &= t_z + v(t_y - \omega_x) \\ {}_x\phi^2 &= \chi t_z - \omega_y \\ {}_y\phi^2 &= v t_z + \omega_x \end{cases} \quad (2.2)$$

Longuet-Higgins et Pradny (1980) ; Longuet-Higgins (1984) ; Waxman et Ullman (1985) ; Ullman (1986) proposent en particulier des algorithmes amenant à la résolution de ces systèmes. Ces méthodes directes ont conduit à une bonne compréhension du problème et de ses ambiguïtés et à la proposition de nombreuses études psychophysiques. Du point de vue de la vision par ordinateur, ces méthodes directes se sont montrées sensibles aux bruits et n'ont pas conduit à des simulations satisfaisantes.

¹³Nous considérons que le système connaît ses paramètres intrinsèques et nous ne discuterons pas de la calibration.

2.2.2 Forme par le mouvement

Pour extraire la forme par le mouvement, il est donc nécessaire de recourir à des informations supplémentaires, des hypothèses ou des heuristiques. C'est le cas pour ces algorithmes. Historiquement la première hypothèse formulée est celle de la *rigidité*. Des données expérimentales supplémentaires ont amené ensuite à la proposition de la *stationnarité*.

Rigidité

Les hypothèses sur lesquelles se basent les algorithmes précédents excluent par principe le cas des objets non rigides. La rigidité permet une grande réduction de la dimensionnalité de la solution (Ullman 1979) et est la base de nombreux travaux (Todd 1982 ; Koenderink 1986).

La rigidité exprime que le mouvement des points est corrélé par la forme et le mouvement de l'objet. Les points ne sont plus indépendants et leurs mouvements relatifs sur l'image peut se déduire de leurs positions et du mouvement de l'objet.

Dans l'espace, la rigidité minimise le *mouvement relatif* entre les points.

On peut écrire les équations donnant la vitesse d'un point de l'image en fonction de son mouvement de l'espace. La rigidité rajoute donc une minimisation à ces équations pour sélectionner une solution particulière. C'est un critère de choix justifiable de manière écologique par le fait qu'une grande partie des objets soumis à notre observation sont rigides ou encourent de faibles déformations.

L'équation 2.2 inclut déjà la rigidité par la référence au mouvement $(\vec{\omega}, \vec{\tau})$ du plan.

L'hypothèse de rigidité présente un certain nombre de problèmes. En premier lieu, pour un même flux optique, différentes réponses sont parfois observées pour des mouvements propres différents, ce que la rigidité ne permet pas de prévoir. Ces différences sont mesurées en particulier entre un observateur statique avec un objet en mouvement et un observateur en mouvement avec un plan statique.¹⁴

Un autre problème est que les objets perçus sont parfois non-rigides ce qui exclut le caractère absolu de la rigidité.¹⁵

Stationnarité

Pour expliquer le rôle du mouvement propre, on a proposé l'hypothèse de stationnarité (Wexler *et al.* 2001b ; Cornilleau-Pérès *et al.* 2002 ; Wexler 2003). Cette hypothèse stipule

¹⁴Rogers et Rogers (1992) montrent que la projection perspective ou le mouvement propre permettent à un observateur de déterminer le caractère concave ou convexe d'une ondulation. Dijkstra *et al.* (1995) obtiennent des résultats similaires pour la perception de sections de sphères en conditions de mouvement propre ou avec un grand champ de vision. Wexler *et al.* (2001b) s'intéressent à la perception de l'orientation 3D de grilles en condition de conflit entre le mouvement propre et la perspective. Panerai *et al.* (2002) ; Peh *et al.* (2002) examinent la contribution du mouvement propre dans la perception de distances absolues.

¹⁵Todd (1982) étudie la perception de mouvements non rigides. Wexler *et al.* (2001a) montrent un conflit entre des objets rigides et stationnaires amenant à la perception d'objets non rigides. Cornilleau-Pérès *et al.* (2002) s'intéressent à la perception de grilles carrées et trapézoïdales.

que l'objet perçu est le plus stationnaire possible dans le référentiel de la scène. Il s'agit d'un critère de choix entre plusieurs solutions des équations du flux optique.

La stationnarité minimise le *mouvement absolu* des points.

Ce critère, par sa référence au repère de la scène, est allocentré. Comme le flux optique ne concerne que le mouvement relatif, ce critère met en jeu le mouvement propre du sujet, à partir duquel le mouvement absolu peut être inféré. Le mouvement propre est donc un ajout important par rapport aux travaux précédents qui ne considèrent que le flux optique pour la perception de la forme. Le mouvement propre induit un mouvement relatif et est particulier en ce qu'il résulte souvent de l'action motrice de l'observateur et qu'il peut être perçu par le système vestibulaire.

La stationnarité et la rigidité sont deux critères distincts permettant de discriminer parmi les multiples solutions des équations du flux optique. Chacun des deux est supporté par des résultats expérimentaux. Chacun s'attache à minimiser une quantité, il arrive que les minima ne correspondent pas. Il y a alors conflit entre ces deux hypothèses ce qui peut amener à la perception d'objets non rigides.

À notre connaissance, il n'existait pas, jusqu'alors, de modèle permettant de concilier ces deux hypothèses.

2.3 Modélisation bayésienne

L'outil utilisé dans ce manuscrit pour construire et utiliser notre modèle est la modélisation bayésienne. Dans un premier temps, nous présentons le formalisme précis de modélisation que nous employons. Nous regardons ensuite l'application de ce type de modèle à des systèmes biologiques.

2.3.1 Outil de modélisation

La modélisation bayésienne s'appuie sur les probabilités pour représenter des connaissances. Nous utilisons le formalisme de la programmation bayésienne, qui donne un cadre pour leur organisation. Le but d'un programme bayésien est d'inférer de nouvelles connaissances à l'aide en particulier de la règle de Bayes.

Probabilités et représentation des connaissances

Nous souhaitons construire un modèle calculant ce à quoi il est raisonnable de s'attendre¹⁶ à partir d'un certain nombre de connaissances préalables et d'observations. Il a été montré et discuté que les probabilités sont à même de représenter la plausibilité de propositions.¹⁷ Les probabilités sont utilisées comme une extension de la logique classique

¹⁶*reasonable expectation*

¹⁷Cox (1946, 1961) montre qu'un système formel manipulant la plausibilité d'une proposition et respectant quelques desiderata est isomorphe au calcul probabiliste. Halpern (1999a,b) critique cette preuve et la restreint aux domaines finis de représentation. Arnborg et Sjödin (2000, 2001, 2003) revient sur les

aristotélicienne.¹⁸ La modélisation bayésienne manipule des variables selon les règles des probabilités à l'image d'un raisonnement logique. Nous présentons ici les probabilités et les règles d'inférence.

Variabes et distributions Une *variable* V possède un *domaine* \mathcal{D}_V dans lequel elle peut prendre une valeur v .¹⁹ Il n'est pas rare de ne pas savoir avec précision quelle valeur possède cette variable.

Dans ce cas, notre connaissance de la valeur de cette variable s'exprime sous la forme d'une *distribution de probabilité* sur son domaine. Une distribution de probabilité sur V est une fonction P de \mathcal{D}_V vers \mathbb{R}^+ d'intégrale (ou somme) unité sur \mathcal{D}_V :

$$\int_{\mathcal{D}_V} P(V) = 1 \quad (2.3)$$

Inférence Pour manipuler ces objets, la théorie des probabilités introduit certains opérateurs. Pour deux variables A et B de domaines respectifs \mathcal{D}_A et \mathcal{D}_B , on peut considérer leur *conjonction* $A \wedge B$ (commutative). Il s'agit d'une nouvelle variable, de domaine $\mathcal{D}_{A \wedge B} = \mathcal{D}_A \times \mathcal{D}_B$, précisant la valeur à la fois de A et de B . On peut considérer la distribution de probabilité $P(A \wedge B)$ que nous noterons $P(A B)$ pour alléger les notations.

Une autre manière de s'intéresser aux relations entre des variables est de considérer sa connaissance (au sens distribution de probabilité) d'une première variable en supposant connaître exactement la valeur d'une seconde. Pour deux variables A et B et pour $b \in \mathcal{D}_B$, on note $P(A \mid [B = b])$, ou plus simplement $P(A \mid b)$, la distribution de probabilité sur A sachant que B a pour valeur b . On appelle *distribution conditionnelle* et on note $P(A \mid B)$, l'ensemble des telles distributions pour tout le domaine \mathcal{D}_B : $P(A \mid B) = \{P(A \mid b) \mid b \in \mathcal{D}_B\}$.

La théorie des probabilités nous fournit des règles pour manipuler ces nouvelles distributions. La première est la règle de Bayes ou règle de la conjonction, énoncée équation 2.4 :

$$P(A B) = P(A \mid B)P(B) = P(B \mid A)P(A) \quad (2.4)$$

Cette règle est importante car elle met en relation la distribution conjointe avec les distributions conditionnelles. On remarque qu'elle est symétrique en A et B du fait de la commutativité de la conjonction.

La seconde règle utile à l'inférence est la marginalisation, présentée équation 2.5. Cette règle découle de la normalisation des distributions 2.3 et de la règle de Bayes 2.4.

$$\int_{\mathcal{D}_A} P(A B) = P(B) \quad (2.5)$$

définitions précises de ces desiderata. La représentation de la plausibilité par les probabilité est reprise entre autres par Jaynes (2003) et van Horn (2003).

¹⁸L'usage des probabilités comme extension de la logique est déjà proposée par de Laplace (1814) et expliquée par Jaynes (2003).

¹⁹Dans tout ce document les variables sont désignées par des lettres majuscules et leurs valeurs par la minuscule correspondante.

Enfin, dans le cas où $\forall b \in \mathcal{D}_B, P(A | b) = P(A)$, c'est-à-dire $P(A | B) = P(A)$ on dit que A et B sont *indépendantes*. On a alors $P(B | A) = P(B)$ et $P(A B) = P(A)P(B)$. On dit aussi que A est indépendante de C sachant B si et seulement si $P(A | B C) = P(A | B)$. Il faut bien noter que ces deux derniers concepts d'indépendance (marginale) et d'indépendance conditionnellement à une autre variable ne s'impliquent pas l'un l'autre. On peut ainsi avoir deux variables indépendantes A et B qui ne sont pas indépendantes sachant une troisième variable C et inversement l'indépendance de deux variables A et B conditionnellement à une troisième variable C , n'implique en rien leur indépendance.

Notions de programmation bayésienne

Il existe plusieurs formalismes de modélisation bayésienne qui ont pour point commun de s'intéresser à la distribution conjointe. Les modèles graphiques, qui recouvrent principalement les processus de Markov, les réseaux bayésiens et les champs de Markov sont la rencontre entre la théorie des probabilités et la théorie des graphes.²⁰

La programmation bayésienne²¹ s'appuie sur une écriture algébrique de la distribution conjointe. Elle s'attache à fournir un cadre pour spécifier des connaissances probabilistes afin de pouvoir répondre à une question qui se présente sous la forme d'une distribution de probabilité. Le cadre se présente comme un processus en 4 parties :

- définition des *variables* sur lesquelles raisonner ;
- *décomposition* de la distribution conjointe en produit de facteurs plus simples ;
- *spécification* de chacun des facteurs de la décomposition ;
- *question* et inférence.

La décomposition est une étape de la spécification de la distribution conjointe portant sur la structure du modèle. Le but est de la simplifier. Pour cela, on utilise la règle de Bayes pour l'écrire sous la forme d'un produit. Le programmeur fait des hypothèses d'indépendance conditionnelle pour simplifier les facteurs de la conjointe. On obtient alors une décomposition de la distribution conjointe.

On spécifie ensuite le moyen de calculer chaque distribution, facteur de cette décomposition. En général, il s'agit d'une forme paramétrique telle une loi uniforme, ou une loi gaussienne et des paramètres de cette loi, tels la moyenne et l'écart-type pour la gaussienne.

De cette manière, on obtient une expression calculable de la distribution conjointe. On pose alors une question probabiliste, c'est-à-dire la distribution sur une partie des variables en connaissant (éventuellement) la valeur de certaines autres variables. En pratique, poser une question revient à partitionner l'ensemble V des variables en trois : $S \subset V, K \subset V, S \cap$

²⁰Pearl (1988) introduit les *bayesian networks*, ou réseaux bayésiens, et des méthodes d'inférence telles que la *belief propagation*. Charniak (1991) présente une introduction didactique à l'utilisation de tels réseaux. Andersson *et al.* (1996) caractérisent les modèles graphiques comportant à la fois des arcs orientés et non orientés (*chain graphs*). Smyth (1997) et Jensen (2000) replacent les réseaux bayésiens, les modèles de Markov cachés (*hidden Markov models*) et les champs de Markov (*Markov random fields*) au sein de la famille générale des modèles graphiques. Murphy (2001) présente la famille des modèles graphiques et les techniques d'inférence et d'apprentissage que l'on peut utiliser principalement dans les réseaux bayésiens.

²¹Lebeltel *et al.* (2000, 2004) présentent le formalisme de la programmation bayésienne et son application à la robotique.

$K = \emptyset, F = V \setminus (S \cup K)$. S est l'ensemble des variables recherchées, K les variables dont la valeur est connue. Les variables restantes F sont appelées variables libres ou variables intermédiaires. La question est :

$$P(S | K) \quad (2.6)$$

Les règles de Bayes et de marginalisation permettent toujours de répondre à cette question à partir de la décomposition. En appliquant successivement 2.4 et 2.5 à la question 2.6, on obtient l'expression 2.7.

$$\begin{aligned} P(S | K) &= \frac{P(S \ K)}{P(K)} \\ P(S | K) &= \frac{\int_F P(S \ F \ K)}{\int_{F \ S'} P(S' \ F \ K)} \end{aligned} \quad (2.7)$$

Le dénominateur de l'expression 2.7 ne dépend pas des valeurs des variables que l'on recherche et n'est donc qu'un facteur de normalisation. En résumé, on obtient l'expression générale 2.8 de la réponse à la question 2.6.

$$P(S | K) \propto \int_F P(V) \quad (2.8)$$

Distributions et maximum d'entropie

Une étape cruciale de la spécification d'un programme bayésien est le choix des distributions de probabilité. Ces choix sont un apport de connaissances par le programmeur.

Dans certains cas, il est possible de construire ces distributions à partir de formes paramétriques et de statistiques exhaustives provenant d'observations. Ce n'est pas toujours le cas, en particulier pour les distributions *a priori*.²²

Un critère de choix de distribution est le *principe d'indifférence*, connu de Bernoulli et nommé par Keynes (1921). Ce critère s'applique en l'absence d'informations sur des variables discrètes et bornées. Dans ce cas, il consiste à considérer toutes les valeurs d'une variable équiprobables. Il s'agit d'un choix conservatif qui ne privilégie aucune possibilité s'il n'y a pas lieu de le faire.

Une généralisation de ce critère se retrouve dans le *principe de maximum d'entropie* (Jaynes 1968, 1982, 2003). Ce principe consiste à utiliser la distribution de maximum d'entropie compatible avec les observables²³ dont on dispose. L'entropie d'une distribution est définie par Shannon (1948). Elle mesure la quantité d'information communicable à l'aide de réalisations de cette distribution. Le but du principe de maximum d'entropie est de rajouter le moins d'information possible. Ainsi choisie, la distribution peut communiquer le plus d'information possible.

²²Une distribution *a priori* est un facteur marginal, c'est-à-dire non conditionnel, de la décomposition. On l'oppose parfois à la distribution *a posteriori*, résultat de l'inférence.

²³Une observable est une fonction d'une distribution dont on connaît la valeur. Elle fournit donc une contrainte que doit satisfaire la distribution.

Lorsqu'il n'y a pas d'observables, la distribution uniforme est de maximum d'entropie et on obtient le principe d'indifférence. La distribution de probabilité de maximum d'entropie lorsque l'on dispose de la moyenne et de l'écart-type comme observables est la distribution gaussienne. Plus généralement, les solutions du maximum d'entropie s'écrivent comme l'exponentielle d'une combinaison linéaire des observables (Jaynes 2003). Ces distributions ont la propriété de posséder des statistiques exhaustives (Jaynes 2003, théorème de Darmois-Koopman-Pitman).

Inférence

L'inférence dans les modèles bayésiens possède une expression simple (voir équation 2.8) qui cache un calcul complexe : l'inférence bayésienne est un problème NP-difficile (Cooper 1990).

Deux pistes sont explorées : les méthodes d'inférence exactes et les méthodes approchées. Parmi les méthodes exactes, on note le *bucket elimination* (Dechter 1996), les méthodes à base de messages : *message passing* (Pearl 1982) et *belief propagation* (Pearl 1988 ; Lauritzen et Spiegelhalter 1988), et l'algorithme de l'arbre de jonction (Jensen *et al.* 1990 ; Jensen et Jensen 1994).

Lorsqu'il y a peu d'indépendances (c'est-à-dire pour un graphe fortement connecté), les complexités de ces méthodes ont tendance à devenir prohibitives. On a alors recourt à des méthodes approchées comme les méthodes de Monte-Carlo (Neal 1993 ; MacKay 1998) ou les méthodes variationnelles (Jaakkola et Jordan 1999).

Certains ensembles d'hypothèses amènent à des inférences polynomiales. C'est le cas du filtre de Kalman (Kalman 1960), de l'algorithme de Viterbi (Forney 1973) ou même des filtres à particules (Arulampalam *et al.* 2002).

Dans ce document nous ne discutons pas de l'inférence. Nous disposons du moteur d'inférence bayésienne ProBT (<http://www.probayes.com/>) qui implémente ces algorithmes. Du point de vue du travail présenté dans ce document, l'inférence est un problème résolu par ce moteur.

2.3.2 Modèles bayésiens de la perception

De nombreux travaux ont été menés sur des modèles bayésiens de la perception (revues par Kersten (1999) ; Schrater *et al.* (2000) ; Mamassian *et al.* (2002) ; Kersten *et al.* (2004)). On note en particulier l'explication des illusions comme des perceptions optimales²⁴ sous certaines hypothèses en utilisant la notion d'observateur idéal, c'est-à-dire combinant au mieux les informations qui lui sont fournies.²⁵

²⁴Optimal est employé ici dans le sens de plus grande probabilité.

²⁵Weiss *et al.* (2002) expliquent l'illusion de déplacement du losange ainsi que des *plaids*, stimulus à base de combinaison de rayures, à l'aide d'une fusion bayésienne. Plus précisément, un losange avec deux côtés horizontaux incliné vers la droite en translation vers la droite, va être perçu, si l'on en cache les sommets, comme se déplaçant en bas à droite. Cet effet est d'autant plus important que le losange est aplati. Weiss *et al.* (2002) généralisent l'approche d'intersection des contraintes du mouvement sur chacun des côtés en

L'usage de modèles bayésiens nécessite la notion d'*a priori*, c'est-à-dire de ce qui est attendu indépendamment d'observations. Des études se sont intéressées à retrouver et quantifier des *a priori* dans certaines tâches de perceptions.²⁶

Les structures classiques des modèles bayésiens trouvent des applications biologiques. La fusion est utilisée dans des modèles d'intégration multisensorielle (Landy *et al.* 1995). Les applications incluent la perception de la profondeur et de l'inclinaison à partir d'indices de texture ou de disparité²⁷, la perception de la couleur²⁸ ou même le couplage visuo-haptique.²⁹

Enfin, suivant des mécanismes de fusion analogues à l'intégration multisensorielle, des modèles bayésien d'intégration sensorimotrice ont été proposés.³⁰

Tous ces modèles ont permis d'expliquer de nouvelles illusions ou de modéliser avec succès des propriétés de la perception sans avoir recours à des modèles artificiellement complexes. Dans cette optique, la modélisation bayésienne se présente comme un cadre général traitant l'incertitude et l'ambiguïté de l'information visuelle, par exemple. De plus, la modélisation bayésienne, issue de la théorie de l'information, fournit la notion d'observateur idéal auquel comparer les performances des observateurs humains.

Au delà de ces modèles de la perception, des travaux s'intéressent à leur implémentation neuronale s'appuyant sur le caractère stochastique des réponses des neurones. On peut noter en particulier les travaux sur le codage par populations³¹, sur les mécanismes de

en faisant une fusion bayésienne ce qui permet de reproduire cette erreur de perception. Geisler et Kersten (2002) utilisent un observateur idéal, au sens statistique, pour la perception de l'inclinaison d'ellipses.

²⁶Mamassian et Landy (1998) étudient des dessins pouvant correspondre à des portions d'ellipsoïdes et de parabolöide hyperbolique (forme de selle) et montrent l'effet de leur orientation sur leur catégorisation. Mamassian et Goutcher (2001) montrent que l'*a priori* sur la direction de l'illumination est en haut à gauche (écarté de 26° par rapport à la verticale). Pour cela, ils utilisent des stimuli représentant une surface plane avec des bandes ondulées en relief. À l'aide de contrastes perpendiculaires à ces bandes, ils modifient ces stimuli pour comparer cet *a priori* avec celui sur la normale d'une surface avec un modèle bayésien de fusion (Mamassian et Landy 2001).

²⁷Jacobs (1999) utilise une fusion naïve entre les indices de texture et de mouvement pour la perception de la profondeur d'un cylindre texturé en rotation. Hillis *et al.* (2004) appliquent le maximum de vraisemblance (absence d'*a priori*) pour la perception de l'inclinaison par la combinaison d'indices de texture (grille) et de disparité en vision binoculaire.

²⁸Maloney (2002) utilise une fusion pour modéliser l'interaction entre l'illumination du fond et les reflets spéculaires pour la perception de la couleur.

²⁹Ernst et Banks (2002) appliquent le maximum de vraisemblance pour répliquer la perception de la hauteur d'une barre à partir d'informations visuelles et haptiques simulées.

³⁰Ghahramani *et al.* (1997) présentent un modèle pour la précision dans une tâche de pointage de stimuli visuels, auditifs ou visuo-auditif. Ils présentent aussi l'application d'un filtre de Kalman à une tâche de suivi de ligne dans des champs de forces avec retour visuel simulé. Körding et Wolpert (2004) reprennent cette dernière expérience en ajoutant de l'incertitude sur la position d'origine (*via* une distribution bimodale) et sur le retour visuel. Ils appliquent la règle de Bayes sur cet *a priori* avec cette vraisemblance pour reproduire la distribution d'arrivée des participants.

³¹Zemel *et al.* (1997) présentent une interprétation probabiliste de l'activité d'une population de neurones en présentant les algorithmes d'encodage et de décodage de distributions de probabilités. Deneve *et al.* (2001) montrent l'inférence avec de telles représentations sur l'exemple de la fusion.

décision³² ou même sur les neurotransmetteurs.³³ On trouve aussi des travaux sur l'intégration multisensorielle au niveau des neurones.³⁴ Enfin, il existe des modèles plus spécifiques d'aires du cortex visuel³⁵ ou sur des populations de l'hippocampe du rat.³⁶

³²Gold et Shadlen (2002) appliquent le Banburismus, théorie de la décision par accumulations de poids d'évidence, pour modéliser l'activité de neurones de l'aire intrapariétale latérale (LIP) du singe.

³³Yu et Dayan (2003) présentent un modèle de Markov caché hiérarchique de l'apprentissage dans lequel ils distinguent deux sortes d'incertitude, attendue et inattendue, qu'ils identifient respectivement au rôle de l'acétylcholine et la norépinéphrine.

³⁴Anastasio *et al.* (2000) ; Colonus et Diederich (2002) utilisent le maximum de vraisemblance sur les neurones du colliculus supérieur pour modéliser l'augmentation des réponses des neurones en perception multisensorielle (*multisensory enhancement*).

³⁵Koechlin *et al.* (1999) présentent un modèle bayésien de l'activité de neurones de l'aire médio-temporale (MT) cohérent avec les propriétés observées dans le cortex. Lee et Mumford (2003) utilisent des filtres à particules pour l'inférence dans un modèle bayésien du cortex visuel.

³⁶Zhang *et al.* (1998) proposent un modèle de décodage des cellules de lieu de l'hippocampe du rat.

Chapitre 3

Exemple introductif

Dans ce chapitre, nous décrivons à la fois, la base des expériences auxquelles nous nous référons pour concevoir et évaluer notre modèle, et notre cadre théorique de travail. Dans un premier temps vient la description d'une expérience de psychophysique simplifiée. Nous illustrons ensuite l'outil de programmation bayésienne à travers la modélisation de cette expérience et nous comparons les résultats de ce modèle à ceux des sujets.

3.1 Hypothèse et protocole

Une expérience est conçue pour tester une hypothèse. Une fois cette hypothèse définie, on précise le protocole expérimental permettant de la tester. Les résultats sont analysés en regard de cette hypothèse et on peut, alors, proposer un modèle explicatif permettant de les reproduire. L'expérience prise comme exemple ici, est celle décrite par van Boxtel *et al.* (2003).

3.1.1 Hypothèse testée

L'ambiguïté concave/convexe décrite au chapitre 1 (voir figure 1.1) peut se produire avec un plan rotatif observé au travers d'une petite ouverture. Le sens de rotation du plan peut alors être difficile à déterminer. C'est la version la plus simple de cette illusion. L'hypothèse testée par cette première expérience sera la présence de cette ambiguïté. Cette expérience sert aussi de condition de référence pour les expériences présentées au chapitre 5.

3.1.2 Protocole expérimental

Stimulus Le stimulus est une animation figurant un plan incliné en rotation. Chaque image est la projection du plan sur un écran en blanc sur fond noir. Pour cette première expérience, l'inclinaison initiale σ du plan est fixe et égale 45° et l'axe de rotation est l'axe vertical. L'orientation τ du plan est tirée uniformément entre 0 et 360° à chaque nouvelle

présentation. La figure 3.1 illustre la position relative du plan et de l'observateur ainsi que les angles d'inclinaison et d'orientation.

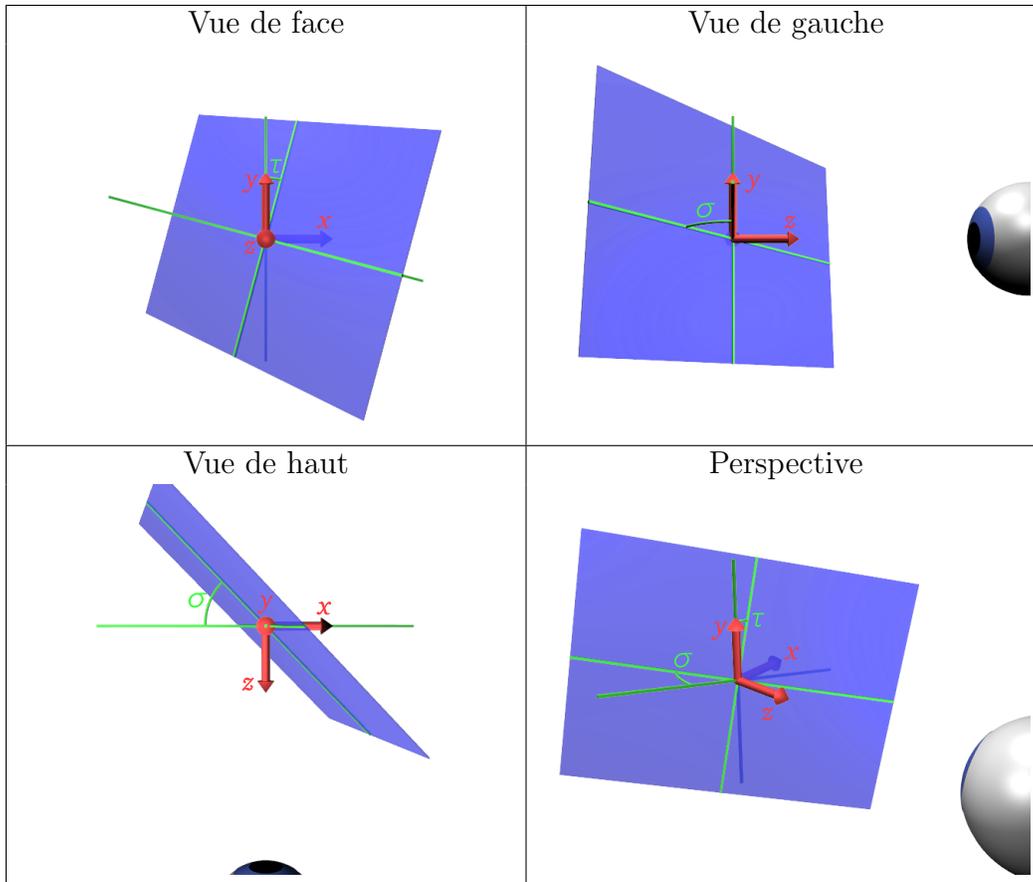


FIG. 3.1 – Position relative du plan et de l'observateur. Le plan objet est figuré en bleu, l'observateur par l'œil en bas à droite situé en $(0, 0, d)$ avec $d > 0$. L'inclinaison du plan, notée σ est l'angle entre le plan objet et le plan fronto-parallèle (xOy). L'orientation τ du plan est l'angle entre l'axe vertical y et l'intersection entre le plan objet et le plan fronto-parallèle. Le plan objet se distingue du plan fronto-parallèle par une rotation de son inclinaison σ autour de l'axe vertical y puis une rotation de son orientation τ autour de l'axe sagittal z .

Le plan n'est figuré qu'à travers une texture aléatoire de points blancs sur fond noir. Les participants ne voient donc que ces points blancs et il n'y a pas d'ombre ou d'éclairage. Les points sont tirés de manière à être répartis uniformément sur l'écran au début de la présentation du stimulus. De cette manière, l'inclinaison ne peut pas être déduite du gradient de densité des points. Les points sont ensuite déplacés à l'écran suivant le mouvement du plan. Un exemple de stimulus affiché à l'écran est montré figure 3.2.

L'observation est monoculaire, c'est-à-dire que l'œil non directeur du sujet est masqué. Pour cette première expérience, il est demandé aux participants de ne pas bouger. On leur demande aussi de focaliser leur regard sur un point de fixation projeté à l'écran. On

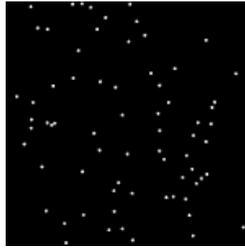


FIG. 3.2 – Exemple d’affichage présenté aux participants.

dispose d’un moyen de mesurer le mouvement de la tête et de l’œil directeur. Dans cette expérience, on s’en sert pour vérifier que les mouvements de l’observateur restent petits. En pratique, il s’agit d’un *tracker* par infrarouge. Un déplacement cumulé de la tête de plus de 3 cm pendant la durée de présentation du stimulus est considéré comme un déplacement et annule la mesure.

La présentation du stimulus a une durée variable comprise entre 150 ms et 850 ms. Durant cette durée, les variations des angles d’inclinaison et d’orientation du plan ne dépassent pas la dizaine de degrés. Enfin, le plan n’est affiché qu’à l’intérieur d’un disque de 5.25 cm de rayon. Ce qui, pour une distance d’observation typique d’environ 60 cm, représente l’intersection avec un cône de demi-angle au sommet d’environ 5° . L’expérience est conduite dans le noir, les bords de l’écran étant cachés pour ne pas donner d’indice visuel extérieur.

Tâche La tâche consiste à retrouver la configuration moyenne (inclinaison et orientation) du plan. Pour cela, après la disparition du stimulus, on présente aux participants une sonde qu’ils doivent aligner avec le plan moyen perçu. La sonde est une grille carrée affichée à l’écran à la place du plan. Les participants utilisent une manette pour modifier et valider l’inclinaison et l’orientation de la sonde. Dans cette première expérience, l’estimation de l’inclinaison n’est pas prise en compte et seule l’orientation nous intéresse. Au final, la question à laquelle cette expérience essaie de répondre est la détermination de l’orientation du plan à l’aide du flux optique.

3.2 Résultats et hypothèses

3.2.1 Résultats expérimentaux

Pour cette expérience, la répartition de l’erreur est la mesure retenue. L’erreur est définie comme étant la différence en degré entre l’orientation choisie par le participant et l’orientation du plan simulé. Il s’agit donc d’une variable d’angle, de topologie circulaire. Pour cette expérience, elle est discrétisée en intervalles de 15° de largeur. La répartition d’erreur est l’histogramme figurant le nombre de réponse pour chacun de ces intervalles.

La figure 3.3 présente la répartition d’erreur par van Boxtel *et al.* (2003).

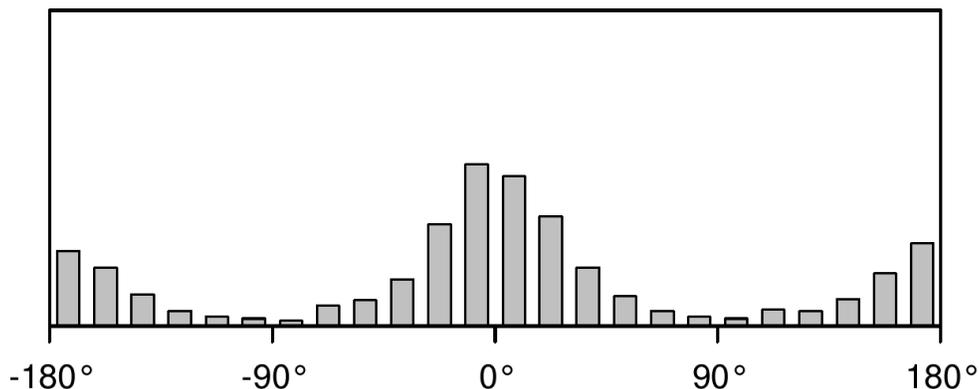


FIG. 3.3 – Répartition expérimentale de l’erreur d’estimation de l’orientation.

Cet histogramme présente deux pics (les bornes -180° et 180° sont confondues). L’un est centré vers 0° alors que l’autre est centré sur 180° . Le nombre d’occurrences rassemblé par le pic autour de 0° est plus important qu’autour de 180° .

Ce graphe montre deux percepts majoritaires. Le premier, celui dont les réponses présentent une erreur proche de 0° , correspond à l’objet simulé. Dans ce cas, les participants ont réussi à identifier la position du plan qui leur était présenté. Le second, représenté par une erreur proche de 180° , correspond à un plan d’orientation symétrique par rapport au plan fronto-parallèle.¹ On qualifie ce second percept d’*inverse*.

3.2.2 Hypothèses explicatives

L’objectif est maintenant de formuler des hypothèses pour construire un modèle reproduisant au mieux les réponses des sujets.

Les hypothèses principales sont issues de la littérature et ont été présentées et discutées en section 2.2. La première hypothèse est la rigidité. Celle-ci stipule que le flux optique perçu est celui d’un objet rigide plutôt que d’une collection de points de mouvements indépendants (voir section 2.2.2). Dans notre contexte d’observation d’un plan, il va donc s’agir d’un plan rigide.

La seconde hypothèse est la stationnarité, qui consiste à considérer que le mouvement du plan dans le référentiel allocentré est plus probablement petit (voir section 2.2.2). Dans cette expérience, le sujet est immobile donc les mouvements du plan, dans les deux référentiels allocentré et égocentré, sont les mêmes. La stationnarité revient à faire l’hypothèse d’un mouvement relatif du plan plus probablement petit.

Enfin, nous supposons que le mouvement et la position du plan sont indépendants. Cela signifie que nous considérons que la vitesse de rotation du plan n’est pas contrainte par

¹Deux plans objets dont l’orientation diffère de 180° sont toujours symétriques par rapport au plan fronto-parallèle (dans lequel est mesuré l’orientation). On sait par ailleurs que le sens de rotation perçu pour les percepts autour de 180° est inversé par rapport au sens de rotation simulé. Il s’agit, là encore, d’une symétrie par rapport au plan fronto-parallèle.

son orientation et réciproquement.

Ainsi, les hypothèses que nous allons implanter dans notre modèle peuvent se résumer ainsi :

1. rigidité : le flux perçu est celui d'un objet rigide ;
2. stationnarité : le mouvement du plan est petit ;
3. mouvement et position indépendants l'un de l'autre.

3.3 Modèle

Nous construisons un modèle pour cette expérience. Cela implique en premier lieu des hypothèses de contexte. En particulier, nous faisons l'hypothèse que nous cherchons à percevoir un plan en mouvement par son flux optique. Ce modèle est construit en suivant le formalisme de la programmation bayésienne abordé en 2.3.1.

3.3.1 Variables

La première étape de spécification d'un programme bayésien est de décider des variables auxquelles il faut s'intéresser. Ces variables sont généralement définies par les hypothèses de contexte. Pour notre modèle de perception d'un plan par le flux optique, on peut distinguer trois groupes de variables par rapport à leur fonctions respectives dans le modèle : les variables d'observation, les variables recherchées et les variables intermédiaires.

Les variables d'observation sont celles dont la valeur est connue au moment de l'interrogation du modèle. Pour cette expérience, il s'agit du flux optique. Le flux optique est le déplacement de l'image. Nous le représentons par ses coefficients de Taylor d'ordre 0, 1 et 2 (voir 2.2.1). Dans le cas présent², seuls trois degrés de liberté subsistent (voir équation 3.8). Les variables d'observation sont donc Φ_0^1 , Φ_1^1 et Φ^2 , toutes trois de domaine $\mathcal{D}_\Phi = \mathbb{R}$.

Nous cherchons à retrouver l'orientation d'un plan. La variable cherchée est donc l'orientation T du plan. Cette orientation est définie comme l'angle entre l'axe vertical et l'intersection du plan objet avec le plan fronto-parallèle (voir figure 3.4). Le domaine de la variable T est $\mathcal{D}_T =]-\pi; \pi]$. Par rapport aux gradients de profondeur χ et ν du plan, on a $\tau = \arctan \frac{\nu}{\chi}$ et $\chi = \cos \tau$ et $\nu = \sin \tau$ (en supposant un plan incliné à 45°).

Le mouvement du plan est à prendre en compte puisque notre modèle s'intéresse à la perception du plan par son mouvement. Nous introduisons donc une variable intermédiaire représentant ce mouvement. Dans cette première expérience, le mouvement du plan est restreint à une rotation autour de l'axe vertical. Notre modèle représente le mouvement du plan par sa vitesse de rotation Ω de domaine $\mathcal{D}_\Omega = \mathbb{R}$.

En résumé, les variables sur lesquelles porte notre modèle sont :

- le flux optique : Φ_0^1 , Φ_1^1 et Φ^2 ;
- l'orientation du plan : T ;
- sa vitesse de rotation : Ω .

²Absence de translation et axe de rotation vertical.

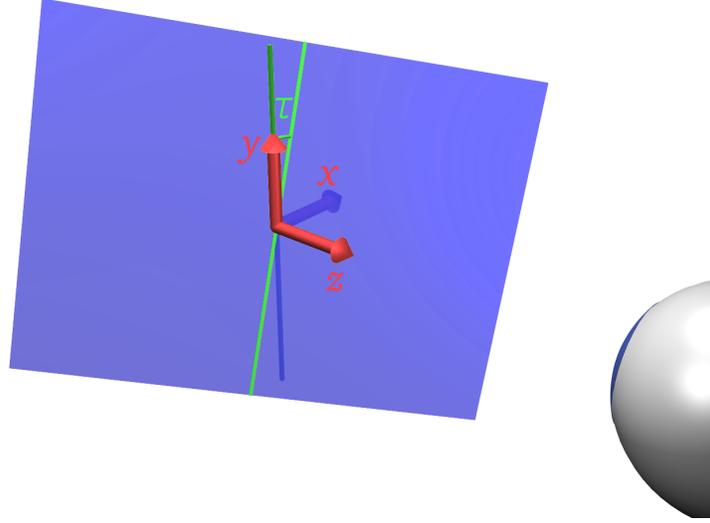


FIG. 3.4 – Définition de la variable T d'orientation du plan. Il s'agit de l'angle entre l'axe vertical et l'intersection du plan objet avec le plan fronto-parallèle.

3.3.2 Décomposition

Les hypothèses de notre modèle (section 3.2.2) se retrouvent en partie dans la décomposition de la distribution de probabilité conjointe sous la forme d'hypothèses d'indépendance conditionnelle.

Une application de la règle de Bayes sur la conjointe sur les variables du modèle mène à l'expression 3.1.

$$P(T \Omega \Phi_0^1 \Phi_1^1 \Phi^2) = P(T \Omega)P(\Phi_0^1 \Phi_1^1 \Phi^2 | T \Omega) \quad (3.1)$$

L'hypothèse de rigidité 1 stipule que le flux optique observé est celui d'un objet rigide. Dans notre contexte d'observation d'un plan, il s'agit donc du flux optique d'un plan rigide. Or, les composantes du flux optique d'un plan sont indépendantes les unes des autres conditionnellement à la position et au mouvement de ce plan. En effet, pour un plan rigide, la connaissance de la position et le mouvement de ce plan suffit à définir chacune des composantes du flux optique indépendamment les unes des autres. Dit autrement, pour un plan rigide, la connaissance de la valeur d'une des composantes du flux optique n'ajoute pas d'information à la connaissance sur une autre composante que l'on peut déduire de la position et du mouvement. Comme il s'agit de distributions de probabilité, nous faisons donc l'hypothèse simplificatrice que, dans tous les cas, les différences entre les flux prévus par la position et le mouvement et les flux observés sont décorrélés. De surcroît, la composante Φ^2 est indépendante de la position T connaissant le mouvement Ω (voir équation 3.8). On écrit donc l'expression 3.2 :

$$P(\Phi_0^1 \Phi_1^1 \Phi^2 | T \Omega) = P(\Phi_0^1 | T \Omega)P(\Phi_1^1 | T \Omega)P(\Phi^2 | \Omega) \quad (3.2)$$

L'hypothèse 3 est que les mouvement et position du plan sont indépendants. En termes probabilistes, cette hypothèse s'écrit :

$$P(T \Omega) = P(T)P(\Omega) \quad (3.3)$$

En substituant les expressions 3.2 et 3.3 dans 3.1 on obtient la décomposition 3.4 de la distribution conjointe.

$$P(T \Omega \Phi_0^1 \Phi_1^1 \Phi^2) = P(T)P(\Omega)P(\Phi_0^1 | T \Omega)P(\Phi_1^1 | T \Omega)P(\Phi^2 | \Omega) \quad (3.4)$$

3.3.3 Distributions

Les hypothèses de notre modèle s'expriment aussi au travers des distributions associées à chacun des facteurs de la distribution conjointe. Une manière de choisir les distributions est d'utiliser le critère du maximum d'entropie (voir section 2.3) en utilisant les hypothèses comme des observables.

L'hypothèse de stationnarité 2 stipule que le mouvement est plus probablement petit. On en déduit que le mouvement le plus probable est d'amplitude nulle. Pour un écart-type donné, la distribution de maximum d'entropie est gaussienne. En notant $\mathcal{N}_{\mu,\sigma}(V)$ la loi normale sur la variable V de moyenne μ et d'écart-type σ , on peut écrire :

$$P(\Omega) = \mathcal{N}_{0,\sigma_\Omega}(\Omega) \quad (3.5)$$

En l'absence d'hypothèse sur l'orientation du plan, la distribution de maximum d'entropie sur la variable T de domaine borné est uniforme. La distribution de maximum d'entropie ajoute le moins d'information possible pour éviter de biaiser l'inférence. Dans le cas présent, cela signifie que toutes les orientations sont équiprobables. Avec $\mathcal{U}(V)$ la loi uniforme sur une variable V , on écrit :

$$P(T) = \mathcal{U}(T) \quad (3.6)$$

Enfin, l'hypothèse de rigidité 1 dans le contexte de notre modèle est que le flux perçu est proche de celui d'un plan. Nous faisons l'hypothèse que la moyenne de la distribution sur le flux sera le flux théorique du plan et que nous pouvons obtenir un écart-type. Ces contraintes se traduisent en distributions gaussiennes pour chacune des composantes du flux optique. On obtient :

$$\forall (\tau, \omega) \in \mathcal{D}_T \times \mathcal{D}_\Omega, \begin{cases} P(\Phi_0^1 | \tau \omega) &= \mathcal{N}_{f_0^1(\tau,\omega),\sigma_{\Phi_0^1}}(\Phi_0^1) \\ P(\Phi_1^1 | \tau \omega) &= \mathcal{N}_{f_1^1(\tau,\omega),\sigma_{\Phi_1^1}}(\Phi_1^1) \\ P(\Phi^2 | \omega) &= \mathcal{N}_{f^2(\omega),\sigma_{\Phi^2}}(\Phi^2) \end{cases} \quad (3.7)$$

avec f_0^1 , f_1^1 et f^2 les expressions du flux optiques dans le cas présent. Elles peuvent s'obtenir à partir des équations 2.2 du flux optique d'un plan rigide en se souvenant que pour cet expérience on considère $t_x = t_y = t_z = 0$ et $\omega_x = \omega_z = 0$. De plus, on a $\chi = \cos \tau$ et

$v = \sin \tau$ (hypothèse d'un plan incliné à 45°). On peut donc reprendre les équations du flux optique pour obtenir :

$$\begin{aligned}\vec{\phi}^0 &= \begin{pmatrix} t_x \\ t_y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \vec{\phi}^1 &= \begin{pmatrix} t_z + \chi(t_x + \omega_y) & -\omega_z + v(t_x + \omega_y) \\ \omega_z + \chi(t_y - \omega_x) & t_z + v(t_y - \omega_x) \end{pmatrix} \\ &= \begin{pmatrix} 0 + \cos \tau \cdot (0 + \omega_y) & -0 + \sin \tau \cdot (0 + \omega_y) \\ 0 + \cos \tau \cdot (0 - 0) & 0 + \sin \tau \cdot (0 - 0) \end{pmatrix} \\ \vec{\phi}^2 &= \begin{pmatrix} \chi t_z - \omega_y \\ vt_z + \omega_x \end{pmatrix} = \begin{pmatrix} \cos \tau \cdot 0 - \omega_y \\ \sin \tau \cdot 0 + 0 \end{pmatrix}\end{aligned}$$

On justifie ainsi notre choix d'éliminer certaines variables du flux optique qui seront constantes (et accessoirement nulles) et on obtient les expressions suivantes de f_0^1 , f_1^1 et f^2 :

$$\begin{cases} f_0^1(\tau, \omega) &= \omega \cos(\tau) \\ f_1^1(\tau, \omega) &= \omega \sin(\tau) \\ f^2(\omega) &= -\omega \end{cases} \quad (3.8)$$

3.3.4 Question

La distribution conjointe du modèle est entièrement spécifiée. Le modèle peut alors répondre à n'importe quelle question probabiliste concernant ses variables. Dans l'expérience, on demande aux participants de donner une estimation de l'orientation après avoir observé le flux optique. En termes probabilistes, cette question est formulée par l'expression 3.9.

$$P(T \mid \Phi_0^1 \Phi_1^1 \Phi^2) \quad (3.9)$$

Les règles d'inférence conduisent à l'expression de cette conditionnelle indiqué par l'équation 3.10 (voir 2.3).

$$P(T \mid \Phi_0^1 \Phi_1^1 \Phi^2) = \frac{P(T) \sum_{\omega \in \mathcal{D}_\Omega} P(\omega) P(\Phi_0^1 \mid T \omega) P(\Phi_1^1 \mid T \omega) P(\Phi^2 \mid \omega)}{\sum_{\tau', \omega \in \mathcal{D}_T \times \mathcal{D}_\Omega} P(\tau') P(\omega) P(\Phi_0^1 \mid \tau' \omega) P(\Phi_1^1 \mid \tau' \omega) P(\Phi^2 \mid \omega)} \quad (3.10)$$

Pour une réalisation donnée de l'expérience, les valeurs du flux optiques sont connues. L'expression 3.10 permet de calculer la distribution de réponse sur l'orientation du plan :

$$\begin{aligned}& P(\tau \mid \phi_0^1 \phi_1^1 \phi^2) \\ &= \frac{P(\tau) \sum_{\omega} P(\omega) P(\phi_0^1 \mid \tau \omega) P(\phi_1^1 \mid \tau \omega) P(\phi^2 \mid \omega)}{\sum_{\tau', \omega} P(\tau') P(\omega) P(\phi_0^1 \mid \tau' \omega) P(\phi_1^1 \mid \tau' \omega) P(\phi^2 \mid \omega)} \\ &= \frac{\sum_{\omega} \mathcal{N}_{0, \sigma_\Omega}(\omega) \mathcal{N}_{f_0^1(\tau, \omega), \sigma_{\Phi_0^1}}(\phi_0^1) \mathcal{N}_{f_1^1(\tau, \omega), \sigma_{\Phi_1^1}}(\phi_1^1) \mathcal{N}_{f^2(\omega), \sigma_{\Phi^2}}(\phi^2)}{\sum_{\tau', \omega} \mathcal{N}_{0, \sigma_\Omega}(\omega) \mathcal{N}_{f_0^1(\tau', \omega), \sigma_{\Phi_0^1}}(\phi_0^1) \mathcal{N}_{f_1^1(\tau', \omega), \sigma_{\Phi_1^1}}(\phi_1^1) \mathcal{N}_{f^2(\omega), \sigma_{\Phi^2}}(\phi^2)} \\ &\propto \sum_{\omega} \mathcal{N}_{0, \sigma_\Omega}(\omega) \mathcal{N}_{f_0^1(\tau, \omega), \sigma_{\Phi_0^1}}(\phi_0^1) \mathcal{N}_{f_1^1(\tau, \omega), \sigma_{\Phi_1^1}}(\phi_1^1) \mathcal{N}_{f^2(\omega), \sigma_{\Phi^2}}(\phi^2) \quad (3.11)\end{aligned}$$

Paramètre	Valeur
σ_{Ω}	$0,2 \text{ rad.s}^{-1}$
$\sigma_{\Phi_0^1}$	0.25 m.s^{-1}
$\sigma_{\Phi_1^1}$	0.25 s^{-1}
σ_{Φ^2}	$1 \text{ m}^{-1}.\text{s}^{-1}$

TAB. 3.1 – Valeurs des paramètres du modèle.

L'écriture mathématique de ce modèle introductif est résumée par la figure 3.5.

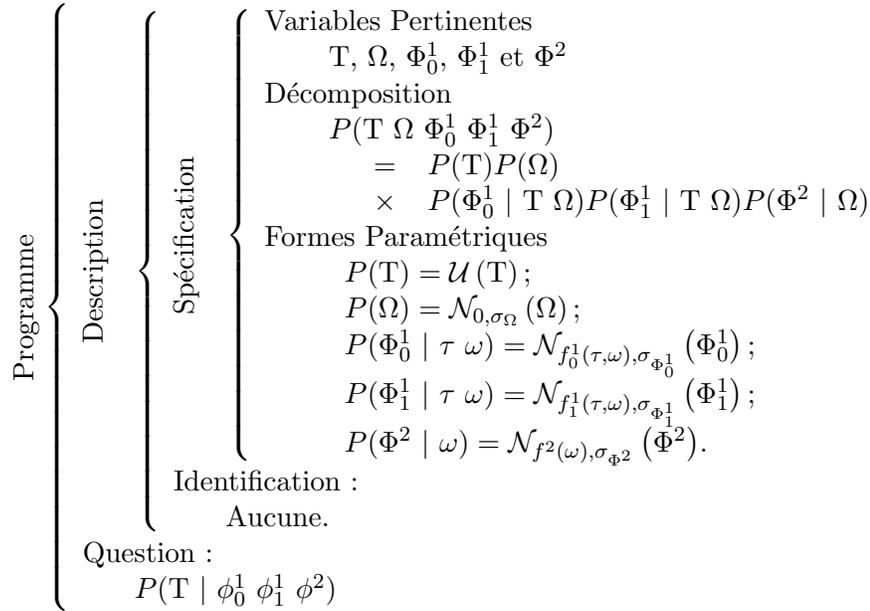


FIG. 3.5 – Résumé du modèle bayésien.

3.4 Résultats

À l'aide de l'expression 3.11, nous pouvons calculer les résultats du modèle et les comparer avec ceux de l'expérience présentés en 3.2.

Les calculs sont réalisés à l'aide du moteur d'inférence ProBT. Les valeurs des paramètres utilisés pour les résultats ci-dessous sont présentés table 3.1.³

3.4.1 Comparaison

De la même manière que pour les résultats expérimentaux, on fait la moyenne de l'écart entre le résultat de notre modèle et le plan simulé. On obtient l'histogramme présenté

³Il est à noter que les résultats sont qualitativement robustes à ces paramètres, voir section 4.5.3.

figure 3.6. Le modèle partage la même bimodalité que les résultats expérimentaux. Les deux pics de probabilité sont situés pour une erreur d'orientation de 0° et 180° . On remarque aussi que ces deux pics présentent une différence de probabilité. Celui centré autour de 0° est plus élevé que celui centré sur 180° , tout comme dans les résultats expérimentaux.

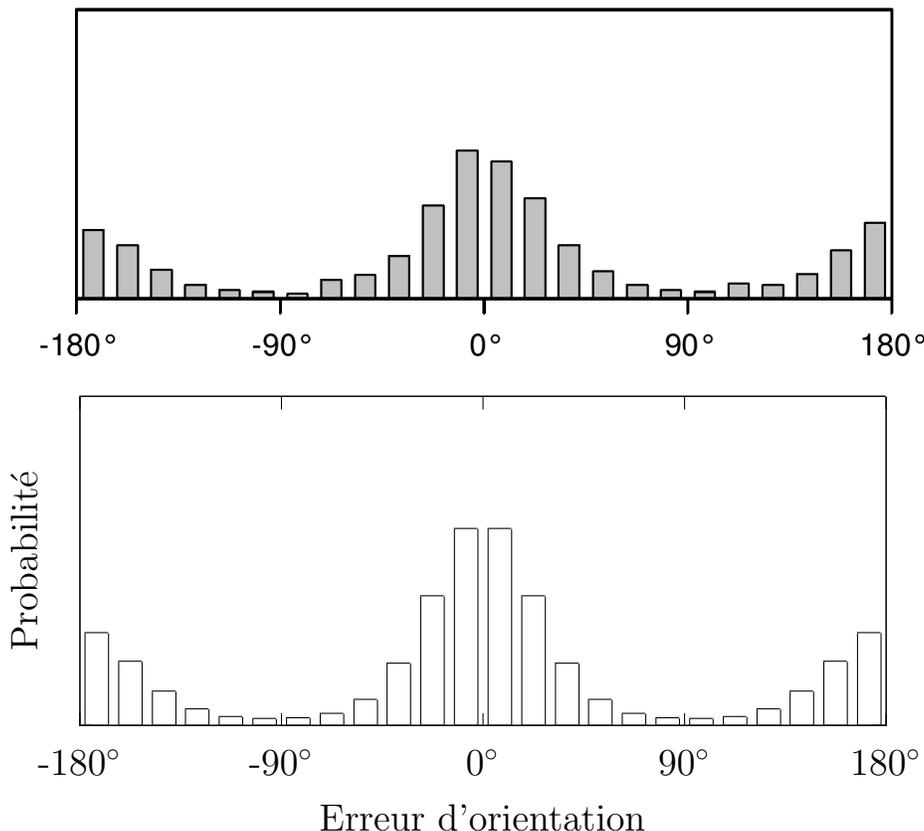


FIG. 3.6 – Comparaison des résultats de l'expérience et du modèle. En haut : répartition de l'erreur des participants (reprise de la figure 3.3). En bas : distribution de probabilité de l'erreur d'estimation de l'orientation par notre modèle.

Grâce au formalisme bayésien structuré autour de l'expression d'hypothèses, nous pouvons explorer l'influence de celles-ci pour expliquer ces résultats.

La rigidité s'énonce dans notre modèle par l'hypothèse que le flux optique observé est plus probablement proche de celui causé par un plan d'une orientation donnée. Or, jusqu'à l'ordre 1, le flux optique est symétrique par rapport au plan fronto-parallèle (voir expression 3.8). Cette symétrie s'exprime par un changement du sens de rotation ($\omega \mapsto -\omega$) et par une orientation du plan opposée ($\tau \mapsto \tau + \pi$). De plus, la distribution sur la vitesse est paire et celle sur l'orientation uniforme. Par conséquent, le seul facteur de la décomposition qui n'est pas invariant par cette symétrie est celui impliquant le flux optique d'ordre 2.

En négligeant ce terme, on aurait une distribution sur l'orientation invariante par translation de 180° ; un pic en 0° entraînant un en 180° . Les pics n'ont pas le même poids,

montrant que le terme impliquant le flux d'ordre deux n'est pas négligeable. Cependant nous avons choisi une valeur de σ_{Φ^2} ⁴ suffisamment grande pour que les deux pics soient comparables (voir table 3.1).

Une distribution uniforme⁵ sur le flux d'ordre deux donnerait deux pics de probabilité aux mêmes valeurs. Comme les résultats de l'expérience montrent une différence dans le nombre d'occurrences d'une perception inverse, nous avons choisi une valeur de cet écart-type en accord avec cette perception plus rare du plan inversé.

3.4.2 Validité des comparaisons

Jusqu'à présent, nous avons comparé nos histogrammes des résultats du modèle avec ceux de l'expérience. Or, notre modèle bayésien calcule une distribution de probabilité alors que les résultats de l'expérience sont une répartition quantifiée des réponses des participants.

Une différence entre notre modèle et l'expérience est le domaine des réponses. À chaque observation, les participants ne donnent qu'une réponse alors qu'*a priori* notre modèle renvoie une distribution de probabilité sur l'ensemble des réponses possibles. Pour ne renvoyer qu'une seule configuration, il est nécessaire de prendre une *décision*, c'est-à-dire choisir une seule valeur à partir de la distribution de probabilité.

Il y a plusieurs choix en la matière. Un choix que l'on trouve souvent dans la littérature est l'utilisation de l'opérateur *argmax*. Il s'agit d'un opérateur qui choisit la réponse de plus grande probabilité. Ce processus est déterministe car l'inférence et l'opérateur *argmax* le sont tous les deux. Cela signifie qu'à un stimulus donné, notre modèle répondrait toujours la même configuration. Il s'avère que ce n'est pas le cas d'observateurs humains. Cette méthode de décision n'est donc pas adaptée à notre modèle et est souvent plus utile lorsque les probabilités sont utilisées en tant que fonction d'évaluation.

Une autre méthode naturelle de décision à partir d'une distribution de probabilité est simplement le tirage selon cette distribution. Il s'agit d'un opérateur non déterministe. Si l'on procède à un ensemble de tirage, l'histogramme des occurrences des réponses choisies grâce à cette méthode va tendre vers la distribution de probabilité, d'après la loi des grands nombres. La distribution de probabilité est donc comparable à un histogramme des réponses de notre modèle, qui est comparable à l'histogramme des réponses des participants.

En allant plus loin, on pourrait supposer qu'un humain utilise des mécanismes analogues à ceux de l'inférence bayésienne de manière à raisonner avec les informations partielles dont il dispose. Dans ce cas, il pourrait calculer une distribution de probabilité et ensuite prendre une décision. L'histogramme des réponses de l'expérience ne serait qu'une mesure de cette distribution et la comparaison de cet histogramme avec notre distribution de probabilité n'en serait que plus justifié.

⁴L'écart-type de la distribution gaussienne sur la variable Φ^2 connaissant le mouvement.

⁵Qui correspond à la limite d'une distribution gaussienne avec un écart-type tendant vers l'infini.

3.5 Résumé

Dans ce chapitre nous avons donc présenté la modélisation d'une expérience de psychophysique au travers d'un exemple simplifié.

Nous avons présenté l'expérience en détaillant le principe, le protocole expérimental et les résultats. Dans la suite, cette expérience sert de condition de référence pour tester des variations dans les conditions expérimentales. Les expériences suivantes gardent le même schéma en ajoutant principalement des degrés de liberté, que ce soit pour le plan ou pour l'observateur.

D'un autre côté, cet exemple illustre notre cadre de modélisation. En particulier, nous avons pu mettre l'accent sur l'importance de la décomposition du point de vue de l'expression des hypothèses de notre modèle. Enfin nous avons montré, lors de la comparaison des résultats, comment la programmation bayésienne se prête à l'interprétation des résultats à la lumière des hypothèses du modèle.

Chapitre 4

Modèle complet

L'exemple précédent illustre l'application de la programmation bayésienne à la modélisation d'une expérience de psychophysique. Nous proposons, dans ce chapitre, un modèle complet de la perception d'un plan par le flux optique. Les différences avec le modèle présenté au dessus se situent principalement dans l'ensemble des degrés de liberté que l'on va envisager pour le plan ou l'observateur.

4.1 Principe

Ce modèle s'appuie sur les hypothèses exprimées dans la littérature et présentées au chapitre 2.2. Un des objectifs de ce chapitre est de continuer à montrer comment ces hypothèses classiques peuvent se formuler dans un modèle avec plus de détails qu'au chapitre précédent. Comme précédemment, nous séparons les hypothèses en deux catégories : les hypothèses de contexte, qui définissent l'ensemble des expériences traitées par le modèle, et les hypothèses propres au modèle, qui en définissent le fonctionnement au sein de notre contexte.

Nous proposons un modèle de perception d'un plan par un observateur *via* le flux optique. Le plan et l'observateur peuvent tous deux être en mouvement. Enfin la taille du champ de vision présenté à l'observateur peut varier. Ces hypothèses définissent le contexte dans lequel nous plaçons notre modèle.

Comme pour l'exemple précédent, la première hypothèse propre à notre modèle est la rigidité (section 2.2.2). Cette hypothèse porte sur la structure de l'objet perçu et stipule que le stimulus correspond à un objet rigide plutôt qu'à une collection d'objets ponctuels. Encore une fois, dans notre contexte, il s'agit donc d'un plan rigide.

La seconde hypothèse de notre modèle est la stationnarité (section 2.2.2). Cette hypothèse porte sur le mouvement de l'objet plutôt que sur sa structure. Elle suppose que l'objet bouge peu dans le référentiel du monde.

Une troisième hypothèse, déjà présente dans le modèle précédent, concerne la relation entre les mouvements, que ce soit celui de l'objet ou de l'observateur, et la position du plan. Nous faisons l'hypothèse que cette position et ces mouvements sont indépendants.

Enfin, pour le modèle complet nous faisons une dernière hypothèse, en accord avec les résultats expérimentaux, qui est que la précision du flux optique perçu dépend de la taille du stimulus présenté à l'observateur.

À part la dernière, ces hypothèses sont les mêmes que pour le modèle introductif du chapitre précédent. Leurs expressions seront donc très proches mais adaptées à la plus grande complexité du modèle.

En résumé, les hypothèses de notre modèle sont :

1. rigidité : le flux perçu est celui d'un objet rigide ;
2. stationnarité : le mouvement de l'objet est petit ;
3. mouvement et position indépendants l'un de l'autre ;
4. précision du flux dépendante de la taille du stimulus.

4.2 Variables

Notre modèle a vocation à représenter un éventail d'expériences plus large que l'exemple précédent. Dans cette section nous présentons les variables sur lesquelles porte notre modèle qui se déduisent des libertés que l'on désire explorer. Nous reprenons les mêmes objets que dans le modèle précédent avec des domaines différents.

4.2.1 Positionnement du plan

Un des éléments fondamentaux des expériences de perception considérées est l'objet qui sert de stimulus. Dans l'expérience précédente, nous ne considérons comme pertinente que son orientation ; l'impact de l'inclinaison sur la perception du flux optique étant ignoré. D'une manière générale, deux paramètres suffisent pour définir un plan passant par un point (en l'occurrence le point de fixation).¹ Il existe plusieurs façons de les choisir. Dans l'exemple du chapitre 3, le plan est représenté par son angle d'orientation τ . Une paramétrisation classique consiste en effet à utiliser l'orientation T et l'inclinaison Σ du plan (voir figure 4.1). Ces deux variables sont des angles et, d'un point de vue calculatoire, nécessitent l'utilisation de fonctions trigonométriques pour être manipulées en relation avec le mouvement ou le flux optique.

Notre choix s'est porté sur une autre paramétrisation classique, utilisant le gradient de profondeur (X, Y) (*chi* et *upsilon* majuscule), illustré figure 4.2 (voir section 2.2.1). En considérant la profondeur² du plan comme une fonction f en coordonnées cartésiennes $z = f(x, y)$ (ce qui est possible sauf si le plan contient l'axe sagittal z), le gradient de f , $\overrightarrow{\text{grad}}f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$ est constant puisque f est linéaire. Graphiquement, les deux composantes

¹Dans tout ce document, nous considérons un modèle instantané de la perception du plan. En particulier ses paramètres d'orientation ou la distance de l'observateur au point de fixation sont fixes, malgré le mouvement relatif. Pour le flux optique, la distance entre l'objet et l'observateur n'est pas pertinente ; le positionnement est donc l'orientation et l'inclinaison du plan.

²Pour un point appartenant au plan, la profondeur est la distance de ce point au plan fronto-parallèle.

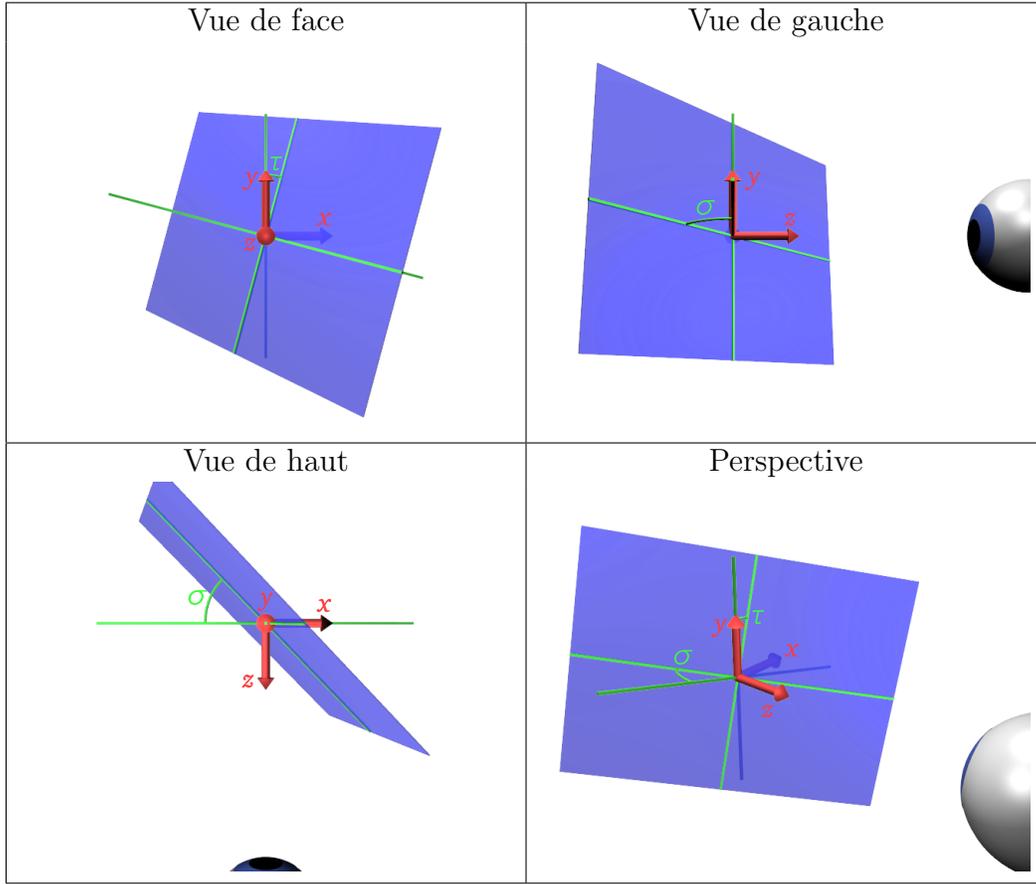


FIG. 4.1 – Illustration de l'inclinaison et de l'orientation d'un plan (rappel de la figure 3.1). Le plan objet, figuré en bleu, se distingue du plan fronto-parallèle par une rotation de son inclinaison σ autour de l'axe vertical y puis d'une rotation de son orientation τ autour de l'axe sagittal z .

X et Y de ce gradient sont les pentes respectives des droites intersections du plan objet avec les plans horizontal et sagittal figurées en vert sur la figure 4.2. On note que $(\chi, v, -1)$ est un vecteur normal au plan.

On peut facilement passer de la paramétrisation (X, Y) à l'orientation et l'inclinaison (T, Σ) grâce aux formules 4.1 et 4.2 similaires (à la tangente près) à un passage de coordonnées cartésiennes à polaires.

$$\begin{cases} \sigma = \arctan\left(\sqrt{\chi^2 + v^2}\right) \\ \tau = \arctan\left(\frac{v}{\chi}\right) \end{cases} \quad (4.1)$$

$$\begin{cases} \chi = \tan(\sigma) \cos(\tau) \\ v = \tan(\sigma) \sin(\tau) \end{cases} \quad (4.2)$$

En conclusion, la position du plan par rapport à l'observateur est représentée par deux variables X et Y de domaines $\mathcal{D}_X = \mathcal{D}_Y = \mathbb{R}$.

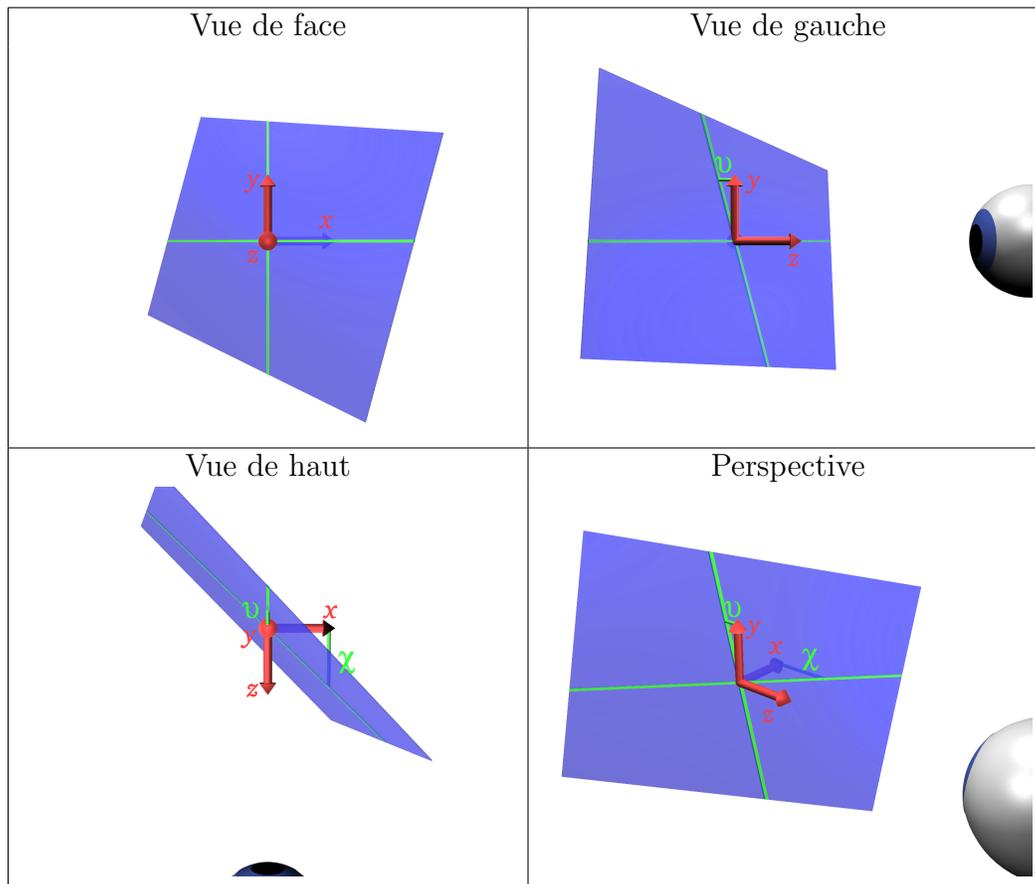


FIG. 4.2 – Illustration des paramètres (X, Υ) caractérisant un plan. L'objet est le plan bleu dans la même position que dans la figure 4.1 et l'observateur est placé en $(0, 0, d > 0)$. En vert sont figurées les intersections du plan avec les plans horizontal et sagittal dont les pentes sont les paramètres X et Υ respectivement.

4.2.2 Mouvement propre

Contrairement à l'exemple simplifié, où le participant devait ne pas bouger, nous considérons dans ce chapitre des expériences dans lesquelles son mouvement sera pris en compte. Plus spécifiquement, certaines conditions vont mettre en jeu un mouvement actif du participant, c'est-à-dire un mouvement propre volontaire et non pas subi.³

Nous introduisons ainsi les vecteurs $\vec{V} = (V_x, V_y, V_z)$ de vitesse de translation et $\vec{R} = (R_x, R_y, R_z)$ de vitesse de rotation de l'œil de l'observateur. Cependant nous n'avons pas de données à notre disposition concernant tous ces degrés de liberté. Dans la suite de ce document, seules certaines de ces variables sont considérées.

Pendant les expériences étudiées, il est demandé aux observateurs de ne pas bouger suivant l'axe vertical. Cependant, l'ensemble des stimuli est globalement invariant par

³Wexler (2003) montre que la perception est moins ambiguë en cas de mouvement volontaire du sujet qu'en cas de mouvement subi.

rotation selon l'axe sagittal. Une translation verticale nous semble donc similaire à une translation transversale et nous nous attendrions à obtenir des résultats qualitativement identiques.

Les données dont nous disposons ne recouvrent que les rotations de l'œil suivant l'axe vertical. Ces mouvements correspondent au suivi d'un objet dans le plan horizontal. La rotation selon l'axe transversal nous paraît similaire à celle suivant l'axe vertical, pour les mêmes raisons de symétrie que pour les translations.

Enfin, la rotation selon l'axe sagittal est assez particulière en ce qu'elle laisse l'image invariante. Nous ne disposons pas non plus de données sur cette rotation.

Au final, pour ne pas surcharger le modèle, nous ne gardons dans la suite que trois des six variables du mouvement propre : les vitesses de translation V_x et V_z suivant les axes transversal et sagittal respectivement et la vitesse de rotation de l'œil R_y autour de l'axe vertical. Ces variables sont illustrées par la figure 4.3 et de domaine réel. En présence de plus de données, le modèle complet inclurait tous ces degrés de liberté.

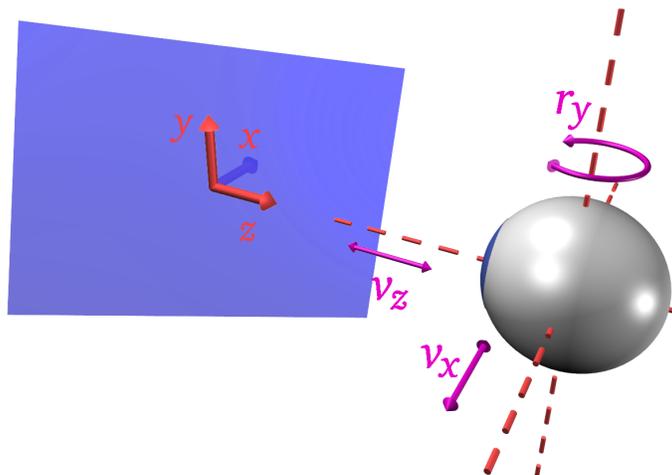


FIG. 4.3 – Illustration des variables retenues pour décrire le mouvement propre de l'observateur.

4.2.3 Mouvement du plan

Comme dans l'exemple précédent, le mouvement du plan est une variable intermédiaire de notre modèle. Au lieu de nous restreindre à une simple rotation autour de l'axe vertical, nous considérons l'ensemble des mouvements possibles. Ce mouvement est exprimé relativement à l'observateur. Le choix du mouvement relatif plutôt qu'absolu est dû au fait que le flux optique est engendré par le mouvement relatif.

Le plan possède trois degrés de liberté pour les translations et trois pour les rotations. Ces mouvements sont représentés par leurs vitesses dans le repère cartésien partageant les axes avec celui de l'observateur, mais centré sur le point de fixation. Les repères dans lesquels sont décrits les mouvements propres de l'observateur et les mouvements relatifs du plan diffèrent donc d'une translation du centre du repère. La translation relative \vec{T} de l'objet par rapport à l'observateur est représentée par trois variables T_x , T_y et T_z de domaine $\mathcal{D}_{\vec{T}} = \mathbb{R}^3$; la rotation relative par $\vec{\Omega} = (\Omega_x, \Omega_y, \Omega_z)$ avec $\mathcal{D}_{\vec{\Omega}} = \mathbb{R}^3$.

4.2.4 Flux optique

Ce modèle s'intéresse à la perception de plans *via* le flux optique. L'ensemble des flux optiques pertinents sont ceux générés par un plan. Nous représentons le flux optique $\vec{\Phi}$ par 8 paramètres (voir section 2.2.1).

Les variables représentant le flux optique dans le modèle sont $\Phi_x^0, \Phi_y^0, \Phi_x^1, \Phi_y^1, \Phi_x^2, \Phi_y^2$. Ces huit variables sont chacune de domaines $\mathcal{D}_{\Phi} = \mathbb{R}$. Ces paramètres peuvent être regroupés en 3 groupes $\vec{\Phi}^0, \vec{\Phi}^1$ et $\vec{\Phi}^2$ selon leur ordre de dérivation.

4.2.5 Champ de vision

Nous introduisons une dernière variable pour définir la perception du stimulus. Nous nous intéressons dans la suite des expériences dans lesquelles la taille du champ de vision du stimulus varie. Pour simplifier, nous considérons deux conditions extrêmes de perception qui correspondent aux conditions expérimentales : un *petit champ* de vision et un *grand champ* de vision.

Un petit champ de vision correspond à une ouverture de 5 à 8°⁴. Un grand champ de vision correspond à une ouverture de 60°.

En conclusion, nous introduisons la variable booléenne Λ . les valeurs 0 et 1 de son domaine \mathcal{D}_{Λ} correspondent respectivement aux conditions petit champ et grand champ.

4.2.6 Récapitulatif

Les variables d'observations de notre modèle sont :

- $\vec{\Phi}$: le flux optique ;
- V_x, V_z et R_y : le mouvement propre de l'observateur ;
- Λ : la taille du champ de vision.

Les variables cherchées X et Y représentent la position du plan. Il s'agit du gradient de profondeur dans le plan fronto-parallèle.

Enfin, nous considérons le mouvement relatif du plan par rapport à l'observateur comme variable intermédiaire. Ce mouvement est séparé en ses composantes de translation $\vec{T} = (T_x, T_y, T_z)$ d'une part et ses composantes de rotation $\vec{\Omega} = (\Omega_x, \Omega_y, \Omega_z)$ d'autre part.

⁴L'ouverture est ici l'angle au sommet du cône de vision.

4.3 Décomposition

Comme dans l'exemple du chapitre 3, nous utilisons les hypothèses du modèle pour spécifier sa structure de dépendance. Cette structure s'exprime au travers de la décomposition de la distribution de probabilité conjointe.

L'hypothèse 3 de notre modèle, précisée section 4.1 est l'indépendance de la position du plan et du mouvement, qu'il s'agisse de celui de l'observateur ou du mouvement du plan lui-même. L'expression probabiliste de cette hypothèse est que la distribution de probabilité sur la position du plan et les mouvements est le produit de la distribution sur la position et de la distribution sur les mouvements (équation 4.3).

$$P(X \Upsilon V_x V_z R_y \vec{T} \vec{\Omega}) = P(X \Upsilon)P(V_x V_z R_y \vec{T} \vec{\Omega}) \quad (4.3)$$

Dans notre contexte d'observation d'un objet plan, l'hypothèse de rigidité 1 stipule que le flux optique perçu sera plus probablement celui d'un plan rigide. Il s'agit d'une hypothèse très importante dans notre modèle et qui a plusieurs implications.

En premier lieu, si le flux perçu est celui du plan, il ne dépend que du mouvement relatif, de la position et de la taille du champ de vision. En particulier, sachant tout cela, il est indépendant du mouvement propre. On écrit donc :

$$P(\vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2 | X \Upsilon V_x V_z R_y \vec{T} \vec{\Omega} \Lambda) = P(\vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2 | X \Upsilon \vec{T} \vec{\Omega} \Lambda) \quad (4.4)$$

Comme pour le modèle didactique du chapitre précédent, nous écrivons que les composantes du flux optiques d'un plan sont indépendantes les unes des autres conditionnellement à la position, le mouvement relatif et la taille du champ de vision (voir section 3.3.2). Il est à noter qu'il s'agit bien d'une indépendance conditionnellement à la connaissance de la position du mouvement et de la taille du champ ; les composantes du flux ne sont pas indépendantes sinon. Mais la connaissance de la valeur d'une composante du flux optique ne modifie pas l'état de connaissance que l'on peut avoir sur une autre composante à partir des seules informations de position, mouvement relatif et taille du champ de vision.

Nous écrivons ainsi l'équation 4.5.

$$\begin{aligned} & P(\vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2 | X \Upsilon \vec{T} \vec{\Omega} \Lambda) \\ &= P(\vec{\Phi}^0 | X \Upsilon \vec{T} \vec{\Omega} \Lambda) \\ &\times P(\vec{\Phi}^1 | X \Upsilon \vec{T} \vec{\Omega} \Lambda) \\ &\times P(\vec{\Phi}^2 | X \Upsilon \vec{T} \vec{\Omega} \Lambda) \end{aligned} \quad (4.5)$$

De la même manière, on écrit que le flux optique d'ordre 0 ne dépend que des translations dans le plan fronto-parallèle (équation 4.6)⁵ et que celui d'ordre 1 ne dépend pas de la taille du champ de vision (équation 4.7).

$$P(\vec{\Phi}^0 | X \Upsilon \vec{T} \vec{\Omega} \Lambda) = P(\vec{\Phi}^0 | T_x T_y) \quad (4.6)$$

⁵Rappelons que la translation relative \vec{T} a trois composantes T_x , T_y et T_z .

$$P(\vec{\Phi}^1 | X \Upsilon \vec{T} \vec{\Omega} \Lambda) = P(\vec{\Phi}^1 | X \Upsilon \vec{T} \vec{\Omega}) \quad (4.7)$$

En résumé, l'expression de la rigidité au sein de la décomposition s'écrit :

$$\begin{aligned} & P(\vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2 | X \Upsilon V_x V_z R_y \vec{T} \vec{\Omega} \Lambda) \\ &= P(\vec{\Phi}^0 | T_x T_y) \\ &\times P(\vec{\Phi}^1 | X \Upsilon \vec{T} \vec{\Omega}) \\ &\times P(\vec{\Phi}^2 | X \Upsilon \vec{T} \vec{\Omega} \Lambda) \end{aligned} \quad (4.8)$$

L'hypothèse 2 de notre modèle est la stationnarité stipulant que le mouvement perçu sera plus probablement d'assez faible amplitude dans le référentiel allocentré. Si le mouvement absolu du plan est petit, cela signifie que son mouvement relatif est proche de l'opposé du mouvement propre de l'observateur et donc qu'il en dépend. Il ne s'agit donc pas d'une hypothèse d'indépendance particulière mais on peut écrire, grâce à la règle de Bayes, l'expression suivante :

$$\begin{aligned} & P(V_x V_z R_y \vec{T} \vec{\Omega}) \\ &= P(V_x V_z R_y)P(\vec{T} \vec{\Omega} | V_x V_z R_y) \end{aligned} \quad (4.9)$$

Pour finir, rajoutons une hypothèse généralement implicite dans les travaux de psychophysique, qui est que la taille du champ est indépendante de la position et des mouvements de l'observateur ou du plan. Il s'agit d'une hypothèse relativement triviale dont l'expression dans notre modèle est montrée par l'équation 4.10.

$$P(X \Upsilon V_x V_z R_y \vec{T} \vec{\Omega} \Lambda) = P(X \Upsilon V_x V_z R_y \vec{T} \vec{\Omega})P(\Lambda) \quad (4.10)$$

Toutes ces hypothèses nous permettent de donner une expression simplifiée de la distribution conjointe. En effet, l'application de la règle de Bayes donne par exemple :

$$\begin{aligned} & P(X \Upsilon V_x V_z R_y \vec{T} \vec{\Omega} \Lambda \vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2) \\ &= P(X \Upsilon V_x V_z R_y \vec{T} \vec{\Omega} \Lambda) \end{aligned} \quad (4.11)$$

$$\times P(\vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2 | X \Upsilon V_x V_z R_y \vec{T} \vec{\Omega} \Lambda) \quad (4.12)$$

En remplaçant respectivement 4.11 et 4.12 par 4.10 et 4.8, on obtient :

$$\begin{aligned} & P(X \Upsilon V_x V_z R_y \vec{T} \vec{\Omega} \Lambda \vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2) \\ &= P(X \Upsilon V_x V_z R_y \vec{T} \vec{\Omega}) \\ &\times P(\Lambda) \\ &\times P(\vec{\Phi}^0 | T_x T_y) \\ &\times P(\vec{\Phi}^1 | X \Upsilon \vec{T} \vec{\Omega}) \\ &\times P(\vec{\Phi}^2 | X \Upsilon \vec{T} \vec{\Omega} \Lambda) \end{aligned} \quad (4.13)$$

On substitue successivement les expressions 4.3 et 4.9 au facteur 4.13 pour obtenir la décomposition finale exprimée par l'équation 4.14.

$$\begin{aligned}
& P(X \Upsilon V_x V_z R_y \vec{T} \vec{\Omega} \Lambda \vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2) \\
= & P(X \Upsilon)P(V_x V_z R_y)P(\Lambda) \\
\times & P(\vec{T} \vec{\Omega} | V_x V_z R_y) \\
\times & P(\vec{\Phi}^0 | T_x T_y) \\
\times & P(\vec{\Phi}^1 | X \Upsilon \vec{T} \vec{\Omega}) \\
\times & P(\vec{\Phi}^2 | X \Upsilon \vec{T} \vec{\Omega} \Lambda)
\end{aligned} \tag{4.14}$$

4.4 Distributions

À partir de l'expression précédente de la distribution conjointe, nous pouvons donner une distribution particulière à chacun des facteurs de la décomposition pour obtenir une expression directement calculable de cette conjointe. Ces choix de distributions constituent la suite de l'expression de nos hypothèses énoncées section 4.1.

4.4.1 Stationnarité

Un des facteurs de la décomposition concerne le lien entre le mouvement relatif et le mouvement propre $P(\vec{T} \vec{\Omega} | V_x V_z R_y)$. Pour le spécifier, nous utilisons l'hypothèse de stationnarité qui stipule que le mouvement absolu est plus probablement faible. Cela signifie que le mouvement relatif du plan par rapport à l'observateur est proche de l'opposé de son mouvement propre.

Nous définissons dans le premier temps l'opposé du mouvement propre dans l'espace des mouvements relatifs. Nous discutons ensuite ce qu'est un mouvement relatif proche d'un autre.

Changement de repère Expliciter le mouvement d'un plan immobile (dans le référentiel global) en fonction du mouvement de l'observateur revient simplement à effectuer un changement de repère $\mathcal{C}_{\mathcal{O} \rightarrow \mathcal{P}}$ en transposant l'opposé du mouvement propre (centré sur l'œil) dans le référentiel choisi pour le mouvement relatif (centré sur le point de fixation) :

$$\begin{pmatrix} t_x \\ t_y \\ t_z \\ \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \mathcal{C}_{\mathcal{O} \rightarrow \mathcal{P}} \begin{pmatrix} v_x \\ v_z \\ r_y \end{pmatrix} \tag{4.15}$$

Lors d'une translation sagittale, le mouvement opposé est une translation sagittale de vitesse opposée (illustré figure 4.4). On a alors :

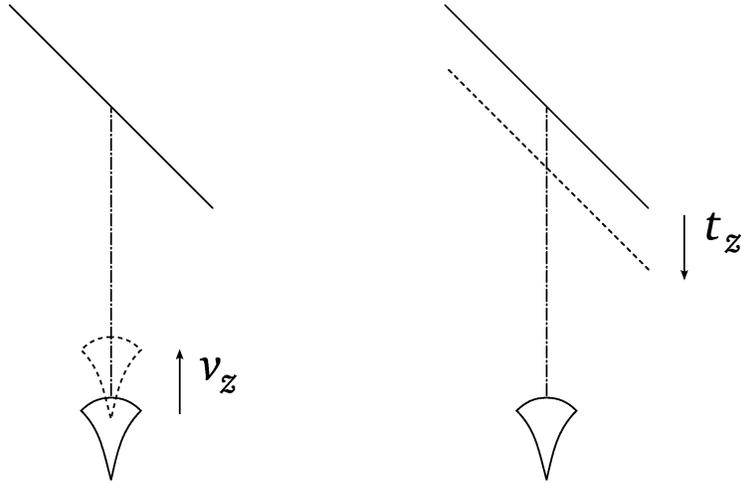


FIG. 4.4 – Illustration du changement de repère de la translation sagittale, exprimée dans le référentiel allocentré à gauche et égocentré à droite.

$$\mathcal{C}_{\mathcal{O} \rightarrow \mathcal{P}} \begin{pmatrix} 0 \\ v_z \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -v_z \\ 0 \\ 0 \end{pmatrix} \quad (4.16)$$

Dans le cas d'une translation transversale, si l'observateur continue de fixer l'origine du repère, le mouvement opposé est une rotation selon l'axe vertical (voir figure 4.5). En appelant d la distance de l'observateur au point de fixation⁶, on obtient :

$$\mathcal{C}_{\mathcal{O} \rightarrow \mathcal{P}} \begin{pmatrix} v_x \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -\frac{v_x}{d} \\ 0 \end{pmatrix} \quad (4.17)$$

Pour une rotation de l'œil autour de l'axe vertical, il y a deux effets sur le plan, dûs au changement de centre de repère. Comme on peut le voir sur la figure 4.6, le plan subit alors une rotation de sens opposé accompagné d'une translation transversale.

$$\mathcal{C}_{\mathcal{O} \rightarrow \mathcal{P}} \begin{pmatrix} 0 \\ 0 \\ r_y \end{pmatrix} = \begin{pmatrix} d * r_y \\ 0 \\ 0 \\ 0 \\ -r_y \\ 0 \end{pmatrix} \quad (4.18)$$

⁶ d est aussi la distance séparant les centres des deux repères

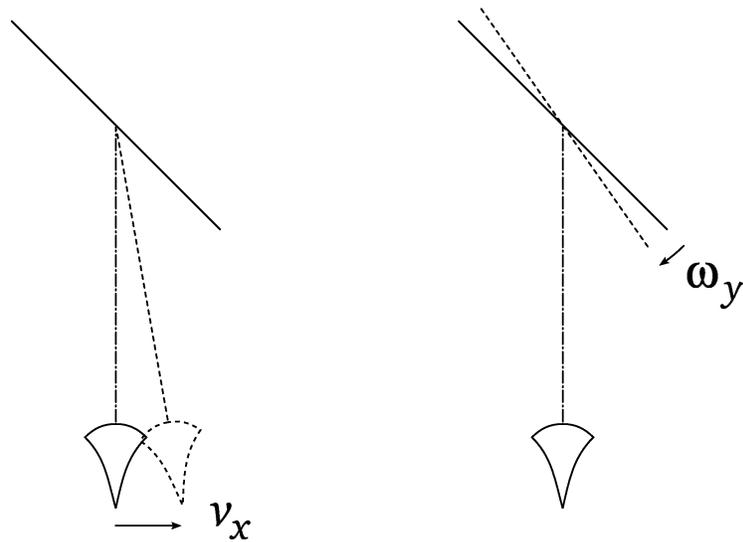


FIG. 4.5 – Illustration du changement de repère de la translation transversale, exprimée dans le référentiel allocentré à gauche et égocentré à droite. On rappelle que, dans ce cas, l'observateur reste sur le point de fixation.

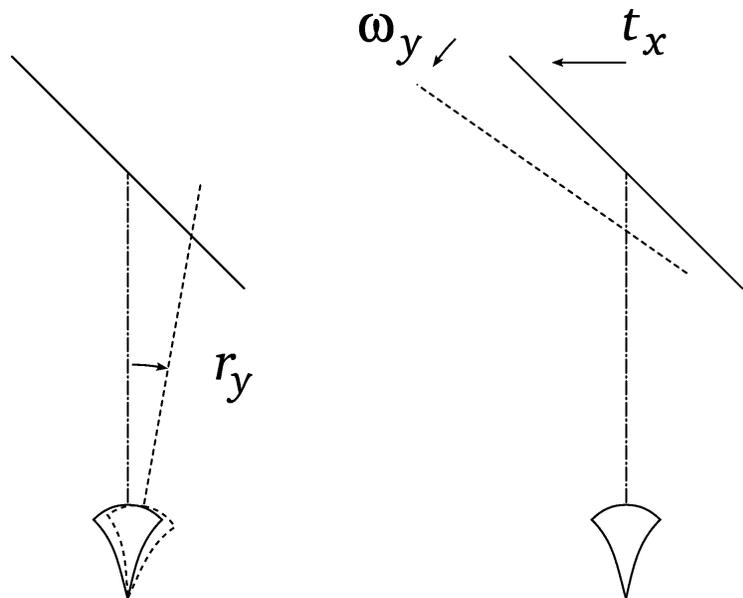


FIG. 4.6 – Illustration du changement de repère de la rotation verticale, exprimée dans le référentiel allocentré à gauche et égocentré à droite.

Dans la suite la distance d entre l'observateur et le point de fixation est considérée égale à l'unité de distance. Au final, comme le changement de repère est linéaire, on obtient

l'expression suivante :

$$\mathcal{C}_{\mathcal{O} \rightarrow \mathcal{P}} \begin{pmatrix} v_x \\ v_y \\ r_y \end{pmatrix} = \begin{pmatrix} r_y \\ 0 \\ -v_z \\ 0 \\ -v_x - r_y \\ 0 \end{pmatrix} \quad (4.19)$$

c'est-à-dire l'ensemble d'équations :

$$\begin{cases} t_x = r_y \\ t_y = 0 \\ t_z = -v_z \\ \omega_x = 0 \\ \omega_y = -v_x - r_y \\ \omega_z = 0 \end{cases} \quad (4.20)$$

Distributions La distribution $P(\vec{T} \vec{\Omega} \mid V_x V_z R_y)$ spécifie dans quelle mesure le mouvement relatif risque de s'éloigner de l'opposé du mouvement propre.

Il y a potentiellement plusieurs manières de la définir. Une première consisterait à discriminer deux cas : plan immobile, pour lequel le mouvement du plan est donné exactement par les équations ci-dessus, et plan mobile pour lequel on n'a aucune information sur le mouvement du plan. La distribution résultante serait la somme d'une uniforme et d'un Dirac. Il y aurait cependant deux problèmes principaux avec cette méthode. Tout d'abord, techniquement, cette distribution est très sensible à la précision avec laquelle le mouvement propre est connu, à cause de sa composante Dirac. De plus, du point de vue des hypothèses, les mouvements du plan ne nous apparaissent pas tous équiprobables. Par exemple, lors des rotations, il semble peu probable que le plan ait effectué un ou plusieurs tours entre deux images.

Une seconde méthode pour choisir la distribution, que nous avons suivie, est de considérer une énergie et d'en déduire la distribution comme inversement proportionnelle à son exponentielle (distribution de Gibbs). Comme nous nous intéressons à la distribution sur le mouvement nous utilisons l'énergie cinétique $E_c = E_t + E_r$ avec E_t l'énergie cinétique de translation $E_t = \frac{1}{2}M.V^2$ et E_r l'énergie cinétique de rotation $E_r = \frac{1}{2}I.\Omega^2$. Notre distribution est donc une Gaussienne centrée sur le mouvement opposé au mouvement propre, et de matrice de covariance σ_M diagonale (équation 4.21).

$$\begin{aligned} & P(\vec{T} \vec{\Omega} \mid V_x V_z R_y) \\ &= \mathcal{N}_{\mathcal{C}_{\mathcal{O} \rightarrow \mathcal{P}}(V_x V_z R_y), \sigma_M}(\vec{T} \vec{\Omega}) \end{aligned} \quad (4.21)$$

Cette distribution est paramétrée par la matrice de covariance σ_M dont nous discuterons section 4.5.3.

4.4.2 Rigidité

La décomposition fait déjà apparaître l'hypothèse 1 de rigidité sous la forme des facteurs mettant en jeu les composantes du flux optique :

$$\begin{cases} P(\vec{\Phi}^0 | T_x T_y) \\ P(\vec{\Phi}^1 | X \Upsilon \vec{T} \vec{\Omega}) \\ P(\vec{\Phi}^2 | X \Upsilon \vec{T} \vec{\Omega} \Lambda) \end{cases} \quad (4.22)$$

Expression du flux optique La rigidité fait l'hypothèse que le flux perçu sera plus probablement celui d'un objet rigide. Dans notre cas, l'objet auquel nous nous intéressons est le plan.

Nous rappelons les équations du flux optiques (voir section 2.2.1 et annexe A) pour un plan de paramètres (χ, v) et de mouvement relatif $(\vec{t}, \vec{\omega})$. Avec nos notations, les flux optiques d'ordres respectifs 0, 1 et 2 sont :

$$\begin{pmatrix} \phi_x^0 \\ \phi_y^0 \end{pmatrix} = \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (4.23)$$

$$\begin{pmatrix} x\phi_x^1 & y\phi_x^1 \\ x\phi_y^1 & y\phi_y^1 \end{pmatrix} = \begin{pmatrix} t_z + \chi \cdot (t_x + \omega_y) & -\omega_z + v \cdot (t_x + \omega_y) \\ \omega_z + \chi \cdot (t_y - \omega_x) & t_z + v \cdot (t_y - \omega_x) \end{pmatrix} \quad (4.24)$$

$$\begin{pmatrix} x\phi^2 \\ y\phi^2 \end{pmatrix} = \begin{pmatrix} \chi \cdot t_z - \omega_y \\ v \cdot t_z + \omega_x \end{pmatrix} \quad (4.25)$$

Distributions L'hypothèse de rigidité que nous discutons ici vise à exprimer que le flux optique perçu est plus probablement celui d'un plan rigide. Pour les distributions $P(\vec{\Phi}^0 | T_x T_y)$, $P(\vec{\Phi}^1 | X \Upsilon \vec{T} \vec{\Omega})$ et $P(\vec{\Phi}^2 | X \Upsilon \vec{T} \vec{\Omega} \Lambda)$, nous choisissons des Gaussiennes centrées respectivement sur les expressions 4.23, 4.24 et 4.25 qui définissent le flux optique d'un plan.

Nous considérons les composantes de chacun des ordres du flux indépendantes les unes des autres. Cela se traduit par des matrices de covariance diagonale pour ces trois distributions. De plus, conformément à l'hypothèse que nous faisons que la taille du champ de vision modifie la qualité de la perception du flux optique d'ordre deux (hypothèse 4), la valeur des écarts-type sur les variables $x\phi^2$ et $y\phi^2$ sera plus grande pour une valeur de Λ correspondant à un petit champ de vision que pour un grand champ.

On obtient les expressions 4.26, 4.27 et 4.28 pour les distributions concernant le flux optique.

$$P(\vec{\Phi}^0 | t_x t_y) = \mathcal{N}_{\mathcal{F}^0(t_x, t_y), \sigma_{\vec{\Phi}^0}}(\vec{\Phi}^0) \quad (4.26)$$

$$P(\vec{\Phi}^1 | \chi v \vec{t} \vec{\omega}) = \mathcal{N}_{\mathcal{F}^1(\chi, v, \vec{t}, \vec{\omega}), \sigma_{\vec{\Phi}^1}}(\vec{\Phi}^1) \quad (4.27)$$

$$P(\vec{\Phi}^2 | \chi v \vec{t} \vec{\omega} \lambda) = \mathcal{N}_{\mathcal{F}^2(\chi, v, \vec{t}, \vec{\omega}), \sigma_{\vec{\Phi}^2}(\lambda)}(\vec{\Phi}^2) \quad (4.28)$$

À l'image de la distribution sur le mouvement relatif, celles-ci sont paramétrées par leur matrices de covariance $\sigma_{\vec{\Phi}0}$, $\sigma_{\vec{\Phi}1}$ et $\sigma_{\vec{\Phi}2}$.

4.4.3 Position

Nous devons spécifier la distribution sur la position du plan $P(X, Y)$. Aucune des hypothèses de notre modèle ne porte dessus. À l'image de l'exemple du chapitre 3, nous la définissons en suivant le principe du maximum d'entropie (section 2.3). Sans observable sur la distribution (comme la moyenne ou l'écart-type), cela signifie une distribution uniforme sur la position du plan.⁷

Par contre, ce choix uniforme est dépendant de l'espace paramétrique par lequel le plan est décrit. Une distribution sur l'espace du gradient (χ, v) est différente dans l'espace des angles d'inclinaison et d'orientation (σ, τ) .⁸ Le choix d'une distribution se fait donc en accord avec une représentation paramétrique donnée.

La signification que nous voulons donner à cette connaissance *a priori* est que la distribution est globalement invariante par une rotation quelconque autour du point de fixation. En conséquence, nous voulons exprimer que la distribution de la normale au plan est uniforme sur une sphère. En coordonnées (Σ, T) cela s'exprime par la densité de probabilité 4.29.⁹

$$\forall(\sigma, \tau) \in \left[0; \frac{\pi}{2}\right] \times [0; 2\pi], P(\sigma, \tau) = \frac{1}{2\pi} \sin \sigma \quad (4.29)$$

En utilisant l'expression du changement de repère \mathcal{C} de (χ, v) vers (σ, τ) (rappelé équation 4.30) et sa matrice jacobienne $J_{\mathcal{C}}$ (expression 4.31) de déterminant $\det J_{\mathcal{C}}$ (expression 4.32), nous transposons cette distribution dans la paramétrisation (X, Y) de notre modèle.

$$\mathcal{C} : \begin{cases} \sigma(\chi, v) = \arctan\left(\sqrt{\chi^2 + v^2}\right) \\ \tau(\chi, v) = \arctan\left(\frac{v}{\chi}\right) \end{cases} \quad (4.30)$$

$$J_{\mathcal{C}} = \begin{pmatrix} \frac{\partial \sigma}{\partial \chi} & \frac{\partial \sigma}{\partial v} \\ \frac{\partial \tau}{\partial \chi} & \frac{\partial \tau}{\partial v} \end{pmatrix} \quad (4.31)$$

$$J_{\mathcal{C}} = \begin{pmatrix} \frac{\chi}{\sqrt{\chi^2 + v^2}(1 + \chi^2 + v^2)} & \frac{v}{\sqrt{\chi^2 + v^2}(1 + \chi^2 + v^2)} \\ \frac{-v}{\chi^2 + v^2} & \frac{\chi}{\chi^2 + v^2} \end{pmatrix}$$

$$\det J_{\mathcal{C}} = \frac{1}{\sqrt{\chi^2 + v^2} (1 + \chi^2 + v^2)} \quad (4.32)$$

⁷Nous négligeons ici le fait qu'on ne peut pas percevoir de plan plus incliné que le demi-angle au sommet du cône de vision. En pratique, les calculs sont effectués sur des domaines bornés qui excluent de fait ces cas de figures (voir 4.5.3).

⁸La différence de domaine permet de s'en convaincre.

⁹Il suffit de considérer que la probabilité d'un ensemble est proportionnel à sa surface et de découper la sphère en bandes définies par σ constant. La surface de chaque bande est $2\pi \sin \sigma d\sigma$ d'où la distribution 4.29.

Nous voulons que notre changement de variable conserve la probabilité de n'importe quel ensemble de configuration. Cela s'écrit :

$$\forall A \subset \mathcal{D}_X \times \mathcal{D}_Y, \int_A P(\chi, v) d\chi dv = \int_{\mathcal{C}(A)} P(\sigma, \tau) d\sigma d\tau \quad (4.33)$$

En utilisant les formules de changement de variables dans une intégrale (variables en dehors de la singularité en $(\chi, v) = (0, 0)$), on souhaite alors :

$$\begin{aligned} \forall A \subset \mathcal{D}_X \times \mathcal{D}_Y \setminus (0, 0), \int_A P(\chi, v) d\chi dv &= \int_A P(\sigma(\chi, v), \tau(\chi, v)) |\det J_{\mathcal{C}}| d\chi dv \\ &= \int_A \frac{\sin\left(\arctan\left(\sqrt{\chi^2 + v^2}\right)\right)}{2\pi\sqrt{\chi^2 + v^2}(1 + \chi^2 + v^2)} d\chi dv \\ \int_A P(\chi, v) d\chi dv &= \int_A \frac{1}{2\pi} (1 + \chi^2 + v^2)^{-\frac{3}{2}} d\chi dv \end{aligned}$$

Une solution de ces contraintes, que nous choisissons pour distribution sur la position, est l'expression 4.34, prolongée par continuité en $(0, 0)$.

$$\forall \chi, v \in \mathbb{R}^2, P(\chi, v) = \frac{1}{2\pi} (1 + \chi^2 + v^2)^{-\frac{3}{2}} \quad (4.34)$$

Cette distribution, image de la distribution uniforme par le changement de repère \mathcal{C} , ne nécessite pas de paramètre.

4.4.4 Distributions inconnues

Les facteurs de la décomposition que nous n'avons pas encore spécifiés concernent le mouvement propre et la taille du champ de vision. Il se trouve que nous n'utiliserons notre modèle qu'en connaissant les valeurs de ces variables. L'inférence bayésienne a la propriété de ne pas dépendre des éventuelles distributions non-conditionnelles portant sur des variables instanciées.¹⁰

En effet, soient A , B et C trois variables, la règle de Bayes donne l'expression 4.35.

$$P(A | b, c) = \frac{P(A, b, c)}{\sum_{a' \in \mathcal{D}_A} P(a', b, c)} \quad (4.35)$$

Dans le cas où la décomposition spécifie une distribution non conditionnelle sur C non nulle en c , on peut alors simplifier l'expression au-dessus par le facteur $P(c)$ pour obtenir l'expression 4.36.

$$\begin{aligned} P(A | b, c) &= \frac{P(c)P(A, b | c)}{\sum_{a' \in \mathcal{D}_A} P(c)P(a', b | c)} \\ &= \frac{P(A, b | c)}{\sum_{a'} P(a', b | c)} \end{aligned} \quad (4.36)$$

¹⁰Sauf aux points où elles s'annulent. Il est cependant toujours possible de trouver une distribution non nulle sur tout le domaine considéré.

En conséquence, nous n'avons pas à effectuer de choix sur les distributions sur le mouvement propre et la taille du champ. Nous ne rajoutons donc pas d'informations inutiles dans notre modèle et les laissons inconnues. Nous vérifions une fois de plus que ces distributions ne sont pas utiles en section 4.5.2.

4.5 Question

Nous pouvons maintenant calculer complètement notre distribution conjointe et donc utiliser notre modèle. Dans le formalisme que nous suivons, cette utilisation passe par le calcul d'une distribution de probabilité particulière appelée *question*.

4.5.1 Formulation de la question

Le but de ce modèle est d'inférer des connaissances probabilistes sur la position du plan par un observateur. Cette perception dépend d'un ensemble d'observations : le flux optique, la taille du champ de vision et le mouvement propre.

La question que nous posons à notre modèle est donc :

$$P(X \Upsilon \mid \vec{\phi} \lambda v_x v_z r_y) \quad (4.37)$$

4.5.2 Inférence

Pour répondre à cette question, nous appliquons les règles de Bayes et de marginalisation (voir section 2.3) à l'expression de conjointe du modèle. Comme pour l'exemple en section 3.3.4, nous obtenons l'expression 4.38 valable sur tout le domaine des variables de position du plan X et Υ .

$$\begin{aligned} & P(\chi v \mid \vec{\phi} \lambda v_x v_z r_y) \\ \propto & P(\chi v) \sum_{(\vec{t}, \vec{\omega}) \in \mathcal{D}_{\vec{T}} \times \mathcal{D}_{\vec{\Omega}}} \times P(\vec{t} \vec{\omega} \mid v_x v_z r_y) \\ & \times P(\vec{\phi}^0 \mid t_x t_y) \\ & \times P(\vec{\phi}^1 \mid \chi v \vec{t} \vec{\omega}) \\ & \times P(\vec{\phi}^2 \mid \chi v \vec{t} \vec{\omega} \lambda) \end{aligned} \quad (4.38)$$

On remarque, sans en être étonné, que l'expression 4.38 ne comporte que des facteurs que nous avons spécifiés complètement. On est donc en présence d'une expression calculable de notre distribution qui peut alors être comparée aux résultats expérimentaux.

Il est à noter que les hypothèses de notre modèle se répartissent entre le choix des variables, l'écriture de la décomposition et la spécification des distributions de probabilité. Au final, la seule chose qui compte pour l'inférence est la distribution conjointe ainsi construite. Il y a, en général, plusieurs manière de spécifier ses hypothèses dans un conjointe mais plus une hypothèse est exprimée tôt (parmi la décomposition plutôt que dans la distribution ou dans le choix des variables plutôt que dans la décomposition), plus elle va

Covariances
$\sigma_{\vec{T}} = 0,3 \cdot \text{Id}_{3 \times 3}$ en m.s^{-1}
$\sigma_{\vec{\Omega}} = 1,2 \cdot \text{Id}_{3 \times 3}$ en rad.s^{-1}
$\sigma_{\Phi^0} = 1 \cdot \text{Id}_{2 \times 2}$ en m.s^{-1}
$\sigma_{\Phi^1} = 0,025 \cdot \text{Id}_{4 \times 4}$ en s^{-1}
$\sigma_{\Phi^2} \mid_{\lambda=0} = 5 \cdot \text{Id}_{2 \times 2}$ en $\text{m}^{-1}.\text{s}^{-1}$
$\sigma_{\Phi^2} \mid_{\lambda=1} = 0,2 \cdot \text{Id}_{2 \times 2}$ en $\text{m}^{-1}.\text{s}^{-1}$

TAB. 4.1 – Matrices de covariance de chaque distribution gaussienne.

être contraignante (une hypothèse d'indépendance conditionnelle ne peut pas être modifiée par les choix des distributions) mais plus il va être possible de faire des simplifications lors du calcul de l'inférence.

4.5.3 Paramètres

Le but de notre modèle est de reproduire un éventail de résultats expérimentaux relatifs à la perception d'un plan par le flux optique. Nous souhaitons donc exhiber un unique jeu de paramètres permettant de reproduire au mieux ces résultats plutôt que de rechercher les paramètres optimaux pour chacune des expériences. Les paramètres de notre modèle sont de deux types : les paramètres des distributions et ceux de l'inférence.

Il y a d'abord les paramètres intrinsèques au modèle. Ceux-ci apparaissent dans les expressions des distributions, telles que les covariances des distributions gaussiennes. Comme nous l'avons vu, notre modèle requiert ainsi les matrices de covariance des distributions sur le mouvement et sur chacun des flux optiques, respectivement $\sigma_M = (\sigma_{\vec{T}}, \sigma_{\vec{\Omega}})$ et σ_{Φ^0} , σ_{Φ^1} et σ_{Φ^2} . Nous avons choisi ces matrices diagonales. Leurs valeurs sont précisées dans la table 4.1.

Ces valeurs sont contraintes par les résultats expérimentaux. Le choix des matrices $\sigma_{\vec{T}}$ et $\sigma_{\vec{\Omega}}$ est discuté aux sections 5.4 et 5.5. Pour les matrices σ_{Φ^0} et σ_{Φ^1} , des valeurs trop faibles peuvent entraîner des artefacts de calcul et des valeurs trop grandes peuvent diluer l'information. L'éventail des valeurs admissibles est large entre ces cas extrêmes. Le choix de la matrice σ_{Φ^2} , qui dépend de la taille du champ de vision Λ est discuté aux sections 5.2 et 5.4.

Notre modèle est implanté à l'aide du moteur d'inférence ProBT. Les algorithmes utilisés permettent de faire de l'inférence exacte ou approchée. Les paramètres de l'inférence concernent la représentation des variables réelles qui sont bornées et discrétisées au besoin.

Le tableau 4.2 présente les bornes minimum et maximum de chaque variable ainsi que le nombre de cas envisagés en cas de sommation. Les variables non listées correspondent à la partie droite de notre question. Elles ne sont donc pas discrétisées et leur domaine n'intervient pas (à partir du moment où il contient les observations). On note que les variables du gradient de profondeur sont maintenant bornées. En termes angulaires, ces valeurs correspondent à ne représenter les plans que jusqu'à une inclinaison d'environ 80° .

Variable	min	max	nombre
X et Υ	-4,75	4,75	19
Ω_x, Ω_y et Ω_z	-1,125 rad.s ⁻¹	1,125 rad.s ⁻¹	9
T_x, T_y et T_z	-1,125 m.s ⁻¹	1,125 m.s ⁻¹	9

TAB. 4.2 – Domaine des variables.

4.6 Test de régression

Commençons par répliquer les résultats de la section 3.3 avec notre modèle complet pour les comparer avec les résultats expérimentaux et vérifier que sa plus grande complexité ne le rend pas moins pertinent.

4.6.1 Rappel de l'expérience

L'expérience est celle décrite à la section 3.1. Elle concerne un plan incliné en rotation figuré par l'entremise d'une texture de points aléatoire. Il est demandé à un observateur monoculaire et immobile d'évaluer l'orientation du plan perçu.

Les résultats de l'expérience sont présentés sous la forme d'une distribution d'erreur sur l'orientation et sont rappelés figure 4.7.

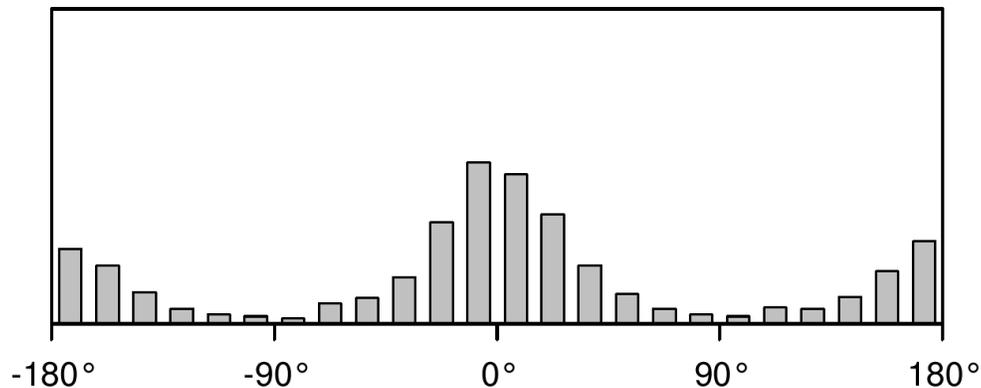


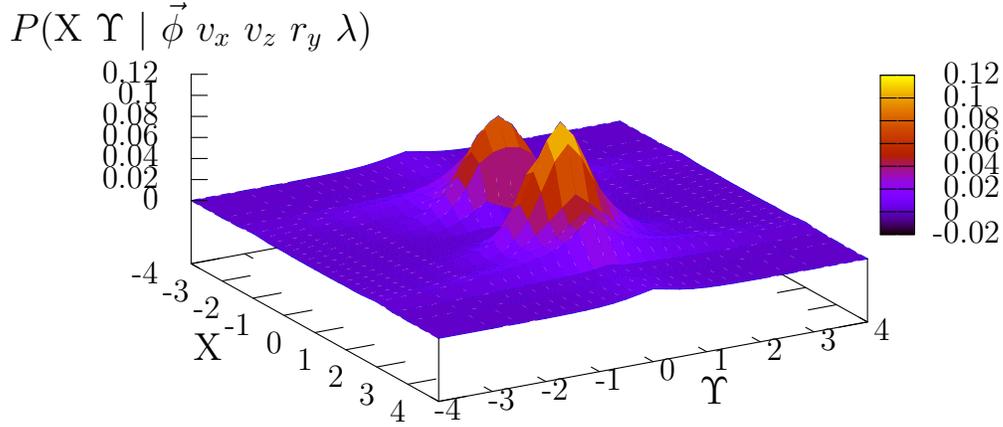
FIG. 4.7 – Répartition de l'erreur d'estimation de l'orientation.

Cet histogramme de réponses présente une forte bimodalité, signe de l'ambiguïté du flux optique.

4.6.2 Obtention de la réponse

Dans notre modèle, nous avons choisi le gradient de profondeur (X, Υ) pour décrire la position spatiale du plan objet perçu. La figure 4.8 présente une réponse de notre modèle, sous la forme d'une distribution dans l'espace (X, Υ).

Il faut à nouveau effectuer un changement de repère pour pouvoir obtenir une distribution sur l'erreur d'orientation comparable avec les résultats expérimentaux. Pour cela,

FIG. 4.8 – Réponse à une question dans l'espace (X, Y) .

nous appliquons le changement de repère inverse à celui nous donnant notre distribution sur la position en procédant de la même manière que pour définir l'*a priori* sur la position en 4.4.3. À partir du changement de repère \mathcal{C}^{-1} de (Σ, T) vers (X, Y) (expression 4.39), on peut écrire l'équation 4.40 pour tout sous-ensemble $A \subset \mathcal{D}_\Sigma \times \mathcal{D}_T$.

$$\mathcal{C}^{-1} : \begin{cases} \chi(\sigma, \tau) = \tan \sigma \cos \tau \\ v(\sigma, \tau) = \tan \sigma \sin \tau \end{cases} \quad (4.39)$$

$$\begin{aligned} \int_A P(\sigma, \tau) d\sigma d\tau &= \int_A P(\chi(\sigma, \tau), v(\sigma, \tau)) |\det J_{\mathcal{C}^{-1}}| d\sigma d\tau \\ &= \int_A P(\chi(\sigma, \tau), v(\sigma, \tau)) \frac{\tan \sigma}{(\cos \sigma)^2} d\sigma d\tau \end{aligned} \quad (4.40)$$

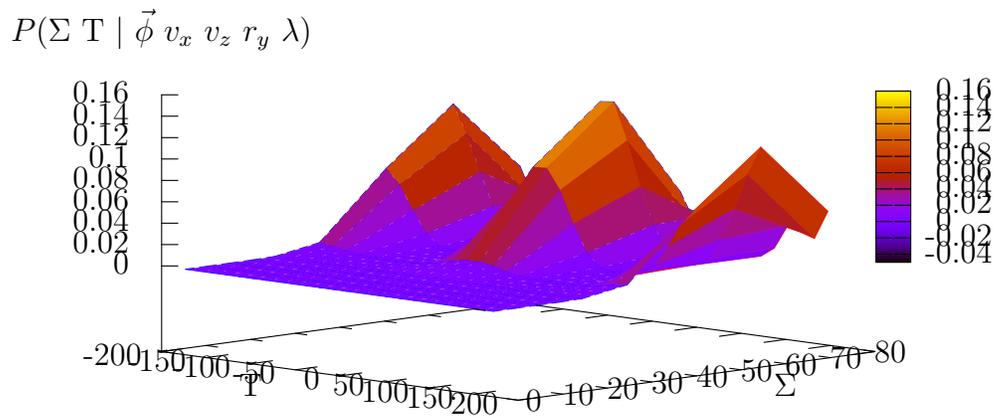
On en déduit l'expression 4.41 de la distribution de probabilité sur l'inclinaison et l'orientation à partir d'une distribution quelconque dans l'espace (X, Y) .

$$\forall (\sigma, \tau) \in \mathcal{D}_\Sigma \times \mathcal{D}_T, P(\sigma, \tau) = \frac{\tan \sigma}{(\cos \sigma)^2} P(\chi(\sigma, \tau), v(\sigma, \tau)) \quad (4.41)$$

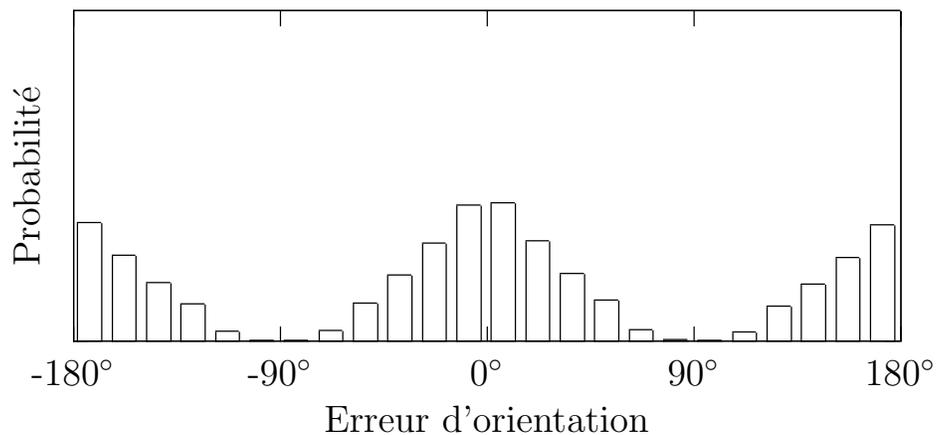
Le graphe 4.9 montre la transposition de la distribution présentée figure 4.8 à l'aide de cette formule.

En utilisant la règle de la marginalisation exposée équation 2.5, on peut déduire l'expression 4.42 de la distribution sur l'orientation, réponse de notre modèle placé dans les conditions de l'expérience.

$$\forall \tau \in \mathcal{D}_T, P(\tau | \vec{\phi}, v_x, v_z, r_y, \lambda) = \sum_{\sigma \in \mathcal{D}_\Sigma} \frac{\tan \sigma}{(\cos \sigma)^2} P(\chi(\sigma, \tau), v(\sigma, \tau) | \vec{\phi}, v_x, v_z, r_y, \lambda) \quad (4.42)$$

FIG. 4.9 – Réponse à une question dans l'espace (Σ, T) .

On peut calculer la distribution sur l'erreur d'estimation et en faire la moyenne d'exac-
ttement la même manière que pour le modèle introductif en 3.4. On obtient finalement la
distribution sur l'erreur d'estimation figurée sur le graphe 4.10.

FIG. 4.10 – Distribution sur l'erreur d'estimation de l'orientation du plan en condition
immobile.

4.6.3 Comparaison

La figure 4.11 compare les résultats de notre modèle avec à la fois les résultats expéri-
mentaux et les résultats du modèle présenté au chapitre 3.

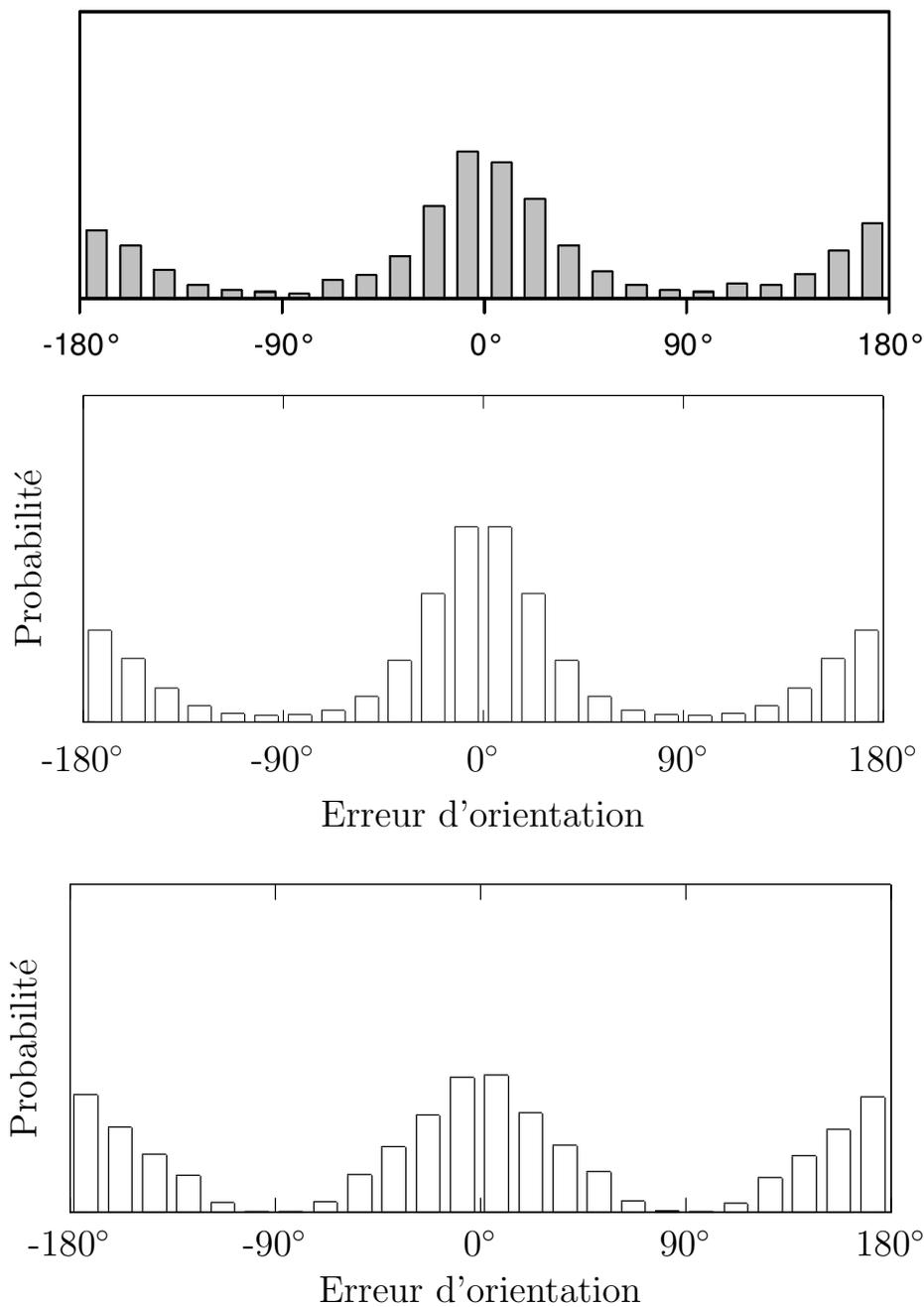


FIG. 4.11 – Comparaison des distributions d’erreur d’estimation de l’orientation. En haut les résultats expérimentaux, au milieu ceux résultats du modèle du chapitre 3 et en bas les résultats du modèle complet.

On remarque la similarité de ces résultats. En particulier, la distribution d’erreur d’estimation donnée par notre modèle présente la bimodalité mettant en valeur l’ambiguïté du stimulus qui peut s’interpréter à la symétrie fronto-parallèle près.

Le modèle présente aussi la différence de probabilité entre les deux pics de cette distribution. Le percept inverse est moins probable dans notre modèle de manière similaire à l'expérience.

On peut noter, tout de même, une moins bonne adéquation aux résultats numériques obtenus en psychophysique (figure 4.11, graphe du haut) pour le modèle complet (figure 4.11 en bas) que pour le modèle simple (figure 4.11 au milieu), présenté au chapitre 3. Dans ce dernier, les paramètres ont été choisis pour ce seul résultat permettant donc sa très bonne reproduction. Par contre, comme indiqué en 4.5.3, tous les résultats présentés dans ce document pour le modèle complet partagent le même jeu de paramètres. Ce jeu de paramètre est un compromis global entre les différents résultats des différentes expériences.

4.7 Résumé

Dans ce chapitre nous avons montré les détails de notre modèle de perception d'un plan par le flux optique. Celui-ci s'appuie sur une liste d'hypothèses issues de la littérature et les implante au sein du cadre théorique de la programmation bayésienne. Ces hypothèses sont les suivantes :

1. rigidité : le flux perçu est plus probablement celui d'un objet rigide ;
2. stationnarité : le mouvement de l'objet est plus probablement petit ;
3. mouvement et position indépendants l'un de l'autre ;
4. précision du flux dépendante de la taille du stimulus.

Dans notre contexte la rigidité est celle d'un plan et son expression se retrouve principalement à deux endroits dans notre modélisation : la décomposition et la spécification des distributions.

Au niveau de la décomposition, la rigidité s'exprime par l'indépendance du flux optique du mouvement propre et la séparation des composantes du flux ce qui nous permet de réécrire l'équation 4.8 :

$$\begin{aligned}
 & P(\vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2 \mid X \Upsilon V_x V_z R_y \vec{T} \vec{\Omega} \Lambda) \\
 = & P(\vec{\Phi}^0 \mid T_x T_y) \\
 \times & P(\vec{\Phi}^1 \mid X \Upsilon \vec{T} \vec{\Omega}) \\
 \times & P(\vec{\Phi}^2 \mid X \Upsilon \vec{T} \vec{\Omega} \Lambda)
 \end{aligned}$$

La spécification de ces trois distributions suit aussi l'hypothèse de rigidité. Nous spécifions ces distributions comme des Gaussiennes centrées autour du flux optique théorique du plan. Nous exprimons bien que le flux optique perçu est plus probablement celui d'un plan.

De la même manière, notre décomposition s'appuie sur l'hypothèse de stationnarité lorsque nous écrivons l'équation 4.9 :

$$\begin{aligned}
 & P(V_x V_z R_y \vec{T} \vec{\Omega}) \\
 = & P(V_x V_z R_y)P(\vec{T} \vec{\Omega} \mid V_x V_z R_y)
 \end{aligned}$$

À ce stade, il ne s'agit encore que d'un choix de décomposition plus que d'une hypothèse simplificatrice.

Ensuite, nous spécifions le facteur $P(\vec{T} \vec{\Omega} | V_x V_z R_y)$ comme une Gaussienne centrée sur l'opposé du mouvement propre. C'est ainsi que nous exprimons que le mouvement du plan est plus probablement petit dans le référentiel allocentré.

Comme pour le chapitre précédent, nous pouvons résumer l'expression bayésienne du modèle par la figure 4.12.

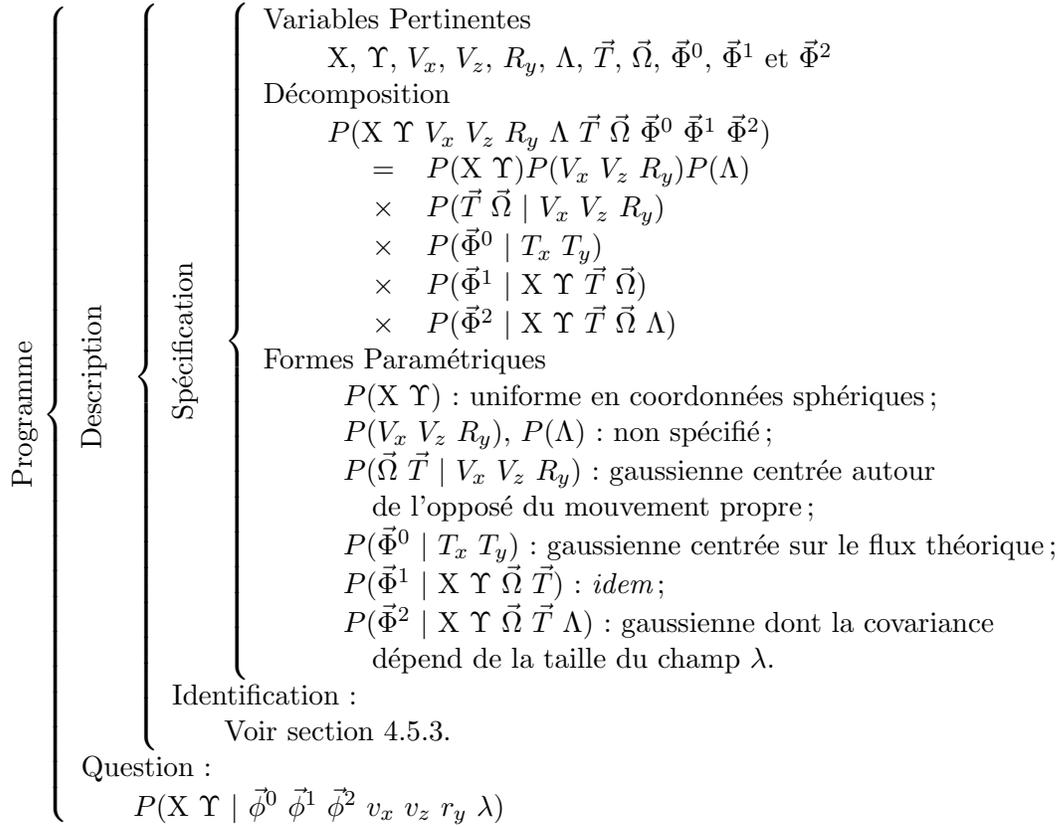


FIG. 4.12 – Résumé du modèle bayésien de perception.

Chapitre 5

Résultats

Dans ce chapitre nous confrontons les résultats de notre modèle, spécifié au chapitre 4, avec ceux de plusieurs expériences de psychophysique issue de la littérature. Nous passons en revue cinq résultats expérimentaux en détaillant à chaque fois l'expérience et ses résultats ainsi que le comportement du modèle et son explication.

5.1 Importance du flux optique

5.1.1 Hypothèse et protocole

Hypothèse testée Nous voulons commencer par tester la perception de l'inclinaison d'un plan par notre modèle en petit champ de vision. Domini et Caudek (1999) ont réalisé une expérience sur la perception de l'inclinaison lors d'une projection orthographique et non plus perspective. E projection orthographique, le flux optique d'ordre deux disparaît. Nous voulons vérifier que la perception de notre modèle en petit champ de vision est similaire à la perception lors d'une projection orthographique.

Protocole expérimental Comme dans l'expérience de perception de l'orientation présentée section 3.1, le stimulus est la projection orthographique d'un plan incliné en rotation. Sa texture est composée de points aléatoires répartis de manière à être distribués uniformément dans un disque après projection à l'écran. Le plan n'est visible qu'à travers une ouverture de 9.6° (petit champ de vision). Les participants ont un œil masqué et restent immobiles pendant la durée de présentation du stimulus, en moyenne 500 ms. Il leur est ensuite demandé d'évaluer l'inclinaison du plan perçu en alignant une sonde en forme d'hémisphère à la position du plan.

Cette expérience teste l'influence sur la perception de l'inclinaison de deux paramètres : l'inclinaison et la vitesse de rotation. L'inclinaison du plan peut donc prendre deux valeurs qui seront notées **petite inclinaison** et **grande inclinaison** qui correspondent respectivement à $\tan \sigma = 1.5$ et $\tan \sigma = 3$ soit des angles d'inclinaison σ d'environ 55° et 70° . Deux vitesses de rotation du plan sont aussi testées : **petite vitesse** et **grande vitesse**,

correspondant respectivement à 0.25 rad.s^{-1} et 0.5 rad.s^{-1} .

Au final, l'expérience porte donc sur quatre combinaisons de conditions expérimentales présentées figure 5.1.

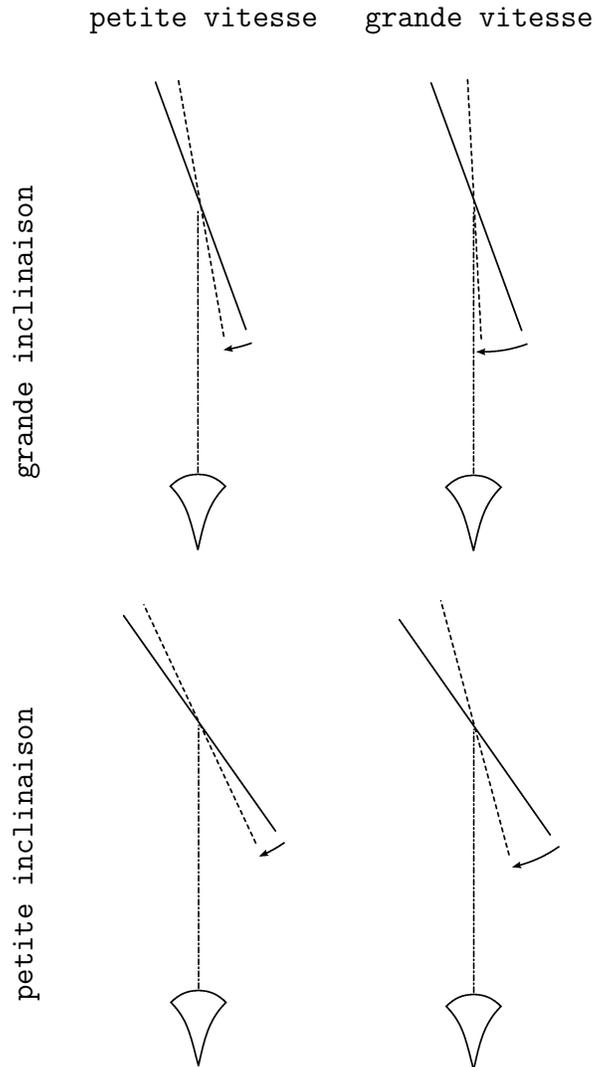


FIG. 5.1 – Résumé des conditions expérimentales de l'expérience 5.1. L'inclinaison est mesurée à partir du plan fronto-parallèle.

5.1.2 Résultats de l'expérience

La figure 5.2, par Domini et Caudek (1999), présente les résultats expérimentaux.

La première remarque est que l'inclinaison perçue est croissante en fonction à la fois de l'inclinaison simulée, ce qui est attendu, mais aussi de la vitesse de rotation à inclinaison

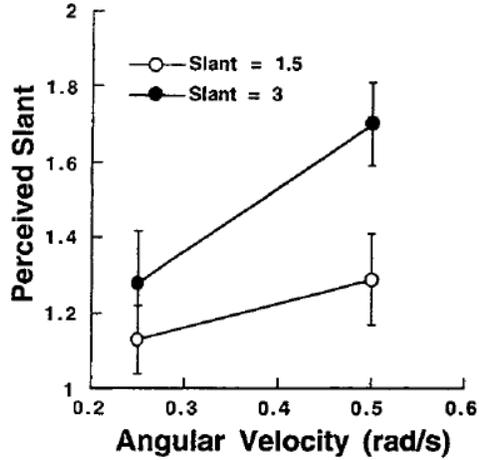


FIG. 5.2 – Valeur estimée de la tangente de l’inclinaison (*perceived slant*) en fonction de la vitesse de rotation pour une **petite inclinaison** en blanc et une **grande inclinaison** en noir.

constante. Autrement dit, pour deux plans inclinés de manière identique, le plan tournant le plus vite sera perçu comme plus incliné.

Ce graphe montre que la perception du plan en **petite inclinaison-grande vitesse** est très proche de celle du plan en **grande inclinaison-petite vitesse**. Cette observation est en accord avec les prévisions des modèles affines qui basent la perception sur le flux optique d’ordre un.

Enfin, en augmentant la tangente de l’inclinaison du plan simulé de 100 %, celle du plan perçu augmente de 23 %. De même, en augmentant la vitesse de rotation de 100 %, la tangente de l’inclinaison augmente elle aussi de 24 %. Ainsi une variation identique de l’inclinaison ou de la vitesse donne une même variation de la perception.

5.1.3 Résultats du modèle

Les résultats de l’expérience sont exprimés sous la forme de la moyenne de la tangente de l’inclinaison. Nous considérons donc la variable Ξ (*xi* majuscule) pour représenter la tangente de l’inclinaison et nous construisons le modèle décrit par la distribution conjointe 5.1.

$$P(\Xi \ X \ \Upsilon \ \vec{\Phi} \ V_x \ V_z \ R_y \ \Lambda) \quad (5.1)$$

$$= P(\vec{\Phi} \ V_x \ V_z \ R_y \ \Lambda) \quad (5.2)$$

$$\times P(X \ \Upsilon \ | \ \vec{\Phi} \ V_x \ V_z \ R_y \ \Lambda) \quad (5.3)$$

$$\times P(\Xi \ | \ X \ \Upsilon) \quad (5.4)$$

Le facteur 5.2 peut rester non spécifié car il n’apparaîtra pas dans l’inférence. En effet, chacune de ces variables sera instanciée lors de la question ; ce facteur sera donc simplifié de la même manière qu’en 4.4.4.

Le facteur 5.3 est la réponse à la question de notre modèle.

Enfin, le facteur 5.4 définit la distribution de probabilité sur la tangente de l'inclinaison en connaissant le gradient de profondeur. Le changement de repère \mathcal{C} défini par l'équation 4.30 donne l'expression $\sigma(\chi, v) = \arctan(\sqrt{\chi^2 + v^2})$. On en déduit que la tangente de l'inclinaison est exactement $\hat{\xi} = \sqrt{\chi^2 + v^2}$. En terme de distribution de probabilité, on utilise donc un Dirac puisqu'il n'y a pas d'incertitude dans le changement de repère. En notant $\delta_h(H)$ le Dirac en h pour la variable H , on écrit :

$$\forall (\chi, v) \in \mathcal{D}_X \times \mathcal{D}_Y, P(\Xi | \chi v) = \delta_{\sqrt{\chi^2 + v^2}}(\Xi) \quad (5.5)$$

L'inférence bayésienne permet de calculer la distribution sur la tangente de l'inclinaison $P(\Xi | \vec{\Phi} V_x V_z R_y \Lambda)$. Nous nous intéressons à sa moyenne $\bar{\Xi}$ donnée par l'expression 5.6, qui nous permet de comparer notre modèle aux résultats expérimentaux.

$$\bar{\Xi} = \int_{\mathcal{D}_X \times \mathcal{D}_Y} \sqrt{\chi^2 + v^2} P(\chi v | \vec{\phi} v_x v_z r_y \lambda) \quad (5.6)$$

La figure 5.3 montre, à droite, les résultats de notre modèle dans les mêmes conditions que l'expérience dont les résultats sont rappelés à gauche.

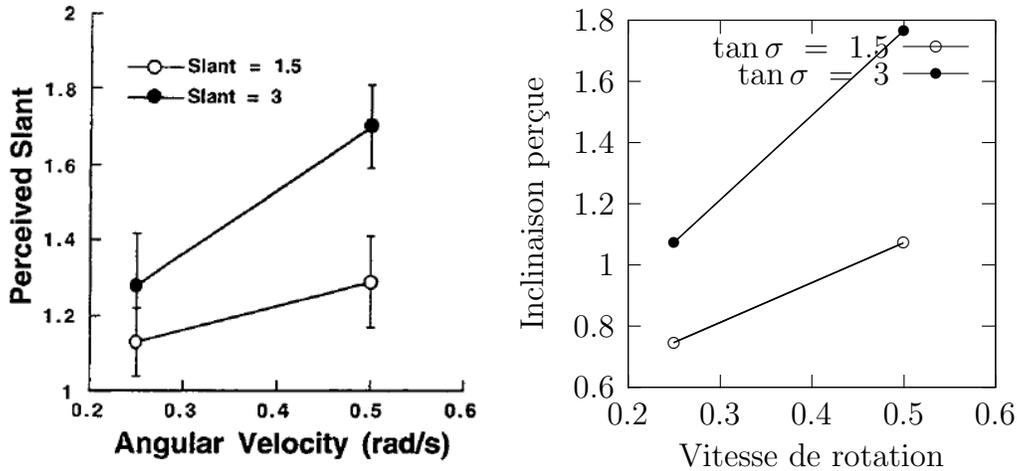


FIG. 5.3 – Comparaison des résultats du modèle à droite avec ceux de l'expérience à gauche. Chaque graphe montre la valeur estimée de la tangente de l'inclinaison en fonction de la vitesse de rotation pour les conditions **petite inclinaison** en blanc et **grande inclinaison** en noir.

L'estimation de Ξ est croissante en fonction à la fois de l'inclinaison simulée et de la vitesse de rotation simulée. En outre, les estimations en **petite inclinaison-grande vitesse** et en **grande inclinaison-petite vitesse** sont, là aussi, très proches l'une de l'autre. Notre modèle, bien que prenant en compte le flux optique d'ordre deux, se

comporte donc comme un modèle affine et réplique les caractéristiques importantes des résultats expérimentaux.

Une augmentation de 100 % de l'inclinaison simulée d'une part, et de la vitesse d'autre part, mènent à une même variation de la moyenne de Ξ . Cette variation est d'environ 60 % pour notre modèle au lieu de 23 % pour les résultats expérimentaux.

Il est à noter que les valeurs de notre modèle sont légèrement inférieures à celles de l'expérience. Nous commentons cette différence dans la section suivante.

5.1.4 Explication

Les hypothèses de notre modèle permettant ce résultat sont principalement la rigidité et l'influence de la taille du champ de vision sur la perception du flux optique d'ordre deux. Bien que notre modèle tienne compte du flux optique complet, l'importance du flux d'ordre deux est fortement réduite en petit champ de vision. Dans cette condition, le flux d'ordre un domine et notre modèle se comporte comme lors d'une perception en projection orthogonale.

Nous examinons les résultats de l'inférence pour justifier la perception non véridique de notre modèle. Tout d'abord, nous nous intéressons aux conditions **petite vitesse-grande inclinaison** et **grande vitesse-petite inclinaison** et à leur perception très proche. Dans un deuxième temps, nous étudions l'évolution de l'estimation de l'inclinaison en fonction de la vitesse.

Perception identique Reprenons les équations du flux optique d'ordre un. Pour simplifier les équations, nous supposons en premier lieu que l'axe de rotation du plan est vertical. Dans ce cas, les composantes suivant l'axe y vertical sont nulles et l'équation 4.24 devient :

$$\begin{cases} x\phi_x^1 = \chi.\omega_y \\ y\phi_x^1 = v.\omega_y \end{cases} \quad (5.7)$$

qui peut se réécrire :

$$\begin{cases} x\phi_x^1 = \xi.\omega_y.\cos\tau \\ y\phi_x^1 = \xi.\omega_y.\sin\tau \end{cases} \quad (5.8)$$

avec toujours $\xi = \tan\sigma$, la tangente de l'inclinaison du plan, τ l'orientation du plan et ω_y sa vitesse de rotation suivant l'axe vertical.

L'équation 5.8 montre que l'inclinaison et la vitesse de rotation n'interviennent dans le flux d'ordre un qu'à travers leur produit $\xi.\omega_y$.¹ On obtient, pour chacune des conditions expérimentales, les flux optiques montrés par la figure 5.4. On vérifie en particulier que le stimulus en **grande vitesse-petite inclinaison** est le même à l'ordre un qu'en **petite vitesse-grande inclinaison**.

Or, la valeur des écarts-type sur les composantes du flux optique d'ordre deux sont très importantes (voir section 4.5.3). Ce choix est justifié par notre hypothèse stipulant que

¹Le produit $\xi.\omega_y$ est parfois nommé *def* dans la littérature (Domini et Caudek 1999, 2003).

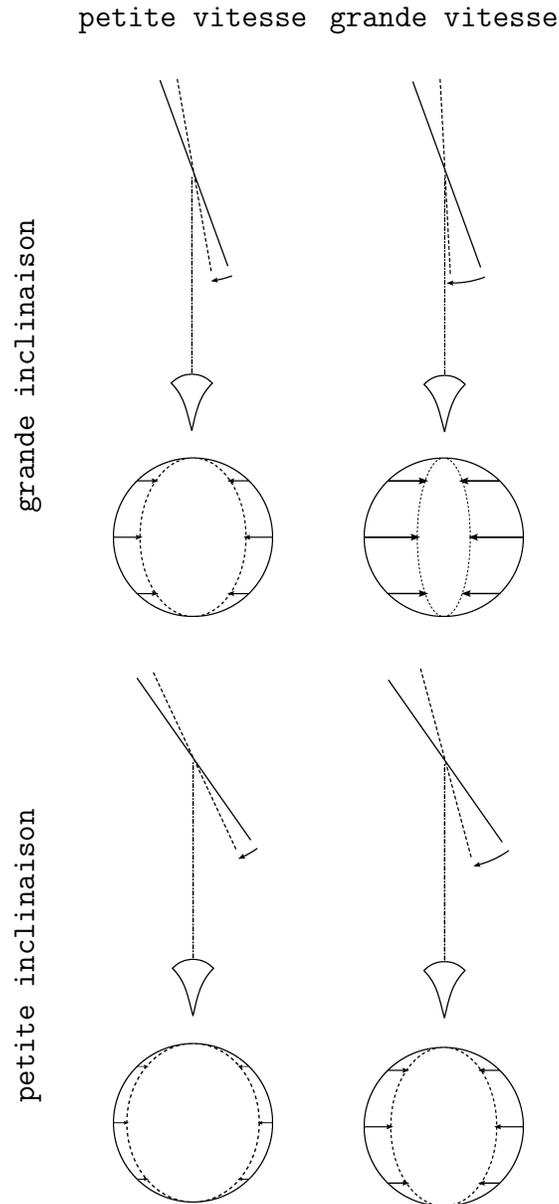


FIG. 5.4 – Flux optiques associés aux conditions de l'expérience 5.1.

la précision du flux optique d'ordre deux diminue avec la taille du champ de vision (voir section 5.2). La perception identique par le modèle de ces deux conditions expérimentales est due à la similarité du flux optique, étant donné que l'influence du flux d'ordre deux est rendue négligeable.

Notre modèle prévoit l'augmentation de la différence de perception entre les conditions *petite vitesse-grande inclinaison* et *grande vitesse-petite inclinaison* avec la taille du champ de vision. Cette expérience n'a pas été réalisée à notre connaissance, il

serait intéressant de le faire pour vérifier cette prévision.

Influence de la vitesse Pour comprendre l'évolution de la moyenne de Ξ en fonction d'un changement d'inclinaison ou de vitesse, nous nous intéressons à la distribution sur Ξ et le module Ω de la vitesse de rotation du plan.

Pour cela nous posons à notre modèle la question $P(X \Upsilon \vec{\Omega} \mid \vec{\varphi} v_x v_z r_y \lambda)$ et nous effectuons le changement de repère 5.9, défini par l'ensemble d'équations 5.10.

$$\begin{pmatrix} X \\ \Upsilon \\ \Omega_x \\ \Omega_y \\ \Omega_z \end{pmatrix} \rightarrow \begin{pmatrix} \Xi \\ T \\ \Omega \\ \Theta \\ \Psi \end{pmatrix} \quad (5.9)$$

$$\begin{cases} \chi &= \xi \cos \tau \\ v &= \xi \sin \tau \\ \omega_x &= \omega \cos \theta \cos \psi \\ \omega_y &= \omega \sin \theta \cos \psi \\ \omega_z &= \omega \sin \psi \end{cases} \quad (5.10)$$

Comme pour les changements de variable précédents, le calcul s'écrit alors :

$$\begin{aligned} & P(\xi \tau \omega \theta \psi \mid \vec{\varphi} v_x v_z r_y \lambda) \\ &= \xi \omega^2 \cos \psi \\ & \times P(\chi(\xi, \tau) v(\xi, \tau) \omega_x(\omega, \theta, \psi) \omega_y(\omega, \theta, \psi) \omega_z(\omega, \theta, \psi) \mid \vec{\varphi} v_x v_z r_y \lambda) \end{aligned} \quad (5.11)$$

En marginalisant 5.11 sur les variables T , Θ et Ψ , on obtient la distribution 5.12 sur les variables Ξ et Ω .

$$\begin{aligned} & P(\xi \omega \mid \vec{\varphi} v_x v_z r_y \lambda) \\ &= \xi \omega^2 \\ & \times \sum_{\tau, \theta, \psi} \cos \psi P(\chi(\xi, \tau) v(\xi, \tau) \omega_x(\omega, \theta, \psi) \omega_y(\omega, \theta, \psi) \omega_z(\omega, \theta, \psi) \mid \vec{\varphi} v_x v_z r_y \lambda) \end{aligned} \quad (5.12)$$

La figure 5.5 présente la distribution sur la tangente de l'inclinaison et le module de la vitesse de rotation en **petite inclinaison** et **grande vitesse**.

D'après la règle de Bayes 2.4, cette distribution est le produit de l'*a priori* de notre modèle pour ce mouvement propre $P(\Xi \Omega \mid v_x v_z r_y)$ et de la vraisemblance du flux $P(\vec{\varphi} \mid \Xi \Omega v_x v_z r_y \lambda)$. Ces distributions sont montrées respectivement figures 5.6 et 5.7.

Dans la figure 5.6, on peut voir la stationnarité qui donne une probabilité plus importante aux vitesses faibles, indépendamment du flux observé. La figure 5.7 montre un ensemble de couples (ξ, ω) vraisemblables pour le flux observé regroupés sur une hyperbole. Cette hyperbole a pour équation $\xi \omega = \left\| \vec{\Phi}^1 \right\|_2$, expression de la rigidité.

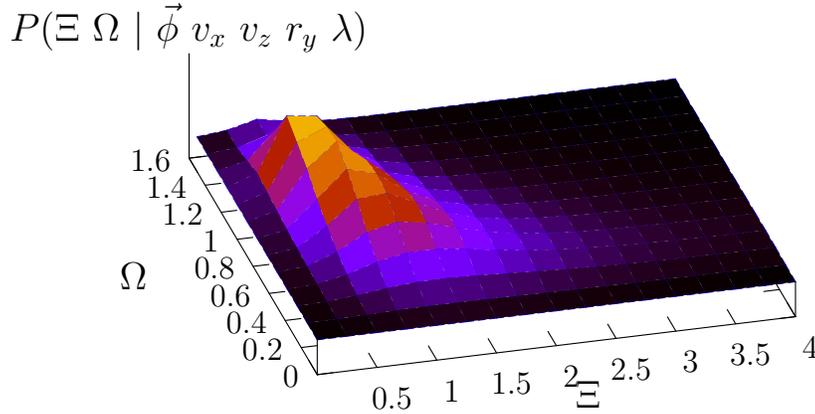


FIG. 5.5 – Distribution sur (Ξ, Ω) en condition **petite inclinaison-grande vitesse**.

L'*a priori* sur (Ξ, Ω) ne change pas en fonction de l'observation du flux, seule la vraisemblance évolue. la figure 5.8 montre la vraisemblance d'un flux en **grande inclinaison** et **grande vitesse**. L'hyperbole des couples d'inclinaison et vitesse est plus éloignée de l'origine, comme prévu par le flux optique. La distribution *a posteriori* correspondante est montrée figure 5.9.

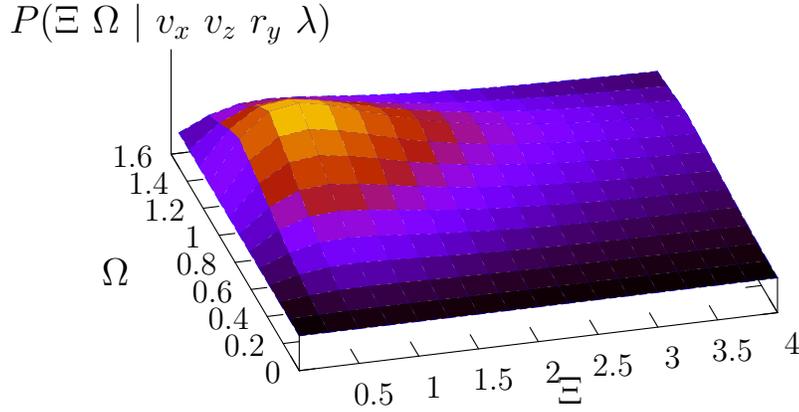
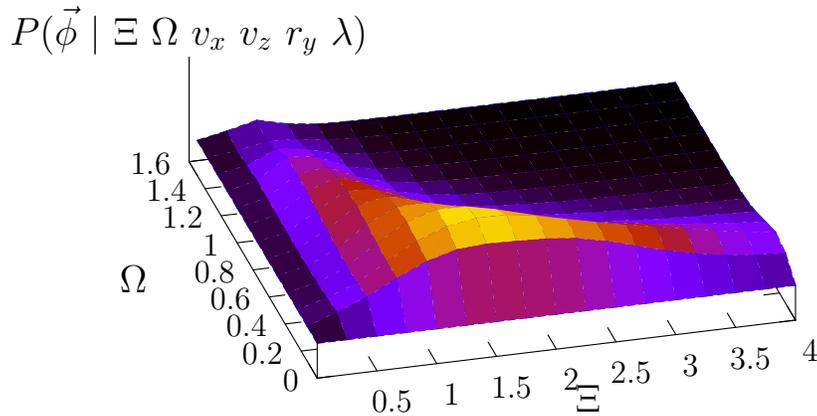
Le pic de distribution s'est déplacé sur l'hyperbole en suivant approximativement la ligne de moins grande pente de l'*a priori*. Cet exemple montre comment évolue l'estimation de l'inclinaison lors d'un changement de l'inclinaison du plan simulé ou de sa vitesse de rotation.

Plus généralement, l'inclinaison perçue résulte d'un compromis entre l'*a priori* sur l'orientation et celui sur la vitesse de rotation. La vitesse de rotation est contrainte par la stationnarité. Il est important de noter que l'inclinaison n'est pas réellement retrouvée et que c'est son évolution en fonction à la fois de l'inclinaison et de la vitesse de rotation simulées qui est reproduite par le modèle. En renforçant la stationnarité dans le modèle², l'adéquation quantitative entre ses résultats et ceux de l'expérience pourrait être meilleure. Rappelons que les paramètres actuels sont le résultat d'un compromis global entre toutes les expériences.

De la même manière, pour les observateurs naïfs³, les effets d'un doublement de la tangente de l'inclinaison puis de la vitesse de rotation sont les mêmes, à savoir une aug-

²Ça peut être fait en diminuant les valeurs des paramètres σ_Ω et σ_T de la matrice de covariance du mouvement relatif sachant le mouvement propre.

³Les résultats expérimentaux par Domini et Caudek (1999, figure 17) présentés figure 5.2 ne concernent qu'une partie des participants. Les résultats des observateurs expérimentés sont très différents (Domini et Caudek 1999, figures 15 et 16).

FIG. 5.6 – Distribution *a priori* sur (Ξ, Ω) (avant d’observer un flux optique).FIG. 5.7 – Vraisemblance de $\vec{\phi}$ sachant (Ξ, Ω) et (v_x, v_z, r_y, λ) en condition **petite inclinaison-grande vitesse**.

mentation d’environ 23 % de l’inclinaison estimée.

Dans notre modèle, le pic de distribution se déplace en suivant la ligne de moins grande pente de l’*a priori*. La croissance identique de l’estimation de la moyenne de l’inclinaison en fonction de la vitesse ou l’inclinaison simulées est alors une propriété locale et contingente de cette distribution *a priori*. Elle est ainsi plus importante dans notre modèle que dans

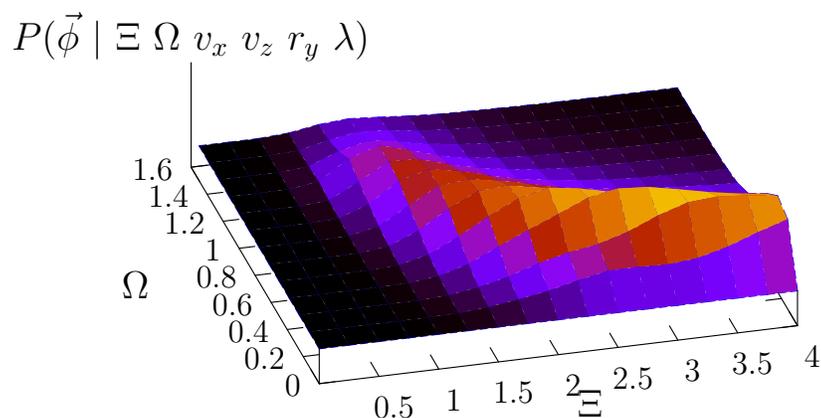


FIG. 5.8 – Vraisemblance de $\vec{\phi}$ sachant (Ξ, Ω) et (v_x, v_z, r_y, λ) en condition grande inclinaison-grande vitesse.

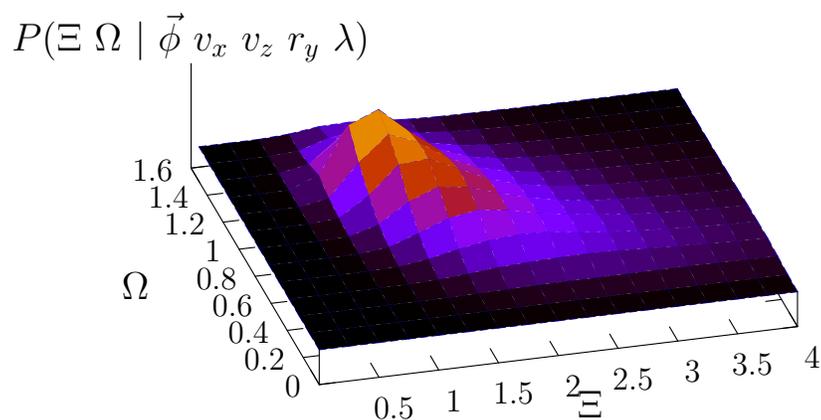


FIG. 5.9 – Distribution sur (Ξ, Ω) en condition grande inclinaison-grande vitesse.

les résultats expérimentaux car l'*a priori* de notre modèle est localement orienté suivant l'inclinaison. Pour s'intéresser plus avant à cette différence, il serait nécessaire de tester si elle s'explique par la variation inter-individus ou non.

Les résultats psychophysiques montrent que la perception humaine ne recouvre pas parfaitement l'inclinaison d'un plan lors d'une projection orthographique. Notre modèle,

bien que prenant en compte le flux optique d'ordre deux, est en accord avec ces résultats. Il explique ce comportement principalement par la rigidité et l'hypothèse que la perception du flux optique est influencée par la taille du champ de vision.

5.2 Influence de la taille du champ de vision

5.2.1 Hypothèse et protocole

Hypothèse testée Une des hypothèses sur lesquelles est bâti notre modèle est que la perception du flux optique est influencée par la taille du champ de vision (hypothèse 4). Nous nous proposons d'en vérifier la pertinence en nous appuyant sur les résultats expérimentaux de Cornilleau-Pérès *et al.* (2002). Le principe de cette expérience est de comparer la perception d'un même flux optique dans des conditions différentes de taille de champ de vision.

Protocole expérimental Plusieurs conditions expérimentales (portant en particulier sur la texture) sont testées dans cet article mais nous n'allons nous intéresser qu'à celles applicables à notre modèle. La condition de base, appelée **petit champ**, est définie avec un observateur immobile à qui est présenté un objet en mouvement dans un petit champ de vision (8°). En ce en quoi notre modèle est concerné, cette expérience est identique à celle de référence présentée en 3.1 et reprise en 4.6.

La seconde condition, **grand champ**, ne diffère que par la taille du champ de vision qui passe alors à 60° . Pour notre modèle, la différence se traduit par une valeur pour la variable Λ observée passant de 0 (**petit champ**) à 1 (**grand champ**).

5.2.2 Résultats de l'expérience

Les résultats expérimentaux sont exprimés en pourcentage d'inversion. Une inversion est définie comme une orientation écarté de celle simulée de plus de 90° . Le pourcentage d'inversion est calculé sur un ensemble d'observations à différentes orientations pour plusieurs personnes. Un pourcentage d'inversion de 0 signifierait donc que les participants ne se trompent jamais sur le sens du plan. Un pourcentage supérieur à 50 indiquerait une préférence pour ne pas percevoir le plan simulé. Dans le cas où un observateur est parfaitement incapable de discriminer le plan simulé du plan inverse, son taux d'inversion est de 50%.

La figure 5.10 présente les taux d'inversion pour chacune des deux conditions **petit champ** (à gauche) et **grand champ** (à droite). En petit champ de vision, les conditions sont les mêmes que pour la première expérience discutée (voir section 4.6). La distribution de probabilité sur l'erreur d'estimation de l'orientation γ montre deux pics de probabilité opposés. Le rapport des masses de probabilité de ces deux parties est le taux d'inversion. On n'est donc pas surpris de constater que le taux d'inversion en **petit champ** est proche

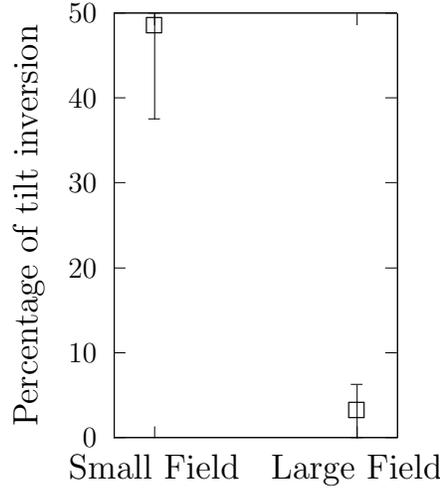


FIG. 5.10 – Comparaison des taux d’inversion pour les conditions **petit champ** (à gauche) et **grand champ** (à droite).

de 50%, légèrement inférieur puisque le pic de probabilité centré en 0° est plus important que celui centré sur 180° .

Le taux d’inversion en condition **grand champ** est fortement réduit. Un grand champ de vision semble donc permettre à un observateur de mieux discriminer les plans.

5.2.3 Résultats du modèle

Pour pouvoir comparer les résultats de notre modèle avec ceux de l’expérience, nous devons calculer le taux d’inversion. Pour cela nous utilisons la définition ci-dessus, c’est-à-dire une erreur d’estimation de l’orientation supérieure à 90° .

Nous introduisons la variable booléenne I (*iota* majuscule) indiquant une inversion. Nous écrivons la décomposition suivante en faisant l’hypothèse d’indépendance conditionnelle que l’inversion ne dépend que de l’orientation :

$$P(I \Sigma T \vec{\Phi} V_x V_z R_y \Lambda) \quad (5.13)$$

$$= P(\vec{\Phi} V_x V_z R_y \Lambda) \quad (5.14)$$

$$\times P(\Sigma T | \vec{\Phi} V_x V_z R_y \Lambda) \quad (5.15)$$

$$\times P(I | T) \quad (5.16)$$

Le facteur 5.14 peut rester non spécifié car il n’apparaît pas dans l’inférence. Le facteur 5.15 se calcule à partir de la réponse de notre modèle grâce à la formule 4.41. Le facteur 5.16 définit la probabilité de percevoir une inversion en fonction de l’orientation perçue. Nous spécifions cette probabilité nulle lorsque l’orientation est à moins de 90° de l’orientation simulée $\hat{\tau}$ et égale à 1 sinon.

La question que l'on pose est $P(\iota | \vec{\phi} v_x v_z r_y \lambda)$. Les règles de Bayes et de marginalisation donnent :

$$P(\iota | \vec{\phi} v_x v_z r_y \lambda) = \sum_{\tau \in \mathcal{D}_T} P(\iota | \tau) P(\tau | \vec{\phi} v_x v_z r_y \lambda) \quad (5.17)$$

Le facteur $P(\tau | \vec{\phi} v_x v_z r_y \lambda)$ peut être calculé à l'aide du changement de variables \mathcal{C}^{-1} par la formule 4.42 (section 4.6.2) :

$$\forall \tau \in \mathcal{D}_T, P(\tau | \vec{\phi} v_x v_z r_y \lambda) = \sum_{\sigma \in \mathcal{D}_\Sigma} \frac{\tan \sigma}{(\cos \sigma)^2} P(\chi(\sigma, \tau) v(\sigma, \tau) | \vec{\phi} v_x v_z r_y \lambda) \quad (5.18)$$

D'après la définition du facteur 5.16, il suffit de considérer le sous-domaine de T correspondant au plan inverse $\mathcal{D}_T^\iota = \mathcal{D}_T \setminus [\hat{\tau} - \frac{\pi}{2}, \hat{\tau} + \frac{\pi}{2}]$.

En résumé, la probabilité d'inversion se calcule par la formule 5.19.

$$P(\iota | \vec{\phi} v_x v_z r_y \lambda) = \int_{\tau \in \mathcal{D}_T^\iota} \int_{\sigma \in \mathcal{D}_\Sigma} \frac{\tan \sigma}{(\cos \sigma)^2} P(\chi(\sigma, \tau) v(\sigma, \tau) | \vec{\phi} v_x v_z r_y \lambda) d\sigma d\tau \quad (5.19)$$

Le graphe 5.11 rappelle les résultats de l'expérience et présente ceux de notre modèle dans les deux conditions **petit champ** et **grand champ**.

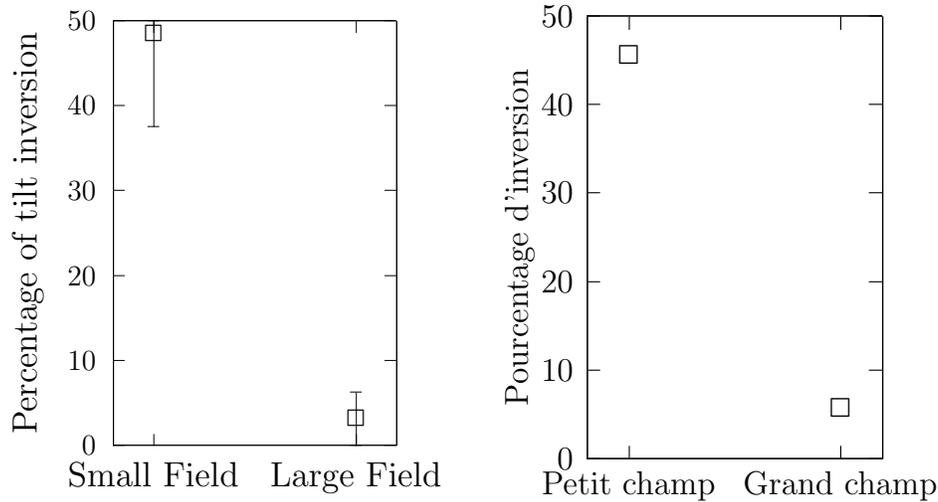


FIG. 5.11 – Résultats du modèle (à droite) en comparaison avec ceux de l'expérience (à gauche). On note la présence de barres d'erreur sur les résultats expérimentaux. Ces barres regroupent à la fois la variation de réponse d'un participant et la variation entre ceux-ci et ne sont donc pas présentes dans les résultats du modèle.

Notre modèle réplique l'influence de la taille du champ sur le taux d'inversion. En condition **petit champ**, ce taux est légèrement inférieur à 0.5 (ou 50 %) conformément aux résultats expérimentaux. Comme rappelé plus haut, cette condition est similaire à l'expérience de la section 4.6. Les résultats de cette expérience sont la distribution sur

l'orientation dont on peut déduire le taux d'inversion. Ainsi donc l'adéquation des résultats de la section 4.6 implique l'adéquation des taux d'inversion en condition **petit champ**.

Pour un grand champ de vision, ce taux est inférieur à 10 %, conformément aux résultats expérimentaux.

5.2.4 Explication

L'influence de la taille du champ de vision sur la perception de formes par le mouvement est entièrement expliquée dans notre modèle par le flux optique d'ordre deux. Dans cette expérience, cette influence est testée par le taux d'inversion. Ce taux ne dépend que du flux d'ordre deux, seule différence entre un plan et son inverse. En effet, mathématiquement, la variable Λ n'apparaît que dans le facteur 4.28 de la décomposition de la conjointe (hypothèse 4) :

$$P(\vec{\Phi}^2 \mid \chi \ v \ \vec{t} \ \vec{\omega} \ \lambda) = \mathcal{N}_{\mathcal{F}^2(\chi, v, \vec{t}, \vec{\omega}), \sigma_{\vec{\Phi}^2}(\lambda)}(\vec{\Phi}^2) \quad (5.20)$$

Seule la matrice de covariance $\sigma_{\vec{\Phi}^2}$ dépend de la taille du champ. Nous avons choisi une covariance diagonale et proportionnelle à la matrice identité (section 4.5.3) :

$$\sigma_{\vec{\Phi}^2}(\lambda) = \sigma_2(\lambda) \cdot \text{Id}_{2 \times 2} \quad (5.21)$$

En résumé, la taille du champ de vision apparaît dans l'expression de la covariance du flux optique d'ordre deux qui influe sur le taux d'inversion. Dans un premier temps, nous étudions l'évolution du taux d'inversion en fonction de la valeur de la covariance. Dans la suite nous nous intéressons à l'évolution de cette covariance par rapport à la taille du champ de vision.

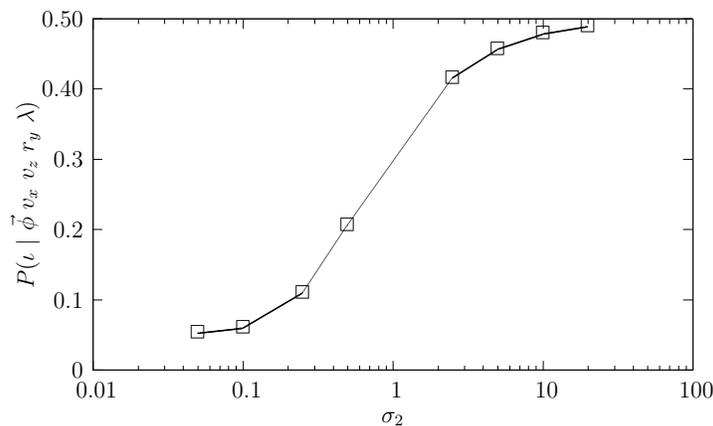


FIG. 5.12 – Évolution du taux d'inversion en fonction de la covariance sur la distribution sur flux optique d'ordre deux (échelle semi-logarithmique).

Le graphe 5.12 montre l'évolution du taux d'inversion en fonction de la valeur σ_2 variant sur une échelle logarithmique. Une valeur très grande pour σ_2 correspond à une faible

sensibilité au flux optique d'ordre deux et, par conséquent, une faible discrimination entre un plan et son inverse qui se traduit par un taux d'inversion proche de 50%. À l'opposé, une faible valeur de σ_2 implique une grande discrimination sur le flux d'ordre deux et donc un faible taux d'inversion. Le taux d'inversion est donc croissant avec σ_2 .

D'un autre côté, le flux optique est décrit par l'équation suivante (voir section 2.2.1) :

$$\vec{\Phi}(x, y) = \vec{\Phi}^0 + \vec{\Phi}^1 \cdot \iota(x, y) + \iota(x, y) \cdot \vec{\Phi}^2 \cdot \iota(x, y) \quad (5.22)$$

L'influence du terme d'ordre deux est quadratique avec la distance au point de fixation alors qu'elle est linéaire pour le terme d'ordre un. Relativement au flux d'ordre un, l'influence du flux d'ordre deux est donc croissante. En **grand champ**, de nombreux points se trouvent éloignés du centre et permettent de mieux percevoir les composantes d'ordre deux du flux optique. L'incertitude σ_2 sur la perception du flux d'ordre deux est donc plus faible en **grand champ** qu'en **petit champ**.

En résumé, en petit champ de vision on a une incertitude importante sur le flux d'ordre deux ce qui implique un taux d'inversion élevé. En grand champ de vision, la perception du flux d'ordre deux est plus précise et donc le taux d'inversion est plus faible.

En conclusion, les résultats de psychophysique montrent que la taille du champ de vision influence la perception de la forme par le mouvement. Le modèle est en accord avec ces résultats. Il explique ce comportement principalement par l'influence du champ de vision sur la covariance du flux optique d'ordre deux (hypothèse 4).

De plus, cet exemple montre comment peut se faire l'identification des paramètres de notre modèle. En effet, les valeurs de la fonction $\lambda \mapsto \sigma_2$ (données en 4.5.3) ont été déterminées à l'aide du graphe 5.12 et des données psychophysiques et sont restées constantes pour toutes les autres expériences. Plus précisément, les résultats de l'expérience présentés figure 5.10 donnent des valeurs que l'on peut reporter sur l'ordonnée du graphe 5.12. Les valeurs en abscisse qui y correspondent sont les paramètres choisis pour les covariances respectives en petit et grand champ de vision.

5.3 Influence du mouvement propre

5.3.1 Hypothèse et protocole

Hypothèse testée L'hypothèse étudiée est que la perception de formes par le mouvement ne s'appuie pas uniquement sur le flux optique mais peut aussi dépendre du mouvement de l'observateur.

Pour le quantifier, van Boxtel *et al.* (2003) ont réalisé une expérience reposant sur la comparaison de la perception d'un même flux optique dans des conditions différentes de mouvement propre.

Protocole expérimental L'appareillage est le même que pour l'expérience présentée en 3.1. Les mouvements des participants sont enregistrés à l'aide d'un *tracker* infrarouge.

Il y a deux conditions expérimentales :

- **actif**, où un observateur en translation suivant l'axe transversal perçoit un plan immobile (dans le référentiel allocentré) ;
- **immobile**, dans laquelle un observateur immobile perçoit un plan en rotation.

Plus précisément, un participant passe tout d'abord l'expérience en condition **actif** pendant laquelle le flux optique, généré par son déplacement, est enregistré. Afin que l'observateur soit exposé au même flux optique, cet enregistrement est rejoué en condition **immobile**. La figure 5.13 résume ces deux conditions expérimentales et leur flux optique associé.

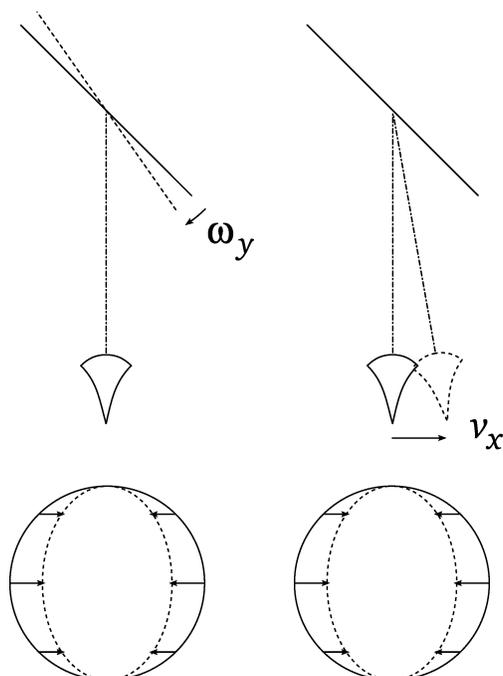


FIG. 5.13 – Illustration comparée des conditions expérimentales : **immobile** à gauche et **actif** à droite. Pour chacune des deux conditions, le flux optique est identique et montré en dessous du mouvement.

Ainsi, il est possible de comparer la perception d'un même flux optique dans des conditions différentes de mouvement propre et d'en étudier l'influence dans la perception des formes par le mouvement.

5.3.2 Résultats de l'expérience

Le graphe 5.14 compare les distributions expérimentales d'erreur d'estimation de l'orientation du plan. Les résultats en condition **immobile** sont les mêmes que pour l'expérience 3.1. On retrouve donc l'inversion figurée par la bimodalité de la distribution d'erreur.

Les résultats en condition **actif** ne présentent pas cette bimodalité. On en conclut que, pour un même stimulus, la forme extraite à partir du flux optique dépend du mouvement propre.

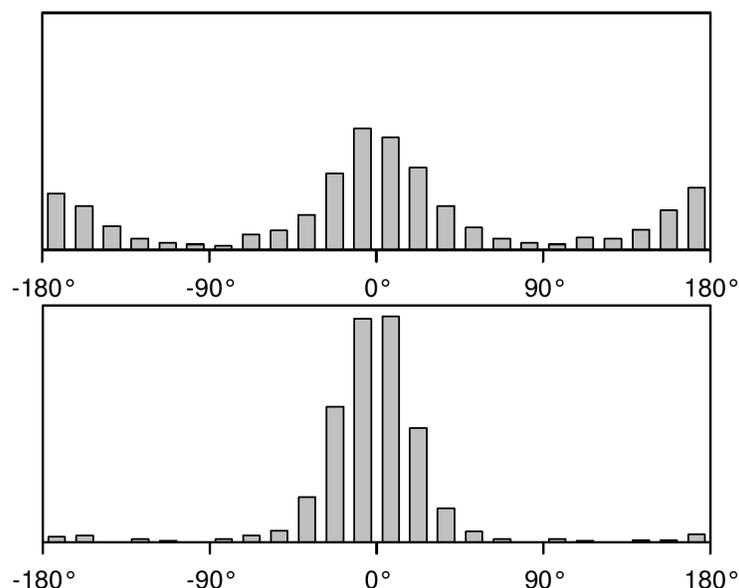


FIG. 5.14 – Distributions d’erreur sur l’estimation de l’orientation pour les conditions immobile (en haut) et actif (en bas), par van Boxtel *et al.* (2003).

5.3.3 Résultats du modèle

En comparaison, la figure 5.15 présente les résultats de notre modèle. Les calculs nécessaires pour obtenir les probabilités sur l’erreur d’orientation sont ceux déjà exposés section 4.6.2.

Comme pour les résultats expérimentaux présentés figure 5.14, la probabilité du percept inverse en condition actif est négligeable.

5.3.4 Explication

L’hypothèse de stationnarité, appliquée aux mouvements des plans inverses, explique ces résultats. En condition immobile, le mouvement absolu du plan inverse est le même que celui du plan simulé. En condition actif, le plan simulé ne bouge pas dans le référentiel allocentré alors que le plan inverse bouge.

La figure 5.16 présente graphiquement ces situations. Elle montre qu’en condition immobile, les vitesses des deux plans, direct et inverse, sont opposées. *A contrario*, en condition actif, le plan inverse possède une grande vitesse alors que le plan direct ne bouge pas.

La différence de probabilité entre le percept direct et inverse est donc due à deux facteurs. Le premier, détaillé section 5.2, est le flux optique d’ordre deux. Dans les deux cas, le percept inverse est moins rigide que le plan simulé, donc sa probabilité est plus faible. On peut le voir, par exemple, en condition immobile car le pic central est plus important que les pics latéraux.

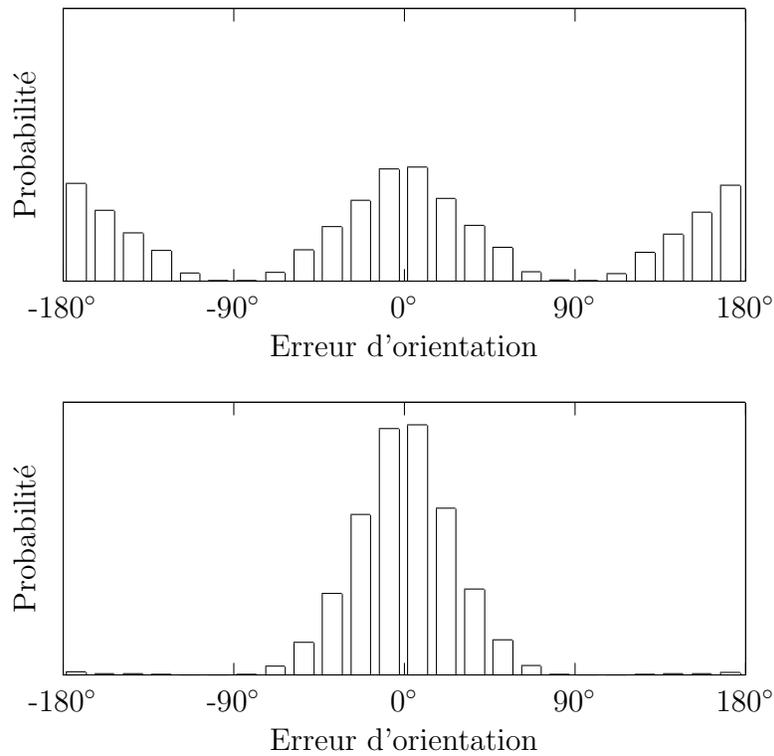


FIG. 5.15 – Distribution de l'erreur sur l'orientation pour le modèle en condition **immobile** en haut et **actif** (en bas).

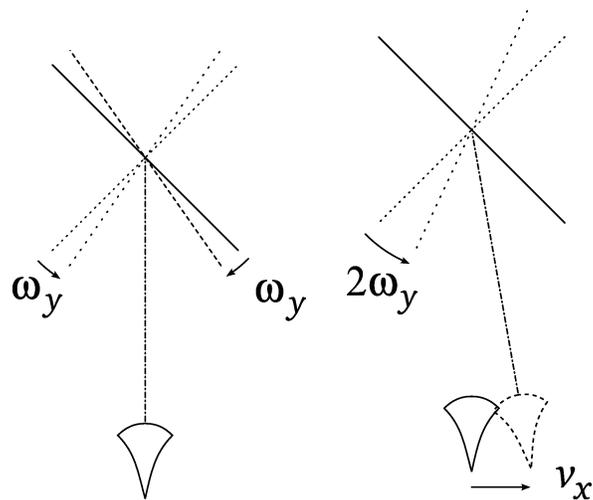


FIG. 5.16 – Illustration comparée des percepts inverses : condition **immobile** à gauche et **actif** à droite. Le plan simulé est figuré en trait plein alors que le plan inverse l'est en pointillés.

Le second facteur est la stationnarité qui s'exprime en condition **actif** par une probabilité plus faible du percept inverse. En condition **actif**, il y a donc concomitance de la stationnarité et de la rigidité. On peut le voir par la quasi disparition des pics latéraux.

Les résultats psychophysiques montrent que le mouvement propre influence la perception de la forme par le mouvement. Le modèle est en accord avec ces résultats. Il explique ce comportement principalement par l'hypothèse de stationnarité (hypothèse 2).

Cependant, la rigidité joue un rôle. Se pose donc la question de son importance relativement à la stationnarité. Cette question pourrait être adressée par notre modèle en modifiant les conditions de cette expérience. En condition **actif**, on peut ainsi simuler un plan en rotation tel que son plan inverse est immobile dans le référentiel allocentré. Dans ce cas, le plan simulé reste plus rigide mais devient moins stationnaire. Les deux hypothèses de rigidité et de stationnarité sont donc en conflit. Malheureusement, cette expérience précise n'a pas été décrite, à notre connaissance, dans la littérature. Nous avons donc choisi de traiter cette question en nous appuyant sur l'expérience de Wexler *et al.* (2001a) étudiée section suivante.

5.4 Influence relative de la stationnarité et de la rigidité

5.4.1 Hypothèse et protocole

Hypothèse testée Cette expérience s'intéresse à vérifier qu'un observateur perçoit parfois un objet plus stationnaire que l'objet simulé rigide (voir section 2.2.2). Pour cela, il faut utiliser un stimulus pour lequel les deux hypothèses de stationnarité et de rigidité induisent des percepts différents. Wexler *et al.* (2001a) ont réalisé cette expérience dans laquelle le flux optique d'un plan simulé en grand mouvement peut s'interpréter comme un plan différent soumis à un mouvement plus faible.

Protocole expérimental En pratique, le plan est observé en petit champ de vision par un observateur étant soit immobile (condition **immobile**) soit en translation sagittale (condition **actif**). Indépendamment de ces deux conditions, le plan est en rotation par rapport à son intersection avec le plan fronto-parallèle. Il est en outre soit en translation relative par rapport à l'observateur (condition **A**) soit à distance fixe (condition **B**). En condition **A**, la vitesse de rotation est fixée par la relation 5.23 avec toujours ω pour le module de la vitesse de rotation, ξ pour la tangente de l'inclinaison et t_z la vitesse de translation sagittale relative entre le plan objet et l'observateur.

$$\omega = -\frac{t_z}{\xi} \quad (5.23)$$

En condition **B**, la vitesse de rotation est reprise de la condition **A**. Dans les deux cas, le flux en condition **actif** est enregistré. Il est ensuite rejoué en condition **immobile** de

manière à comparer la perception du même flux optique dans des conditions différentes de mouvement.

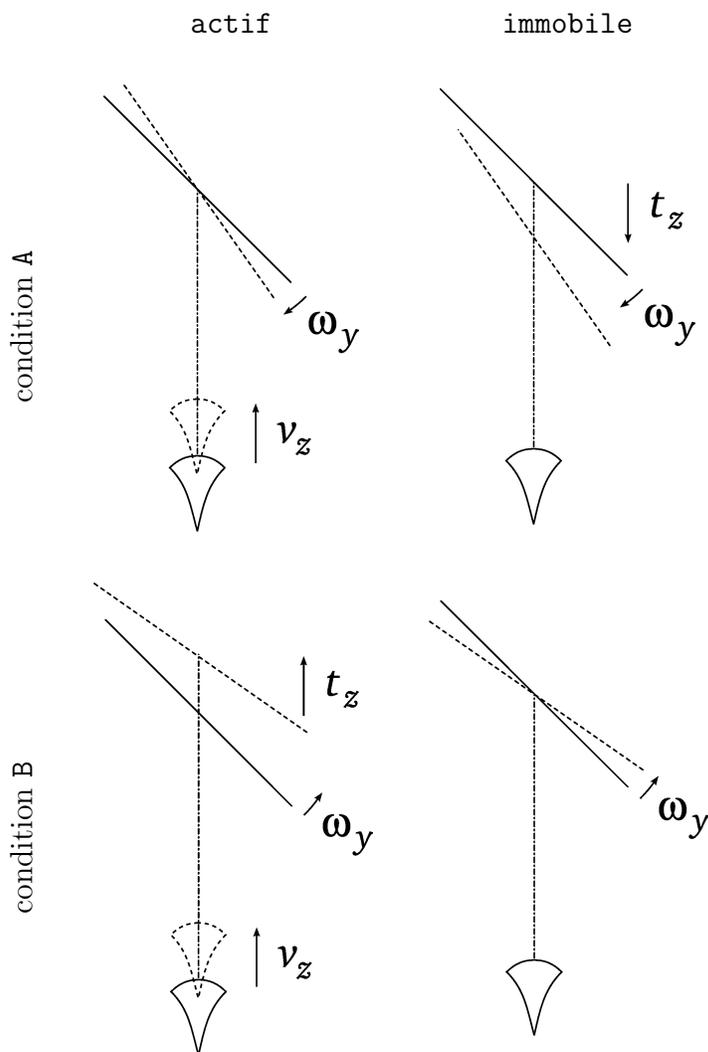


FIG. 5.17 – Conditions de l'expérience 5.4. La ligne du haut montre la condition A, correspondant à une translation relative du plan par rapport à l'observateur. La ligne du bas présente la condition B, correspondant à une absence de translation relative entre le plan et l'observateur. La colonne de gauche présente la condition **actif**, où le participant est en translation sagittale. La colonne de droite montre la condition **immobile**, où l'observateur est immobile.

La figure 5.17 résume ces quatre conditions expérimentales. L'expérience de référence (section 3.1) se retrouve en **B-immobile**, lorsque ni le plan ni l'observateur ne sont en translation sagittale dans le référentiel allocentré.

Les conditions **A-immobile** et **B-actif** mettent en œuvre un objet en translation absolue,

en plus d'être en rotation. La stationnarité d'un tel objet est donc plus faible que celle d'un plan seulement en rotation avec la même vitesse angulaire. Dans l'hypothèse où la rigidité domine la stationnarité, on s'attend à avoir un faible taux de non-rigidité dans les réponses pour toutes les conditions. Au contraire, si c'est la stationnarité qui domine, le pourcentage de réponses non-rigides doit être élevé dans les conditions A-immobile et B-actif.

5.4.2 Résultats de l'expérience

Les résultats sont exprimés en termes de pourcentage de réponse rigide ou non. Pour cela, on définit une réponse rigide comme étant proche de la configuration (objet et mouvement) simulée. Dans cette expérience, il est demandé aux participants d'évaluer l'orientation de l'axe de rotation. Une réponse est dite rigide si et seulement si l'orientation perçue par l'observateur est à moins de 45° de celle de l'axe de rotation simulée.

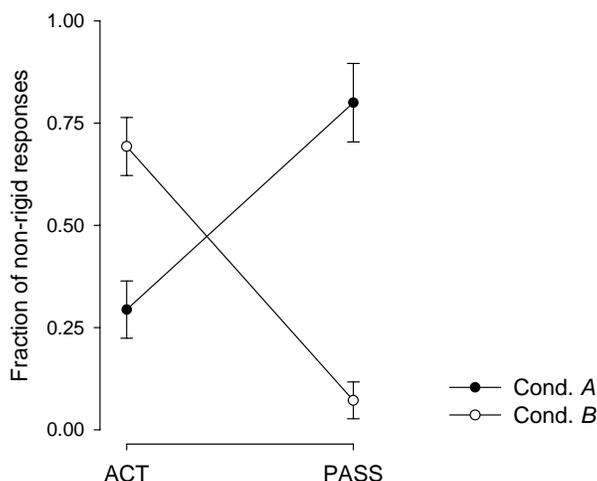


FIG. 5.18 – Comparaison des pourcentages de réponse non-rigide (*fraction of non-rigid response*) pour les quatre conditions expérimentales. En noir sont les résultats de la conditions A, correspondant à une translation relative et en blanc les résultats de la condition B, correspondant à une absence de translation relative.

La figure 5.18 montre les pourcentages de réponse non-rigide pour les quatre conditions expérimentales. On y remarque que le taux de non-rigidité est faible pour les conditions A-actif et B-immobile et qu'il est élevé dans les conditions A-immobile et B-actif. Ces résultats plaident donc en faveur d'une primauté de la stationnarité sur la rigidité.

5.4.3 Résultats du modèle

Nous définissons pour les résultats de notre modèle un percept rigide par rapport à l'orientation du plan perçu. Un percept est dit rigide si son orientation est éloignée de celle du plan simulée de moins de 45° , à 180° près.

En pratique, soit N (*nu* majuscule) la variable booléenne représentant un percept non rigide. De la même manière que pour le calcul de la probabilité d'inversion en 5.2.3, le calcul de la probabilité de réponse non rigide se fait en sommant la distribution sur l'ensemble des angles définissant une réponse non rigide $\mathcal{S}_N = \mathcal{D}_T \setminus ([\hat{\tau} - \frac{\pi}{4}, \hat{\tau} + \frac{\pi}{4}] \cup [\hat{\tau} + \frac{3\pi}{4}, \hat{\tau} + \frac{5\pi}{4}])$. La probabilité de réponse non rigide est donnée simplement par l'expression 5.24.⁴

$$P(\nu \mid \vec{\phi} v_x v_z r_y \lambda) = \int_{\tau \in \mathcal{S}_N} \int_{\sigma \in \mathcal{D}_\Sigma} \frac{\tan \sigma}{(\cos \sigma)^2} P(\chi(\sigma, \tau) v(\sigma, \tau) \mid \vec{\phi} v_x v_z r_y \lambda) d\sigma d\tau \quad (5.24)$$

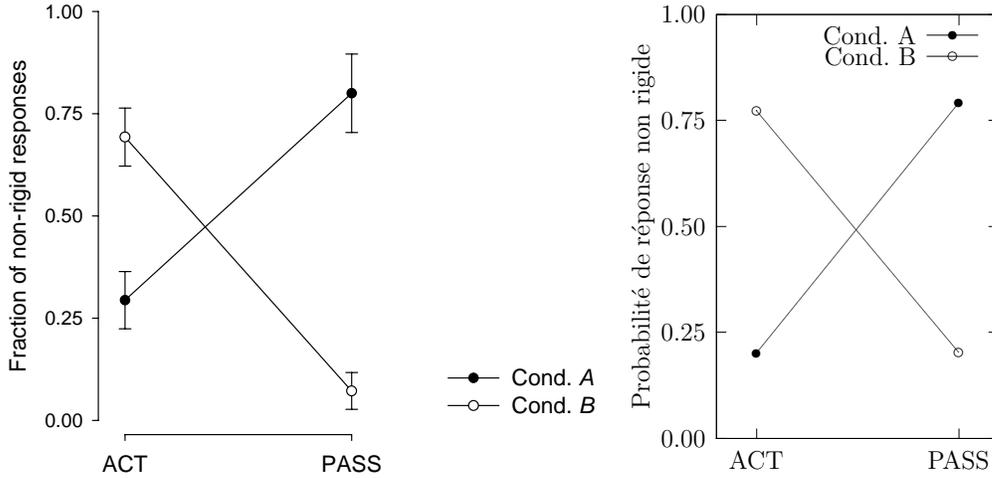


FIG. 5.19 – Comparaison des probabilités de réponse non rigide de notre modèle (à droite) avec les résultats expérimentaux (à gauche).

La figure 5.19 compare les probabilités de réponse non-rigide dans chacune des conditions avec les résultats expérimentaux. Les probabilités de réponse non rigide sont faibles en condition A-actif et condition B-immobile. Dans les conditions A-immobile et B-actif, notre modèle montre une probabilité de réponse non-rigide élevée. Ces résultats sont similaires à ceux de l'expérience de psychophysique.

5.4.4 Explication

Dans notre modèle, ces résultats s'expliquent par la comparaison des hypothèses de stationnarité et de rigidité. Dans les deux conditions A-immobile et B-actif, la stationnarité et la rigidité induisent des percepts différents mais l'influence de la rigidité est plus faible que celle de la stationnarité.

Plus précisément, la rigidité porte sur chacune des composantes du flux optique. Nous avons vu que les influences relatives des diverses composantes varie (section 5.2). En petit champ, comme c'est le cas dans cette expérience, la perception du flux d'ordre deux est bien moins précise que celle du flux d'ordre un. Nous montrons ici que le conflit entre la rigidité et la stationnarité porte principalement sur le flux d'ordre deux.

⁴Voir la démonstration de l'expression 5.19, section 5.2.3.

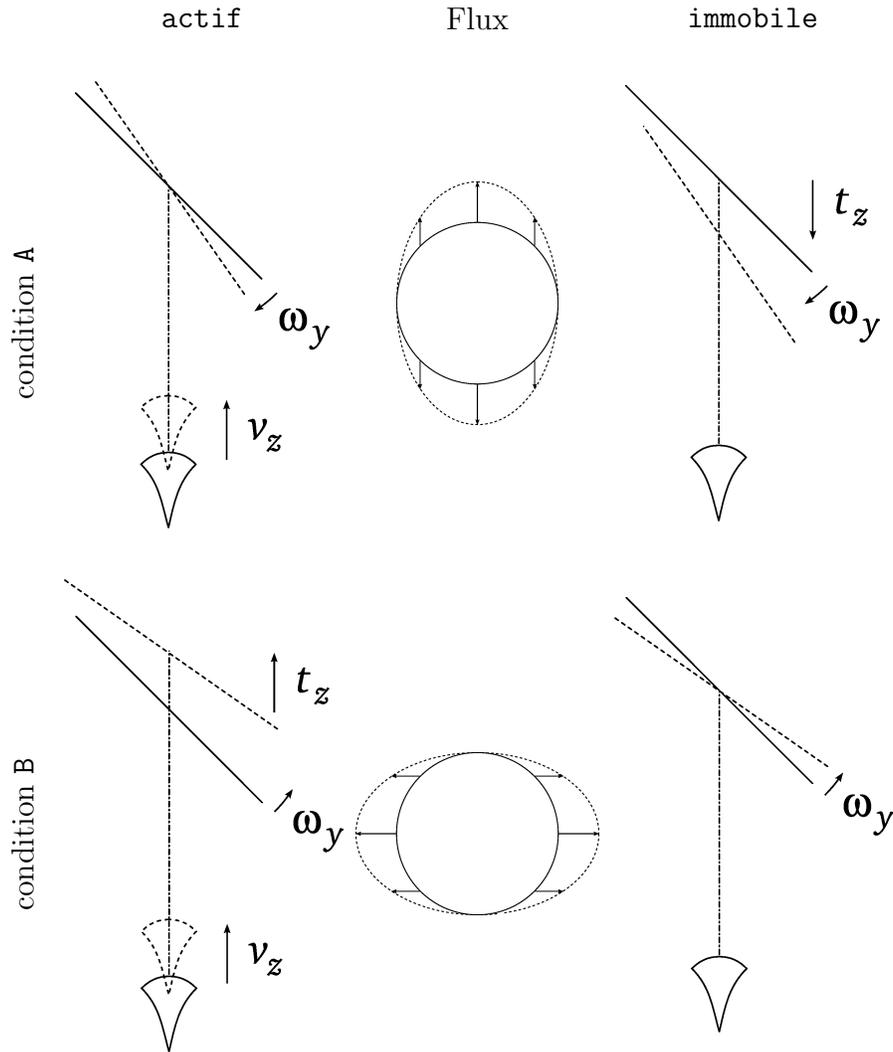


FIG. 5.20 – Flux optiques des différentes conditions de l’expérience 5.4.

La figure 5.20 montre les flux optiques à l’ordre un correspondant à chacune des conditions expérimentales. Ces flux optiques peuvent être générés par plusieurs combinaisons d’objets et de mouvements. Ceux présentés figure 5.20 sont les plus rigides, par définition.

En condition A, le flux optique est parallèle à l’axe de rotation. Cela est dû au choix de la vitesse de rotation : $\omega = \frac{t_z}{\xi}$ (formule 5.23).

La figure 5.21 présente, en condition A, les configurations les moins rigides⁵ satisfaisant le flux d’ordre un. On remarque que les plans non-rigides ne sont plus en translation relative par rapport à l’observateur. D’autre part, leurs vitesses de rotation sont les mêmes que celles des plans simulés.

⁵Les configurations les moins rigides correspondent à une orientation orthogonale à celle simulée, soit d’erreur maximale modulo π .

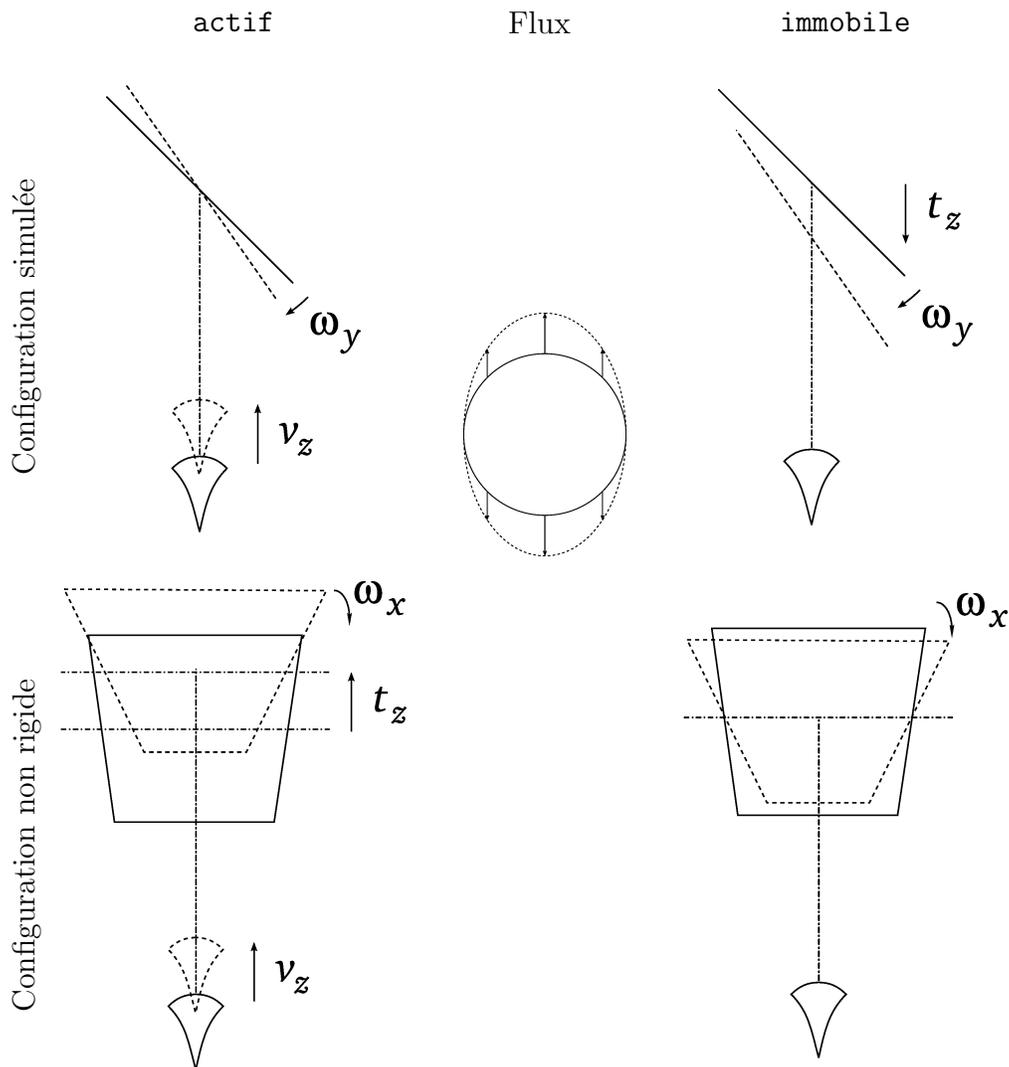


FIG. 5.21 – Comparaison des configurations amenant le flux optique d’ordre un en condition A pour l’expérience 5.4. La ligne du haut montre les objets simulés donc les plus rigides. La ligne du bas montre le plan le moins rigide, c’est-à-dire d’orientation 90° par rapport au plan simulé.

En condition **actif**, le plan non-rigide subit une translation dans le référentiel allocentré, alors que ce n’est pas le cas pour le plan rigide. La stationnarité et la rigidité préfèrent donc toutes deux le plan simulé, d’où le faible taux de non-rigidité. En condition **immobile**, le plan non-rigide n’est pas en translation (dans le référentiel du monde) alors que le plan simulé est en translation sagittale. Dans ce cas, le plan le plus stationnaire est le plan non-rigide. Comme la probabilité de non-rigidité est importante en condition **A-immobile**, on déduit que la stationnarité l’emporte sur la rigidité du flux d’ordre deux.

L’explication est exactement la même pour la condition B. La figure 5.22 montre les

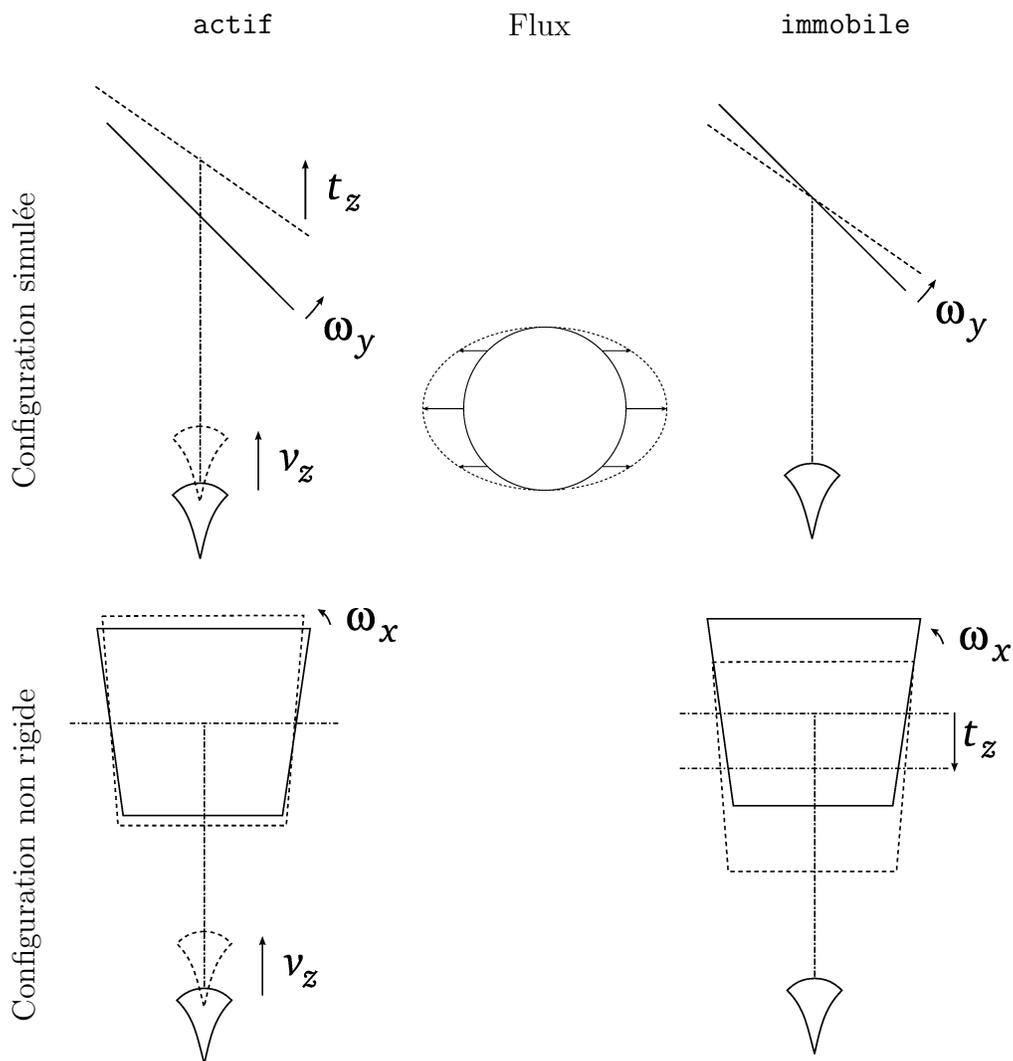


FIG. 5.22 – Comparaison des configurations amenant le flux optique d’ordre un en condition B pour l’expérience 5.4. La ligne du haut montre les objets simulés donc les plus rigides. La ligne du bas montre le plan le moins rigide, c’est-à-dire d’orientation 90° par rapport au plan simulé.

configurations rigides et non-rigides pour les conditions B-actif et B-immobile. On vérifie que l’objet non-rigide en condition B-actif est plus stationnaire, justifiant le taux de non-rigidité élevé.

Les résultats psychophysiques montrent que la stationnarité peut l’emporter sur la rigidité. Le modèle est en accord avec ces résultats. Il explique ce comportement par la relation entre la contrainte de stationnarité et la rigidité dans la perception du flux d’ordre deux.

En outre, cette expérience a été réalisée en petit champ de vision. Comme

la taille du champ de vision influe sur la perception du flux d'ordre deux, on pourrait s'intéresser aux résultats de cette expérience en grand champ de vision. À notre connaissance, les résultats de cette expérience ne sont pas disponibles dans la littérature. Il a cependant été noté une forte diminution de la perception non rigide en grand champ (Wexler *et al.* 2001a, expérience 1). Notre modèle est consistant avec ces résultats et prévoit, par exemple, un pourcentage de non-rigidité de 20 % en condition **A-immobile** en grand champ de vision, en comparaison de près de 80 % en petit champ de vision.

5.5 Influence du *shear*

5.5.1 Hypothèse et protocole

Hypothèse testée Il a été remarqué (Cornilleau-Pérès *et al.* 2002 ; van Boxtel *et al.* 2003) une différence de précision dans l'estimation de l'orientation en fonction de l'axe de rotation. Nous consacrons cette section à étudier cet effet à l'aide de notre modèle.

L'expérience de van Boxtel *et al.* (2003) vise à mesurer l'influence de l'angle de *shear* caractérisant la position relative du plan et de son axe de rotation.

Protocole expérimental Ces expériences concernent un plan incliné en profondeur, d'orientation quelconque, pouvant être en rotation autour d'un axe situé dans le plan fronto-parallèle. Il y a deux conditions expérimentales :

- **actif**, où un observateur en translation suivant l'axe transversal perçoit un plan immobile ;
- **immobile**, dans laquelle un observateur immobile perçoit un plan en rotation autour d'un axe fronto-parallèle.

On appelle *shear* ou *winding angle* l'angle entre un plan et son axe de rotation. Plus précisément, l'orientation d'un plan est définie par l'angle entre la verticale et la droite intersection du plan objet avec le plan fronto-parallèle. Le *shear* est l'angle entre cette droite et l'axe de rotation. Cet angle est noté H (*eta* majuscule) et est illustré par la figure 5.23. le *shear* est non orienté. Il est nul lorsque l'axe de rotation est contenu dans le plan objet et vaut 90° lorsque l'axe de rotation est la projection de la normale de ce plan dans le plan fronto-parallèle.

5.5.2 Résultats de l'expérience

Les résultats expérimentaux sont exprimés sous la forme de la distribution d'erreur absolue d'estimation de l'orientation. Cet erreur tient compte de l'inversion c'est-à-dire par exemple que des erreurs de 20° , -20° , 160° ou -160° correspondent toutes à 20° d'erreur. La distribution porte donc sur le domaine $[0^\circ; 90^\circ]$.

La figure 5.24 présente les distributions d'erreur pour les deux valeurs extrêmes du *shear* : $\eta = 0^\circ$ et $\eta = 90^\circ$ et pour les deux conditions **actif** et **immobile**. En haut, les graphes montrent les distributions d'erreur en **actif** pour un petit *shear* à gauche et pour

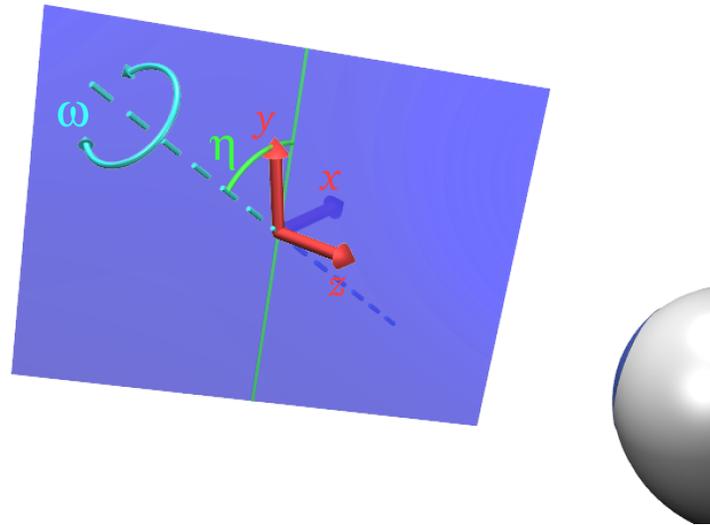


FIG. 5.23 – Définition du *shear* H pour un plan en rotation autour d'un axe du plan fronto-parallèle.

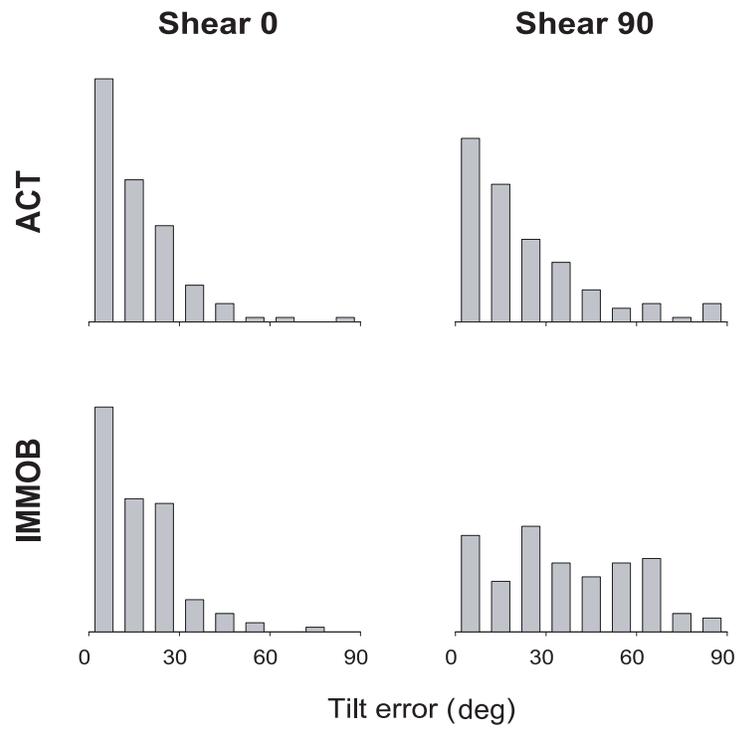


FIG. 5.24 – Distributions d'erreur du *shear* pour l'expérience 5.5.

un grand *shear* à droite. On remarque que l'erreur moyenne est légèrement supérieure pour

$\eta = 90^\circ$ (20° d'erreur) que pour $\eta = 0^\circ$ (15°).

Les graphiques en bas de la figure 5.24 présentent les résultats en condition **immobile**. Pour un petit *shear*, on retrouve une erreur moyenne comparable à la condition active. Cependant, le graphe en bas à droite montre une dispersion beaucoup plus importante de l'estimation de l'orientation avec une erreur moyenne de plus de 35°.

5.5.3 Résultats du modèle

Nous devons calculer la distribution sur l'erreur absolue en prenant en compte l'inversion. Nous considérons la variable Δ de domaine $\mathcal{D}_\Delta = [0; \pi]$ et reprenons l'expression de la distribution sur l'angle T d'orientation.⁶ Notons $\hat{\tau}$ l'orientation simulée. Comme une erreur absolue δ donnée correspond aux 4 orientations $\hat{\tau} + \delta$, $\hat{\tau} - \delta$, $\hat{\tau} + \delta + \pi$ et $\hat{\tau} - \delta + \pi$, on obtient l'expression 5.25.

$$\begin{aligned}
\forall \delta \in \mathcal{D}_\Delta, \quad & P(\delta \mid \vec{\phi} v_x v_z r_y \lambda) \\
= & \sum_{\sigma \in \mathcal{D}_\Sigma} \frac{\tan \sigma}{(\cos \sigma)^2} P(\chi(\sigma, \hat{\tau} + \delta) v(\sigma, \hat{\tau} + \delta) \mid \vec{\phi} v_x v_z r_y \lambda) \\
+ & \sum_{\sigma \in \mathcal{D}_\Sigma} \frac{\tan \sigma}{(\cos \sigma)^2} P(\chi(\sigma, \hat{\tau} - \delta) v(\sigma, \hat{\tau} - \delta) \mid \vec{\phi} v_x v_z r_y \lambda) \\
+ & \sum_{\sigma \in \mathcal{D}_\Sigma} \frac{\tan \sigma}{(\cos \sigma)^2} P(\chi(\sigma, \hat{\tau} + \pi - \delta) v(\sigma, \hat{\tau} + \pi - \delta) \mid \vec{\phi} v_x v_z r_y \lambda) \\
+ & \sum_{\sigma \in \mathcal{D}_\Sigma} \frac{\tan \sigma}{(\cos \sigma)^2} P(\chi(\sigma, \hat{\tau} + \delta - \pi) v(\sigma, \hat{\tau} + \delta - \pi) \mid \vec{\phi} v_x v_z r_y \lambda)
\end{aligned} \tag{5.25}$$

La figure 5.25 montre les distributions sur Δ dans les mêmes conditions que la figure 5.24.

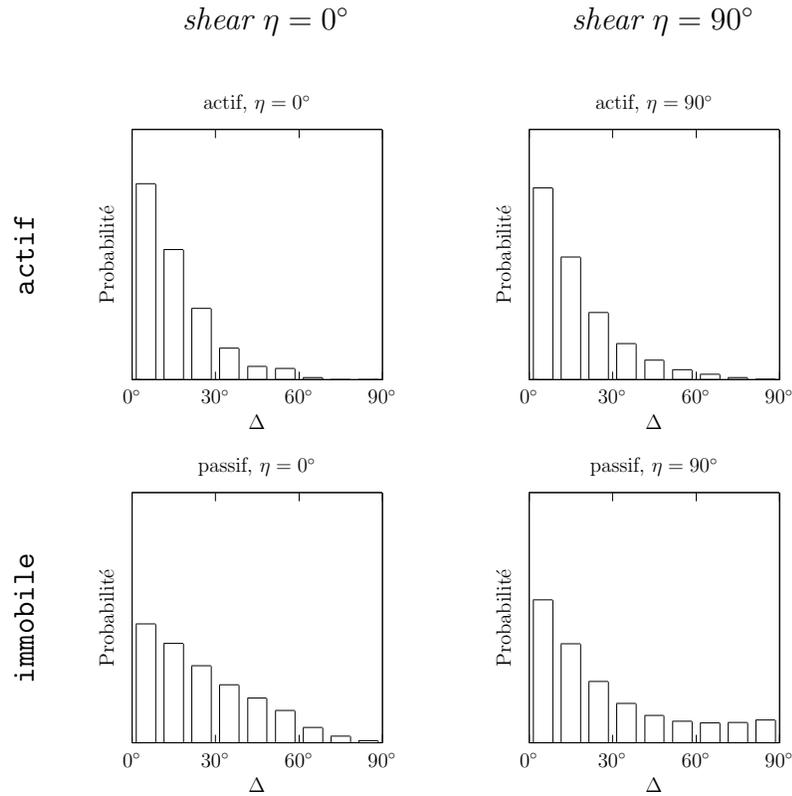
En condition **actif**, l'erreur moyenne passe de 15° pour $\eta = 0^\circ$ à 18° pour $\eta = 90^\circ$. Comme pour les résultats expérimentaux, une augmentation de l'angle de *shear* mène à une légère augmentation de la dispersion.

En condition **immobile**, l'erreur moyenne passe de 25° pour un petit *shear* à 37° pour un *shear* élevé. La différence de dispersion est plus importante pour un observateur immobile que pour un observateur en mouvement. Ces résultats sont compatibles avec les résultats expérimentaux.

Les résultats du modèle en condition **immobile** pour un grand angle de *shear* sont continus. On remarque une décroissance rapide de la probabilité d'erreur entre 0 et 40° puis un plateau avec une légère remontée pour une erreur proche de 90°.

Les résultats expérimentaux dans la même condition présentent un profil beaucoup plus irrégulier. Les moyennes de ces deux distributions sont toutes deux autour de 35°. Il y a donc une différence sur la répartition des erreurs, mais pas sur le comportement du modèle, exprimé en termes d'augmentation de la dispersion avec le *shear*. Il serait

⁶Formule 4.42 section 4.6.2 et 5.18 section 5.2.3.

FIG. 5.25 – Distribution sur l'erreur Δ pour le modèle.

intéressant d'examiner plus en détail ces résultats pour déterminer en particulier si, pour un échantillon plus grand, la répartition garderait son irrégularité.

5.5.4 Explication

Dans notre modèle, ce comportement s'explique à nouveau par la combinaison des hypothèses de rigidité et de stationnarité. Plus précisément, l'ensemble des mouvements les plus rigides sont moins stationnaires pour un *shear* élevé que pour un faible *shear*.

Pour illustrer ce point, nous comparons les configurations correspondant à des erreurs de 0 ou 90° dans les deux conditions de *shear* : $\eta = 0^\circ$ et $\eta = 90^\circ$. Nous commençons par la condition **immobile** puis nous vérifions la différence d'influence en condition **actif**.

Comme pour l'expérience précédente, l'observation se fait en petit champ de vision. La rigidité s'exprime donc principalement par le flux d'ordre un. La figure 5.26 montre les flux optiques à l'ordre un associés aux stimulus dans des conditions extrêmes du *shear* avec l'exemple d'un plan vertical.

La figure 5.27 montre les percepts correspondant aux erreurs $\delta = 0^\circ$ et $\delta = 90^\circ$ pour les deux valeurs du *shear* considérées. Ces configurations partagent le flux optique d'ordre un. Pour une erreur nulle, on retrouve les plans simulés. Pour $\eta = 0^\circ$, la configuration simulée est la même qu'en condition **B-immobile** de l'expérience 5.4. La configuration d'orientation

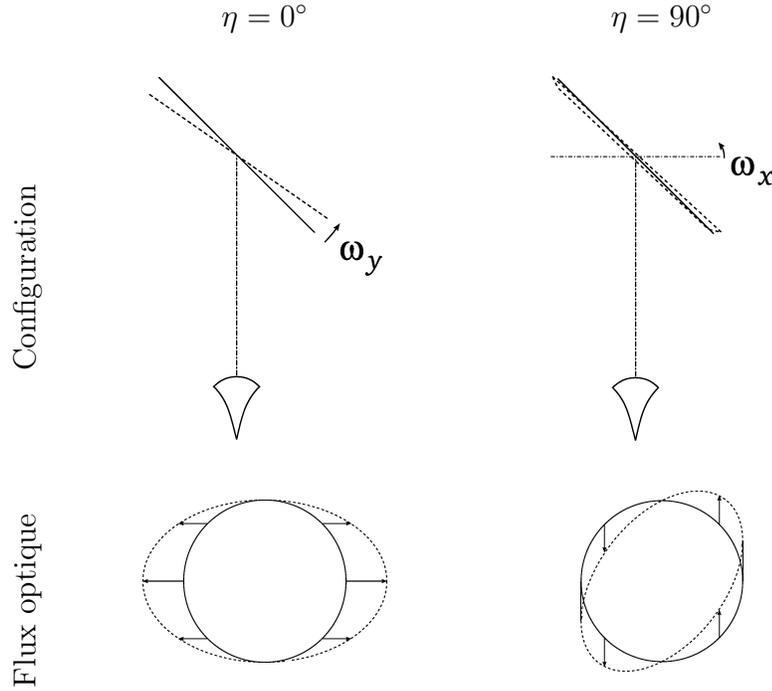


FIG. 5.26 – Configurations et flux optiques associés à un plan vertical pour deux valeurs du *shear* en condition *immobile*.

orthogonale est, elle aussi, la même.

Pour un petit *shear*, la configuration alternative fait intervenir une translation ainsi qu'une rotation de même vitesse que celle d'origine (que l'on va noter ω_0). Pour $\eta = 90^\circ$, le mouvement du plan se compose de deux rotations de même vitesse que celle de référence.

Mathématiquement, soit $\vec{\mu}_\eta(\delta)$ le mouvement relatif $(\vec{\Omega}, \vec{T})$ le plus probable pour un *shear* η et une position du plan correspondant à une erreur d'orientation δ . Dans les quatre conditions de la figure 5.27, on a :

$$\begin{aligned} \delta = 0^\circ \quad \vec{\mu}_0(0) &= \left(\begin{pmatrix} 0 \\ \omega_0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right) & \vec{\mu}_{90}(0) &= \left(\begin{pmatrix} -\omega_0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right) \\ \delta = 90^\circ \quad \vec{\mu}_0(90) &= \left(\begin{pmatrix} \omega_0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \xi\omega_0 \end{pmatrix} \right) & \vec{\mu}_{90}(90) &= \left(\begin{pmatrix} 0 \\ -\omega_0 \\ \xi\omega_0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right) \end{aligned}$$

On peut comparer les probabilités de percevoir une erreur d'orientation $\delta = 90^\circ$ entre les conditions $\eta = 0^\circ$ et $\eta = 90^\circ$. Comme nous sommes en petit champ de vision, nous négligeons l'influence du flux optique d'ordre deux. Les flux d'ordre zéro et un sont identiques et les orientations sont équiprobables. La comparaison des probabilités de perception d'erreur se fait donc sur la probabilité du mouvement.

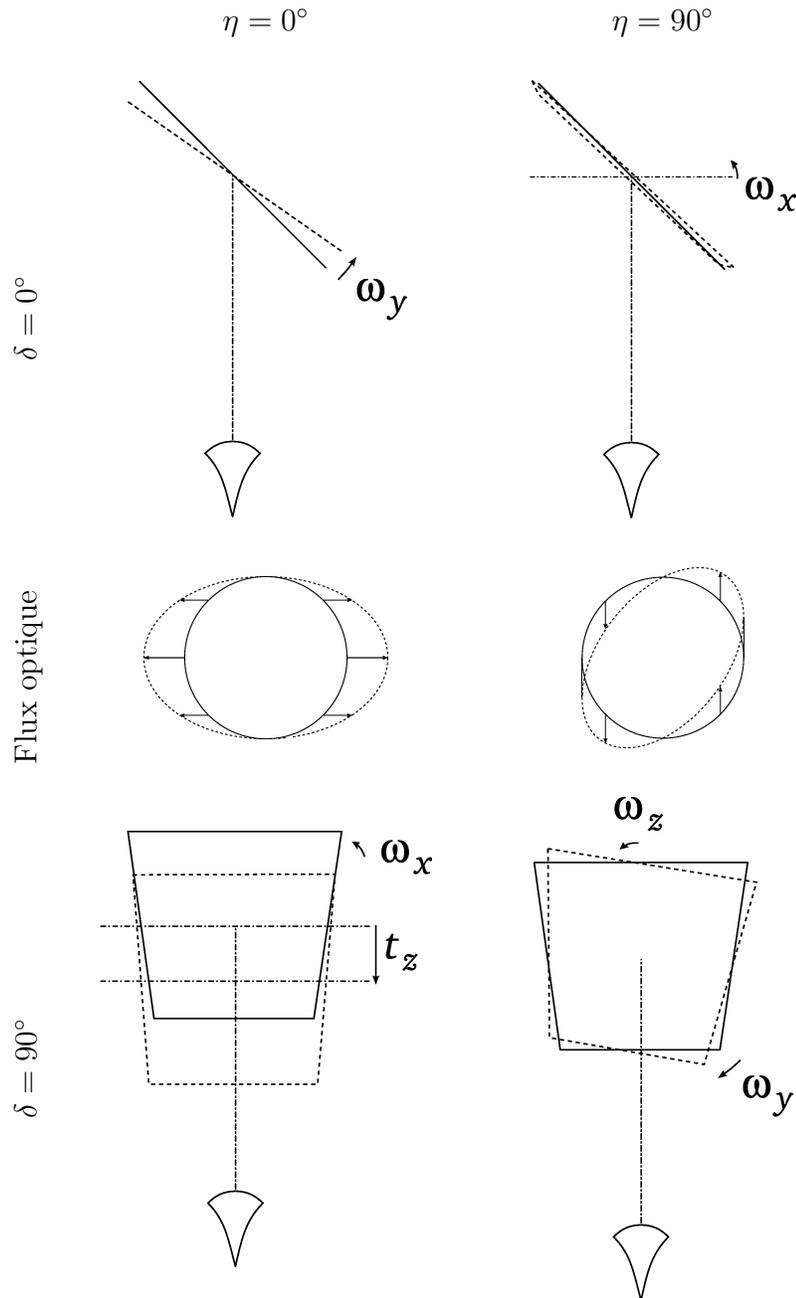


FIG. 5.27 – Percepts correspondant aux erreurs $\delta = 0^\circ$ et $\delta = 90^\circ$ pour les deux conditions de *shear* : $\eta = 0^\circ$ et $\eta = 90^\circ$.

Avec nos notations $\sigma_{\vec{\Omega}} = \sigma_{\Omega} \cdot \text{Id}_{3 \times 3}$ et $\sigma_{\vec{T}} = \sigma_T \cdot \text{Id}_{3 \times 3}$ pour les matrices de covariances respectives pour les rotations et les translations, on peut écrire l'inégalité 5.26 en condition

immobile (c'est-à-dire pour un mouvement propre nul ($v_x = v_z = r_y = 0$)) :

$$\begin{aligned}
P(\vec{\mu}_0(90) \mid v_x v_z r_y) &< P(\vec{\mu}_{90}(90) \mid v_x v_z r_y) \\
\iff \mathcal{N}_{\bar{0},\sigma_{\bar{\Omega}}}((\omega_0, 0, 0)) \mathcal{N}_{\bar{0},\sigma_{\bar{T}}}((0, 0, \xi\omega_0)) &< \mathcal{N}_{\bar{0},\sigma_{\bar{\Omega}}}((0, -\omega_0, \xi\omega_0)) \mathcal{N}_{\bar{0},\sigma_{\bar{T}}}((0, 0, 0)) \\
&\iff \mathcal{N}_{0,\sigma_T}(\xi\omega_0) < \mathcal{N}_{0,\sigma_{\Omega}}(\xi\omega_0) \\
&\iff \sigma_T < \sigma_{\Omega}
\end{aligned} \tag{5.26}$$

Ainsi, le résultat selon lequel la probabilité de percevoir une grande erreur d'orientation est plus faible pour un petit *shear* que pour un grand *shear* est équivalente dans le modèle à avoir la covariance sur la translation inférieure à celle sur la rotation. C'est une contrainte dont nous avons tenu compte pour le choix des paramètres (section 4.5.3).

On peut vérifier que, pour un *shear* élevé, les probabilités de petite ou grande erreur sont proches. On écrit :

$$\begin{aligned}
&\frac{P(\vec{\mu}_{90}(90) \mid v_x v_z r_y)}{P(\vec{\mu}_{90}(0) \mid v_x v_z r_y)} \\
&= \frac{\mathcal{N}_{(0,0,0),\sigma_{\bar{\Omega}}}((0, -\omega_0, \xi\omega_0)) \mathcal{N}_{(0,0,0),\sigma_{\bar{T}}}((0, 0, 0))}{\mathcal{N}_{(0,0,0),\sigma_{\bar{\Omega}}}((-\omega_0, 0, 0)) \mathcal{N}_{(0,0,0),\sigma_{\bar{T}}}((0, 0, 0))} \\
&= \frac{\mathcal{N}_{0,\sigma_{\Omega}}(\xi\omega_0)}{\mathcal{N}_{0,\sigma_{\Omega}}(0)} \\
&= \exp -\frac{1}{2} \left(\frac{\xi\omega_0}{\sigma_{\Omega}} \right)^2
\end{aligned} \tag{5.27}$$

Pour ces simulations, le rapport $\frac{\xi\omega_0}{\sigma_{\Omega}}$ est toujours inférieur à 1, ce qui explique les résultats.

En condition **actif**, dans les conditions de l'expérience, le mouvement propre est $(v_x, v_z, r_y) = (\omega_0, 0, 0)$. L'expression 5.27 devient 5.28.⁷

$$\begin{aligned}
&\frac{P(\vec{\mu}_{90}(90) \mid v_x v_z r_y)}{P(\vec{\mu}_{90}(0) \mid v_x v_z r_y)} \\
&= \frac{\mathcal{N}_{(0,-\omega_0,0),\sigma_{\bar{\Omega}}}((\omega_0, 0, \xi\omega_0)) \mathcal{N}_{(0,0,0),\sigma_{\bar{T}}}((0, 0, 0))}{\mathcal{N}_{(0,-\omega_0,0),\sigma_{\bar{\Omega}}}((0, -\omega_0, 0)) \mathcal{N}_{(0,0,0),\sigma_{\bar{T}}}((0, 0, 0))} \\
&= \frac{\mathcal{N}_{(0,-\omega_0,0),\sigma_{\bar{\Omega}}}((\omega_0, 0, \xi\omega_0))}{\mathcal{N}_{(0,-\omega_0,0),\sigma_{\bar{\Omega}}}((0, -\omega_0, 0))} \\
&= \exp -\frac{1}{2} \frac{\xi^2\omega_0^2 + 2\omega_0^2}{\sigma_{\Omega}^2}
\end{aligned} \tag{5.28}$$

Le rapport 5.28 entre la probabilité du mouvement pour une grande erreur et celle pour une petite erreur d'estimation de l'orientation en grand *shear* et en **actif** est inférieur à

⁷Nous avons changé l'orientation du plan conformément à notre représentation paramétrique sans perte de généralité.

celui en `immobile` exprimé en 5.27 (d'un facteur $\exp\left(\frac{\omega_0^2}{\sigma_\Omega^2}\right)$). Cela explique la moins grande influence du *shear* en condition `actif` qu'en condition `immobile`.

En conclusion, les résultats psychophysiques montrent que le *shear* influence la perception de la forme par le mouvement. Le modèle est en accord avec ces résultats. Il explique ce comportement par la plus faible stationnarité des perceptions erronées pour un petit *shear*.

En outre, cette expérience donne des informations sur les paramètres de notre modèle. En effet, la relation 5.26 reliant les paramètres σ_T et σ_Ω est une contrainte que doit vérifier notre modèle pour présenter une plus grande incertitude pour un plus grand *shear*. Une autre contrainte est le rapport $\frac{\xi\omega_0}{\sigma_\Omega}$. L'effet du *shear* sur la dispersion des erreurs d'estimation est d'autant plus élevé que ce rapport est petit. Connaissant la vitesse et l'inclinaison simulées, nous pouvons choisir l'écart-type sur les rotations pour obtenir une certaine dispersion. Pour que la dispersion augmente avec l'angle de *shear*, nous pouvons alors choisir un écart-type sur les translations plus petit que celui sur les rotations.

5.6 Résumé

Dans ce chapitre, nous avons montré l'utilisation de notre modèle et l'exploitation de ses résultats.

Nous nous sommes, en premier lieu, intéressés au fonctionnement de notre modèle jusqu'à l'ordre un. Pour cela, nous avons utilisé des résultats d'expériences menées avec une projection orthographique. Nous avons montré que la perception de l'inclinaison d'un plan par notre modèle en petit champ de vision est consistante avec la perception orthographique malgré la prise en compte du flux optique d'ordre deux.

Nous avons ensuite suivi l'évolution de la perception avec la taille du champ de vision. En petit champ de vision, les participants ne perçoivent l'orientation d'un plan qu'avec une forte ambiguïté. Cette ambiguïté disparaît en grand champ de vision. Notre modèle reproduit ces résultats en s'appuyant sur l'hypothèse que la taille du champ de vision modifie l'influence relative des diverses composantes du flux optique.

La troisième expérience à laquelle nous nous sommes intéressés concernait l'influence du mouvement propre dans la perception du flux optique. Nous avons vérifié que notre modèle bayésien est capable de reproduire la disparition de l'ambiguïté sur l'orientation lors d'un mouvement de l'observateur. Ce résultat est la conséquence de l'emploi de l'hypothèse de stationnarité stipulant qu'un plus faible mouvement absolu de l'objet sera plus probable qu'un mouvement plus important.

Nous avons comparé l'influence des hypothèses de stationnarité et de rigidité. En nous appuyant sur un stimulus pour lequel les interprétations de la stationnarité et de la rigidité sont différentes, nous avons constaté que notre modèle se comporte de la même manière que les participants à l'expérience en préférant le percept statique. Notre modèle prédit en outre une diminution de l'influence de la stationnarité avec l'agrandissement du champ de vision.

Enfin nous avons étudié, à l'aide de notre modèle bayésien, la variation de précision dans la perception de l'orientation d'un plan en fonction de l'angle de son axe de rotation. Notre modèle permet d'expliquer cette observation en se basant sur l'hypothèse de stationnarité.

Ces expériences et leurs résultats sont tirés de la littérature du domaine. On peut noter que leurs origines sont diverses et que nous ne nous sommes pas restreints à des résultats d'une provenance unique. Cela nous a permis de tester une large palette des variables de notre modèle dans des conditions d'utilisation variées.

Notre modèle, basé principalement sur les hypothèses de stationnarité et de rigidité, permet une bonne reproduction qualitative des résultats des expériences, c'est-à-dire qu'il respecte les propriétés générales des courbes (relations entre les modes, sens de variation...).

Rappelons que ces multiples résultats ne font appel qu'à un seul modèle et un seul jeu de paramètres communs à toutes les expériences. Ce choix vient au prix de l'adéquation numérique à cause, en particulier, des différences de protocoles entre les équipes et de participants aux expériences. Les résultats des différentes expériences sont numériquement inconsistants vis-à-vis du modèle. Un exemple visible est le taux d'inversion calculable en condition `immobile` pour les résultats de van Boxtel *et al.* (2003) qui varie de quelques unités de celui en `petit champ` par Cornilleau-Pérès *et al.* (2002). Ces deux résultats sont indiscernables pour notre modèle ; par contre, les variations du taux d'inversion entre les conditions `actif` et `immobile` d'une part, et `petit champ` et `grand champ` d'autre part, sont correctement reproduites par le modèle.

Chapitre 6

Utilisations avancées

Nous n'avons, jusqu'à présent, utilisé notre modèle que directement et indépendamment d'autres systèmes. Nous examinons dans ce chapitre les possibilités d'extensions et d'utilisation de notre modèle.

Nous commençons par l'utilisation de notre modèle à la perception d'objets ondulés. Dans un second temps, nous présentons un modèle de reparamétrisation du flux optique susceptible de fournir ses entrées à notre modèle de perception du plan. Enfin nous récapitulons les interfaces de notre modèle de perception et la manière d'utiliser séquentiellement plusieurs modèles.

6.1 Expérience approchée

Nous désirons nous intéresser à l'effet de la poursuite sur la perception d'un plan par le flux optique. Or, à notre connaissance, il n'y a pas d'expérience de poursuite de plans dans la littérature. L'expérience a été réalisée avec un objet ondulé par Naji et Freeman (2004). La problématique de cette section est double : la transposition de notre modèle plan à un objet ondulé et l'influence de la poursuite sur la perception de la forme par le flux optique.

6.1.1 Hypothèse et protocole

Hypothèse testée La poursuite est le suivi par l'œil d'un objet. Cette expérience s'intéresse à son influence sur la perception de la forme par le flux optique. Ainsi, en nous appuyant sur la littérature, nous allons comparer à la fois la perception d'un même flux optique en situation de poursuite ou non et la perception d'une même scène avec ou sans poursuite.

Protocole expérimental L'expérience de Naji et Freeman (2004) porte sur un objet ondulé représenté figure 6.1. L'objet a une texture de points aléatoires et est montré en petit champ de vision.

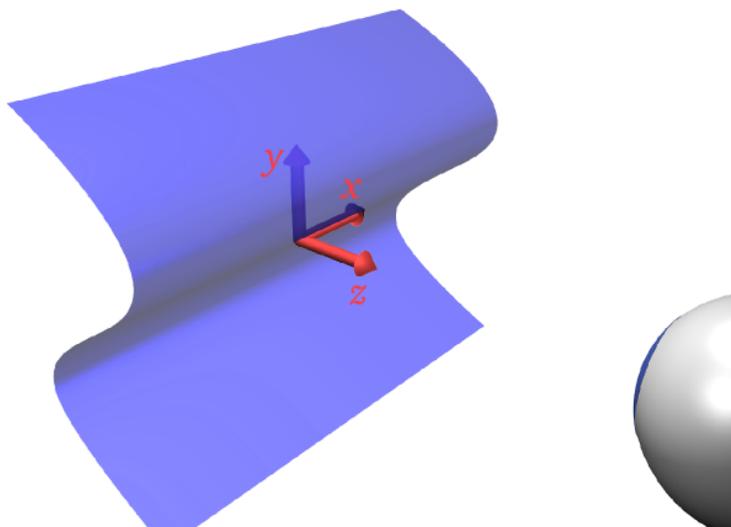


FIG. 6.1 – Objet présenté dans l’expérience 6.1.3 réalisée par Naji et Freeman (2004).

Trois conditions de perception sont comparées.¹ La première, condition A, porte sur la perception de l’objet en translation, sans suivi par l’observateur. Dans une deuxième condition, libellée B, l’objet ne subit qu’une rotation suivant l’axe vertical. Enfin, la troisième condition, appelée C, porte sur la perception lors d’une translation avec suivi du regard par l’observateur.

Pour chacune de ces conditions, l’amplitude et le sens de l’ondulation varient en passant par une amplitude nulle, c’est-à-dire le plan fronto-parallèle. La figure 6.2 montre les divers objets simulés en rapport avec l’amplitude signée S du stimulus.

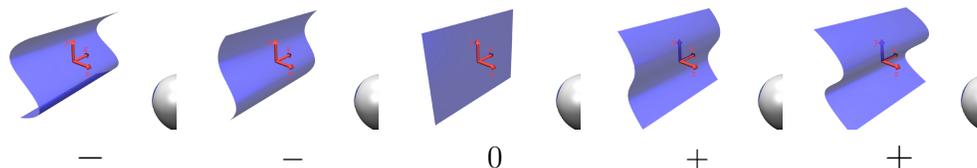


FIG. 6.2 – Objet simulé en fonction de l’amplitude du stimulus. On remarque que le haut de l’objet est plus éloigné de l’observateur à mesure que le stimulus est plus positif.

L’évaluation de la perception se fait par l’intermédiaire de deux questions posées à l’observateur. La première s’intéresse à la perception du relief de l’objet. La question posée est de savoir si le participant a vu un objet plat (*flat*) ou non. La seconde porte sur la phase de l’ondulation. Il est demandé à l’observateur si la partie haute de l’objet est éloignée (*top-far*) ou proche, relativement au reste de l’objet. La réponse *top-far* correspond à la

¹Nous utilisons les dénominations de Naji et Freeman (2004), qui sont les mêmes que pour l’expérience 5.4. Les conditions sont différentes mais nous réemployons ces noms en référence à l’expérience.

perception d'un objet associé à un stimulus positif comme on peut le voir dans la figure 6.2.

Les trois conditions expérimentales sont résumées par la figure 6.3. On s'aperçoit que la condition B est la même que la condition **immobile** de référence des précédentes expériences. On remarque en outre que les conditions A et C partagent le même mouvement de l'objet dans le repère allocentré alors que les conditions B et C partagent le mouvement de l'objet dans le référentiel de la rétine et donc le même flux optique.

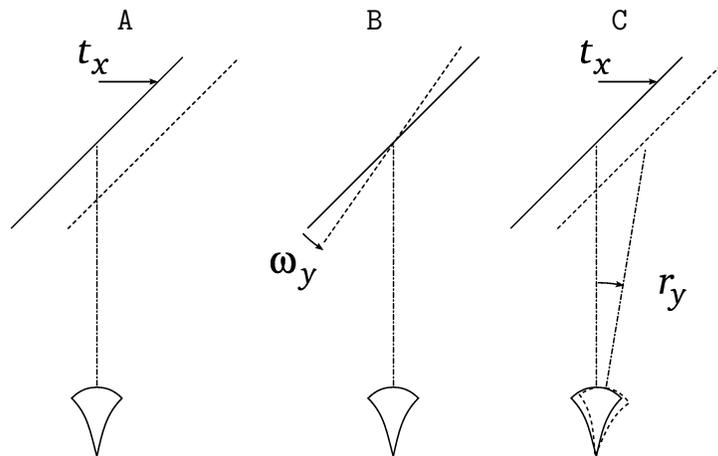


FIG. 6.3 – Conditions de l'expérience 6.1.3. La condition A met en jeu un objet en translation transversale sans suivi de l'observateur. L'objet est en rotation dans la condition B. Enfin la condition C concerne un objet en translation transversale suivi par le regard du participant.

6.1.2 Transposition du modèle

Cette expérience met en jeu le flux optique d'un objet ondulé alors que nous modélisons la perception d'un plan incliné. Nous montrons d'abord que l'inclinaison et l'ondulation sont soumises aux mêmes ambiguïtés avant de nous intéresser à la transposition des questions posées aux observateurs.

Stimulus

Le stimulus réellement présenté aux participants est une approximation de la vitesse obtenue en cas de rotation d'une surface ondulée en considérant une projection orthographique. Pour nous comparer aux résultats expérimentaux, nous partons donc du stimulus de l'article original, bien qu'appauvri.

En notant f la fréquence du stimulus, t le temps écoulé depuis le début de la présentation, f_s la fréquence spatiale de l'ondulation et S son amplitude signée, la composante horizontale $v_x(x, y, t)$ du flux optique au point de l'image de coordonnées (x, y) à l'instant

t est donnée par l'équation 6.1 (Naji et Freeman 2004, section 2.1.1).

$$v_x(x, y, t) = 2\pi f \cos(2\pi ft) S \sin(2\pi f_s y) \quad (6.1)$$

On remarque en particulier que, comme la série de Taylor de la fonction sinus ne contient que des composantes impaires, les flux d'ordre 0 et 2 sont constamment nuls et ne permettent donc pas de lever l'ambiguïté.

En notant $\omega_0 = 2\pi f$ la vitesse de rotation relative de l'objet autour de l'axe vertical, $\chi_{\omega_0}(t) = \cos(\omega_0 t)$ l'inclinaison de l'objet à l'instant t (qui dépend de la vitesse de rotation) et $p(y) = S \sin(2\pi f_s y)$ le profil de l'ondulation, on obtient alors l'équation 6.2.

$$v_x(x, y, t) = \omega_0 \chi_{\omega_0}(t) p(y) \quad (6.2)$$

On remarque ainsi que le stimulus est invariant par symétrie fronto-parallèle \mathcal{S} . En effet, une telle symétrie change le sens de rotation et le profil comme indiqué par l'équation 6.3.

$$\begin{cases} \omega & \mapsto -\omega \\ p(y) & \mapsto -p(y) \end{cases} \quad (6.3)$$

Cette symétrie est celle qui définit l'inversion dans le cas de la perception de l'orientation d'un plan comme étudiée sections 4.6 et 5.2. Ainsi, la perception de l'ambiguïté de l'ondulation peut être transposée à celle de l'orientation.

Nous simulons ainsi un plan incliné à gauche ($\chi > 0$ et $v = 0$) pour figurer un stimulus positif et incliné à droite ($\chi < 0$ et $v = 0$) pour un stimulus négatif.

De même, les variations d'amplitude du stimulus sont simulées à l'aide de l'inclinaison. Un stimulus d'amplitude nulle est donc simulé par le plan fronto-parallèle (comme dans l'expérience originale) et une plus grande amplitude est transposée en une plus grande inclinaison. Le choix de correspondance que nous avons fait est :

$$\chi \propto S \quad (6.4)$$

La figure 6.4 montre l'équivalence entre les divers stimuli de l'expérience et les plans simulés pour notre modèle.

Il est à noter que l'angle entre le mouvement et l'orientation d'une surface influe sur la précision de la perception (voir l'effet du *shear* section 5.5). C'est aussi le cas de la magnitude de l'inclinaison et de la vitesse de rotation. C'est donc le facteur de proportionnalité entre χ et S qui va permettre de normaliser les pourcentages d'inversion entre l'expérience originale et nos simulations transposées à un plan. Il serait intéressant de poursuivre ce travail en regardant l'influence de l'angle de *shear* sur ces résultats et en particulier un *shear* de 90° , correspondant à un plan incliné suivant le même axe que l'objet ondulé servant de stimulus.

Questions

Puisque nous simulons un stimulus différent de celui utilisé pour l'expérience, nous devons adapter les questions posées aux observateurs. La première question posée aux

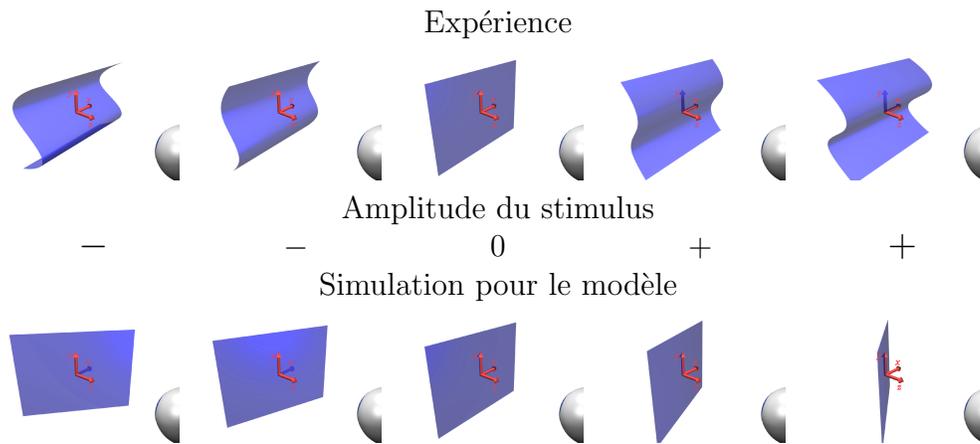


FIG. 6.4 – Correspondance entre les objets présentés lors de l'expérience (en haut) et ceux simulés pour notre modèle (en bas) d'après l'équation 6.4.

participants est de savoir si l'objet perçu est plat ou non. Un objet plat correspond à un stimulus d'amplitude nulle pour l'expérience. Un stimulus d'amplitude nulle pour le modèle correspond à un plan non incliné, c'est-à-dire le plan fronto-parallèle. La transposition de la question *flat* à notre modèle est donc de savoir si le plan est fronto-parallèle ou non.

La seconde question posée lors de l'expérience, *top-far*, s'intéresse à la phase de l'ondulation. Un objet est dit *top-far* lorsque sa partie haute est éloignée de l'observateur. Cela correspond à une amplitude positive du stimulus. Pour notre modèle, un stimulus positif correspond à un plan incliné à gauche. La transposition de la question *top-far* dans notre modèle est donc de savoir si le plan perçu est incliné à gauche ou non.

En termes d'orientation, un stimulus négatif correspond à une position non *top-far*. En conséquence, la probabilité de réponse *top-far* est la probabilité de se tromper de 180° sur l'orientation et donc est la probabilité d'inversion. Pour un stimulus positif, la probabilité de réponse *top-far* est donc la probabilité de ne pas percevoir d'inversion.

6.1.3 Résultats de l'expérience

La figure 6.5 présente les résultats expérimentaux sous la forme des pourcentages de réponse *top-far* et *flat*.

Tout d'abord, si l'on s'intéresse aux réponses à la question *top-far*, on remarque que les profils sont différents. Dans les conditions A et C, les stimuli plus positifs sont perçus comme étant probablement plus *top-far*, comme attendu (figure 6.2). Au contraire, en condition B, le profil de réponse est quasi symétrique. On remarque ainsi que pour un stimulus de grande amplitude, les participants préfèrent répondre que l'objet n'est pas en configuration *top-far*, indépendamment de la phase de l'ondulation simulée. Cette incapacité à distinguer la phase est l'équivalent de l'inversion de la perception de l'orientation d'un plan.

On remarque en outre, à la pente des courbes *top-far* en conditions A et C, que la perception est plus précise en cas de poursuite.

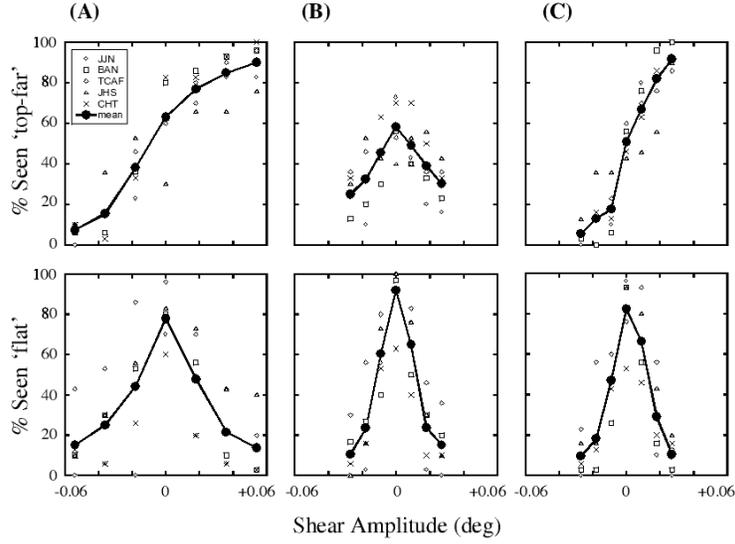


FIG. 6.5 – Résultats de l’expérience 6.1.3 (Naji et Freeman 2004, figure 3). En haut sont présentés les pourcentages de perception *top-far* en fonction de l’amplitude du stimulus pour les trois conditions A, B et C. De la même manière, les graphes du bas présentent les pourcentages de perception *flat*.

Cette différence est confirmée par les courbes de pourcentage de réponse *flat* en bas de la figure 6.5. En effet, ces graphes présentent un pic centré autour d’une amplitude nulle comme attendu. Plus précisément, la largeur du pic est plus importante en condition A que dans les conditions B et surtout C, indiquant une plus grande incertitude.

6.1.4 Sous-modèle initial

Obtention des réponses

Comme pour les expériences précédentes, nous construisons des modèles probabilistes auxiliaires pour calculer les réponses à chacune de ces deux questions.

Sous-modèle *top-far* Nous introduisons la variable booléenne Θ décidant si une perception est *top-far* ou non. Nous définissons le sous-modèle avec la décomposition 6.5.

$$P(\Theta \ X \ \Upsilon \ \vec{\Phi} \ V_x \ V_z \ R_y \ \Lambda) \quad (6.5)$$

$$= P(\vec{\Phi} \ V_x \ V_z \ R_y \ \Lambda) \quad (6.6)$$

$$\times P(X \ \Upsilon \ | \ \vec{\Phi} \ V_x \ V_z \ R_y \ \Lambda) \quad (6.7)$$

$$\times P(\Theta \ | \ X) \quad (6.8)$$

Comme d’habitude, le facteur 6.6 peut rester non spécifié puisqu’il ne concerne que des variables qui seront instanciées lors de la question. Le facteur 6.7 est la réponse donnée

par le modèle défini au chapitre 4. Le facteur 6.8 est la définition d'une réponse *top-far* sachant l'inclinaison horizontale X du plan perçu. D'après la définition de la section 6.1.2, nous choisissons de lui donner une probabilité unité lorsque le plan est incliné à gauche ($\chi > 0$) et une probabilité nulle sinon (expression 6.9).

$$\forall \chi \in \mathcal{D}_X, P(\theta | \chi) = \begin{cases} 1 & \text{si } \chi > 0 \\ 0 & \text{sinon} \end{cases} \quad (6.9)$$

L'inférence bayésienne donne l'expression 6.10 pour la probabilité d'une réponse *top-far*.

$$P(\theta | \vec{\phi} v_x v_z r_y \lambda) = \int_{\chi \in \mathbb{R}_+^*} \int_{v \in \mathcal{D}_R} P(\chi v | \vec{\phi} v_x v_z r_y \lambda) d\chi dv \quad (6.10)$$

La construction de ce modèle est directe et repose sur la définition d'une réponse *top-far* (6.1.2).

Sous-modèle *flat* Le deuxième modèle à construire concerne la probabilité de réponse *flat*. Nous commençons par construire un modèle naïf et montrons qu'il ne correspond pas à l'expérience à décrire. Nous montrons enfin une définition utilisable pour ce modèle.

Nous introduisons la variable binaire Ψ décidant si une réponse est *flat* ou non et nous construisons le modèle à partir de la décomposition de la conjointe 6.11 :

$$P(\Psi X \Upsilon \vec{\Phi} V_x V_z R_y \Lambda) \quad (6.11)$$

$$= P(\vec{\Phi} V_x V_z R_y \Lambda) \quad (6.12)$$

$$\times P(X \Upsilon | \vec{\Phi} V_x V_z R_y \Lambda) \quad (6.13)$$

$$\times P(\Psi | X \Upsilon) \quad (6.14)$$

Le facteur 6.12 ne concerne que des variables qui sont instanciées lors d'une question et peut donc rester non spécifié. Le facteur 6.13 est la réponse donnée par notre modèle du chapitre 4. Le facteur 6.14 est la définition du stimulus d'amplitude nulle. D'après la définition donnée en 6.1.2, nous choisissons dans un premier temps de fixer la probabilité d'une réponse *flat* à 1 si le plan est le plan fronto-parallèle et à 0 sinon :

$$\forall \chi, v \in \mathcal{D}_X \times \mathcal{D}_R, P(\psi | \chi v) = \begin{cases} 1 & \text{si } \chi = v = 0 \\ 0 & \text{sinon} \end{cases} \quad (6.15)$$

L'inférence bayésienne donne alors, pour ces formes paramétriques, l'expression 6.16.

$$P(\psi | \vec{\phi} v_x v_z r_y \lambda) = \int_{(\chi, v) = (0, 0)} P(\chi v | \vec{\phi} v_x v_z r_y \lambda) d\chi dv \quad (6.16)$$

Or, le domaine d'intégration de l'expression 6.16 est de mesure nulle. L'intégrale est donc nulle elle aussi.

Avec ce choix de décomposition et de formes paramétriques, la probabilité d'une réponse *flat* est toujours nulle. Ce choix est dû à une définition très restrictive d'une réponse

plate. L'expression 6.15 signifie qu'une telle réponse n'est donnée que lorsque l'observateur est *certain* que le plan est *exactement* fronto-parallèle. Cette contrainte est, de manière évidente, trop forte et nous devons la relâcher en spécifiant une probabilité non nulle sur un support de mesure non nulle. La problématique qui se pose alors est de choisir la dispersion de la famille de distributions choisies.

Pour cette expérience, nous choisissons une forme paramétrique proportionnelle à une gaussienne centrée en $(0, 0)$. Cette distribution privilégie les configurations (χ, v) du plan proche du plan fronto-parallèle. Elle est paramétrée par la matrice de covariance $\vec{\sigma}_\Psi$ de la gaussienne et un facteur \beth (*beth*) compris entre 0 et 1. Pour ne pas privilégier de direction, nous choisissons une expression invariante par changement d'orientation. Cela se traduit par une matrice de covariance proportionnelle à l'identité : $\vec{\sigma}_\Psi = \sigma_\psi \cdot \text{Id}_{2 \times 2}$.

En résumé, nous écrivons la forme paramétrique 6.17.

$$\forall \chi, v \in \mathcal{D}_X \times \mathcal{D}_Y, P(\psi | \chi v) = \beth \exp\left(-\frac{1}{2} \frac{\chi^2 + v^2}{\sigma_\psi^2}\right) \quad (6.17)$$

L'inférence bayésienne sur ce modèle donne l'expression 6.18 de la probabilité d'une réponse *flat* pour une simulation.

$$P(\psi | \vec{\phi} v_x v_z r_y \lambda) = \int_{\chi, v \in \mathcal{D}_X \times \mathcal{D}_Y} \beth \exp\left(-\frac{1}{2} \frac{\chi^2 + v^2}{\sigma_\psi^2}\right) P(\chi v | \vec{\phi} v_x v_z r_y \lambda) d\chi dv \quad (6.18)$$

Il est important de noter que pour traiter les résultats de notre modèle de perception du plan par le flux optique, le modèle *flat* nécessite l'introduction de paramètres (\beth et σ_ψ). Ce n'est pas le cas du modèle de réponse *top-far* ou des traitements du chapitre 5 donnant les résultats du modèle.

Résultats du modèle

Les résultats de notre modèles sont présentés figure 6.6. Les valeurs choisies des paramètres du modèle *flat* sont $\beth = 2.5$ et $\sigma_\psi = 0.2$.

Les résultats expérimentaux sont rappelés figure 6.7.

On remarque en premier lieu que les probabilités de réponse *top-far* sont bien croissantes dans les conditions A et C alors qu'en comparaison ce n'est pas le cas de la condition B. De plus, la transition entre une faible probabilité de réponse *top-far*, pour un stimulus négatif et une forte probabilité, pour un stimulus positif est plus aiguë.

Pour les réponses *flat*, on remarque que les courbes de probabilité sont symétriques et que le pic est plus large en condition A qu'en condition C.

Ces caractéristiques sont les mêmes que pour les résultats expérimentaux.

Cependant, on peut noter des différences entre les résultats de notre modèle et ceux de l'expérience. En premier lieu, la réponse *top-far* en condition B est presque constante à 50 % alors que les résultats expérimentaux présentent un biais en défaveur d'une réponse *top-far*. On remarque que cette réponse est très légèrement croissante. Ensuite, une autre différence importante consiste en la faiblesse des probabilités de réponse *flat* en condition B.

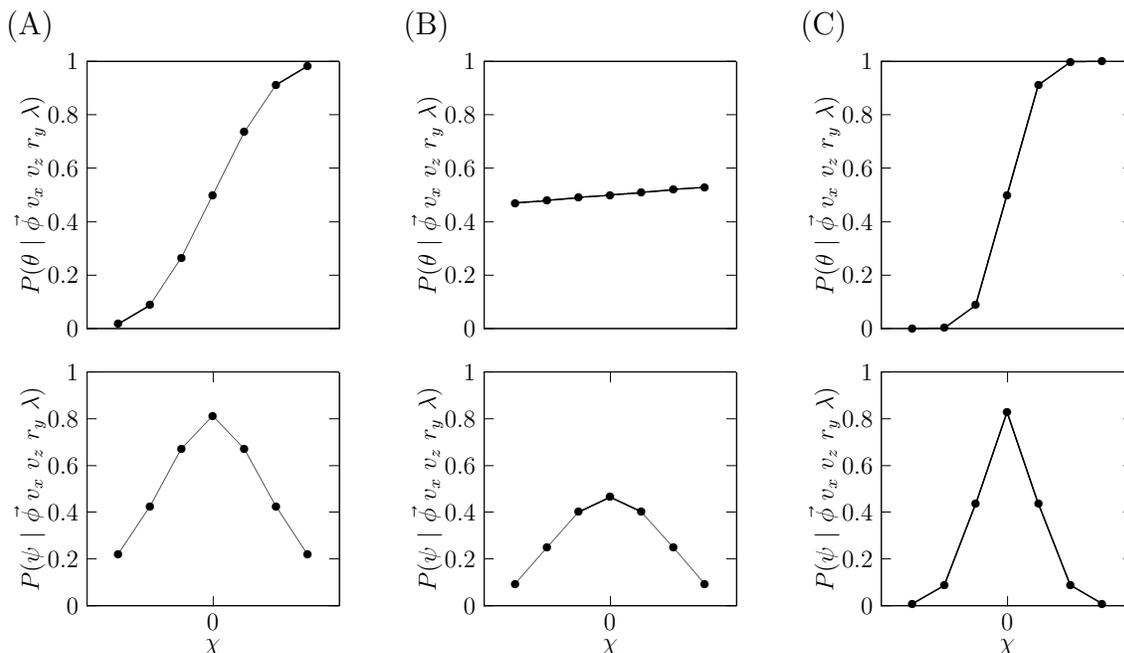


FIG. 6.6 – Résultats du modèle pour l'expérience 6.1.3. La ligne du haut présente la probabilité de réponse *top-far* pour chacune des conditions A, B et C.

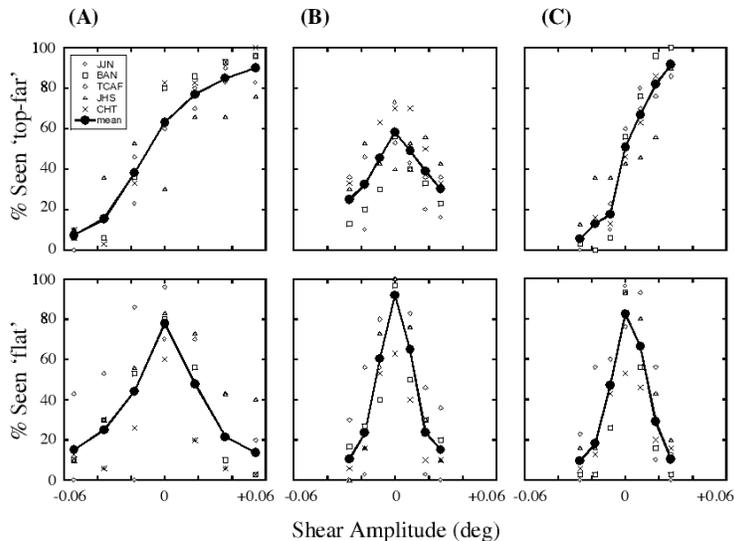


FIG. 6.7 – Rappel des résultats expérimentaux (figure 6.5). En haut sont présentés les pourcentages de perception *top-far* en fonction de l'amplitude du stimulus pour les trois conditions A, B et C. De la même manière, les graphes du bas présentent les pourcentages de perception *flat*.

Explication

Nous expliquons d'abord les réponses en condition B par la rigidité à l'ordre un, puisque nous sommes en petit champ de vision. Ensuite nous discutons de la plus grande précision

des réponses *top-far* en condition C qu'en A et l'expliquons par la plus grande rigidité des réponses d'inclinaison élevé en condition C qu'en condition A.

Condition B Nous examinons tout d'abord la condition B avant de comparer les conditions A et C. Pour étudier la condition B, nous nous référons aux expériences précédentes. En effet dans ce cas, le stimulus figure un plan incliné en rotation autour de l'axe vertical. Il est présenté à un observateur immobile. Cette condition est la même que notre expérience de référence (3.1 et 4.6) ou que la condition *immobile* de l'expérience 5.3.

Les résultats de cette dernière expérience sont la distribution d'erreur sur l'orientation. Cette distribution est bimodale avec deux pics autour d'une erreur nulle ou 180° . Le rapport de probabilité de ces deux pics est le taux d'inversion qui est légèrement inférieur à 50% (expérience 5.2).

Pour la question *flat*, la symétrie des réponses provient simplement de la symétrie du flux optique d'ordre 1. En effet, la distribution sur l'orientation pour un stimulus négatif est la symétrique, par rapport à une orientation nulle de celle obtenue pour un stimulus positif de même amplitude. Comme la probabilité de réponse *flat* est l'intégrale du produit de cette distribution par une fonction symétrique par rapport à l'orientation nulle, la courbe de probabilité en fonction du stimulus est symétrique. En notant $P(\chi v | s)$ la distribution de réponse $P(\chi v | \vec{\phi} v_x v_z r_y \lambda)$ de notre modèle pour le stimulus d'amplitude signée s , on peut écrire :

$$\begin{aligned}
& P(\psi | -s) \\
= & \int_{\chi, v \in \mathcal{D}_X \times \mathcal{D}_Y} P(\psi | \chi v) P(\chi v | -s) d\chi dv \quad \text{par définition} \\
= & \int_{\chi, v \in \mathcal{D}_X \times \mathcal{D}_Y} P(\psi | \chi v) P(-\chi -v | s) d\chi dv \quad \text{par symétrie de notre modèle} \\
= & \int_{\chi, v \in \mathcal{D}_X \times \mathcal{D}_Y} P(\psi | -\chi -v) P(-\chi -v | s) d\chi dv \quad \text{par symétrie du modèle } flat \\
= & \int_{\chi, v \in \mathcal{D}_X \times \mathcal{D}_Y} P(\psi | \chi v) P(\chi v | s) d\chi dv \quad \text{par changement de variable dans l'intégrale} \\
= & P(\psi | s)
\end{aligned}$$

Ce raisonnement est valide aussi pour les conditions A et C, ce qui explique les symétries des trois graphes de réponses *flat*.

Conditions A et C Nous nous intéressons maintenant à la différence de perception entre les conditions A et C. Pour cela, nous rappelons dans un premier temps les mouvements du plan avant de comparer les probabilités de perception de chacune de ces deux conditions en fonction de l'inclinaison perçue. Nous allons montrer que relativement à la condition C, l'inclinaison perçue est sous-estimée en condition A ; c'est ce qui explique le profil moins marqué en condition A qu'en condition C.

La figure 6.8 rappelle les stimulus des conditions A et B. En condition A, l'observateur

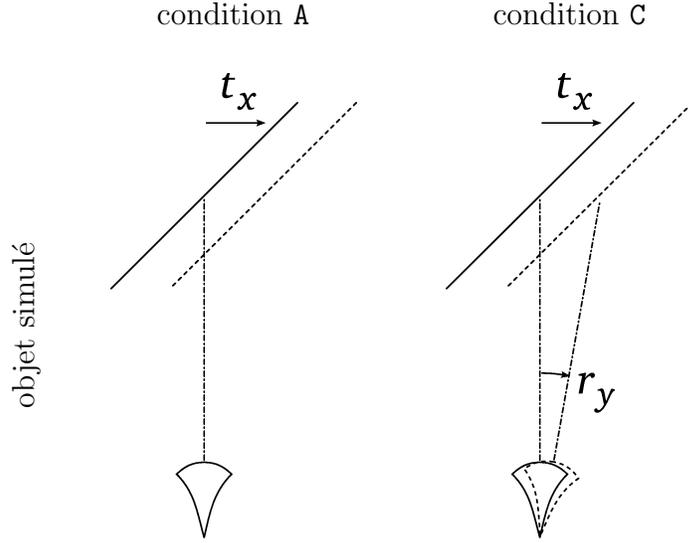


FIG. 6.8 – Configurations des objets simulés pour les conditions A et C.

	\vec{T}	$\vec{\Omega}$
A	$\begin{pmatrix} v_0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$
C	$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ v_0 \\ 0 \end{pmatrix}$

TAB. 6.1 – Mouvements relatifs pour les conditions A et C.

est immobile. Son mouvement propre est donc nul. Mathématiquement, cela correspond à $v_x^A = v_z^A = r_y^A = 0$.

En condition C, le seul mouvement de l'observateur est le suivi de l'objet du regard. Cela s'écrit $v_x^C = v_z^C = 0$ et $r_y^C = -v_0$ avec v_0 la vitesse de translation du plan.²

Les mouvements relatifs respectifs des conditions A et C sont résumé par le tableau 6.1.

Enfin, la table 6.2 présente le mouvement relatif attendu d'après le mouvement propre. Il s'agit du mouvement relatif d'un objet immobile dans le référentiel allocentré.

Le flux optique d'ordre 0, se restreint aux composantes du plan fronto-parallèle de la translation relative. Il est donc nul en condition C. En condition A, sa composante selon l'axe transversal est $\phi_x^0 = v_0$. Le flux optique d'ordre 1 est identique dans les deux cas et se réduit à la seule composante ${}_x\Phi_x^1$. Pour une inclinaison simulée $\hat{\chi}$, sa valeur est ${}_x\phi_x^1 = \chi \cdot v_0$.

²Rappelons que l'observateur est à distance unité du plan.

$$\begin{array}{cc}
& \vec{T} & \vec{\Omega} \\
\mathbf{A} & \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\mathbf{C} & \begin{pmatrix} -v_0 \\ 0 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ v_0 \\ 0 \end{pmatrix}
\end{array}$$

TAB. 6.2 – Mouvements attendus pour les conditions **A** et **C**.

On remarque qu'en condition **A**, le flux optique d'ordre un est le produit de l'inclinaison par la translation relative. En condition **C**, c'est le produit de l'inclinaison par la rotation relative.

Intéressons-nous à la probabilité de perception d'une inclinaison $(\chi, v) = (\alpha \hat{\chi}, 0)$ donnée, avec $\alpha \in \mathbb{R}_+^*$. En condition **A**, le mouvement relatif devient une translation de vitesse $\frac{v_0}{\alpha}$. En condition **C**, il s'agit d'une rotation de vitesse $\frac{v_0}{\alpha}$. Le rapport $\gamma(\alpha)$ des probabilités de perception de la position $(\alpha \hat{\chi}, 0)$ en conditions **A** et **C** pour ces mouvements peut s'écrire :

$$\begin{aligned}
& \gamma(\alpha) \\
&= \frac{P^A(\chi, v \mid \vec{\phi}, v_x, v_z, r_y, \lambda)}{P^C(\chi, v \mid \vec{\phi}, v_x, v_z, r_y, \lambda)} \\
&= \frac{\mathcal{N}_{0, \sigma_T} \left(\frac{v_0}{\alpha} \right) \mathcal{N}_{0, \sigma_\Omega} (0) \mathcal{N}_{v_0, \sigma_{\Phi^0}} \left(\frac{v_0}{\alpha} \right)}{\mathcal{N}_{-v_0, \sigma_T} (0) \mathcal{N}_{v_0, \sigma_\Omega} \left(\frac{v_0}{\alpha} \right) \mathcal{N}_{0, \sigma_{\Phi^0}} (0)} \\
&= \frac{\exp -\frac{1}{2} \left[\left(\frac{v_0}{\alpha \sigma_T} \right)^2 + \left(\frac{v_0 \cdot (1-\alpha)}{\alpha \sigma_{\Phi^0}} \right)^2 \right]}{\exp -\frac{1}{2} \left[\left(\frac{-v_0}{\sigma_T} \right)^2 + \left(\frac{v_0 \cdot (1-\alpha)}{\alpha \sigma_\Omega} \right)^2 \right]} \\
&= \exp -\frac{1}{2} \left[\left(\frac{v_0}{\alpha} \right)^2 \left(\frac{1}{\sigma_T^2} + \left(\frac{(1-\alpha)}{\sigma_{\Phi^0}} \right)^2 \right) - \left(\frac{\alpha}{\sigma_T} \right)^2 - \left(\frac{(1-\alpha)}{\sigma_\Omega} \right)^2 \right] \\
&= \exp -\frac{v_0^2}{2\alpha^2} \left[\left(\frac{1}{\sigma_{\Phi^0}^2} - \frac{1}{\sigma_\Omega^2} - \frac{1}{\sigma_T^2} \right) \alpha^2 - 2 \left(\frac{1}{\sigma_{\Phi^0}^2} - \frac{1}{\sigma_\Omega^2} \right) \alpha + \frac{1}{\sigma_T^2} + \frac{1}{\sigma_{\Phi^0}^2} - \frac{1}{\sigma_\Omega^2} \right]
\end{aligned}$$

On remarque donc que le rapport $\gamma(\alpha)$ est l'exponentielle du produit d'un terme monotone par une forme quadratique en α .

Le terme en α^2 est négatif (car $\sigma_T < \sigma_{\Phi^0}$). Cela signifie que le rapport $\gamma(\alpha)$ est très grand pour les valeurs extrêmes de la variable α . Ainsi pour une inclinaison très faible ou très élevée, la probabilité de perception sera plus importante en condition **A** qu'en condition **C**.

De plus, $\alpha = 1$ est une racine de ce polynôme et la demi-somme des racines est

supérieure à 1. En effet, il suffit de faire le demi-rapport des coefficients des termes en α et α^2 . Ainsi, la forme quadratique est positive pour des valeurs de α immédiatement supérieures à un. Cela se traduit par une probabilité de perception d'une inclinaison légèrement supérieure à l'inclinaison simulée plus importante en condition C qu'en condition A.

Ainsi, la moyenne de l'inclinaison perçue est plus faible en condition A qu'en condition C. La discrimination est donc plus rapide dans cette dernière condition ce qui explique les profils plus marqués des courbes de réponse. Dit autrement, la courbe de réponse en condition A est dilatée selon l'axe des abscisses par rapport à la condition C.

6.1.5 Amélioration du modèle *top-far*

Dans cette section, nous souhaitons améliorer la réponse du modèle à la question *top-far* en condition B principalement.

Sous-modèle *top-far*

Pour améliorer le modèle *top-far*, nous commençons par le spécifier de manière inverse. Nous vérifions que nous pouvons obtenir une expression équivalente avant d'utiliser une nouvelle forme paramétrique.

Inversion du modèle Le modèle *top-far* présenté en 6.1.4 repose sur une expression directe de la réponse connaissant la position du plan : $P(\Theta | X)$. La spécification actuelle est celle d'une fonction de Heaviside :

$$\forall \chi \in \mathcal{D}_X, P(\theta | \chi) = \begin{cases} 1 & \text{si } \chi > 0 \\ 0 & \text{sinon} \end{cases} \quad (6.19)$$

Le modèle inverse se décompose de la manière suivante :

$$P(\Theta | X) = P(\Theta)P(X | \Theta) \quad (6.20)$$

Nous choisissons de faire l'hypothèse que le plan a autant de chance d'être en position *top-far* ou pas. Ainsi l'*a priori* sur Θ est uniforme ($P(\theta) = P(\bar{\theta}) = \frac{1}{2}$).

Pour la position du plan sachant qu'il est *top-far*, nous choisissons une distribution uniforme sur un quart de sphère. Cela revient à ne considérer que la moitié de la distribution sur (X, Y) . On écrit ainsi la distribution 6.21.

$$\begin{aligned} \forall \chi \in \mathcal{D}_X, P(\chi | \theta) &= \int_{v \in \mathcal{D}_Y} P(\chi | v | \theta) dv \\ &= \begin{cases} \int_{v \in \mathcal{D}_Y} \frac{1}{\pi} (1 + \chi^2 + v^2)^{-\frac{3}{2}} dv & \text{si } \chi > 0 \\ 0 & \text{sinon} \end{cases} \\ &= \begin{cases} \frac{2}{\pi} \frac{1}{1 + \chi^2} & \text{si } \chi > 0 \\ 0 & \text{sinon} \end{cases} \end{aligned} \quad (6.21)$$

On définit la distribution sur la position du plan dans le cas non *top-far* de manière symétrique :

$$\forall \chi \in \mathcal{D}_X, P(\chi | \bar{\theta}) = \begin{cases} 0 & \text{si } \chi > 0 \\ \frac{2}{\pi} \frac{1}{1+\chi^2} & \text{sinon} \end{cases} \quad (6.22)$$

Nous vérifions, à l'aide de l'inférence bayésienne, que cette spécification permet de retrouver l'expression 6.9 :

$$\begin{aligned} \forall \chi \in \mathcal{D}_X, P(\theta | \chi) &= \frac{P(\theta)P(\chi | \theta)}{P(\theta)P(\chi | \theta) + P(\bar{\theta})P(\chi | \bar{\theta})} \\ &= \begin{cases} \frac{\frac{1}{2} \frac{2}{\pi} \frac{1}{1+\chi^2}}{\frac{1}{2} \frac{2}{\pi} \frac{1}{1+\chi^2} + \frac{1}{2} \times 0} & \text{si } \chi > 0 \\ \frac{\frac{1}{2} \times 0}{\frac{1}{2} \times 0 + \frac{1}{2} \frac{2}{\pi} \frac{1}{1+\chi^2}} & \text{sinon} \end{cases} \\ &= \begin{cases} 1 & \text{si } \chi > 0 \\ 0 & \text{sinon} \end{cases} \end{aligned}$$

Ce modèle est très précis. Il indique qu'il est impossible qu'un observateur se trompe, même lorsque le plan perçu est très proche du plan fronto-parallèle. On remarque aussi sa spécification très discontinue. Nous modifions le modèle inverse de manière à régler ces insuffisances.

Sigmoïde Nous choisissons, dans un premier temps, un modèle différent pour la probabilité de la position sachant *top-far* ou non. Ces distributions peuvent être considérées comme les définitions, par un observateur, de ce qu'est un plan *top-far* ou non.

Nous construisons notre modèle en supposant qu'un plan *top-far* est orienté vers la droite et suffisamment incliné. Nous choisissons, par exemple, une distribution gaussienne centrée sur l'équivalent d'un plan incliné à droite d'un angle de $\frac{\pi}{4}$, c'est-à-dire d'une valeur $\chi = 1$. On obtient :

$$\forall \chi \in \mathcal{D}_X, P(\chi | \theta) = \mathcal{N}_{1, \sigma_\theta}(\chi) \quad (6.23)$$

Sans autre information, nous supposons une symétrie entre les positions *top-far* et non *top-far*. La distribution sur la position du plan sachant qu'il n'est pas *top-far* s'écrit alors :

$$\forall \chi \in \mathcal{D}_X, P(\chi | \bar{\theta}) = \mathcal{N}_{-1, \sigma_\theta}(\chi) \quad (6.24)$$

Nous calculons la réponse à la question $P(\Theta | X)$.

$$\begin{aligned} \forall \chi \in \mathcal{D}_X, P(\theta | \chi) &= \frac{P(\theta)P(\chi | \theta)}{P(\theta)P(\chi | \theta) + P(\bar{\theta})P(\chi | \bar{\theta})} \\ &= \frac{\frac{1}{2} \mathcal{N}_{1, \sigma_\theta}(\chi)}{\frac{1}{2} \mathcal{N}_{1, \sigma_\theta}(\chi) + \frac{1}{2} \mathcal{N}_{-1, \sigma_\theta}(\chi)} \\ &= \frac{1}{1 + \frac{\mathcal{N}_{-1, \sigma_\theta}(\chi)}{\mathcal{N}_{1, \sigma_\theta}(\chi)}} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{1 + \exp\left(-\frac{1}{2\sigma_\theta^2}(\chi + 1)^2 + \frac{1}{2\sigma_\theta^2}(\chi - 1)^2\right)} \\
&= \frac{1}{1 + \exp\left(-\frac{4\chi}{2\sigma_\theta^2}\right)} \\
&= \text{Sigmoid}_{0,1/2\sigma_\theta^2}(\chi)
\end{aligned}$$

où $\text{Sigmoid}_{\alpha,\beta}(V)$ est la fonction sigmoïde de paramètre α et β sur la variable V :

$$\forall v \in \mathcal{D}_V, \text{Sigmoid}_{\alpha,\beta}(v) = \frac{1}{1 + \exp[-4\beta(v - \alpha)]} \quad (6.25)$$

La réponse à la question $P(\Theta | X)$ est une sigmoïde de point d'inflexion $\alpha = 0$.³ L'intérêt d'inverser le modèle est de pouvoir définir la conditionnelle $P(\Theta | X)$ en spécifiant ce à quoi un observateur s'attend dans chacun des cas *top-far* ou non. Nous pensons que cette écriture est plus facile à concevoir et spécifier (voir section 6.1.6).

Résultats du modèle

La figure 6.9 compare les réponses expérimentales à la question *top-far* avec les résultats de notre modèle.

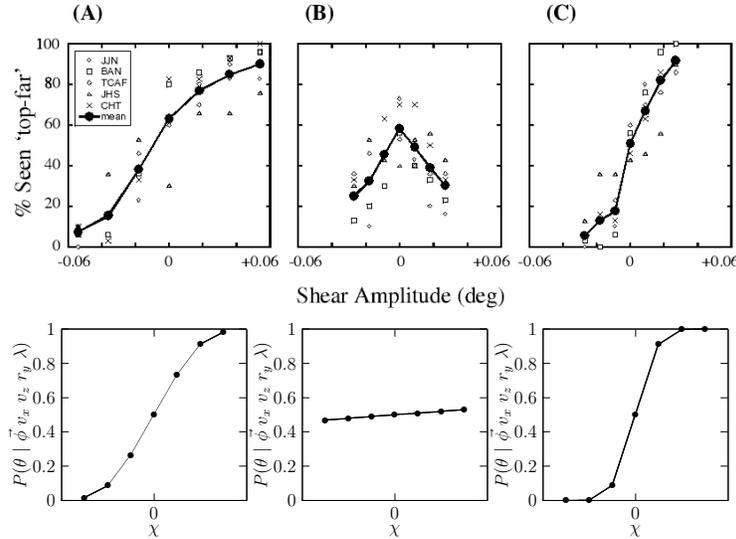


FIG. 6.9 – Réponses à la question *top-far*. En haut, résultats expérimentaux. En bas, résultats du modèle bayésien avec le nouveau modèle *top-far*.

Ce modèle est plus satisfaisant du point de vue de sa spécification. Cependant ses résultats sont toujours très proches des résultats initiaux.

³La limite de la fonction sigmoïde lorsque le paramètre *beta* tend vers l'infini est une fonction de Heaviside.

Explication

Nous montrons que cette spécification ne peut pas reproduire les résultats expérimentaux. Plus précisément, une sigmoïde ne permet pas d'obtenir des résultats pour lesquels la probabilité de réponse *top-far* est supérieure puis inférieure à 50% à mesure de l'augmentation de la valeur absolue de l'amplitude du stimulus.

Rappelons que cette expérience est effectuée en petit champ de vision et donc, si l'on néglige l'effet du flux d'ordre deux, la distribution sur la position (X, Y) du plan est symétrique par rapport à l'origine $(0, 0)$. On peut montrer que, avec la spécification actuelle, la probabilité de réponse *top-far* est égale à $\frac{1}{2}$, indépendamment du signe ou de l'amplitude du stimulus.

$$\begin{aligned}
& P(\theta \mid \vec{\phi} v_x v_z r_y \lambda) \\
&= \int_{\chi \in \mathcal{D}_X} \text{Sigmoid}_{0,1/2\sigma_\theta^2}(\chi) \int_{v \in \mathcal{D}_Y} P(\chi v \mid \vec{\phi} v_x v_z r_y \lambda) d\chi dv \\
&= \int_{\chi \in \mathbb{R}_+} \text{Sigmoid}_{0,1/2\sigma_\theta^2}(\chi) \int_{v \in \mathcal{D}_Y} P(\chi v \mid \vec{\phi} v_x v_z r_y \lambda) d\chi dv \\
&\quad + \int_{\chi \in \mathbb{R}_-} \text{Sigmoid}_{0,1/2\sigma_\theta^2}(\chi) \int_{v \in \mathcal{D}_Y} P(\chi v \mid \vec{\phi} v_x v_z r_y \lambda) d\chi dv \\
&= \int_{\chi \in \mathbb{R}_+} \text{Sigmoid}_{0,1/2\sigma_\theta^2}(\chi) \int_{v \in \mathcal{D}_Y} P(\chi v \mid \vec{\phi} v_x v_z r_y \lambda) d\chi dv \\
&\quad + \int_{\chi \in \mathbb{R}_+} \text{Sigmoid}_{0,1/2\sigma_\theta^2}(-\chi) \int_{v \in \mathcal{D}_Y} P(-\chi -v \mid \vec{\phi} v_x v_z r_y \lambda) d\chi dv \\
&= \int_{\chi \in \mathbb{R}_+} \text{Sigmoid}_{0,1/2\sigma_\theta^2}(\chi) \int_{v \in \mathcal{D}_Y} P(\chi v \mid \vec{\phi} v_x v_z r_y \lambda) d\chi dv \\
&\quad + \int_{\chi \in \mathbb{R}_+} \text{Sigmoid}_{0,1/2\sigma_\theta^2}(-\chi) \int_{v \in \mathcal{D}_Y} P(\chi v \mid \vec{\phi} v_x v_z r_y \lambda) d\chi dv \\
&= \int_{\chi \in \mathbb{R}_+} \left(\text{Sigmoid}_{0,1/2\sigma_\theta^2}(\chi) + \text{Sigmoid}_{0,1/2\sigma_\theta^2}(-\chi) \right) \int_{v \in \mathcal{D}_Y} P(\chi v \mid \vec{\phi} v_x v_z r_y \lambda) d\chi dv \\
&= \int_{\chi \in \mathbb{R}_+} 1 \times \int_{v \in \mathcal{D}_Y} P(\chi v \mid \vec{\phi} v_x v_z r_y \lambda) d\chi dv \\
&= \frac{1}{2} \int_{\chi \in \mathcal{D}_X} \int_{v \in \mathcal{D}_Y} P(\chi v \mid \vec{\phi} v_x v_z r_y \lambda) d\chi dv \\
&= \frac{1}{2}
\end{aligned}$$

La sigmoïde de paramètre $\alpha = 0$ ne permet donc pas de reproduire les résultats expérimentaux. De même, un paramètre α négatif (correspondant par exemple à un *a priori* en faveur de *top-far*) résultera en une probabilité de réponse *top-far* toujours supérieure à

$\frac{1}{2}$. En effet, on a :

$$\forall \chi \in \mathcal{D}_X, \begin{cases} \alpha < 0 \Rightarrow \text{Sigmoid}_{\alpha,\beta}(\chi) + \text{Sigmoid}_{\alpha,\beta}(-\chi) > 1 \\ \alpha = 0 \Rightarrow \text{Sigmoid}_{\alpha,\beta}(\chi) + \text{Sigmoid}_{\alpha,\beta}(-\chi) = 1 \\ \alpha > 0 \Rightarrow \text{Sigmoid}_{\alpha,\beta}(\chi) + \text{Sigmoid}_{\alpha,\beta}(-\chi) < 1 \end{cases} \quad (6.26)$$

Comme l'influence du flux optique d'ordre 2 est croissante, il est impossible avec une sigmoïde d'obtenir les résultats expérimentaux de la condition B pour lesquels la probabilité de réponse *top-far* est décroissante avec la valeur absolue de l'amplitude signée du stimulus en passant de plus de 50% pour une amplitude nulle aux alentours de 30%.

6.1.6 Sous-modèle final

Obtention des réponses

Nous modifions une dernière fois notre modèle de manière à pouvoir reproduire les réponses *top-far* de la condition B de l'expérience.

On remarque qu'une spécification symétrique des termes $P(\chi | \theta)$ et $P(\chi | \bar{\theta})$ ne mène qu'à une probabilité uniforme pour la question $P(\theta | \vec{\phi} v_x v_z r_y \lambda)$ puisque la distribution $P(\chi v | \vec{\phi} v_x v_z r_y \lambda)$ est symétrique.⁴ En conséquence, nous modifions la spécification précédente en la rendant dissymétrique.

En partant de la spécification 6.1.5, il y a plusieurs paramètres qui peuvent briser la symétrie. Il y a en premier temps l'*a priori* $P(\Theta)$. Cependant le changer ne modifie dans l'inférence que le paramètre α de la sigmoïde pour la question $P(\theta | \chi)$. De même, changer les moyennes des Gaussiennes ne fait que décaler la position du point d'inflexion (paramètre α) sans changer de forme paramétrique.

Nous choisissons donc de changer les valeurs des écarts-type des Gaussiennes. On note σ_θ et $\sigma_{\bar{\theta}}$ les écarts-type respectifs des distributions $P(\chi | \theta)$ et $P(\chi | \bar{\theta})$. L'inférence s'écrit :

$$\begin{aligned} \forall \chi \in \mathcal{D}_X, P(\theta | \chi) &= \frac{P(\theta)P(\chi | \theta)}{P(\theta)P(\chi | \theta) + P(\bar{\theta})P(\chi | \bar{\theta})} \\ &= \frac{1}{1 + \frac{\mathcal{N}_{-1,\sigma_\theta}(\chi)}{\mathcal{N}_{1,\sigma_{\bar{\theta}}}(\chi)}} \\ &= \frac{1}{1 + \frac{\sigma_{\bar{\theta}}}{\sigma_\theta} \exp\left(-\frac{1}{2\sigma_{\bar{\theta}}^2}(\chi + 1)^2 + \frac{1}{2\sigma_\theta^2}(\chi - 1)^2\right)} \\ &= \frac{1}{1 + \frac{\sigma_{\bar{\theta}}}{\sigma_\theta} \exp -\frac{1}{2} \left[\left(\frac{1}{\sigma_{\bar{\theta}}^2} - \frac{1}{\sigma_\theta^2} \right) \chi^2 + 2 \left(\frac{1}{\sigma_{\bar{\theta}}^2} + \frac{1}{\sigma_\theta^2} \right) \chi + \left(\frac{1}{\sigma_{\bar{\theta}}^2} - \frac{1}{\sigma_\theta^2} \right) \right]} \end{aligned}$$

Dans le cas $\sigma_{\bar{\theta}} > \sigma_\theta$, on peut noter $\hat{\sigma}^2 = \frac{\sigma_\theta^2 \sigma_{\bar{\theta}}^2}{\sigma_{\bar{\theta}}^2 - \sigma_\theta^2}$ et $\hat{\mu} = \frac{\sigma_\theta^2 + \sigma_{\bar{\theta}}^2}{\sigma_{\bar{\theta}}^2 - \sigma_\theta^2}$. On obtient :

$$\forall \chi \in \mathcal{D}_X, P(\theta | \chi) = \frac{1}{1 + \frac{\sigma_{\bar{\theta}}}{\sigma_\theta} \exp \frac{1}{2\hat{\sigma}^2} [(\chi - \hat{\mu})^2 + (1 - \hat{\mu}^2)]}$$

⁴Pour simplifier les calculs nous négligeons le flux optique d'ordre deux.

$$= \frac{1}{1 + \frac{\Upsilon}{\mathcal{N}_{\hat{\mu}, \hat{\sigma}}(\chi)}} \quad (6.27)$$

avec $\Upsilon = \frac{\sigma_{\bar{\theta}}}{\sqrt{2\pi}\hat{\sigma}\sigma_{\theta}} \exp \frac{1-\hat{\mu}^2}{2\hat{\sigma}^2}$.

Résultats du modèle

Nous pouvons alors ajuster les moyennes et écarts-type des distributions Gaussiennes pour construire le modèle auxiliaire et reprendre les résultats de notre modèle global. La figure 6.10 compare les réponses expérimentales à la question *top-far* avec les résultats de notre modèle. Les paramètres utilisés pour ces courbes sont précisés dans le tableau 6.3.

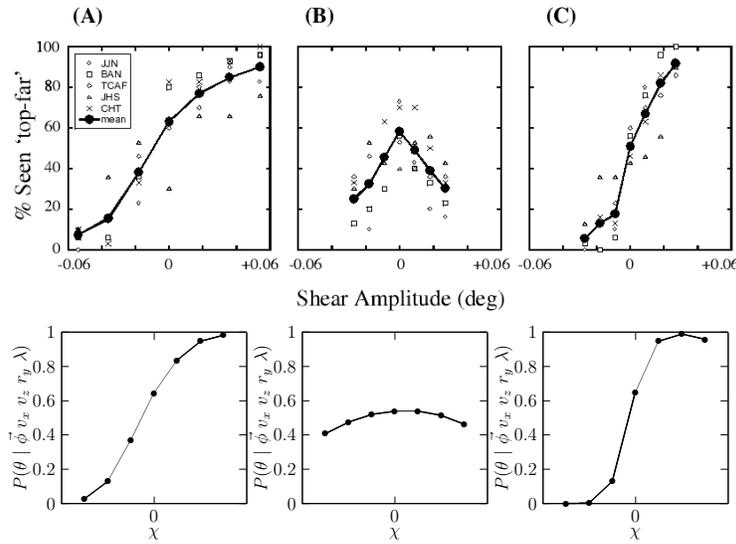


FIG. 6.10 – Réponses à la question *top-far*. En haut, résultats expérimentaux. En bas, résultats du modèle bayésien avec le nouveau modèle *top-far*.

Paramètre	Valeur
$P(\theta)$	$\frac{1}{2}$
μ_{θ}	0.39
$\mu_{\bar{\theta}}$	-1.00
σ_{θ}	0.16
$\sigma_{\bar{\theta}}$	0.40

TAB. 6.3 – Paramètres du modèle auxiliaire *top-far*.

On remarque sur ces derniers résultats que le profil des réponses *top-far* correspond aux résultats expérimentaux. Ceci est vrai à la fois pour les conditions A et C, qui étaient déjà correctement modélisées, que pour la condition B.

Explication

L'inférence $P(\Theta | X)$ du sous-modèle est exprimé par l'équation :

$$\forall \chi \in \mathcal{D}_X, P(\theta | \chi) = \frac{1}{1 + \frac{1}{\mathcal{N}_{\hat{\mu}, \hat{\sigma}}(\chi)}} \quad (6.28)$$

L'inférence globale permettant d'arriver à la probabilité d'une réponse *top-far* s'écrit :

$$P(\theta | \vec{\phi} v_x v_z r_y \lambda) = \sum_{\chi, v \in \mathcal{D}_X \times \mathcal{D}_v} P(\theta | \chi) P(\chi v | \vec{\phi} v_x v_z r_y \lambda) \quad (6.29)$$

On sait que la Gaussienne prend des valeurs négligeables loin de la moyenne $\hat{\mu}$. C'est donc aussi le cas de la probabilité $P(\theta | X)$. La figure 6.11 montre la forme paramétrique de cette fonction.

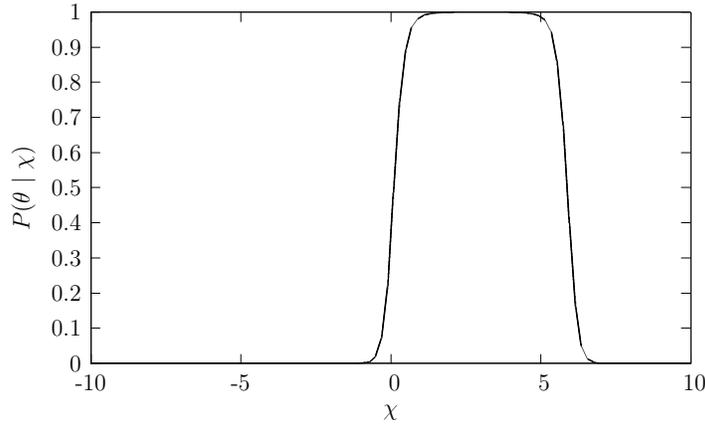


FIG. 6.11 – Probabilité $P(\theta | X)$ d'une réponse *top-far* en fonction de la position X du plan. Les paramètres pour cette courbe sont $\sigma_\theta = \frac{1}{2}$ et $\sigma_{\bar{\theta}} = \frac{\sqrt{2}}{2}$.

Sur cette figure, on voit que si une distribution sur X symétrique possède une fraction non négligeable de sa masse de probabilité à des valeurs importantes, on obtient une probabilité de réponse *top-far* inférieure à 50%. Si, par contre, la masse de la distribution est localisée autour de l'origine $\chi = 0$, alors il est possible d'avoir une probabilité de réponse *top-far* légèrement supérieure à 50%.

Ces résultats se basent sur notre modèle avec une hypothèse supplémentaire portant sur les définitions respectives d'un plan *top-far* et non *top-far*. Les résultats en condition B nécessitent une spécification asymétrique de ces distributions. Dans ce cas précis l'hypothèse que nous avons exprimé est que la définition *a priori* d'un plan *top-far* est plus restreinte en extension que celle d'un plan non *top-far*. Dit autrement, on peut dire qu'un plan non *top-far* est défini de manière moins précise qu'un plan *top-far*.⁵ À ce titre, une perception qui ne correspond tout-à-fait ni à un plan *top-far*, ni à un plan non *top-far*, sera plus probablement attribuée à un plan non *top-far*.

⁵Cette asymétrie est étonnante. Un exemple d'explication pourrait prendre en compte les conditions

6.2 Modèle de flux optique

Une des entrées de notre modèle est l'information sur le flux optique sous une forme paramétrique particulière (2.2.1). Nous proposons dans cette section un modèle bayésien de reparamétrisation du flux optique et discutons de son intégration avec notre modèle de perception du plan. L'objectif est de rester dans un cadre de modélisation unique de manière à pouvoir utiliser des outils généraux de composition de modèles comme on le verra section 6.3.

6.2.1 Principe

La plupart des travaux d'extraction du flux optique de la littérature (voir section 2.1) se basent sur plusieurs images et cherchent à calculer un vecteur vitesse en chaque point de l'image (ou un sous-ensemble généralement dense).

Notre objectif, pour obtenir l'entrée de notre modèle, est de représenter le flux optique par ses composantes d'ordre 0, 1 et 2 définies d'après l'expression suivante :

$$\begin{aligned} \vec{\Phi}(x, y) &= \vec{\Phi}^0 + \vec{\Phi}^1 \cdot {}^t(x, y) + {}^t(x, y) \cdot {}^t\vec{\Phi}^2 \cdot {}^t(x, y) \\ \begin{pmatrix} u_x \\ u_y \end{pmatrix} &= \begin{pmatrix} \Phi_x^0 + {}_x\Phi_x^1 \cdot x + {}_y\Phi_x^1 \cdot y + {}_x\Phi^2 \cdot x^2 + {}_y\Phi^2 \cdot xy \\ \Phi_y^0 + {}_x\Phi_y^1 \cdot x + {}_y\Phi_y^1 \cdot y + {}_x\Phi^2 \cdot xy + {}_y\Phi^2 \cdot y^2 \end{pmatrix} \end{aligned} \quad (6.30)$$

dans laquelle (u_x, u_y) est la vitesse dans l'image du point de coordonnées (x, y) et $\vec{\Phi}^0$, $\vec{\Phi}^1$ et $\vec{\Phi}^2$ les composantes recherchées du flux optique aux ordre 0, 1 et 2 respectivement.

Cette section, consacrée à la génération de l'entrée du modèle du chapitre 4, ne considère que huit composantes du flux optique mais le modèle présenté peut être adapté sans mal aux douze composantes d'un flux quelconque en modifiant simplement les formules du flux d'ordre deux.

L'entrée de ce modèle est un ensemble de points distribués autour de l'origine et dont la vitesse instantanée est connue. Ces données sont des observations du flux optique telles que définies en vision par ordinateur (voir section 2.1). Notre modèle de flux optique procède donc à son prétraitement en vue d'être utilisé par le modèle de perception du chapitre 4.

Le principe de ce modèle est d'utiliser la fusion bayésienne sur chaque point pour recouvrer le flux optique. La fusion consiste principalement en la réutilisation d'un même modèle pour plusieurs variables de même domaine (des observations par exemple) en vue d'inférer des connaissances sur une autre variable. Dans notre cas, chaque vitesse est une observation particulière du flux optique, qui peut être déduit de leur recoupement.

d'éclairage. Un plan *top-far* est mieux illuminé donc possiblement mieux perçu qu'un plan non *top-far* si la lumière vient d'en haut. Des études montrent qu'un *a priori* courant est que la lumière vient plus probablement d'en haut à gauche (Mamassian et Goutcher 2001). Une marginalisation de la précision de la perception sur cet *a priori* donne une précision plus faible de perception d'un plan non *top-far* par rapport à un plan *top-far*.

6.2.2 Modèle naïf

Nous présentons dans un premier temps l'outil de fusion bayésienne avant de l'appliquer à notre modèle.

Fusion Naïve La fusion bayésienne (parfois qualifiée de *naïve*) est une forme de décomposition particulière (Lebeltel *et al.* 2004 ; Pradalier *et al.* 2003). Elle concerne un ensemble $\{O^i\}_{i \in [1;N]}$ de N observations O_i de domaine \mathcal{D}_O et V une variable sur laquelle on veut fusionner. La décomposition s'écrit alors :

$$P(V O^1 \dots O^N) = P(V) \prod_{i \in [1;N]} P(O^i | V) \quad (6.31)$$

dans laquelle $P(V)$ est un *a priori* sur la variable à fusionner et $P(O^i | V)$ une distribution conditionnelle indépendante de l'indice i de l'observation. Cela signifie que le modèle est symétrique par rapport aux observations et qu'en particulier leur ordre n'est pas pertinent.

La question posée est alors $P(V | O^1 \dots O^N)$. Et l'inférence s'écrit :

$$P(V | O^1 \dots O^N) \quad (6.32)$$

$$= \frac{P(V O^1 \dots O^N)}{P(O^1 \dots O^N)} \quad (6.33)$$

$$= \frac{P(V) \prod_{i \in [1;N]} P(O^i | V)}{\sum_{v \in \mathcal{D}_V} P(v) \prod_{i \in [1;N]} P(O^i | v)} \quad (6.34)$$

$$\propto P(V) \prod_{i \in [1;N]} P(O^i | V) \quad (6.35)$$

Variation Dans le cas présent, chaque observation est la vitesse de déplacement d'un point particulier. On a donc $O^i = ((X^i, Y^i), (U_x^i, U_y^i))$.

D'autre part, la variable sur laquelle nous souhaitons fusionner est le flux optique. Comme pour le modèle principal, nous le représentons par sa forme paramétrique $V = \vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2$ avec :

$$\begin{aligned} \vec{\Phi}^0 &= \begin{pmatrix} \Phi_x^0 \\ \Phi_y^0 \end{pmatrix} \\ \vec{\Phi}^1 &= \begin{pmatrix} x \Phi_x^1 & y \Phi_x^1 \\ x \Phi_y^1 & y \Phi_y^1 \end{pmatrix} \\ \vec{\Phi}^2 &= \begin{pmatrix} x \Phi^2 \\ y \Phi^2 \end{pmatrix} \end{aligned}$$

Décomposition Ainsi la décomposition s'écrit simplement :

$$\begin{aligned} &P(\vec{\Phi}^0 \vec{\Phi}^2 \vec{\Phi}^2 X^1 Y^1 U_x^1 U_y^1 \dots X^N Y^N U_x^N U_y^N) \\ &= P(\vec{\Phi}^0 \vec{\Phi}^2 \vec{\Phi}^2) \prod_{i \in [1;N]} P(X^i Y^i U_x^i U_y^i | (\vec{\Phi}^0 \vec{\Phi}^2 \vec{\Phi}^2)) \end{aligned}$$

Formes paramétriques Le facteur important dans la fusion est la distribution $P(O^i | V)$ reliant chaque observation à la variable fusionnée, c'est-à-dire dans notre cas :

$$P(X^i Y^i U_x^i U_y^i | \vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2)$$

Nous utilisons tout d'abord la règle de Bayes et faisons ensuite l'hypothèse que la position d'un point dont on observe la vitesse ne dépend pas du flux optique :

$$\begin{aligned} & P(X^i Y^i U_x^i U_y^i | \vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2) \\ &= P(X^i Y^i | \vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2) P(U_x^i U_y^i | X^i Y^i \vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2) \\ &= P(X^i Y^i) P(U_x^i U_y^i | X^i Y^i \vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2) \end{aligned}$$

Le facteur $P(X^i Y^i)$ est marginalisé lors de la question et peut donc rester non spécifié. Pour le reste, nous supposons que la vitesse observée est proche de la vitesse théorique au point considéré.

Le choix de la distribution précise est à faire en fonction de l'application et éventuellement des caractéristiques du processus fournissant les vitesses en entrée de ce modèle.

Dans notre cas, en l'absence de telles informations, nous choisissons des distributions gaussiennes. Nous écrivons ainsi :

$$P(U_x^i U_y^i | X^i Y^i \vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2) = \mathcal{N}_{(\hat{u}_x^i, \hat{u}_y^i), \vec{\sigma}_U} (U_x^i, U_y^i) \quad (6.36)$$

avec $\vec{\sigma}_U$ une matrice de covariance et

$$\begin{aligned} \begin{pmatrix} \hat{u}_x^i \\ \hat{u}_y^i \end{pmatrix} &= \vec{\Phi}^0 + \vec{\Phi}^1 \cdot {}^t(x^i, y^i) + {}^t(x, y) \cdot {}^t\vec{\Phi}^2 \cdot {}^t(x, y) \\ &= \begin{pmatrix} \Phi_x^0 + \Phi_x^1 \cdot x^i + \Phi_x^2 \cdot x^{i2} + \Phi_x^3 \cdot x^i y^i \\ \Phi_y^0 + \Phi_y^1 \cdot x^i + \Phi_y^2 \cdot x^i y^i + \Phi_y^3 \cdot y^{i2} \end{pmatrix} \end{aligned}$$

d'après la définition de notre paramétrisation du flux optique 6.30.

Pour compléter notre modèle, il reste à spécifier une distribution *a priori* sur les paramètres du flux optique. Nous choisissons d'écrire cette distribution sous la forme d'une gaussienne centrée sur une valeur nulle des paramètres du flux. Par ce choix nous spécifions que les vitesses atteintes par les points sont moins probablement très élevées.

Question et inférence Au final, l'inférence donne la distribution de probabilité sur le flux optique connaissant les vitesses de déplacement d'un ensemble de points de l'image. La question est donc :

$$P(\vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2 | X^1 Y^1 U_x^1 U_y^1 \dots X^N Y^N U_x^N U_y^N) \quad (6.37)$$

Et l'inférence s'écrit, avec notre choix de distribution :

$$\begin{aligned} & P(\vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2 | X^1 Y^1 U_x^1 U_y^1 \dots X^N Y^N U_x^N U_y^N) \\ & \propto P(\vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2) \prod_{i \in [1; N]} P(U_x^i U_y^i | \vec{\Phi}^0 \vec{\Phi}^1 \vec{\Phi}^2) \end{aligned} \quad (6.38)$$

Complexités La complexité porte principalement sur deux aspects : le temps de calcul et l'espace mémoire. Les deux variables que nous considérons pour l'expression de ces complexités sont le nombre d'observations : N , et la taille de l'ensemble des flux optiques considérés : $\text{Card}(\mathcal{D}_{\vec{\Phi}})$. Le domaine de définition du flux optique est théoriquement infini. En pratique, nous ne pouvons représenter qu'un sous-ensemble discret et borné de valeurs et c'est ce sous-ensemble dont nous considérons la taille.

Si l'on s'intéresse à la probabilité d'un flux optique particulier (à un facteur multiplicatif près), la complexité temporelle de l'expression 6.38 est linéaire en nombre d'observations. Pour l'ensemble de la distribution de sortie, le calcul est en $\Theta(N \cdot \text{Card}(\mathcal{D}_{\vec{\Phi}}))$.

La complexité spatiale de l'inférence est en $\Theta(\text{Card}(\mathcal{D}_{\vec{\Phi}}) + N)$.

De ces deux complexités on remarque qu'un facteur important est la taille de l'espace paramétrique du flux optique considéré. Dans notre cas, où il s'agit du flux optique d'un plan, celui-ci est de dimension 8. L'espace à explorer est donc lui aussi de dimension 8 ce qui en empêche une recherche systématique dès que l'on discrétise suffisamment ses composantes.

À noter que, si l'on utilise une telle distribution de probabilité en tant qu'entrée du modèle principal, il devient nécessaire de sommer sur tout cet espace, ce qui est impraticable. Il est donc nécessaire de prendre une décision sur cette sortie pour en faire une entrée de notre modèle de perception du flux optique (voir section 6.3).

Dans le cas particulier d'une vraisemblance gaussienne sur la vitesse connaissant le flux optique et la position, et un *a priori* gaussien sur les composantes du flux optique, l'inférence est plus simple et donne une distribution gaussienne. Son maximum, qui en est aussi la moyenne, peut être obtenu par une méthode des moindres carrés en ajoutant la contrainte quadratique de l'*a priori*. Cette dernière simplification n'est possible qu'avec des distributions gaussiennes.

6.2.3 Modèle découplé

Une solution pour réduire la complexité générale de cette inférence est de découpler autant que possible les composantes du flux. On les rend ainsi indépendantes les unes des autres. D'un point de vue calculatoire, l'espace de sortie devient alors l'union des espaces de chacune des composantes au lieu d'en être le produit cartésien.

En pratique, nous proposons un modèle par ordre du flux. Chaque modèle prend les mêmes observations en entrée et considère les composantes qu'il manipule indépendamment des autres.

Flux d'ordre 0

À chacune des deux coordonnées x et y est associée une composante d'ordre 0 du flux optique. Les modèles associés sont symétriques et nous ne détaillons dans cette partie que le modèle de reparamétrisation pour la composante Φ_x^0 .

Ce modèle est un modèle de fusion bayésienne sur le flux Φ_x^0 avec pour seules observations les vitesses horizontales des points de l'image $\{U_x^i\}_{i \in [1;N]}$. La décomposition est

alors :

$$P(\Phi_x^0 U_x^1 \dots U_x^N) = P(\Phi_x^0) \prod_{i \in [1;N]} P(U_x^i | \Phi_x^0) \quad (6.39)$$

Nous considérons dans ce modèle que le flux d'ordre 0 est la vitesse moyenne des points de l'image. C'est une approximation négligeant l'influence d'une partie du flux d'ordre deux.

En faisant des statistiques sur les vitesses des points, on peut obtenir l'écart-type entre la vitesse et le flux d'ordre 0. Le principe du maximum d'entropie nous permet alors d'écrire la distribution $P(U_x^i | \Phi_x^0)$ de la vitesse connaissant le flux d'ordre 0 comme une distribution gaussienne centrée en cette valeur.

Comme pour le modèle naïf 6.2.2, nous spécifions la distribution $P(\Phi_x^0)$ comme une gaussienne centrée en 0 et d'écart-type $\sigma_{\Phi_x^0}$.

L'inférence s'écrit alors :

$$\begin{aligned} \forall \phi_x^0 \in \mathcal{D}_{\Phi_x^0}, \quad & P(\phi_x^0 | u_x^1 \dots u_x^N) \\ & \propto P(\phi_x^0) \prod_{i \in [1;N]} P(u_x^i | \phi_x^0) \\ & \propto P(\phi_x^0) \exp \left(-\frac{1}{2\sigma_{\phi_x^0}^2} \sum_{i \in [1;N]} (u_x^i - \phi_x^0)^2 \right) \end{aligned} \quad (6.40)$$

Avec ce choix de Gaussiennes, la vraisemblance opère à nouveau une méthode des moindres carrés. Elle calcule alors une distribution gaussienne centrée sur la moyenne des observations.

Dans le cas général de distributions quelconques, la complexité en temps de cette inférence est $\Theta(N \cdot \text{Card}(\mathcal{D}_{\Phi_x^0}))$.

Flux d'ordre 1

Comme pour le flux d'ordre 0, nous considérons les vitesses selon x et y de façon indépendante. L'indépendance entre les flux optiques d'ordre 1 et 0 sachant les vitesses repose sur l'élimination des termes constants. Pour deux points (x^1, y^1) et (x^2, y^2) d'un plan, leurs vitesses selon l'axe x s'écrivent :

$$\begin{aligned} u_x^1 &= \phi_x^0 + {}_x\phi_x^1 \cdot x^1 + {}_y\phi_x^1 \cdot y^1 + {}_x\phi_x^2 \cdot (x^1)^2 + {}_y\phi_x^2 \cdot (y^1)^2 \\ u_x^2 &= \phi_x^0 + {}_x\phi_x^1 \cdot x^2 + {}_y\phi_x^1 \cdot y^2 + {}_x\phi_x^2 \cdot (x^2)^2 + {}_y\phi_x^2 \cdot (y^2)^2 \end{aligned}$$

En faisant la différence, on obtient :

$$\begin{aligned} u_x^2 - u_x^1 &= {}_x\phi_x^1 \cdot (x^2 - x^1) + {}_y\phi_x^1 \cdot (y^2 - y^1) \\ &\quad + {}_x\phi_x^2 \cdot ((x^2)^2 - (x^1)^2) + {}_y\phi_x^2 \cdot ((y^2)^2 - (y^1)^2) \end{aligned} \quad (6.41)$$

$$\begin{aligned} u_x^2 &= u_x^1 + {}_x\phi_x^1 \cdot (x^2 - x^1) + {}_y\phi_x^1 \cdot (y^2 - y^1) \\ &\quad + {}_x\phi_x^2 \cdot ((x^2)^2 - (x^1)^2) + {}_y\phi_x^2 \cdot ((y^2)^2 - (y^1)^2) \end{aligned} \quad (6.42)$$

À partir de cette constatation, nous construisons un modèle portant sur les composantes ${}_x\Phi_x^1$ et ${}_y\Phi_x^1$ ainsi que les observations groupées par paires de points. Pour la suite, nous supposons N pair, sans perte de généralité. La décomposition s'écrit :

$$\begin{aligned} & P({}_x\Phi_x^1 \ {}_y\Phi_x^1 \ U_x^1 \ X^1 \ Y^1 \ \dots \ U_x^N \ X^N \ Y^N) \\ = & P({}_x\Phi_x^1 \ {}_y\Phi_x^1) \prod_{i \in [1; N/2]} \left\{ \begin{array}{l} P(X^{2i} \ X^{2i+1} \ Y^{2i} \ Y^{2i+1} \ U_x^{2i}) \\ P(U_x^{2i+1} \mid {}_x\Phi_x^1 \ {}_y\Phi_x^1 \ X^{2i} \ X^{2i+1} \ Y^{2i} \ Y^{2i+1} \ U_x^{2i}) \end{array} \right. \quad (6.43) \end{aligned}$$

Comme précédemment, le facteur $P({}_x\Phi_x^1 \ {}_y\Phi_x^1)$ est spécifié comme un *a priori* gaussien centré en $(0, 0)$. Le facteur $P(X^{2i} \ X^{2i+1} \ Y^{2i} \ Y^{2i+1} \ U_x^{2i})$ peut rester non spécifié car les valeurs de ces variables seront connues au moment de la question. Enfin, le facteur $P(U_x^{2i+1} \mid {}_x\Phi_x^1 \ {}_y\Phi_x^1 \ X^{2i} \ X^{2i+1} \ Y^{2i} \ Y^{2i+1} \ U_x^{2i})$ peut s'écrire comme une Gaussienne centrée sur l'expression 6.44 :

$$\hat{u}_x^{2i+1} = u_x^{2i} + {}_x\phi_x^1 \cdot (x^{2i+1} - x^{2i}) + {}_y\phi_x^1 \cdot (y^{2i+1} - y^{2i}) \quad (6.44)$$

Après simplifications, l'inférence s'écrit :

$$\begin{aligned} & \forall {}_x\phi_x^1, {}_y\phi_x^1 \in \mathcal{D}_{{}_x\Phi_x^1} \times \mathcal{D}_{{}_y\Phi_x^1}, \\ & P({}_x\phi_x^1 \ {}_y\phi_x^1 \mid u_x^1 \ x^1 \ y^1 \ \dots \ u_x^N \ x^N \ y^N) \\ \propto & P({}_x\phi_x^1 \ {}_y\phi_x^1) \prod_{i \in [1; N/2]} P(u_x^{2i+1} \mid {}_x\Phi_x^1 \ {}_y\Phi_x^1 \ x^{2i} \ x^{2i+1} \ y^{2i} \ y^{2i+1} \ u_x^{2i}) \\ \propto & P({}_x\phi_x^1 \ {}_y\phi_x^1) \exp \left(-\frac{1}{2\sigma_{\phi_x^1}^2} \sum_{i \in [1; N/2]} (u_x^{2i+1} - \hat{u}_x^{2i+1})^2 \right) \end{aligned}$$

Le principe de fonctionnement de ce modèle est d'évaluer chaque valeur du flux optique. Cette évaluation se base sur l'erreur réalisée sur la prédiction d'une vitesse en un point en ce basant sur celle d'un autre point et la valeur considérée du flux.

Ce modèle de prédiction est incomplet, puisqu'il ne prend pas en compte le flux d'ordre deux. Cependant, nous traduisons cette incomplétude en incertitude sur la prédiction pour l'intégrer au modèle.

Il est à noter que, pour une telle décomposition, les deux composantes ${}_x\Phi_x^1$ et ${}_y\Phi_x^1$ ne peuvent pas être considérées séparément. En effet, dans ce cas, l'hypothèse serait que la deuxième composante du flux ne serait pas nécessaire pour définir la vitesse à l'ordre un. Les résultats de ce modèle confirment l'impossibilité de séparer ces composantes.

Au final, ce modèle du flux d'ordre 1 a une complexité en temps de l'ordre de : $\Theta \left(N \cdot \text{Card} \left(\mathcal{D}_{{}_x\Phi_x^1} \times \mathcal{D}_{{}_y\Phi_x^1} \right) \right)$.

Flux d'ordre 2

Le flux d'ordre 1 est la matrice jacobienne du champ de vitesse, c'est-à-dire les dérivées selon chaque dimension de chacune des composantes de vitesse. Cette notion de dérivée

se retrouve dans l'expression 6.41 sous la forme de la différence de deux observations de vitesses en des points différents.

De la même manière, les composantes du flux d'ordre 2 sont des dérivées partielles de deuxième ordre de la vitesse. Il est donc possible de calculer la vitesse d'un point donné à partir de celles de deux autres points en sachant le flux optique. À l'image du modèle précédent, nous choisissons la probabilité d'une valeur du flux optique en fonction de l'erreur qu'elle entraîne sur les vitesses prévues.

Le modèle porte donc sur les deux composantes ${}_x\Phi^2$ et ${}_y\Phi^2$ du flux optique ainsi que sur les observations, groupées par trois. Nous supposons ainsi, sans perte de généralité, N multiple de 3.

La décomposition s'écrit :

$$\begin{aligned}
& P({}_x\Phi^2 \ {}_y\Phi^2 \ U_x^1 \ U_y^1 \ X^1 \ Y^1 \ \dots \ U_x^N \ U_y^N \ X^N \ Y^N) \\
&= P({}_x\Phi^2 \ {}_y\Phi^2) \\
&\times \prod_{i \in [1; N/3]} \left\{ \begin{array}{l} P(X^{3i} \ X^{3i+1} \ X^{3i+2} \ Y^{3i} \ Y^{3i+1} \ Y^{3i+2} \ U_x^{3i} \ U_x^{3i+1} \ U_y^{3i} \ U_y^{3i+1}) \\ P(U_x^{3i+2} \mid {}_x\Phi^2 \ {}_y\Phi^2 \ X^{3i} \ X^{3i+1} \ X^{3i+2} \ Y^{3i} \ Y^{3i+1} \ Y^{3i+2} \ U_x^{3i} \ U_x^{3i+1}) \\ P(U_y^{3i+2} \mid {}_x\Phi^2 \ {}_y\Phi^2 \ X^{3i} \ X^{3i+1} \ X^{3i+2} \ Y^{3i} \ Y^{3i+1} \ Y^{3i+2} \ U_y^{3i} \ U_y^{3i+1}) \end{array} \right.
\end{aligned} \tag{6.45}$$

Les formes paramétriques sont similaires à celles du modèle de reparamétrisation pour le flux d'ordre 1 :

- $P(X^{3i} \ X^{3i+1} \ X^{3i+2} \ Y^{3i} \ Y^{3i+1} \ Y^{3i+2} \ U_x^{3i} \ U_x^{3i+1} \ U_y^{3i} \ U_y^{3i+1})$: non spécifié (car simplifié à l'inférence) ;
- $P(U_x^{3i+2} \mid {}_x\Phi^2 \ {}_y\Phi^2 \ X^{3i} \ X^{3i+1} \ X^{3i+2} \ Y^{3i} \ Y^{3i+1} \ Y^{3i+2} \ U_x^{3i} \ U_x^{3i+1})$: distribution gaussienne centrée sur la valeur théorique \hat{u}_x^{3i+2} et d'écart-type σ_{ϕ^2} ;
- $P(U_y^{3i+2} \mid {}_x\Phi^2 \ {}_y\Phi^2 \ X^{3i} \ X^{3i+1} \ X^{3i+2} \ Y^{3i} \ Y^{3i+1} \ Y^{3i+2} \ U_y^{3i} \ U_y^{3i+1})$: distribution gaussienne centrée sur la valeur théorique \hat{u}_y^{3i+2} et d'écart-type σ_{ϕ^2} .

La question que l'on pose à ce modèle est simplement :

$$P({}_x\Phi^2 \ {}_y\Phi^2 \mid u_x^1 \ u_y^1 \ x^1 \ y^1 \ \dots \ u_x^N \ u_y^N \ x^N \ y^N) \tag{6.46}$$

L'inférence est semblable à celle du modèle concernant le flux d'ordre 1.

$$\begin{aligned}
& \forall {}_x\phi^2, {}_y\phi^2 \in \mathcal{D}_{{}_x\Phi^2} \times \mathcal{D}_{{}_y\Phi^2}, \\
& P({}_x\phi^2 \ {}_y\phi^2 \mid u_x^1 \ u_y^1 \ x^1 \ y^1 \ \dots \ u_x^N \ u_y^N \ x^N \ y^N) \\
& \propto P({}_x\phi^2 \ {}_y\phi^2) \exp \left(-\frac{1}{2\sigma_{\phi^2}^2} \sum_{i \in [1; N/3]} \left[(u_x^{3i+2} - \hat{u}_x^{3i+2})^2 + (u_y^{3i+2} - \hat{u}_y^{3i+2})^2 \right] \right)
\end{aligned}$$

Cette expression est linéaire en nombre d'observations pour chaque valeur $({}_x\phi^2, {}_y\phi^2)$ du flux optique. Comme pour le modèle du flux optique d'ordre 1, la complexité dans le cas général est en $\Theta \left(N \cdot \text{Card} \left(\mathcal{D}_{{}_x\Phi^2} \times \mathcal{D}_{{}_y\Phi^2} \right) \right)$.

Considérations de complexité

Ce nouveau modèle a pour but de réduire la complexité, à la fois spatiale et temporelle du processus de reparamétrisation du flux optique.

Ce module se décompose en trois sous-modèles indépendants, chacun relatifs à un ordre du flux optique. Le sous-modèle proposé pour l'ordre 0 traite indépendamment une des deux composantes Φ_x^0 ou Φ_y^0 . Il est donc nécessaire de le déployer deux fois. La sortie du sous-modèle proposé pour l'ordre 1 est une distribution sur les deux composantes du flux relatives à une même composante de vitesse. Comme pour le sous-modèle à l'ordre 0, deux instances de ce modèle sont nécessaires pour obtenir les quatre composantes du flux optique d'ordre 1. Enfin, le modèle concernant le flux optique d'ordre 2 traite conjointement les deux composantes ${}_x\Phi^2$ et ${}_y\Phi^2$.

La complexité totale est donc la somme des complexités des 5 sous-systèmes :

$$\begin{aligned} & \Theta \left(N \cdot \left[2 \cdot \text{Card}(\mathcal{D}_{\Phi_x^0}) + 2 \cdot \text{Card}(\mathcal{D}_{x\Phi_x^1} \times \mathcal{D}_{y\Phi_x^1}) + \text{Card}(\mathcal{D}_{x\Phi^2} \times \mathcal{D}_{y\Phi^2}) \right] \right) \\ &= \Theta(N \cdot N_{\Phi}^2) \end{aligned} \quad (6.47)$$

avec N_{Φ} le nombre de cas considéré par composante du flux optique. L'expression 6.47 est à mettre en relation avec la complexité du modèle naïf qui est de l'ordre de $N \cdot N_{\Phi}^8$.

De la même manière, la complexité spatiale passe de $\Theta(N + N_{\Phi}^8)$ pour le modèle naïf à seulement $\Theta(N + N_{\Phi}^2)$.

Notons, encore une fois, qu'avec des vraisemblances gaussiennes, il existe des formules analytiques donnant le résultat de l'inférence. Cependant les distributions d'erreurs des vitesses fournies en entrée ne sont en général pas gaussiennes et dépendent des algorithmes utilisés.

Validation

Dans cette section nous commençons par vérifier la précision de cette reparamétrisation du flux optique avant de l'appliquer comme entrée de notre modèle de perception présenté au chapitre 4.

Précision Nous nous plaçons en petit champ de vision, et nous simulons une ouverture de 8° . Les entrées de ce modèle de reparamétrisation du flux optique sont un ensemble de points répartis dans ce champ de vision et les vitesses associées. Le nombre typique de points présentés lors d'une expérience de psychophysique est de l'ordre de quelques centaines. Dans les simulations suivantes nous considérons $N = 84$ points.

Soit un plan objet d'orientation $(\hat{\chi}, \hat{v})$ et de mouvement relatif $(\hat{T}, \hat{\Omega})$. Le flux optique observé est $\hat{\Phi} = \left(\hat{\Phi}^0, \hat{\Phi}^1, \hat{\Phi}^2 \right)$.

Soit un ensemble \mathcal{P} de N points $(x^i, y^i)_{i \in [1;N]}$, tirées uniformément dans le disque visible⁶, les projections respectives sur le plan des vitesses de déplacement des points du plan

⁶Ce disque est l'intersection du plan fronto-parallèle et du cône de révolution de visibilité.

objet dont ces points sont la projection sont données par l'ensemble $\dot{\mathcal{P}} = (u_x^i, u_y^i)_{i \in [1;N]}$ avec :

$$\begin{pmatrix} u_x^i \\ u_y^i \end{pmatrix} = \hat{\Phi}(x^i, y^i) = \begin{pmatrix} \hat{\Phi}_x^0 + \hat{\Phi}_x^1 \cdot x^i + \hat{\Phi}_x^2 \cdot (x^i)^2 + \hat{\Phi}_y^1 \cdot y^i + \hat{\Phi}_y^2 \cdot (y^i)^2 \\ \hat{\Phi}_y^0 + \hat{\Phi}_y^1 \cdot x^i + \hat{\Phi}_y^2 \cdot x^i y^i + \hat{\Phi}_x^1 \cdot y^i + \hat{\Phi}_x^2 \cdot (y^i)^2 \end{pmatrix} \quad (6.48)$$

Nous appliquons le modèle de reparamétrisation du flux optique sur ces observations pour obtenir une distribution de probabilité sur le flux optique. La figure 6.12 montre un exemple de distribution sur le flux optique d'ordre un obtenu à partir de l'observation des N points.

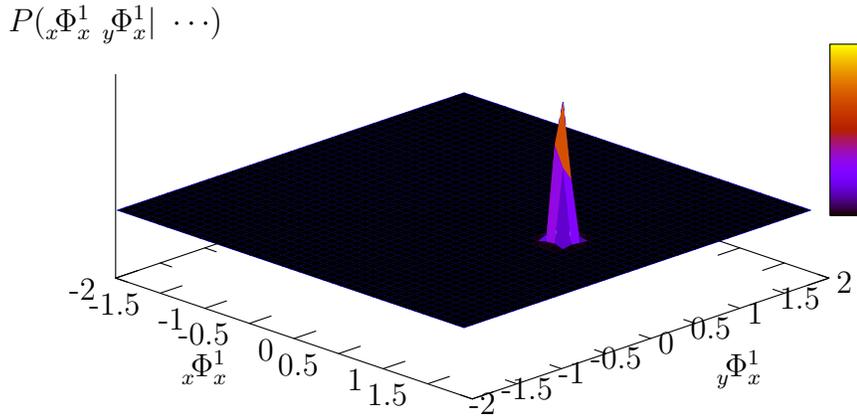


FIG. 6.12 – Exemple de distribution obtenue conjointement sur les composantes Φ_x^1 et Φ_y^1 du flux optique d'ordre un.

	Flux théorique	Moyenne du modèle
$\vec{\Phi}^0$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$
$\vec{\Phi}^1$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1,0111 & 0,0063 \\ 0,0001 & -0,0072 \end{pmatrix}$
$\vec{\Phi}^2$	$\begin{pmatrix} -1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} -1 \\ 0 \end{pmatrix}$

TAB. 6.4 – Comparaison entre les valeurs moyennes des composantes du flux calculées par le modèle (à 10^{-4} près) et le flux théorique. Le plan est vertical et incliné à 45° par rapport au plan fronto-parallèle ($(\chi, \nu) = (1, 0)$) et est en rotation autour de l'axe vertical ($\vec{\Omega} = {}^t(0, 1, 0)$ et $\vec{T} = \vec{0}$).

	Différence moyenne	Écart-type	Erreur maximale	Covariances
Ordre 0	0,0017	0,0113	0,0396	$\sigma_{\Phi^0} = 1 \text{ m.s}^{-1}$
Ordre 1	0	0,0099	0,0416	$\sigma_{\Phi^1} = 0,025 \text{ s}^{-1}$
Ordre 2	0,0035	0,016	0,0792	$\sigma_{\Phi^2 \lambda=0} = 5 \text{ m}^{-1}.\text{s}^{-1}$

TAB. 6.5 – Précision du modèle bayésien de reparamétrisation du flux optique (valeurs à 10^{-4} près). Ces moyennes ont été calculées pour plusieurs mouvements et pour 16 positions différentes du plan, réparties uniformément sur l'hémisphère unité. Chacune des configurations a été simulée 50 fois avec des points différents. La colonne de droite rappelle les valeurs des covariances du modèle de perception du flux optique. L'unité de distance est la distance entre l'observateur et le point de fixation. Dans le cas présent, nous avons simulé une distance de 1 m, d'où les unités présentées. Quelque soit la distance, on peut se ramener à ce cas par une simple homothétie.

Le tableau 6.4 compare, sur un exemple, les valeurs des composantes d'un flux optique simulé avec les valeurs moyennes des distributions obtenues par notre modèle de reparamétrisation. Sur cet exemple, l'erreur sur les flux d'ordre 0 et 2 est inférieure à 10^{-14} ; pour le flux d'ordre 1, l'erreur est de l'ordre du centième.

Le tableau 6.5 présente, entre autres, l'écart-type d'estimation de notre modèle pour un ensemble de présentation du stimulus et rappelle les valeurs des covariances sur chacune des distributions sur le flux optique spécifiées dans notre modèle de perception de la position du plan par le flux optique. Ces valeurs de covariances sont l'incertitude attendue sur les composantes du flux optique. L'écart-type de la reparamétrisation du flux optique par ce modèle est compatible, puisqu'inférieur, avec celui de notre modèle de perception.

Intégration Nous pouvons utiliser notre modèle de reparamétrisation du flux optique avec notre modèle de perception de l'orientation d'un plan à l'aide du flux optique. Pour cela, nous construisons un modèle portant sur les points et leurs vitesses, le flux optique, le mouvement propre et l'orientation du plan. La décomposition est la suivante :

$$P(X \Upsilon \vec{\Phi} V_x V_z R_y \Lambda X^1 Y^1 U_x^1 U_y^1 \dots X^N Y^N U_x^N U_y^N) \quad (6.49)$$

$$= P(X^1 Y^1 U_x^1 U_y^1 \dots X^N Y^N U_x^N U_y^N) P(V_x V_z R_y \Lambda) \quad (6.50)$$

$$\times P(\vec{\Phi} | X^1 Y^1 U_x^1 U_y^1 \dots X^N Y^N U_x^N U_y^N) \quad (6.51)$$

$$\times P(X \Upsilon | \vec{\Phi} V_x V_z R_y \Lambda) \quad (6.52)$$

Les facteurs 6.50 concernent les observations de notre modèle et sont donc laissés non spécifiés. Le facteur 6.51 est une question posée au présent modèle de reparamétrisation du flux optique. Le facteur 6.52 est la question posée à notre modèle de perception présenté au chapitre 4.

La question posée est :

$$P(X \Upsilon | V_x V_z R_y \Lambda X^1 Y^1 U_x^1 U_y^1 \dots X^N Y^N U_x^N U_y^N) \quad (6.53)$$

L'inférence bayésienne donne l'expression :

$$\begin{aligned}
& P(X \Upsilon \mid V_x V_z R_y \Lambda X^1 Y^1 U_x^1 U_y^1 \dots X^N Y^N U_x^N U_y^N) \\
= & \sum_{\vec{\phi} \in \mathcal{D}_{\vec{\phi}}} P(\vec{\phi} \mid X^1 Y^1 U_x^1 U_y^1 \dots X^N Y^N U_x^N U_y^N) P(X \Upsilon \mid \vec{\phi} V_x V_z R_y \Lambda) \quad (6.54)
\end{aligned}$$

L'espace de sommation $\mathcal{D}_{\vec{\phi}}$ est de dimension 8. Comme le modèle de perception réalise déjà une intégration sur l'espace $\mathcal{D}_{\vec{T}} \times \mathcal{D}_{\vec{\Omega}}$ de dimension 6, la complexité finale du calcul est $\Theta(N.N_{\Phi}^{10}.N_{\Omega}^3.N_T^3)$.⁷

Pour réduire cette complexité, on peut approximer le résultat ci-dessus en prenant une décision intermédiaire sur le flux optique. Cette décision peut être le flux de maximum de probabilité pour le flux moyen. La complexité devient $\Theta(N.N_{\Phi}^2 + N_{\Omega}^3.N_T^3)$.⁸

Pour vérifier l'intégration de ces deux modèles on peut en comparer ses résultats avec les résultats du modèle de perception seul. En prenant l'exemple de l'ambiguïté sur l'orientation (section 4.6), on obtient la figure 6.13 qui compare les précédents résultats à ceux du modèle intégré. On remarque que les deux histogrammes sont très proches : l'erreur relative maximale est inférieure à 0,1 %. Ce résultat confirme que ce modèle de reparamétrisation du flux optique est compatible avec le modèle de perception du plan.

Conclusion

Nous avons proposé plusieurs modèles de reparamétrisation du flux optique à partir d'un ensemble de points et de leurs vitesses respectives dans notre forme paramétrique. Après un modèle naïf d'introduction, nous avons présenté un modèle dans lequel les composantes du flux sont découplées. Ce découplage permet une réduction sensible des complexités du modèle.

Enfin, nous avons vérifié la pertinence de ce modèle en l'intégrant avec le modèle de perception de l'orientation d'un plan par le flux optique présenté au chapitre 4.

6.3 Interfaces du modèle de perception

Dans cette section, nous généralisons l'agencement de modèles bayésiens en série. Nous prenons l'exemple de notre modèle de perception du chapitre 4 pour montrer comment brancher un modèle en amont puis en aval en utilisant le concept d'interfaces.

6.3.1 Rappel du modèle

La figure 6.14 rappelle la spécification de notre modèle bayésien de perception de l'orientation par le mouvement. On y retrouve la liste de ses variables $X, \Upsilon, V_x, V_z, R_y, \Lambda, \vec{T}, \vec{\Omega}, \vec{\Phi}^0, \vec{\Phi}^1$ et $\vec{\Phi}^2$ et la question $P(X \Upsilon \mid \vec{\phi}^0 \vec{\phi}^1 \vec{\phi}^2 v_x v_z r_y \lambda)$.

⁷Il s'agit du produit des complexités des trois modèles.

⁸La complexité suite à une décision est la somme des complexité de chacun des modèles.

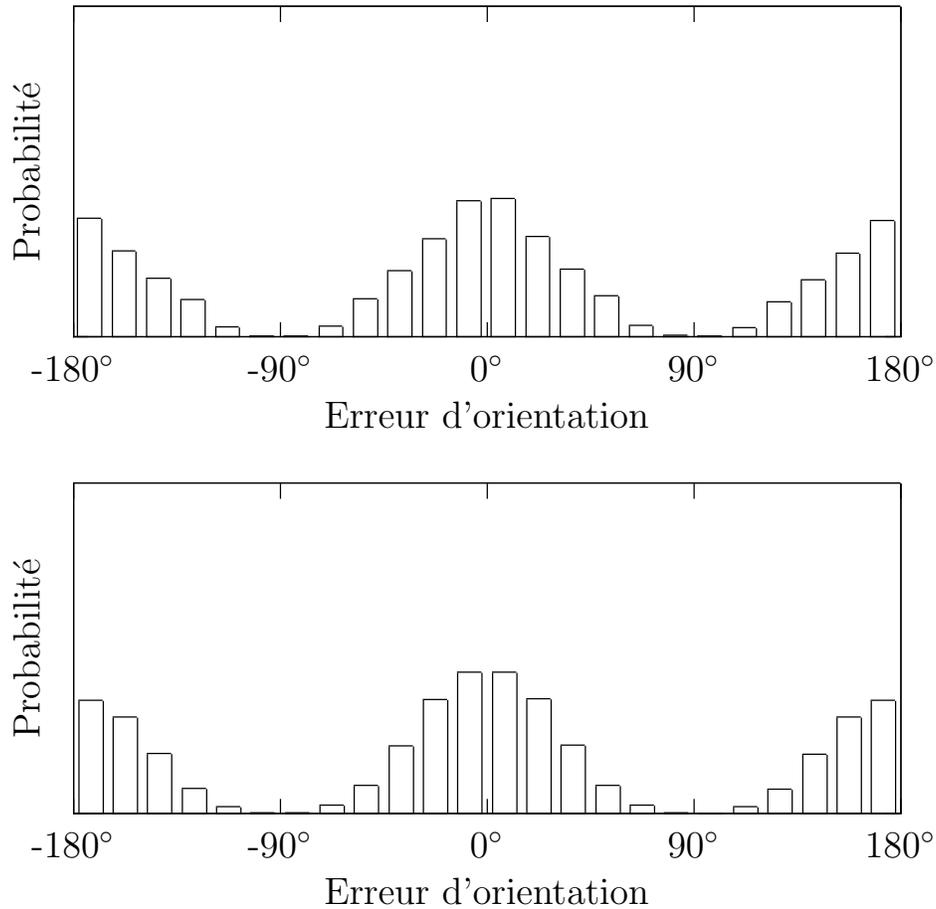


FIG. 6.13 – Comparaison des distributions d’erreur d’estimation de l’orientation. Le graphe du haut montre les résultats du modèle de perception de plan avec le flux théorique en entrée (reproduite de la section 4.6). En bas, l’entrée est la moyenne de l’estimation par notre modèle de reparamétrisation du flux optique.

Dans la suite, les variables de droite sont appelées les entrées et les variables de gauche les sorties de notre modèle. Ces variables sont les interfaces avec lesquelles un modèle bayésien peut s’intégrer dans un système plus complexe sans modification.

D’une façon plus générale, la question que l’on pose dans un programme bayésien est de la forme $P(C | E)$ où C et E sont des parties de l’ensemble des variables. La variable C , conjonction de variables de ce modèle, est la **sortie**. La variable E est l’**entrée**.⁹ La variable restante D est réputée libre, au sens où on ne connaît ni ne cherche sa valeur.

Dans le cas de notre modèle, $C = X \wedge Y$, $D = \vec{\Omega} \wedge \vec{T}$ et $E = \vec{\Phi} \wedge V_x \wedge V_z \wedge R_y \wedge \Lambda$.

⁹Dans cette partie, la couleur d’une variable indique son rôle au sein du modèle en question. Nous notons les variables cherchées en **rouge**; les variables connues en **bleu** et les variables libres en noir.

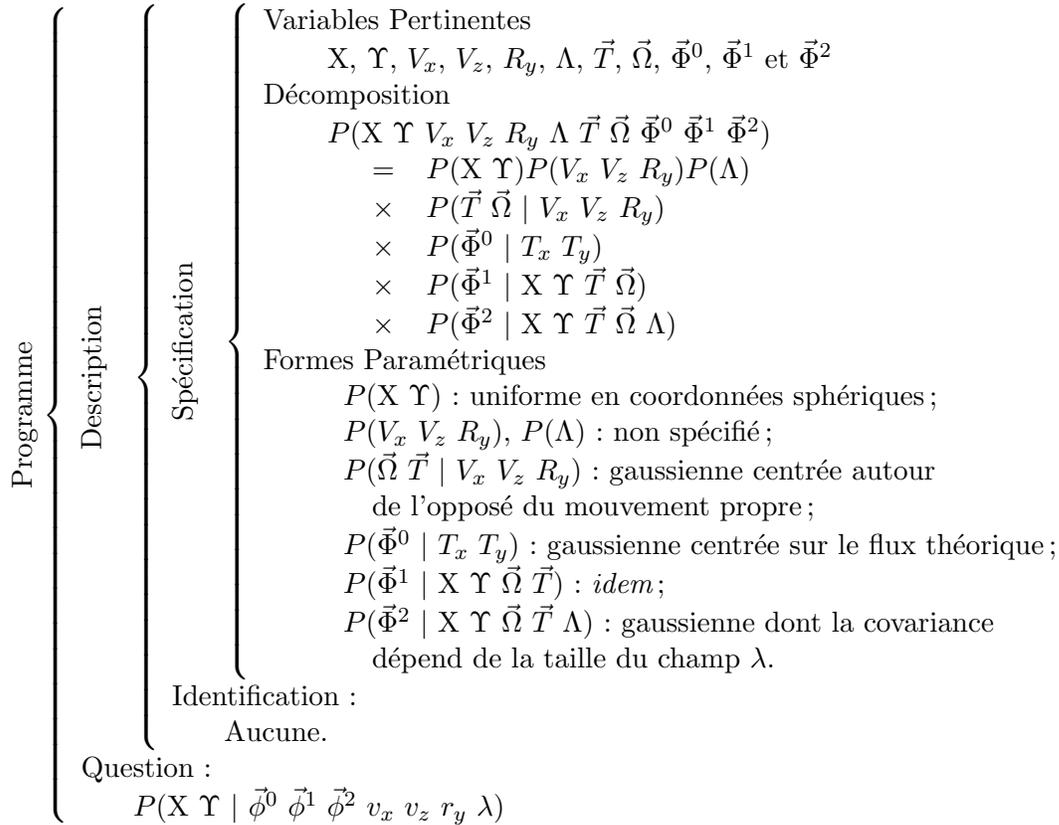


FIG. 6.14 – Rappel du modèle bayésien de perception.

6.3.2 Entrées

Ces variables sont appelées variables d'entrée car c'est sur celles-ci que l'on exprime une connaissance sur une expérience particulière lors de l'utilisation du modèle.

Valeurs connues L'utilisation classique d'un tel modèle est de fournir des valeurs pour les variables d'entrée. On instancie le modèle avec des valeurs connues de ces variables. L'inférence bayésienne donne l'expression familière 6.55.

$$P(\mathbf{C} | \mathbf{e}) = \frac{\sum_{d \in \mathcal{D}_D} P(\mathbf{C} \ d \ \mathbf{e})}{\sum_{d, c \in \mathcal{D}_D \times \mathcal{D}_C} P(c \ d \ \mathbf{e})} \quad (6.55)$$

Cette expression donne l'état de connaissance sur \mathbf{C} de ce modèle étant donnée la valeur \mathbf{e} de la variable \mathbf{E} . Cet état de connaissance sur \mathbf{C} s'exprime sous la forme d'une distribution de probabilité. L'instanciation par des valeurs connues de certaines variables est l'utilisation nominale d'un modèle bayésien. C'est ce que nous avons fait tout au long de ce document.

Néanmoins, cette utilisation présuppose l'observabilité de ces variables. Dit autrement, le résultat de l'inférence est la distribution de probabilité lors d'une certitude absolue sur la valeur des variables de droite.

Distribution de probabilité Une telle certitude n'est pas toujours possible. C'est le cas par exemple lorsque ces variables ne sont accessibles qu'à partir d'un autre modèle probabiliste. La connaissance sur ces variables est en général incertaine et ne peut être utilisée directement dans un autre modèle.

Une telle distribution de probabilité sur une variables d'entrée d'un modèle est nommée *soft evidence* dans le domaine des réseaux bayésiens en opposition avec une *hard evidence* qu'est une certitude sur la valeur d'une variable.

Pour injecter une connaissance incertaine sur une variable d'entrée, il suffit de considérer un modèle auxiliaire regroupant cette connaissance avec celles du modèle d'origine.

Soit la distribution $P(E)$.¹⁰ On écrit un modèle auxiliaire portant sur les variables C et E avec comme décomposition :

$$\begin{aligned} & P(C \ E) \\ = & P(E) \end{aligned} \tag{6.56}$$

$$\times P(C \mid E) \tag{6.57}$$

Le facteur 6.56 est la distribution $P(E)$. Le facteur 6.57 est la question posée au modèle principal.

On pose alors la question $P(C)$. On remarque qu'elle n'est pas conditionnée. En effet, dans ce cas, il n'y a pas de variable dont la valeur est connue qui pourrait aller en partie droite. L'inférence s'écrit :

$$P(C) = \sum_{e \in \mathcal{D}_E} P(e)P(C \mid e) \tag{6.58}$$

Cela revient simplement à calculer la distribution de probabilité moyenne à partir de la conditionnelle $P(C \mid E)$.

Cette expression peut se généraliser à n'importe quelle distribution sur E . Dans le cas où cette distribution est elle-même conditionnée, ce conditionnement est propagé à la variable C . Avec un modèle amont $P(E \ F \ G)$ et la question $P(E \mid g)$, la connaissance cherchée sur C devient conditionnée par la valeur g . On écrit alors :

$$\begin{aligned} & P(C \mid g) \\ = & \frac{\sum_{e \in \mathcal{D}_E} P(C \mid e)P(e \mid g)}{\sum_{c, e \in \mathcal{D}_C \times \mathcal{D}_E} P(c \mid e)P(e \mid g)} \\ = & \sum_{e \in \mathcal{D}_E} P(C \mid e)P(e \mid g) \end{aligned} \tag{6.59}$$

¹⁰Cette distribution peut être obtenue à partir d'un modèle probabiliste amont $P(E \ F \ G)$ auquel cas il s'agit de l'instanciation $P(E \mid g)$ de la distribution conditionnelle $P(E \mid G)$.

Cette expression est familière. Nous l'avons déjà utilisée pour intégrer le modèle de reparamétrisation du flux optique au modèle de perception. En effet, l'expression 6.59 est une généralisation de l'expression 6.54 :

$$\begin{aligned} & P(\mathbf{X} \Upsilon \mid V_x V_z R_y \Lambda X^1 Y^1 U_x^1 U_y^1 \dots X^N Y^N U_x^N U_y^N) \\ = & \sum_{\vec{\phi} \in \mathcal{D}_{\vec{\phi}}} P(\mathbf{X} \Upsilon \mid \vec{\phi} V_x V_z R_y \Lambda) P(\vec{\phi} \mid X^1 Y^1 U_x^1 U_y^1 \dots X^N Y^N U_x^N U_y^N) \end{aligned}$$

6.3.3 Sorties

Les variables de gauche dans une question sont appelées des sorties car le but d'un programme bayésien est de calculer sa connaissance sur ces variables.

Propagation de l'incertitude Une manière d'utiliser cette connaissance est de propager l'incertitude à un autre modèle bayésien. Pour cela on procède de la même manière que pour le traitement d'une *soft evidence*.

Soit le modèle aval $P(A B C)$ auquel on pose la question $P(A \mid C)$, et le modèle principal $P(C D E)$ auquel on pose la question $P(C \mid E)$. On construit le modèle auxiliaire :

$$\begin{aligned} & P(A C E) \\ = & P(E) \end{aligned} \tag{6.60}$$

$$\times P(C \mid E) \tag{6.61}$$

$$\times P(A \mid C) \tag{6.62}$$

Le facteur 6.60 ne concerne que des variables instanciées dans la question finale et peut donc rester non spécifié. Les facteurs 6.61 et 6.62 sont les questions posées respectivement aux modèles principal et aval.

Ce nouveau modèle s'intéresse à la connaissance sur A . Les connaissances dont il dispose sont les entrées du modèle principal. L'inférence bayésienne sur le nouveau modèle s'écrit :

$$\begin{aligned} & P(A \mid e) \\ = & \frac{\sum_{c \in \mathcal{D}_c} P(A \mid c) P(c \mid e)}{\sum_{a, c \in \mathcal{D}_a \times \mathcal{D}_c} P(a \mid c) P(c \mid e)} \\ = & \sum_{c \in \mathcal{D}_c} P(A \mid c) P(c \mid e) \end{aligned} \tag{6.63}$$

Nous avons déjà utilisé cette expression. Par exemple lorsque nous avons mis en place un modèle de décision en aval de notre modèle, nous avons écrit l'équation 6.29 :

$$P(\theta \mid \vec{\phi} v_x v_z r_y \lambda) = \sum_{\chi, v \in \mathcal{D}_X \times \mathcal{D}_v} P(\theta \mid \chi) P(\chi v \mid \vec{\phi} v_x v_z r_y \lambda) \tag{6.64}$$

Une telle construction peut se voir comme un changement de variable pour le modèle principal opéré par le modèle aval. De nombreux exemples de tels changements de variable

parcèlement ce document et en particulier le chapitre 5 dans lequel nous utilisons des modèles auxiliaires pour exprimer la sortie de notre modèle dans le même espace que les résultats expérimentaux.

On remarque que cette inférence fait intervenir une sommation sur tout le domaine de la variable intermédiaire sur laquelle porte l'incertitude à propager.

Décision Cette sommation peut s'avérer trop coûteuse pour certaines applications lorsque le domaine de la variable intermédiaire est très grand. C'est le cas, par exemple, pour propager l'incertitude du modèle de reparamétrisation du flux optique au modèle de perception (voir 6.2.3).

De plus, il est parfois plus utile d'obtenir une valeur particulière d'une variable plutôt qu'une distribution sur tout son domaine.

Dans ces deux cas, on peut alors utiliser une procédure de décision sur la distribution de probabilité pour obtenir une ou plusieurs valeurs. Il existe plusieurs méthodes parmi lesquelles l'extraction du maximum de la distribution. C'est ce qui s'appelle le *maximum a posteriori*. Une autre méthode est le tirage d'une valeur suivant cette distribution ce qui permet de faire de l'échantillonnage.

À partir de cette valeur, on peut interroger le modèle aval de manière directe comme en section 6.3.2.

6.4 Résumé

Nous nous sommes intéressés dans cette partie à l'utilisation du modèle bayésien de perception d'un plan par le flux optique.

Dans un premier temps, nous avons examiné l'applicabilité de notre modèle à une expérience portant sur un objet ondulé pour l'étude de la poursuite. Nous avons montré que les symétries de perception sont identiques et mis en avant une transposition du stimulus et des réponses. À partir de là, nous avons étudié la poursuite et proposé une explication aux résultats de cette expérience comme pour les expériences du chapitre 5.

Ensuite, nous avons présenté un modèle de reparamétrisation du flux optique. Ce modèle est destiné à fournir son entrée principale au modèle de perception. Nous avons proposé en premier lieu un modèle naïf basé sur l'outil de la fusion bayésienne. Ce premier essai présente l'inconvénient d'une complexité trop élevée. Nous avons alors séparé le problème en composantes indépendantes pour proposer un traitement parallèle et moins coûteux. Nous avons validé ce modèle en étudiant la précision et en l'appliquant conjointement avec le modèle de perception de plan.

Nous avons terminé ce chapitre en revenant sur les concepts d'entrée et sortie du modèle. Nous avons montré, de manière générale, que l'inférence dans un modèle bayésien peut s'effectuer par l'intermédiaire d'une question conditionnée par une valeur connue précise ou bien alternativement par une sommation sur une distribution de probabilité sur l'entrée. De manière symétrique, un programme bayésien calcule une distribution de probabilité à l'aide de laquelle on peut, au besoin, décider d'une valeur particulière des variables de

sortie. Cette discussion montre comment mettre des programmes bayésiens en série. De cette manière, on peut intégrer un programme bayésien, tel que notre modèle de perception, à un système de plus haut niveau sans en changer la spécification.

Chapitre 7

Conclusion

Nous nous sommes intéressé, dans ce manuscrit, à un problème de perception de plans par le flux optique. Il s'agit d'un problème inverse, mal-posé et incomplet. Nous avons proposé un modèle psychophysique simple de cette perception.

Nous rappelons dans un premier temps ses hypothèses et ses résultats. Nous reviendrons ensuite sur les utilisations avancées que nous avons proposées pour ce modèle avant de nous intéresser, plus généralement, aux perspectives ouvertes par ce travail.

7.1 Hypothèses

Pour décrire notre modèle, nous avons choisi le formalisme de la programmation bayésienne. Grâce à l'utilisation des probabilités et de la règle de Bayes, ce formalisme permet de gérer les caractères inverse, mal-posé et incertain des problèmes de perception en plus de permettre une expression claire des hypothèses et la compréhension de son fonctionnement.

Notre modèle rassemble au sein de ce formalisme un ensemble d'hypothèses issues de la littérature psychophysique. Historiquement, le premier principe proposé pour résoudre le problème de perception de formes à partir du flux optique est la rigidité. Cette hypothèse suppose que le flux optique est plus probablement engendré par un objet rigide plutôt que par un ensemble de points de mouvements indépendants. Il s'agit de considérer invariante la forme générale de l'objet. C'est le cas de nombreux objets de la vie quotidienne, mais cette hypothèse est mise en défaut par des objets articulés ou du tissu par exemple. Dans notre modèle, la rigidité est probabiliste, au sens où la perception est plus probablement celle d'un objet rigide mais peut tout de même correspondre à un objet subissant de légères déformations.

Des expériences récentes ont montré que, dans certains cas, le percept dominant ne correspond pas à l'objet le plus rigide et que, pour un même flux optique, des conditions différentes de mouvement propre amènent à des jugements différents. L'hypothèse de stationnarité a été proposée pour pallier ces problèmes. Elle stipule que le mouvement absolu est petit et, par là même, prend en compte le mouvement propre de l'observateur. Cette supposition correspond à une invariance de la position de l'objet. Comme pour la rigidité,

cette invariance est courante mais pas absolue. La stationnarité proposée dans notre modèle est donc probabiliste. Les objets perçus sont plus probables lorsque leur mouvement est plus restreint.

La troisième hypothèse de notre modèle est d'ordre structurelle. On considère la position du plan indépendante des mouvements de l'objet et du sujet. C'est une sorte d'hypothèse d'holonomie sur les mouvements des objets. Pour un objet connu, il est parfois possible de connaître une dépendance entre sa position et son mouvement : le côté d'un train se déplace très probablement selon l'axe de ce train. Il s'agit en général d'une connaissance de plus haut niveau que notre modèle n'exploite pas à l'heure actuelle.

Une dernière hypothèse dans notre modèle est que la précision du flux optique perçu dépend de la taille du champ de vision disponible. Cette hypothèse porte plus précisément sur le flux d'ordre deux et se justifie par les dérivations successives.

7.2 Résultats

Pour valider notre modèle et ses hypothèses, nous en avons comparé les résultats avec ceux de participants à des expériences de psychophysique.

Nous avons passé en revue cinq expériences de la littérature s'intéressant chacune à des points différents. Nous avons ainsi examiné la perception de l'inclinaison, l'influence de la taille du champ de vision et du mouvement propre sur la perception de l'orientation, l'importance relative de la stationnarité par rapport à la rigidité et enfin l'effet de l'angle de *shear* sur la précision de la perception.

Pour chacune de ces cinq expériences, présentées au chapitre 5, notre modèle exhibe une bonne reproduction qualitative des réponses des participants et en particulier dans leurs erreurs d'appréciation les plus communes.

Les résultats numériques présentés ont été réalisés à l'aide d'un seul jeu de paramètres pour montrer la généralité du modèle. En contrepartie, les valeurs trouvées ne correspondent pas exactement aux résultats expérimentaux. Ces différences peuvent s'expliquer par des variations dans les conditions expérimentales et la variabilité des sujets (voir discussion 5.6).

Pour chaque expérience, nous sommes partis de notre modèle et de son fonctionnement pour proposer des explications aux perceptions des observateurs humains. Le choix du formalisme de la programmation bayésienne ainsi que le nombre réduit des hypothèses de notre modèle en font un outil adéquat à la compréhension de ces expériences.

Ainsi, nous avons montré comment réconcilier les interprétations euclidiennes et affines de la perception de l'orientation et de l'inclinaison grâce à la variation de la covariance sur le flux optique d'ordre deux en fonction de la taille du champ. Nous avons expliqué, sur une expérience, la prééminence de la stationnarité sur la rigidité par la comparaison des covariances du mouvement et du flux optique d'ordre deux. Enfin, nous avons extrait des relations précises entre les covariances des translations et des rotations attendues suite à un mouvement propre.

7.3 Utilisations avancées

Nous avons proposé des extensions au domaine d'application de notre modèle. En premier lieu, nous avons examiné une expérience de perception d'objets ondulés. Nous avons montré que les ambiguïtés étudiées sont de même nature que celles de la perception de l'orientation du plan. Nous avons transposé les conditions de cette expérience vers notre modèle puis les résultats de notre modèle dans le contexte de l'expérience. Pour cela, nous avons introduit des modèles de décision qui se placent en aval de notre modèle de perception pour répondre aux questions définies dans le protocole expérimental. Grâce à cela, notre modèle de perception peut reproduire partiellement des résultats concernant des objets ondulés.

Dans un second temps, nous nous sommes intéressés au flux optique d'entrée de notre modèle. Notre modèle utilise une forme paramétrique du flux décomposé en dérivées d'ordre 0, 1 et 2. Les algorithmes d'extraction du flux optique de la littérature produisent généralement un champ de vecteurs de déplacement à partir d'une suite d'images. Nous avons proposé un traitement d'un tel champ de vitesses afin d'en extraire l'entrée de notre modèle. Ce traitement se base sur le mécanisme de la fusion bayésienne et nous avons proposé une simplification pour en faciliter l'inférence.

Enfin, nous avons présenté une généralisation de ces deux processus en considérant les interfaces de notre modèle. Nous dissociions les interfaces d'entrée et de sortie d'un modèle bayésien et avons montré, dans les deux cas, comment peut s'opérer la transmission de connaissances probabilistes entre deux modèles.

Un caractère important de notre travail est la généralité de l'approche. Le modèle présenté est l'exemple d'un canevas de modèle de perception. Il est facile de modifier une partie du modèle, de rajouter des hypothèses ou de l'intégrer dans un système plus complet sans en altérer les propriétés.

7.4 Perspectives

Un tel modèle bayésien ouvre de nombreuses perspectives de travail. Un programme bayésien porte sur des variables, définit la distribution conjointe sur ces variables et pose une question à cette conjointe. À partir de là, on peut définir des perspectives par ordre croissant d'importance des modifications.

Une première piste d'extension du travail porte sur la question posée. Dans ce document nous nous sommes intéressé presque exclusivement à connaître la forme de l'objet à partir du flux optique et du mouvement propre. Une autre question que l'on peut poser au modèle concerne simplement le mouvement de l'objet. Nous l'avons fait section 5.1 pour expliquer le fonctionnement du modèle mais on pourrait s'y intéresser en temps que perception à part entière.

Une autre question pourrait porter sur le mouvement propre. Il est considéré comme une entrée du programme provenant, par exemple, du système vestibulaire mais c'est une distinction qui vient de la question posée. Il peut être intéressant de se demander la connais-

sance que le flux optique apporte sur le mouvement propre. C'est la question mise en avant par lavection, illustrée par l'illusion du train : vous êtes dans un train en gare ; vous voyez le train à côté du vôtre bouger ; lequel démarre ? Dans cette situation, le flux optique vous informe d'un mouvement relatif ; le système vestibulaire n'a pas relevé de mouvement particulier mais n'est pas très sensible aux accélérations d'un train. Il faut alors trouver d'autres sources d'information telles que des amers visuels pour décider de cette question.

La décomposition d'un modèle bayésien en définit la structure. Une perspective intéressante serait la construction de modèles plus élaborés par la multiplication de la structure élémentaire du modèle dans le temps ou l'espace. Un premier exemple est le filtrage temporel. Dans ce cas, on considère l'évolution dans le temps et on recopie la décomposition du modèle à chaque pas de temps. Un tel modèle permet en général d'augmenter la précision de la réponse grâce à l'accumulation d'observations.

Le filtrage est la multiplication du modèle dans le temps ; il est aussi possible d'en faire la multiplication dans l'espace. On peut, par exemple, faire observer le modèle en plusieurs endroits de l'image et faire la fusion des résultats pour obtenir une information sur toute l'image. De cette façon, il est possible de percevoir correctement des formes plus complexes qu'un plan.

Une autre manière d'étendre le modèle à d'autres formes est de modifier les variables. Dans le présent document, la variable de forme est un couple de variables réelles décrivant un plan. En utilisant trois paramètres supplémentaires on peut représenter les coniques, par exemple. On peut aussi considérer une carte de profondeur pour augmenter les degrés de liberté d'objets reconnaissables.

Enfin, une perspective qui retient particulièrement notre attention est l'échange de connaissances probabilistes entre différents systèmes partageant les mêmes variables. Nous avons présenté l'assemblage en série avec l'exemple du flux optique généré par un modèle d'extraction et utilisé par le modèle de perception. Les probabilités nous permettent de concevoir aisément des mécanismes plus riches du point de vue de l'échange de connaissances entre sous-modèles.

D'une manière générale, nous pensons qu'il est possible d'appliquer la méthodologie employée dans ce document pour la perception de plan par le flux optique à d'autres problèmes de modélisation psychophysique.

Bibliographie

- E. H. ADELSON et J. R. BERGEN : Spatiotemporal energy models for the perception of motion. *Journal of Optical Society of America, A*, 2(2):248–99, 1985.
- G. ADIV : Determining three-dimensional motion and structure from optical flow generated by several moving objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(4):384–401, 1985.
- G. ADIV : Inherent ambiguities in recovering 3-D motion and structure from a noisy flow field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(5):477–489, 1989.
- P. ANANDAN : A computational framework and an algorithm for the measurement of visual motion. *International Journal of Computer Vision*, 2(3):283–310, 1989.
- T. J. ANASTASIO, P. E. PATTON et K. BELKACEM-BOUSSAID : Using Bayes' rule to model multisensory enhancement in the superior colliculus. *Neural Computation*, 12(5):1165–87, 2000.
- S. A. ANDERSSON, D. MADIGAN et M. D. PERLMAN : An alternative Markov property for chain graphs. In E. HORVITZ et F. V. JENSEN, édés : *Proceedings of the 12th Annual Conference on Uncertainty in Artificial Intelligence (UAI-96)*, p. 40–8. Morgan Kaufmann Publishers, 1996.
- S. ARNBORG et G. SJÖDIN : Bayes rules in finite models. In *Proc. European Conference on Artificial Intelligence*, p. 571–575, 2000. URL <ftp://ftp.nada.kth.se/pub/documents/Theory/Stefan-Arnborg/foc1.ps>.
- S. ARNBORG et G. SJÖDIN : On the foundations of bayesianism. In A. MOHAMMAD-DJAFARI, éd. : *Bayesian Inference and Maximum Entropy Methods in Science and Engineering, 20th International Workshop, Gif-sur-Yvette, 2000*, p. 61–71. American Institute of Physics, 2001.
- S. ARNBORG et G. SJÖDIN : What is the plausibility of probability?, 2003. URL <ftp://ftp.nada.kth.se/pub/documents/Theory/Stefan-Arnborg/m2001.pdf>.

- M. S. ARULAMPALAM, S. MASKELL, N. GORDON et T. CLAPP : A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–88, 2002.
- S. T. BARNARD et W. B. THOMPSON : Disparity analysis of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(4):333–40, 1980.
- J. L. BARRON, S. S. BEAUCHEMIN et D. J. FLEET : On optical flow. In *6th International Conference on Artificial Intelligence and Information-Control Systems of Robots*, p. 3–14, Bratislava, Slovakia, 1994a.
- J. L. BARRON, D. J. FLEET et S. S. BEAUCHEMIN : Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, 1994b.
- S. S. BEAUCHEMIN et J. L. BARRON : The computation of optical flow. *ACM Computing Surveys*, 27(3):433–67, 1996.
- M. L. BRAUNSTEIN : Depth perception in rotating dot patterns: effects of numerosity and perspective. *Journal of Experimental Psychology*, 64:415–20, 1962.
- E. CHARNIAK : Bayesian networks without tears. *AI Magazine*, 12(4):50–63, 1991.
- H. COLONIUS et A. DIEDERICH : A maximum-likelihood approach to modeling multi-sensory enhancement. In T. G. DIETTERICH, S. BECKER et Z. GHAHRAMANI, édés : *Advances in Neural Information Processing Systems 14*, Cambridge, MA, 2002. MIT Press.
- G. F. COOPER : The computational complexity of probabilistic inference using bayesian belief networks. *Artificial Intelligence*, 42(2-3):393–405, 1990.
- V. CORNILLEAU-PÉRÈS et J. DROULEZ : Visual perception of surface curvature: psychophysics of curvature detection induced by motion parallax. *Perception and Psychophysics*, 46(4):351–64, 1989.
- V. CORNILLEAU-PÉRÈS, M. WEXLER, J. DROULEZ, E. MARIN, C. MIÈGE et B. BOURDONCLE : Visual perception of planar orientation: dominance of static depth cues over motion cues. *Vision Research*, 42:1403–1412, 2002.
- R. T. COX : Probability, frequency and reasonable expectation. *American Journal of Physics*, 14(1):1–13, 1946.
- R. T. COX : *The Algebra of Probable Inference*. Johns Hopkins University Press, Baltimore, MD, 1961.
- P. S. de LAPLACE : Essai philosophique sur les probabilités. In *Œuvres complètes de Laplace (1878)*. Gauthier-Villars, Paris, France, 1814. URL <http://gallica.bnf.fr/Catalogue/noticesInd/FRBNF30739022.htm>.

- R. DECHTER : Bucket elimination: a unifying framework for probabilistic inference. *In* E. HORVITZ et F. V. JENSEN, édés : *Proceedings of the 12th Annual Conference on Uncertainty in Artificial Intelligence (UAI-96)*, p. 211–219. Morgan Kaufmann Publishers, 1996.
- S. DENEVE, P. E. LATHAM et A. POUGET : Efficient computation and cue integration with noisy population codes. *Nature Neuroscience*, 4(8):826–31, 2001.
- T. M. H. DIJKSTRA, V. CORNILLEAU-PÉRÈS, C. C. A. M. GIELEN et J. DROULEZ : Perception of three-dimensional shape from ego- and object-motion: comparison between small- and large-field stimuli. *Vision Research*, 35(4):453–462, 1995.
- F. DOMINI et C. CAUDEK : Perceiving surface slant from deformation of optic flow. *Journal of Experimental Psychology: Human Perception and Performance.*, 25(2):426–44, 1999.
- F. DOMINI et C. CAUDEK : 3-D structure perceived from dynamic information: a new theory. *Trends in Cognitive Sciences*, 7(10):444–9, 2003.
- M. O. ERNST et M. S. BANKS : Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870):429–33, 2002.
- O. FAUGERAS : What can be seen in three dimensions with an uncalibrated stereo rig? *In European Conference on Computer Vision*, p. 563–78, 1992.
- O. FAUGERAS : *Three-dimensional computer vision: a geometric viewpoint*. MIT Press, 1993.
- C. FERMÜLLER et Y. ALOIMONOS : Observability of 3D motion. *International Journal of Computer Vision*, 27(1):43–63, 2000.
- D. J. FLEET et A. D. JEPSON : Computation of component image velocity from local phase information. *International Journal of Computer Vision*, 5(1):77–104, 1990.
- G. D. FORNEY : The Viterbi algorithm. *Proceedings of the IEEE*, 61:268–78, 1973.
- D. A. FORSYTH et J. PONCE : *Computer vision: a modern approach*. Prentice Hall, 2002.
- W. S. GEISLER et D. KERSTEN : Illusions, perception and Bayes. *Nature Neuroscience*, 5(6):598–604, 2002.
- Z. GHAHRAMANI, D. M. WOLPERT et M. I. JORDAN : Computational models of sensorimotor integration. *In* P. G. MORASSO et V. SANGUINETI, édés : *Self-organization, computational maps and motor control*, p. 117–47. Elsevier, 1997.
- E. J. GIBSON, J. J. GIBSON, O. W. SMITH et A. FLOCK : Motion parallax as a determinant of perceived depth. *Journal of Experimental Psychology*, 54:40–51, 1959.

- J. J. GIBSON : The perception of visual surfaces. *American Journal of Psychology*, 63 (3):367–84, 1950.
- J. I. GOLD et M. N. SHADLEN : Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron*, 36:299–308, 2002.
- B. F. J. GREEN : Figure coherence in the kinetic depth effect. *Journal of Experimental Psychology*, 62:262–72, 1961.
- J. Y. HALPERN : A counterexample to theorems of Cox and Fine. *Journal of Artificial Intelligence Research*, 10:67–85, 1999a.
- J. Y. HALPERN : Cox’s theorem revisited. *Journal of Artificial Intelligence Research*, 11:429–435, 1999b.
- R. HARTLEY, R. GUPTA et T. CHANG : Stereo from uncalibrated cameras. In *IEEE conference on Computer Vision*, p. 761–4, 1992.
- D. J. HEEGER : Optical flow using spatiotemporal filters. *International Journal of Computer Vision*, 1(4):279–302, 1988.
- J. M. HILLIS, S. J. WATT, M. S. LANDY et M. S. BANKS : Slant from texture and disparity cues: optimal cue combination. *Journal of Vision*, 4:967–992, 2004.
- B. K. P. HORN et B. G. SCHUNCK : Determining optic flow. *Artificial Intelligence*, 17:185–204, 1981.
- T. S. HUANG et A. N. NETRAVALI : Motion and structure from feature correspondences: a review. *Proceedings of the IEEE*, 82(2):252–68, 1994.
- T. S. JAAKKOLA et M. I. JORDAN : Variational probabilistic inference and the QMR-DT network. *Journal of Artificial Intelligence Research*, 10:291–322, 1999.
- R. A. JACOBS : Optimal integration of texture and motion cues to depth. *Vision Research*, 39:3621–9, 1999.
- E. T. JAYNES : Prior probabilities. *IEEE Transactions on Systems Science and Cybernetics*, 4(3):227–41, 1968.
- E. T. JAYNES : On the rationale of maximum-entropy methods. *Proceedings of the IEEE*, 70(9):939–52, 1982.
- E. T. JAYNES : *Probability Theory: The Logic of Science*. Cambridge University Press, 2003.
- F. V. JENSEN : *Bayesian Graphical Models*. Wiley, Sussex, UK, 2000.

- F. V. JENSEN et F. JENSEN : Optimal junction trees. In R. L. MANTARAS et D. POOLE, édés : *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence. Seattle, Washington*, p. 360–366. Morgan Kaufmann Publishers, 1994.
- F. V. JENSEN, S. L. LAURITZEN et K. OLESEN : Bayesian updating in recursive graphical models by local computations. *Computational Statistical Quarterly*, 4:269–82, 1990.
- F. KAHL et B. TRIGGS : Critical motions in euclidean structure from motion. In *International Conference on Computer Vision and Pattern Recognition*, Fort Collins, Colorado, 1999.
- R. E. KALMAN : A new approach to linear filtering and prediction problems. *Transactions of the ASME—Journal of Basic Engineering*, 82(Series D):35–45, 1960.
- D. KERSTEN : High-level vision as statistical inference. In M. S. GAZZANIGA, éd. : *The new cognitive neurosciences—second edition*, chap. 25. MIT Press, 1999.
- D. KERSTEN, P. MAMASSIAN et A. YUILLE : Object perception as bayesian inference. *Annual Review of Psychology*, 55:271–304, 2004.
- J. M. KEYNES : *Fundamental Ideas*, chap. 4. Macmillan, 1921.
- E. KOECHLIN, J. L. ANTON et Y. BURNOD : Bayesian inference in populations of cortical neurons: a model of motion integration and segmentation in area MT. *Biological Cybernetics*, 80:25–44, 1999.
- J. J. KOENDERINK : Optic flow. *Vision Research*, 26(1):161–179, 1986.
- J. J. KOENDERINK et A. J. van DOORN : Affine structure from motion. *Journal of Optical Society of America, A*, 8(2):377–85, 1991.
- K. P. KÖRDING et D. M. WOLPERT : Bayesian integration in sensorimotor learning. *Nature*, 427:244–7, 2004.
- M. S. LANDY, L. T. MALONEY, E. B. JOHNSTON et M. YOUNG : Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research*, 35:389–412, 1995.
- M. LAPPE, F. BREMMER et A. V. van den BERG : Perception of self-motion from visual flow. *Trends in Cognitive Sciences*, 3(9):329–36, 1999.
- S. L. LAURITZEN et D. SPIEGELHALTER : Local computations with probabilities on graphical structures and their application to expert systems (with discussion). *Journal of the Royal Statistical Society, Series B*, 50:157–224, 1988.
- O. LEBELTEL, P. BESSIÈRE, J. DIARD et E. MAZER : Bayesian robots programming. Rap. tech., Laboratoire Leibniz-IMAG, 2000.

- O. LEBELTEL, P. BESSIÈRE, J. DIARD et E. MAZER : Bayesian robot programming. *Advanced Robotics*, 16(1):49–79, 2004.
- D. N. LEE : Plummeting gannets: a paradigm of ecological optics. *Nature*, 293(5830):293–4, 1981.
- T. S. LEE et D. MUMFORD : Hierarchical bayesian inference in the visual cortex. *Journal of Optical Society of America, A*, 20(7):1434–48, 2003.
- H. C. LONGUET-HIGGINS : The visual ambiguity of a moving plane. *Proceedings of the Royal Society of London B, Biological Sciences*, 223:165–175, 1984.
- H. C. LONGUET-HIGGINS et K. PRADZNY : The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B, Biological Sciences*, 208(1173):385–97, 1980.
- B. D. LUCAS et T. KANADE : An iterative image registration technique with an application to stereo vision. *In Proceedings of Imaging Understanding Workshop*, p. 121–130, 1981.
- D. J. C. MACKAY : Introduction to Monte Carlo methods. *In M. I. JORDAN, éd. : proceedings of an Erice summer school*, 1998.
- L. T. MALONEY : Illuminant estimation as cue combination. *Journal of Vision*, 2:493–504, 2002.
- P. MAMASSIAN et H. H. BÜLTHOFF : Active kinetic depth effect. Rap. tech. 27, Max Planck Institute for Biological Cybernetics, Tübingen, Germany, feb 1996.
- P. MAMASSIAN et R. GOUTCHER : Prior knowledge on the illumination position. *Cognition*, 81:B1–B9, 2001.
- P. MAMASSIAN et M. S. LANDY : Observer biases in the 3D interpretation of line drawings. *Vision Research*, 38:2817–32, 1998.
- P. MAMASSIAN et M. S. LANDY : Interaction of visual prior constraints. *Vision Research*, 41:2653–68, 2001.
- P. MAMASSIAN, M. S. LANDY et L. T. MALONEY : *Bayesian modelling of visual perception*, p. 13–36. Rao, R. P. N. and Olshausen, B. A. and Lewicki, M. S., 2002.
- P. F. McLAUHLAN et D. W. MURRAY : A unifying framework for structure and motion recovery from image sequences. *In Proceedings of the International Conference on Computer Vision*, p. 314–320, 1995.
- J. B. MULLIGAN : Anisotropy in an ambiguous kinetic depth effect. *Journal of the Optical Society of America A*, 9(4):521–9, 1992.

- K. P. MURPHY : An introduction to graphical models. Rap. tech., 2001. URL http://www.cs.ubc.ca/~murphyk/Papers/intro_gm.pdf.
- H.-H. NAGEL : On the estimation of optical flow: relations between different approaches and some new results. *Artificial Intelligence*, 33:299–324, 1987.
- J. J. NAJI et T. C. A. FREEMAN : Perceiving depth order during pursuit eye movement. *Vision Research*, 44:3025–34, 2004.
- R. M. NEAL : Probabilistic inference using markov chain monte carlo methods. Rap. tech. CRG-TR-93-1, Dept. of Computer Science, University of Toronto, 1993.
- J. F. NORMAN et J. T. TODD : The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations. *Perception and Psychophysics*, 53(3):279–91, 1993.
- J. OLIENSIS : The least-squares error for structure from infinitesimal motion. *International Journal of Computer Vision*, 61(3):259–99, 2005.
- F. PANERAI, V. CORNILLEAU-PÉRÈS et J. DROULEZ : Contribution of extraretinal signals to the scaling of object distance during self-motion. *Perception and Psychophysics*, 64:717–731, 2002.
- J. PEARL : Reverend Bayes on inference engine: a distributed hierarchical approach. In D. WALTZ, éd. : *Proceedings of the AAAI National Conference on AI*, p. 133–6. AAAI Press, 1982.
- J. PEARL : *Probabilistic reasoning in intelligent systems*. Morgan Kauffmann, San Mateo, California, 1988.
- C.-H. PEH, F. PANERAI, J. DROULEZ, V. CORNILLEAU-PÉRÈS et L.-F. CHEONG : Absolute distance perception during in-depth head movement: calibrating optic flow with extra-retinal information. *Vision Research*, 42:1991–2003, 2002.
- Z. PIZLO : Perception viewed as an inverse problem. *Vision Research*, 41(24):3141–61, 2001.
- T. POGGIO : Vision by man and machine. *Scientific American*, 250:106–116, 1984.
- C. PRADALIER, F. COLAS et P. BESSIÈRE : Expressing bayesian fusion as a product of distributions: application to randomized Hough transform. In *Proceedings of the Conference on Bayesian Methods and Maximum Entropy in Science and Engineering*, Jackson Hole, WY (US), August 2003. URL <http://emotion.inrialpes.fr/bibemotion/2003/PCB03a>.
- B. ROGERS et M. GRAHAM : Motion parallax as an independent cue for depth perception. *Perception*, 8:125–134, 1979.

- B. J. ROGERS et S. ROGERS : Visual and nonvisual information disambiguate surfaces specified by motion parallax. *Perception and Psychophysics*, 52:446–452, 1992.
- P. R. SCHRATER, D. C. KNILL et E. P. SIMONCELLI : Mechanisms of visual motion detection. *Nature Neuroscience*, 3:64–8, 2000.
- C. E. SHANNON : A mathematical theory of communication. *The Bell System Technical Journal*, 27:623–656, Oct 1948.
- L. G. SHAPIRO et G. STOCKMAN : *Computer vision*. Prentive Hall, 2001.
- P. SMYTH : Belief network, hidden Markov models, and Markov random fields: a unifying view. *Pattern Recognition Letters*, 18(11-13):1261–1268, 1997.
- G. SPERLING, M. S. LANDY, B. A. DOSHER et M. E. PERKINS : Kinetic depth effect and identification of shape. *Journal of Experimental Psychology: Human Perception and Performance*, 15(4):826–40, 1989.
- J. T. TODD : Visual information about rigid and non-rigid motion: a geometric analysis. *Journal of Experimental Psychology: Human Perception and Performance*, 8(2):238–252, 1982.
- J. T. TODD et J. F. NORMAN : The visual perception of smoothly curved surfaces from minimal apparent motion sequences. *Perception and Psychophysics*, 50(6):509–23, 1991.
- B. TRIGGS : Factorization methods for projective structure and motion. *In International Conference on Computer Vision and Pattern Recognition*, p. 845–851, San Francisco, CA, 1996.
- S. ULLMAN : *The interpretation of visual motion*. Cambridge: MIT Press, 1979.
- S. ULLMAN : The optical flow of planar surfaces. *Spatial Vision*, 1(4):263–76, 1986.
- J. J. A. van BOXTEL, M. WEXLER et J. DROULEZ : Perception of plane orientation from self-generated and passively observed optic flow. *Journal of Vision*, 3(5):318–32, 2003. URL <http://journalofvision.org/3/5/1/>.
- K. S. van HORN : Constructing a logic of plausible inference: a guide to Cox’s theorem. *International Journal of Approximate Reasoning*, 34(1):3–24, 2003. URL <http://www.cs.ndsu.nodak.edu/~vanhorn/tr/rcox.ps>.
- T. VIÉVILLE, C. ZELLER et L. ROBERT : Using collineations to compute motion and structure in an uncalibrated image sequence. *International Journal of Computer Vision*, 20:213–42, 1996.
- H. von HELMHOLTZ. : *Handbuch der Physiologischen Optik*. Hamburg: Voss, 1867.

- H. WALLACH et D. N. O'CONNEL : The kinetic depth effect. *Journal of Experimental Psychology*, 45:205–217, 1953.
- H. WALLACH, J. STANTON et D. BECKER : The compensation for movement-produced changes in object orientation. *Perception and Psychophysics*, 15:339–343, 1974.
- W. H. WARREN : Self-motion: visual perception and visual control. In *Perception of space and motion*, vol. 5 de *Handbook of perception and cognition*. Academic Press, 1995.
- A. M. WAXMAN et S. ULLMAN : Surface structure and three-dimensional motion from image flow kinematics. *International Journal of Robotics Research*, 4(3):72–94, 1985.
- Y. WEISS, E. P. SIMONCELLI et E. H. ADELSON : Motion illusions as optimal percepts. *Nature Neuroscience*, 5(6):598–604, 2002.
- M. WEXLER : Voluntary head movement and allocentric perception of space. *Psychological Science*, 14:340–346, 2003.
- M. WEXLER, I. LAMOURET et J. DROULEZ : The stationarity hypothesis: an allocentric criterion in visual perception. *Vision Research*, 41:3023–3037, 2001a.
- M. WEXLER, F. PANERAI, I. LAMOURET et J. DROULEZ : Self-motion and the perception of stationary objects. *Nature*, 409:85–88, 2001b.
- A. YU et P. DAYAN : Expected and unexpected uncertainty: ACh and NE in the neocortex. In S. BECKER, S. THRUN et K. OBERMAYER, édés : *Advances in Neural Information Processing Systems 15*, p. 157–164, Cambridge, MA, 2003. MIT Press.
- R. S. ZEMEL, P. DAYAN et A. POUGET : Probabilistic interpretation of population codes. *Neural Computation*, 10(2):403–30, 1997.
- K. ZHANG, I. GINZBURG, B. L. MCNAUGHTON et T. J. SEJNOWSKI : Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. *Journal of Neurophysiology*, 79:1017–44, 1998.

Annexe A

Équations du flux optique d'un plan

Dans ce chapitre, nous démontrons les équations que nous utilisons, qui donnent le flux optique d'un plan en fonction de son mouvement et de sa position.

Principe général Nous considérons un plan \mathcal{P} , un point \widetilde{M} de ce plan dans le repère 3D et sa projection M dans le plan image. Le plan est soumis à une translation et une rotation. Le flux optique s'obtient en dérivant les coordonnées du point de l'image par rapport au temps. Pour cela, nous utilisons la composition de la projection sur le plan image et du mouvement 3D du point du plan objet.

Formalisation Plus précisément, soit \mathcal{P} un plan objet de paramètres (χ, v) . $(\chi, v, -1)$ est un vecteur normal de ce plan.

On note $\vec{t} = (t_x, t_y, t_z)$ le vecteur vitesse du plan \mathcal{P} et $\vec{\omega} = (\omega_x, \omega_y, \omega_z)$ son vecteur de vitesse de rotation.

Soit \widetilde{M} , de coordonnées $(\tilde{x}, \tilde{y}, \tilde{z})$, un point du plan dans l'espace de coordonnée 3D. Il satisfait l'équation du plan :

$$\begin{aligned}(\tilde{x}, \tilde{y}, \tilde{z}) \cdot (\chi, v, -1) &= 0 \\ \tilde{x}\chi + \tilde{y}v - \tilde{z} &= 0\end{aligned}\tag{A.1}$$

Soit P l'application qui projette¹ un point \widetilde{M} en un point M de l'image de coordonnées (x, y) . On a :

$$P : \begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R}^2 \\ \begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} & \mapsto & \begin{pmatrix} x = \frac{\tilde{x}}{1-\tilde{z}} \\ y = \frac{\tilde{y}}{1-\tilde{z}} \end{pmatrix} \end{array}\tag{A.2}$$

En résumé, \widetilde{M} est un point du plan en mouvement, M est sa projection sur l'image et P l'application de projection de \widetilde{M} en M . En considérant les points comme des fonctions

¹ P n'est pas une projection au sens d'application linéaire.

du temps, on a :

$$M(t) = P \circ \widetilde{M}(t) \quad (\text{A.3})$$

Le flux optique $\vec{\Phi} = (u_x, u_y)$ est le déplacement des points de l'image, c'est-à-dire la dérivée des coordonnées des points par rapport au temps :

$$\vec{\Phi} = \frac{dM}{dt} \quad (\text{A.4})$$

En appliquant la relation A.3 et la règle de dérivation de la composition, on peut écrire :

$$\vec{\Phi} = \frac{dP}{d\widetilde{M}}(\widetilde{M}) \times \frac{d\widetilde{M}}{dt} \quad (\text{A.5})$$

L'équation A.5 fait apparaître la dérivée de la projection, $\frac{dP}{dM}$, et le mouvement du point 3D, $\frac{d\widetilde{M}}{dt}$.

Dérivée de la projection Nous commençons par calculer le terme $\frac{dP}{dM}$. Il s'agit de la matrice jacobienne de P , il faut donc calculer les dérivées partielles selon chacune des trois composantes de \widetilde{M} des deux composantes de M .

On obtient alors l'expression :

$$\begin{aligned} \frac{dP}{d\widetilde{M}}(\widetilde{M}) &= \begin{pmatrix} \frac{\partial x}{\partial \tilde{x}} & \frac{\partial x}{\partial \tilde{y}} & \frac{\partial x}{\partial \tilde{z}} \\ \frac{\partial y}{\partial \tilde{x}} & \frac{\partial y}{\partial \tilde{y}} & \frac{\partial y}{\partial \tilde{z}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{1-\tilde{z}} & 0 & \frac{\tilde{x}}{(1-\tilde{z})^2} \\ 0 & \frac{1}{1-\tilde{z}} & \frac{\tilde{y}}{(1-\tilde{z})^2} \end{pmatrix} \end{aligned} \quad (\text{A.6})$$

Mouvement du point 3D Nous calculons maintenant le mouvement du point 3D du plan. Le plan \mathcal{P} est soumis à un mouvement $\vec{t}, \vec{\omega}$. Le mouvement $\frac{d\widetilde{M}}{dt}$ d'un point \widetilde{M} de ce plan est donc :

$$\begin{aligned} \frac{d\widetilde{M}}{dt} &= \vec{t} + \vec{\omega} \wedge \widetilde{M} \\ &= \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} + \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \wedge \begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \chi\tilde{x} + \nu\tilde{y} \end{pmatrix} \\ \frac{d\widetilde{M}}{dt} &= \begin{pmatrix} t_x + \chi\omega_y\tilde{x} + (\nu\omega_y - \omega_z)\tilde{y} \\ t_y + (\omega_z - \chi\omega_x)\tilde{x} - \nu\omega_x\tilde{y} \\ t_z + \omega_x\tilde{y} - \omega_y\tilde{x} \end{pmatrix} \end{aligned} \quad (\text{A.7})$$

Équations du flux optique Nous substituons les expressions A.6 et A.7 dans l'équation A.5 pour obtenir :

$$\begin{aligned}
\vec{\Phi} &= \frac{dP}{d\widetilde{M}}(\widetilde{M}) \times \frac{d\widetilde{M}}{dt} \\
&= \begin{pmatrix} \frac{1}{1-\bar{z}} & 0 & \frac{\bar{x}}{(1-\bar{z})^2} \\ 0 & \frac{1}{1-\bar{z}} & \frac{\bar{y}}{(1-\bar{z})^2} \end{pmatrix} \times \begin{pmatrix} t_x + \chi\omega_y\bar{x} + (v\omega_y - \omega_z)\bar{y} \\ t_y + (\omega_z - \chi\omega_x)\bar{x} - v\omega_x\bar{y} \\ t_z + \omega_x\bar{y} - \omega_y\bar{x} \end{pmatrix} \\
\vec{\Phi} &= \begin{pmatrix} \frac{t_x + \chi\omega_y\bar{x} + (v\omega_y - \omega_z)\bar{y}}{1-\bar{z}} + \frac{\bar{x}}{1-\bar{z}} \times \frac{t_z + \omega_x\bar{y} - \omega_y\bar{x}}{1-\bar{z}} \\ \frac{t_y + (\omega_z - \chi\omega_x)\bar{x} - v\omega_x\bar{y}}{1-\bar{z}} + \frac{\bar{y}}{1-\bar{z}} \times \frac{t_z + \omega_x\bar{y} - \omega_y\bar{x}}{1-\bar{z}} \end{pmatrix} \quad (\text{A.8})
\end{aligned}$$

Or, d'après la définition de P (équation A.2), on a : $\frac{\bar{x}}{1-\bar{z}} = x$, $\frac{\bar{y}}{1-\bar{z}} = y$ et aussi $\frac{1}{1-\bar{z}} = 1 + \chi x + v y$. Nous utilisons ces expressions dans l'équation A.8 pour obtenir les équations A.9 du flux optique généré par un plan :

$$\begin{aligned}
\vec{\Phi} &= \begin{pmatrix} \frac{t_x + \chi\omega_y\bar{x} + (v\omega_y - \omega_z)\bar{y}}{1-\bar{z}} + \frac{\bar{x}}{1-\bar{z}} \times \frac{t_z + \omega_x\bar{y} - \omega_y\bar{x}}{1-\bar{z}} \\ \frac{t_y + (\omega_z - \chi\omega_x)\bar{x} - v\omega_x\bar{y}}{1-\bar{z}} + \frac{\bar{y}}{1-\bar{z}} \times \frac{t_z + \omega_x\bar{y} - \omega_y\bar{x}}{1-\bar{z}} \end{pmatrix} \\
\vec{\Phi} &= \begin{pmatrix} t_x + x [t_z + \chi(t_x + \omega_y)] + y [-\omega_z + v(t_x + \omega_y)] + x^2(\chi t_z - \omega_y) + xy(vt_z + \omega_x) \\ t_y + x[\omega_z + \chi(t_y - \omega_x)] + y[t_z + v(t_y - \omega_x)] + xy(\chi t_z - \omega_y) + y^2(vt_z + \omega_x) \end{pmatrix} \\
\vec{\Phi} &= \vec{\phi}^0 + \vec{\phi}^1 \cdot {}^t(x, y) + \vec{\phi}^2 \cdot {}^t(x, y) \quad (\text{A.9})
\end{aligned}$$

avec :

$$\begin{aligned}
\vec{\phi}^0 &= \begin{pmatrix} \phi_x^0 \\ \phi_y^0 \end{pmatrix} = \begin{pmatrix} t_x \\ t_y \end{pmatrix} \\
\vec{\phi}^1 &= \begin{pmatrix} x\phi_x^1 & y\phi_x^1 \\ x\phi_y^1 & y\phi_y^1 \end{pmatrix} = \begin{pmatrix} t_z + \chi(t_x + \omega_y) & -\omega_z + v(t_x + \omega_y) \\ \omega_z + \chi(t_y - \omega_x) & t_z + v(t_y - \omega_x) \end{pmatrix} \\
\vec{\phi}^2 &= \begin{pmatrix} x\phi^2 \\ y\phi^2 \end{pmatrix} = \begin{pmatrix} \chi t_z - \omega_y \\ vt_z + \omega_x \end{pmatrix} \blacksquare
\end{aligned}$$

Abstract. Perception can be seen as collecting and confronting various pieces of information in order to understand the environment. Man uses many sensory modalities such as vision, touch, audition, or proprioception. Various kind of information can be conveyed by each sense. For example vision is about texture, colour, shapes or movement. Optic flow is the displacement of the image on the retina. Man can perceive part of the shapes from optic flow alone. However, extracting structure from motion is a complicated issue. First, perception is an *inverse* problem in general. Geometry can give the optic flow given the shape and movement of the objects of the scene. Then, many configurations can lead to the same optic flow. Such a problem is said *ill-posed*. Last, one can seldom be sure of some perception with an arbitrary precision. Perception is therefore an *uncertain* problem. In order to solve these issue, we use the *Bayesian Programming* formalism. It is designed to reason with uncertainty using probabilities. Multiple solutions of an ill-posed problem are taken into account using multi-modal probability distributions. Finally, the symmetry of Bayes' rule allows for the same inference for either direct or inverse problems. We propose a Bayesian model of the perception of planes from the optic flow based on some explicit assumptions. The main hypotheses involved in our model are *rigidity* and *stationarity*. Rigidity is a constraint that minimises the relative 3D motion of the observed dots. Stationarity minimises the absolute 3D motion of the dots. We validate our model with five experiments from the literature. For each experiment, the model reproduces qualitative features of the results and allows for an explanation of the experimental result based on the functioning of the model. Finally, we show how to adapt our model to an experiment of perception of a corrugated surface by optic flow. We also present a Bayesian model to generate the input of our perception model and generalise the integration of our model into a more complete system.

Keywords : Bayesian model, perception, structure from motion, optic flow.

Résumé. La perception est le regroupement et recouplement d'un ensemble d'informations en vue de comprendre son environnement. À cette fin, l'homme utilise plusieurs modalités sensorielles telles que la vision, le toucher, l'ouïe ou même la proprioception. Ces sens eux-mêmes peuvent convoyer différents types d'information. La vision par exemple, propose des indices de texture, de couleurs, de formes ou de mouvement. Le flux optique est ainsi le déplacement de l'image sur la rétine. Avec lui seul, un homme est capable d'appréhender une partie des formes observées. Or, extraire la forme à partir du mouvement est un problème très complexe. D'une part, la perception en général est un problème *inverse*. La géométrie permet de déterminer le flux optique à partir de la forme et du mouvement des objets. Ensuite, plusieurs configurations peuvent engendrer un même flux optique. Dans ce cas, on dit que le problème est *mal-posé*. Enfin, il est rarement possible d'être certain d'une perception à une précision arbitraire. Toute perception est donc entachée d'*incertitude*. Pour résoudre ces difficultés, nous utilisons le formalisme de la *programmation bayésienne*. Celui-ci permet de raisonner sur l'incertitude à l'aide des probabilités. Les multiples solutions d'un problème mal-posé sont prises en compte par la multi-modalité des distributions de probabilité. Enfin, la symétrie de la règle de Bayes permet de faire l'inférence de la même manière que ce soit sur les problèmes inverses et directs. Nous proposons un modèle bayésien de la perception de plans par le flux optique, basé sur quelques hypothèses explicites. Les principales hypothèses mises en œuvre dans ce modèle sont la *rigidité* et la *stationnarité*. La rigidité porte sur le mouvement des points et exprime que le flux optique est plus probablement dû à un objet rigide qu'à une collection de points de mouvements indépendants. La stationnarité porte sur le mouvement de l'objet en indiquant qu'il est plus probablement petit. Nous validons notre modèle en nous basant sur cinq expériences issues de la littérature. Pour chacune d'elles, nous expliquons les détails du fonctionnement du modèle et les hypothèses qui conditionnent la bonne reproduction des résultats. Enfin, nous montrons d'une part comment intégrer notre modèle au sein de systèmes plus complexes et, d'autre part, l'applicabilité du modèle bayésien à des expériences de perception du flux optique différentes de part leurs protocoles, mais similaires dans leurs ambiguïtés.

Mots clé : modèle bayésien, perception, structure par le mouvement, flux optique.