

Développement d'un système passif de suivi 3D du mouvement humain par filtrage particulaire

Jamal Saboune

▶ **To cite this version:**

Jamal Saboune. Développement d'un système passif de suivi 3D du mouvement humain par filtrage particulaire. Informatique [cs]. Université de Technologie de Troyes, 2008. Français. NNT : . tel-00598867

HAL Id: tel-00598867

https://theses.hal.science/tel-00598867

Submitted on 7 Jun 2011

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.

Université de Technologie de Troyes (UTT)

Ecole Doctorale « Sciences des Systèmes Technologiques et Organisationnels »

THESE

Présentée en vue d'obtenir le grade de docteur de l'université

Spécialité «Optimisation et Sûreté des Systèmes»

Par :

Jamal SABOUNE

Développement d'un système passif de suivi 3D du mouvement humain par filtrage particulaire

Composition du jury

Rapporteurs :	Rachid ALAMI	(Directeur de Recherche CNRS, LASS – Toulouse)
	Jean-Louis VERCHER	(Directeur de Recherche CNRS, Laboratoire de Mouvement et Perception, Marseille)
Examinateurs :	James CROWLEY	(Professeur, INPG – Grenoble)
	David HEWSON	(Enseignant-chercheur, Université de Technologie de Troyes)
	Jean PAYSANT	(Professeur, Faculté de Médecine - UHP, Nancy)
	Hichem SNOUSSI	(Maître de conférences, Université de Technologie de Troyes)
Directeurs :	François CHARPILLET	(Directeur de Recherche, INRIA – Nancy)
	Jacques DUCHENE	(Professeur, Université de Technologie de Troyes)

Remerciements

Je tiens tout d'abord à remercier mon directeur de thèse M. François Charpillet avec qui j'ai eu la chance et le plaisir de travailler durant ces quatre années. Ses conseils aux niveaux scientifique et personnel ont toujours été clairs et m'ont permis d'aboutir à la production de ce travail. Je le remercie aussi pour sa disponibilité et les discussions régulières que nous avons eues. Je remercie également mon co-directeur de thèse M. Jacques Duchêne, pour ses conseils et ses remarques concernant ce manuscrit.

Je remercie M. Rachid Alami et M. Jean Louis Vercher pour avoir accepté de rapporter ma thèse et pour les efforts que leur a demandés ce travail. Leurs critiques étaient constructives et leurs indications m'étaient utiles.

Je présente mes vifs remerciements à M. James Crowley qui m'a fait l'honneur de présider le jury et pour sa lecture attentive du manuscrit.

Je remercie également les membres du jury M. David Hewson et M. Hichem Snoussi pour l'intérêt qu'ils ont eu pour mon travail.

J'exprime mes sincères remerciements à M. Jean Paysant pour sa collaboration, ses conseils, son enthousiasme et sa disponibilité qui m'ont permis d'approfondir l'étude du comportement locomoteur humain et de valider le travail effectué. Je tiens à lui exprimer ma profonde reconnaissance et je le remercie vivement pour sa participation au jury.

Je remercie Cédric Rose pour sa contribution scientifique dans ce travail de thèse et pour son aide. Je le remercie également ainsi que Daniel Szer et Sylvain Castagnos, pour l'ambiance de travail détendue qu'ils ont fait régner dans le bureau, pour leur vraie amitié, pour leur écoute et pour leur soutien dans les moments difficiles. Sans eux, ces quatre années passées n'auraient pas été aussi agréables.

Je tiens aussi à remercier Maan el Badaoui el Najjar pour son amitié et ses conseils ainsi que Vincent Thomas pour l'aide et la motivation qu'il m'a transmis.

Je remercie tous les membres de l'équipe MAIA qui m'ont permis d'effectuer ce travail dans la bonne humeur. Je dois un grand merci aux membres (actuels et passés) du LORIA que j'ai pu côtoyer durant mon séjour. Dans l'ordre alphabétique : Yoann Bertrand-Pierron, Geoffray Bonnin, Luc Buatois, Rodolphe Charrier, Gabriel Corona, Abdallah Dib, Ilham Esslimani, Arnaud Glad, Régis Lhoste, Sergio Nogueira, Loïc Pelissier, Céline Simon, Olivier Simonin, Julien Thomas, Cherif Smaili, François Rispal. J'en oublie sûrement, qu'ils veulent bien m'en excuser.

Je pense aussi à mes amis avec qui j'ai eu des moments de complicité durant cette période passée à Nancy. Dans l'ordre alphabétique : Ahmad Dib, Toufik El Khatib, George Habib, Rimond Hamia, Mohamad Hajj Chehadeh, Jamil Houhou, Mohamad Nassar, Nazih Ouwayed, Rami Saad.

Enfin, je veux exprimer toute ma gratitude à mes parents et à ma famille qui m'ont soutenu et encouragé depuis toujours et sans lesquels je ne serais pas parvenu à accomplir ce travail.

Résumé

Les chutes des personnes âgées à domicile présentent un problème important à cause de leur fréquence et de leurs conséquences médicales, psychologiques et économiques. Pour cela, l'idée d'une prévention des chutes s'avère nécessaire pour réduire leur nombre et leurs conséquences indésirables. Cette idée constitue la motivation principale des projets ANR PARAChute et Predica dont le but est de détecter une dérive vers la chute chez les personnes âgées à domicile. Ce travail de thèse s'inscrit dans le cadre de ce projet et a comme objectif de concevoir un système d'analyse de la marche qui devrait s'intégrer dans la méthode à proposer pour évaluer un risque de chute chez la personne âgée.

Une étude des approches déjà adoptées pour évaluer l'équilibre dynamique d'une personne nous a permis d'identifier un certain nombre de paramètres et de tests validés comme étant pertinents pour la détection d'un risque de chute. Or, ces approches s'appuient sur des tests cliniques effectués par les gériatres ou sur l'utilisation de capteurs portables encombrants d'où la nécessité de développer un nouvel outil autonome et non encombrant pour extraire ces paramètres. Notre solution a été de concevoir un système de suivi 3D complet du mouvement humain, capable de nous fournir les positions de quelques points clé du corps durant la marche et à partir desquelles nous pouvons évaluer les paramètres recherchés.

Le système de suivi développé utilise des images vidéo obtenues à l'aide d'un caméscope numérique grand public et ne nécessite pas l'intervention d'un opérateur ou l'usage de marqueurs portables. N'utilisant aucun modèle temporel de marche notre méthode est générique et simple. Elle est fondée sur un modèle 3D virtuel humanoïde et sur un nouvel algorithme de filtrage particulaire que nous avons développé dans le but d'estimer la configuration la plus probable de ce modèle vis à vis de l'image vidéo du corps réel. En raison de la nature non gaussienne et multimodale des densités de probabilité impliquées, l'usage d'une approche particulaire était nécessaire.

L'algorithme développé, appelé 'Interval Particle Filtering' (IPF), réorganise l'espace de recherche des configurations probables du modèle virtuel d'une façon déterministe, efficace et optimisée. Dans le but de réduire la complexité de cet algorithme, une version factorisée utilisant les réseaux bayésiens dynamiques a été ensuite proposée. Cette version améliorée profite de la

modélisation de la marche en une chaîne cinématique pour factoriser la construction du vecteur d'état et la fonction d'observation.

En terme de reconstruction 3D, les résultats obtenus à l'aide de notre système de suivi sont assez satisfaisants. La comparaison des valeurs des paramètres recherchés, calculées à partir des positions 3D fournies par un système de référence et par l'algorithme (IPF), montre une erreur moyenne de mesure de 5%. Une étude du suivi longitudinal chez les personnes âgées sera alors nécessaire pour choisir les paramètres qui permettent de déceler une évolution vers un risque de chute chez les personnes âgées.

Plan de la thèse

Le plan du présent manuscrit de thèse est le suivant :

Le premier chapitre forme une introduction générale et recadre les motivations et les objectifs des travaux menés dans le cadre de cette thèse. En effet, dans cette introduction nous exposerons la problématique de la chute chez les personnes âgées à domicile, sous l'angle des facteurs précipitants et de leurs conséquences de ce type de chutes. Cette problématique est la motivation principale de l'étude PARAChute (Personnes Agées et Risque de Chute) dans le cadre de laquelle est menée cette thèse. L'objectif de cette thèse peut se résumer à la conception et la validation d'un système d'analyse de la marche. Ce capteur 'passif' devrait servir ensuite pour construire un indicateur de risque de chute intégré dans un système de suivi longitudinal pour la prévention des chutes chez les personnes âgées.

Dans le deuxième chapitre nous détaillerons toutes les méthodes déjà utilisées pour évaluer les risques de chutes et les différents systèmes d'analyse de la marche commercialisés. En effet nous décrirons les tests gériatriques cliniques, les approches à base de capteurs portables et les parcours instrumentés utilisés dans le cadre d'une évaluation de la qualité d'équilibre. Nous montrerons pour chacune de ces méthodes les avantages et les inconvénients. Le choix des paramètres à évaluer pour former l'indicateur de risque sera exposé et la nécessité du développement d'un nouveau système d'analyse sera également discutée.

Le troisième chapitre sera consacré à l'étude des différentes techniques utilisées pour effectuer un suivi 3D complet du mouvement humain à partir d'une vidéo. La solution retenue étant fondée sur l'usage d'un modèle 3D humanoïde virtuel, une description des méthodes permettant d'estimer la configuration de ce modèle à partir d'une observation donnée, sera faite. Le choix d'adopter une approche particulaire de ce problème d'estimation sera aussi justifié.

La description détaillée de la solution retenue sera présentée dans le chapitre 4. Ainsi, nous détaillerons le modèle virtuel utilisé, la fonction d'observation utilisée et l'algorithme 'Interval

Particle Filtering' que nous avons développé. Les améliorations que nous avons apportées à cet algorithme pour réduire sa complexité, seront aussi introduites. Ces améliorations sont d'ordre algorithmique et d'ordre technique.

Le chapitre 5 décrira la mise en œuvre de l'algorithme développé, et présentera les résultats de reconstruction 3D de quelques mouvements complexes. Aussi, nous dresserons une comparaison portant sur les paramètres de la marche calculés à partir de notre système développé et à partir d'un système de suivi 3D de référence.

Nous conclurons dans le chapitre 6 en citant les contributions principales de ce travail de thèse et en décrivant les perspectives à suivre.

Table de matières :

Résumé 1
Plan de la thèse 3
Chapitre 1 : Cadre, objectifs et motivations 11
1.1 Objectifs et contraintes du projet PARAChute 12
1.2 Objectifs de la thèse 13
1.3 Description de la marche 14
1.4 Causes des chutes 15
1.4.1 Facteurs de risque de chutes chroniques 15
1.4.2 Facteurs précipitants intrinsèques 17
1.4.3 Facteurs précipitants extrinsèques 19
1.5 La chute chez la personne âgée 19
1.5.1 Statistiques 19
1.5.2 Conséquences des chutes chez les personnes âgées 20
1.6 Vieillissement de la population en France et conséquences 21
1.7 Conclusion 23
Bibliographie 24
Chapitre 2 : Différentes approches pour l'évaluation des risques de chute chez les personnes âgées 25
2.1 Tests cliniques usuels 26
2.1.1 Test de Berg 27
2.1.2 L'indice 'DGI' (Dynamic Gait Index) 27
2.1.3 Le test de '8 foot Up and Go' 28
2.1.4 Le test de Tinetti 28
2.1.5 Tests divers 30
2.1.6 Conclusion 31
2.2 Approches ambulatoires par capteurs portables 32
2.2.1 Electromyographie 32
2.2.2 Semelles instrumentées et capteurs FSR 33
2.2.3 Goniométrie 34
2.2.4 Utilisation d'accéléromètres 35
2.2.5 Utilisation de gyroscopes 37
2.3 Approches passives par parcours instrumentés 38

2.3.1 Plate-forme de force 38
2.3.2 Tapis actimétriques 39
2.3.3 Télémètre 40
2.3.4 Systèmes de suivi du mouvement 3D 41
2.4 Discussion et conclusion 43
2.4.1 Incompatibilité des méthodes et des technologies utilisées avec les contraintes du projet PARAChute 44
2.4.2 Récapitulation des paramètres identifiés comme pertinents pour l'évaluation de l'équilibre dynamique 45
2.4.3 Quel type de capteurs adopter ? 46
Bibliographie 47
Chapitre 3 : Suivi 3D du mouvement humain et filtrage particulaire 51
3.1 Analyse globale du mouvement humain par stéréovision 52
3.2 Analyse fine du mouvement humain par utilisation de modèles humanoïdes 54
3.2.1 Le système w4 55
3.2.2 Utilisation des modèles mathématiques de mouvement 56
3.2.3 Méthode de Gavrila et Davis 57
3.2.4 Suivi 2D et reconstruction 3D à l'aide de modèles appris 58
3.2.5 Discussion et conclusion 59
3.3 Formulation du problème 60
3.4 La théorie de l'estimation dynamique 61
3.5 Filtres Bayésiens optimaux 64
3.5.1 Filtre de Kalman 64
3.5.2 Approche par maillage de l'espace d'état 66
3.5.3 Conclusion 66
3.6 Approches particulaires pour l'estimation dynamique 67
3.6.1 Principe et procédure de base (dite *'a priori'*) 67
3.6.2 Algorithme de Condensation 69
3.6.3 Application du filtrage particulaire pour le suivi du mouvement humain 71
3.7 Conclusion 73
Bibliographie 74
Chapitre 4 : Approche proposée 77
4.1 Le modèle 3D virtuel utilisé 78
4.2 La fonction de vraisemblance utilisée 79

4.2.1 Filtre seuil 81
4.2.2 Filtre Médian 81
4.2.3 Filtre HSV 82
4.3 Quel estimateur bayésien utiliser ? 83
4.4 L'algorithme 'Interval Particle Filtering' (IPF) 85
4.4.1 Différences avec l'algorithme de Condensation 85
4.4.2 Evolution dynamique des particules dans IPF 87
4.4.3 Structure générale de l'algorithme IPF 88
4.4.3.1 L'étape de 'Sélection' 88
4.4.3.2 L'étape de 'Diffusion' 88
4.4.3.2 L'étape de 'Mesure' 90
4.4.4 Discussion 92
4.5 L'algorithme 'Factored Interval Particle Filtering' (FIPF) 92
4.5.1 Réseaux bayésiens 93
4.5.2 Réseaux bayésiens dynamiques 94
4.5.3 Inférence dans les réseaux bayésiens et utilisation du filtrage particulaire 97
4.5.4 Structure du réseau bayésien utilisé 97
4.5.4.1 Factorisation du vecteur d'état à l'aide d'une chaîne cinématique 98
4.5.4.2 Factorisation de la fonction d'observation ('poids') 99
4.5.4.3 Représentation de la dynamique des particules 100
4.5.5 Intérêt de la factorisation 100
4.5.6 Algorithme d'inférence utilisé dans FIPF 102
4.6 Conclusion 103
Bibliographie 105
Chapitre 5 : Mise en oeuvre et résultats 107
5.1 Dispositif expérimental 108
5.1.1 Description du matériel utilisé 108
5.1.2 Description de l'ensemble des sujets tests et du protocole adopté 109
5.2 Paramètres et réglages des algorithmes utilisées 110
5.2.1 Création de l'image 'silhouette' 110
5.2.2 Réglage et calibration de la caméra virtuelle 110
5.2.3 Choix des paramètres de l'algorithme 113
5.2.4 Initialisation de l'algorithme 'IPF' 114
5.2.5 Détermination des coordonnées du point 'origine' et de l'orientation du corps 116

5.2.6 Schéma général de la méthode 118
5.3 Extraction des paramètres 119
5.3.1 Vitesse et accélération 120
5.3.2 Longueurs et largeur des pas 121
5.3.3 Hauteurs des pas gauche et droit 123
5.3.4 Régularité et symétrie des pas 123
5.3.5 Oscillation du tronc dans le plan medio-lateral, angles de flexion aux genoux et chevilles 124
5.3.6 Temps d'appui et du double appui 124
5.3.7 Déviation du trajet 127
5.4 Résultats et comparaisons des paramètres 127
5.4.1 Longueur du pas gauche 127
5.4.2 Longueur du pas droit 128
5.4.3 Largeur du pas 129
5.4.4 Vitesse instantanée 131
5.4.5 Temps d'appui gauche, droit et de double appui 132
5.4.6 Comparaison des mouvements de quelques points clés 133
5.4.7 Discussion 134
5.5 Reconstruction 3D 135
5.5.1 Marche rapide avec occlusion 135
5.5.2 Levée d'une chaise 136
5.5.3 Attraper un objet par terre 136
5.5.4 Suivi en plein air 137
5.5.5 Simulation d'une amputation 138
5.6 Complexité de l'algorithme et temps de calcul 139
5.6.1 Amélioration au niveau des technologies 140
5.6.2 Amélioration au niveau de l'algorithme développé 141
5.7 Conclusion 141
Chapitre 6 : Conclusions générales et perspectives 143
6.1 Motivations, objectifs et contraintes 143
6.2 Synopsis de la thèse 144
6.3 Contributions essentielles 148
6.3.1 Conception d'un nouveau système de suivi 3D du mouvement humain 148
6.3.2 Conception d'une méthode passive pour l'analyse fine de la marche humaine 148

6.3.3 Développement d'un nouvel algorithme de filtrage particulaire IPF 149
6.3.4 Conception d'une approche factorisée du filtrage particulaire..... 150
6.4 Perspectives..... 151
6.4.1 Vers un système de suivi 3D en temps réel ?..... 151
6.4.2 Vers un indicateur fiable de risque de chute chez les personnes âgées 152
Annexe A : Formation de l'image et calibration d'une caméra..... 155
Bibliographie..... 162
Annexe B: Description du modèle 3D virtuel utilisé..... 163
Publications..... 171

Chapitre 1 : Cadre, objectifs et motivations

Ce travail de thèse s'inscrit dans le cadre du projet national RNTS PARAChute. Le but de ce projet est de développer une méthode et un outil pour détecter une dérive vers la chute chez la personne âgée à domicile, en adoptant un suivi longitudinal de l'évolution du risque de chute. Cette méthode doit se fonder sur l'utilisation de capteurs non encombrants à domicile, d'où l'objectif principal de cette thèse qui est la conception et le développement d'un système passif d'analyse de la marche. Ce système ainsi développé fera partie de la méthode globale à proposer.

Dans ce chapitre nous présentons les objectifs généraux du projet PARAChute qui s'inscrivent dans la logique de prévention primaire de chute des personnes âgées à domicile. En effet, une fois la tendance à la chute détectée, une intervention d'ordre médical ou kinésithérapeutique pourrait réduire le risque de déséquilibre. Nous exposerons aussi les critères auxquels doivent se conformer les méthodes et les capteurs utilisés dans le but d'évaluer ce risque, tels que l'objectivité, la simplicité et l'adaptation à la vie quotidienne. Ensuite, nous détaillerons les étapes menant à réaliser l'objectif final de ce travail de thèse qui est de développer un système non intrusif et non encombrant d'analyse de la marche, pouvant s'intégrer facilement au domicile.

Le système à concevoir porte sur l'évaluation de l'équilibre dynamique. Par conséquent, après une brève description biomécanique du mouvement de la marche chez l'homme et des systèmes musculaires, proprioceptifs et cérébraux qui y interviennent, nous détaillerons les facteurs chroniques, intrinsèques et extrinsèques précipitant la chute durant ce mouvement. Cette étude portant sur la marche et la perte de l'équilibre dynamique, aide à comprendre le risque important et la fréquence des chutes chez une personne âgée à domicile.

Le problème que posent les chutes des personnes âgées chez elles, sera exposé à travers les statistiques qui montrent leur nombre important et leur coût élevé, et à travers leurs conséquences indésirables au niveau sociologique, médical et psychologique. D'autre part, en dressant un état des lieux du vieillissement de la population en France nous remarquons un

accroissement permanent de la population âgée. Dans ce contexte, le problème de la chute des personnes âgées devient un problème de santé publique, d'où la motivation et l'idée derrière le projet PARAChute.

1.1 Objectifs et contraintes du projet PARAChute

Avec le vieillissement de la population française et le nombre important de chutes chez les personnes âgées et leurs conséquences, l'idée de prédire le risque de ces chutes dans la perspective de réduire leur nombre, devient intéressante et utile. D'où l'objectif de l'étude PARAChute (Personnes Agées et Risque de Chute), dans le cadre de laquelle s'inscrit ce travail de thèse, qui est de proposer une méthodologie et une technologie permettant de déceler chez la personne âgée à domicile ou en environnement contrôlé une évolution du comportement locomoteur, prédictif du risque de chute. Cette étude consiste à réaliser un suivi longitudinal de l'évolution du risque de chute chez la personne âgée en utilisant un dispositif situé à son domicile (ou en institution) ; ce suivi comporte une phase d'apprentissage d'un indicateur individuel (composé à partir de paramètres caractérisant l'équilibre) et une phase d'évaluation de cet indicateur pour estimer la déviation par rapport à la référence apprise. Les patients sont alors suivis dans les conditions habituelles de leur vie quotidienne et non pas dans des conditions modifiées (comme dans le cas des tests cliniques). Le système de décision doit utiliser la personne comme sa propre référence ; en effet, il n'existe pas de groupes de référence définissant d'une façon absolue des personnes à faible ou à fort risque de chute, quelques soient les critères utilisés pour la classification. La variance intra - classes de ces groupes n'a aucune chance de produire une sensibilité et une spécificité suffisantes.

L'évaluation du risque de chute peut être considérée comme préventive, vu le fait que ce risque peut être réduit grâce à des interventions d'ordre médical ou kinésithérapeutique. En effet des programmes d'exercices, soit en groupe [1,2], soit à domicile [3,4] sont considérés comme les interventions les plus efficaces. Les exercices de renforcement musculaire et de rééducation des équilibres (statique et dynamique) sont plus efficaces s'ils sont personnalisés aux principaux facteurs de risque [5]. Pour que le coût de ces exercices ne soit pas exorbitant il faut que les personnes qui en ont besoin soient bien ciblées. L'évaluation devient alors essentielle ; elle doit être simple et écologique dans l'objectif d'un véritable 'dépistage' de personnes à risque de chutes. L'idée d'observer le comportement de marche d'une personne au quotidien, à son domicile, a donc trouvé un écho très favorable chez les médecins qui voient dans la conception d'un système d'analyse de la marche un véritable outil de prévention primaire.

Dans le but de proposer un système simple adapté à la vie quotidienne des personnes âgées, le projet PARAChute a fixé un certain nombre de contraintes à prendre en compte pour l'élaboration du système de suivi :

- le système de suivi doit pouvoir s'adapter au domicile de la personne âgée de façon élémentaire, sans perturbation de son environnement usuel ;
- le protocole doit s'appuyer sur des gestes quotidiens ;
- le protocole ne doit nécessiter la présence d'aucune tierce personne ;
- l'ensemble du système doit préserver de façon absolue l'intimité de la personne. Quelles que soient les données captées, l'information disponible en sortie du système ne doit faire apparaître que ce qui est nécessaire à l'évaluation du risque de chute ;
- la personne doit pouvoir vivre seule ou non ;
- le système doit pouvoir fonctionner aussi bien seul que comme élément d'information dans un réseau de soins (ou de veille) à domicile.

Le projet PARAChute inclut trois sous projets :

1. Le premier sous projet a pour but de produire une information pertinente sur la qualité de l'équilibre statique. En d'autres termes, il s'intéresse à proposer une méthode d'évaluation de l'équilibre statique;
2. Le deuxième sous projet a pour but de produire une information pertinente sur la qualité de la marche. La proposition d'une méthode d'analyse de la marche habituelle fait alors l'objet de ce sous projet ;
3. Le troisième et dernier sous projet a pour but de produire les méthodologies d'auto apprentissage, d'intégration des informations fournies par les sous projets précédents et de prise de décision. Ce sous projet s'intéresse à la proposition des méthodologies de traitement de l'information pour la décision.

1.2 Objectifs de la thèse

Ce travail de thèse s'inscrit dans le cadre du deuxième sous projet PARAChute. Le but de ce sous projet étant d'évaluer la qualité de la marche, l'objectif principal de cette thèse est alors de développer un système non intrusif capable de fournir une analyse de la marche d'une personne

âgée à partir de ses activités quotidiennes. Pour parvenir à cet objectif, plusieurs tâches seront à effectuer :

- étudier les tests et les méthodes utilisées dans la littérature pour l'évaluation du risque de chute chez les personnes âgées ;
- choisir parmi ces méthodes celles qui peuvent s'adapter au cadre de l'étude selon leur objectivité, simplicité, passivité et possibilité d'être automatisées ;
- identifier les capteurs et les technologies permettant de reproduire ces méthodes ou d'extraire les paramètres sur lesquels elles se basent pour fournir une décision sur la qualité de l'équilibre ;
- concevoir et développer, si nécessaire, les outils pour estimer ces paramètres ;
- tester et valider ces outils développés.

1.3 Description de la marche [6]

Il est intéressant d'étudier et de décrire le mouvement de la marche. En effet, la marche est une modalité particulière du mouvement nécessitant l'intégrité des voies motrices, cérébelleuses, vestibulaires et des afférences proprioceptives. Les voies motrices comprennent les systèmes pyramidal (relatif au cortex moteur) et extra-pyramidal (relatif au tronc cérébral), les voies cérébelleuses sont impliquées dans les réactions d'équilibre statique et dynamique, les voies vestibulaires participent au tonus postural (maintenir une posture) et à l'information sur la direction et la vitesse du mouvement et les afférences proprioceptives informent de la position des articulations et des membres dans l'espace. Les afférences prenant leurs origines de la face plantaire sont particulièrement importantes pour la marche.

La marche est une activité alternée des membres inférieurs qui permet le déplacement du corps tout en assurant le maintien de son équilibre en orthostatisme (position debout immobile). C'est une activité motrice automatique qui demande un apprentissage dès l'enfance (automatisme acquis). Nous décrivons trois systèmes impliqués dans la marche :

- le système anti-gravitaire : Ce système sert au maintien de l'attitude érigée en s'opposant à l'effet de la pesanteur. Il est sous la dépendance du tonus des muscles anti-gravitaires qui sont, chez l'homme, les muscles extenseurs des membres inférieurs et les muscles paravertébraux. Les afférences de ce système proviennent de la plante des pieds, du labyrinthe de l'oreille interne et des récepteurs musculo-tendineux ;

- le système de production du pas : La marche pourrait être décrite comme une suite de ruptures d'équilibre conduisant à une chute suivie d'une réaction "parachute" d'un des membres inférieurs en alternance. La production du pas est donc une activité rythmique au cours de laquelle le poids du corps alterne d'un membre inférieur à l'autre, par l'intermédiaire d'un appui unipodal. Chez le nouveau né, il existe une marche automatique provoquée par le maintien en orthostatisme et par la stimulation podale. Cette marche automatique disparaît en quelques semaines, puis un apprentissage produit les séquences d'activités musculaires nécessaires au déplacement : c'est un automatisme acquis ;
- le système d'équilibre et d'adaptation posturale : La station debout définit la posture ; la fonction d'équilibration vise au maintien de la posture. Ce système est complexe et intègre les informations de quatre modes de perception (la vue, le système vestibulaire, les voies sensitives afférentes proprioceptives, les voies sensitives afférentes tactiles). Ce système est statique (tonus postural) et dynamique (marche). La réaction d'équilibration se définit comme l'ensemble des mouvements compensatoires automatiques qui permet l'adaptation posturale pour des débattements posturaux (mouvement oscillatoire du centre de gravité).

Les modifications apportées par le vieillissement n'ont pas été étudiées de façon systématique. Plusieurs arguments tendent à montrer que les réactions posturales se maintiennent efficacement au cours du vieillissement physiologique mais la vitesse des réactions et les capacités d'adaptation aux situations extrêmes sont moindres.

1.4 Causes des chutes [6]

Une étude des facteurs favorisant la chute, nous aidera à mieux comprendre ce phénomène. De manière didactique, seules les chutes dans le contexte des troubles de la marche sont envisagées. En pratique, plusieurs facteurs peuvent être à l'origine d'une chute. Dans ce qui suit nous allons exposer les différentes causes de chutes répertoriées en trois grandes catégories.

1.4.1 Facteurs de risque de chutes chroniques

Les facteurs de risque chroniques prédisposant à la chute sont une résultante de pathologies chroniques ou des conséquences du vieillissement. En effet ces facteurs sont neurologiques, neuromusculaires, ostéo-articulaires et visuels. Parmi les affections neurologiques, les atteintes vasculaires sont les plus fréquentes, que ce soit à l'étage cortical ou sous-cortical, ou au niveau du tronc cérébral. Les maladies dégénératives corticales (maladie d'Alzheimer) ou sous-corticales représentent le deuxième grand groupe d'affections neurologiques. La majorité des

atteintes neurologiques périphériques est associée à des modifications de la statique rachidienne, conséquences de l'arthrose et de l'ostéopénie (fragilisation progressive du tissu osseux due à une diminution de sa densité).

Les atteintes musculaires chez le sujet âgé doivent évoquer en priorité une étiologie endocrinienne, en particulier les hyperthyroïdies. Les présentations atypiques des déficits du système locomoteur apparaissant à l'effort chez le sujet âgé, imposent que cette maladie soit recherchée. Les atteintes ostéo-articulaires sont dominées par la limitation d'amplitude des articulations coxo-fémorales, qui ne peuvent assurer leur fonction d'adaptation de l'équilibre en situation extrême. Les interventions chirurgicales (prothèse totale de hanche), réalisées avant l'apparition d'une amyotrophie (diminution du volume musculaire), sont efficaces pour réduire l'effet indésirable de ces atteintes. Des travaux récents ont insisté sur l'importance de la limitation de la dorsiflexion de la cheville comme facteur prédictif de chute. Les pieds doivent faire l'objet d'une attention particulière et peuvent justifier un examen podologique spécialisé. A côté de la baisse de l'acuité visuelle, il convient de rechercher des altérations périphériques du champ visuel qui favorisent souvent les chutes sur obstacle. Le syndrome dépressif par l'inhibition motrice et psychique qu'il entraîne et par la prescription de psychotropes qu'il impose est un facteur fréquent prédisposant aux chutes. De nombreuses affections peuvent réduire l'adaptation à l'effort et les mouvements compensatoires d'adaptation posturale. Enfin, la dénutrition protéino-énergétique et son retentissement sur la force musculaire et sur les structures neurologiques périphériques et centrales favorise le risque de chute. Les facteurs de risque chroniques prédisposant à la chute peuvent être classées en quatre grandes catégories :

1. Affections neurologiques :

- **Affections neurologiques centrales :**
  - Atteintes corticales : vasculaire, dégénérative, tumorale.
  - Atteintes sous-corticales dégénérative : maladie de Parkinson, Steele-Richardson.
  - Atteintes sous-corticales vasculaire : artériopathies hypertensives, hydrocéphalie à pression normale.
  - Atteintes du tronc cérébral, cérébelleuse et vestibulaire : insuffisance vertébro-basilaire, affections de l'oreille interne.

- **Affections neurologiques périphériques :**
  - Etroitesses canalaires : myélopathie cervicarthrosique, canal lombaire étroit.
  - Myélopathies : sclérose combinée de la moëlle.

- Affections radiculaires (sciatique) et tronculaires (paralysie du sciatique poplité externe).
- Polyneuropathies : toxique, médicamenteuse.

2. Affections neuromusculaires :

- Myopathies thyroïdiennes, cortisoniques, ostéomalaciques.
- Pseudopolyarthrite rhizomélique.
- Myasthénie.

3. Affections ostéo-articulaires :

• **Affections du rachis :** cyphose dorsale ostéoporotique, arthrose cervicale.

• **Affections articulaires des membres inférieurs :**

- coxofémorales : coxarthrose.
- genoux : gonarthrose avec laxité capsuloligamentaire, chondrocalcinose.
- cheville : atteintes musculotendineuses responsables d'une limitation de la dorsiflexion.
- pied : pathologies de la statique (hallux valgus, pied creux antérieur, avant-pied plat), arthropathies microcristallines.

4. Autres affections :

• **Atteintes visuelles :** diminution de l'acuité ou du champ visuel (cataracte, glaucome chronique, dégénérescence maculaire liée à l'âge) ;

• inhibition motrice des syndromes dépressifs ;

• diminution de l'adaptation à l'effort : insuffisance cardiaque, broncho-pneumopathie chronique, anémie ;

• dénutrition protéinoénergétique sévère ;

• diminution de la vigilance (psychotropes).

1.4.2 Facteurs précipitants intrinsèques

Sont considérés comme tels tous les agents qui déclenchent la chute. Ces facteurs sont souvent multiples et associés entre eux. Ils peuvent être mineurs lorsque les facteurs de risque chroniques prédominent. Tout événement pathologique peut favoriser la chute. L'ensemble des étiologies des malaises et pertes de connaissance doit être recherché. Les causes cardiaques dominent, notamment les troubles du rythme supra-ventriculaires et toutes les affections pouvant occasionner un bas débit cérébral telles les troubles de conduction, l'infarctus du myocarde,

l'embolie pulmonaire ou la sténose aortique serrée. Parmi les causes vasculaires, l'hypotension orthostatique, impliquée dans 10 à 15 % des chutes, est de loin la plus fréquente. Elle est habituellement multi-factorielle (diminution du volume total du sang circulant, insuffisance veineuse des membres inférieurs, dysfonctionnement du système nerveux autonome, désadaptation à l'effort). Les causes neurologiques sont dominées par les accidents vasculaires cérébraux. Les épisodes confusionnels exposent à un risque de chute par l'altération de la vigilance. Si la crise comitiale est souvent rapportée par l'entourage, les états de mal épileptiques peuvent être symptomatiques chez le sujet âgé. L'hématome sous dural justifie un scanner cérébrale lorsque la chute est précédée de troubles du comportement d'apparition récente. Enfin, les causes métaboliques sont représentées par les hypo et les hyperkaliémies (taux de potassium plasmatique) responsables de troubles paroxystiques du rythme cardiaque. Les hypoglycémies peuvent aussi s'observer lors de cachexie (dégradation profonde de l'état général, accompagnée d'une maigreur importante). L'hypercalcémie (concentration anormalement élevée du calcium) s'exprime souvent au grand âge par ses manifestations neuro-psychiques, et doit faire rechercher une hyperparathyroïdie. Les facteurs intrinsèques précipitants de la chute sont répertoriés en quatre grandes catégories :

1. **Causes cardiaques :**
    - Troubles du rythme paroxystiques supra-ventriculaires et ventriculaires.
    - Troubles de la conduction.
    - Infarctus du myocarde.
    - Embolie pulmonaire.
    - Sténose aortique serrée.

2. **Causes vasculaires :**
    - Hypotension orthostatique.
    - Malaise vago-vagal.
    - Hypersensibilité sino-carotidienne.

3. **Causes neurologiques :**
    - Accident vasculaire cérébral.
    - Etat confusionnel.
    - Hématome sous-dural.
    - Crise convulsive.

4. **Causes métaboliques :**

- Hypo et hyper kaliémie.
- Hypoglycémie.
- Hypercalcémie.

1.4.3 Facteurs précipitants extrinsèques

Les facteurs extrinsèques sont souvent des facteurs liés à l'environnement de vie de la personne âgée. Le plus fréquemment mis en cause est le port de chaussures inadaptées, trop lâches, maintenant mal le pied, à semelle usée ou glissante. Les chaises bancales, les tapis mal fixés, les objets mal rangés ou traînant au sol, les animaux domestiques, le sol mouillé ou glissant ne sont que des exemples parmi les plus fréquents des facteurs environnementaux précipitant la chute de la personne âgée. Parmi ces autres facteurs nous notons l'habillement (vêtements trop longs), le mobilier (fauteuil, lit trop haut ou trop bas), les obstacles au sol (fils électriques, carrelage ou revêtement de sol irrégulier ou décollé), le mauvais éclairage, les baignoires glissantes et les toilettes inadaptées.

1.5 La chute chez la personne âgée

Après avoir énuméré tous les facteurs précipitant la chute, nous pouvons déduire que ces facteurs sont d'origines variées et que l'occurrence, chez la personne âgée (à domicile), d'au moins l'un d'eux est presque inévitable. Compte tenu de ce constat, nous pouvons déduire un risque important de perte d'équilibre pour tout sujet âgé. Cette déduction est justifiée par les statistiques que nous allons exposer ci dessous. D'autre part, nous allons présenter les conséquences indésirables d'une chute qui en font un problème majeur de santé.

1.5.1 Statistiques

La chute des personnes âgées dans leur environnement quotidien est un problème majeur de santé publique en raison de leur fréquence et de leurs conséquences médicales et sociales. Il est estimé qu'un tiers des sujets de plus de 65 ans et la moitié de ceux de plus de 85 ans font une ou plusieurs chutes par an. En France, les chutes seraient responsables à court terme de 12 000 décès par an. Elles représentent également un facteur d'entrée dans la dépendance : 40 % des sujets âgés hospitalisés pour chute sont orientés ensuite vers une institution. Ce nombre des chutes est probablement sous estimé, tant au domicile qu'en institution, compte tenu de la non déclaration des accidents (seules 5 à 10 % des chutes sont portées à la connaissance médicale [7]. Le risque apparaît encore plus important en institution qu'au domicile (1,7 chute par résident et

par an) [8]. Ces chutes sont responsables d'une morbidité importante (50 000 fractures de l'extrémité supérieure du fémur, par an) [9], et d'une mortalité évaluée à plus de 12 000 décès par an en France, soit plus de morts que par accident de la route. Le risque de mortalité est multiplié par quatre dans l'année qui suit la chute par rapport à la population du même âge, quelle que soit la gravité de la chute. D'autre part Le coût annuel médical des chutes est considérable : il est estimé à environ deux milliards d'euros pour la France [10].

1.5.2 Conséquences des chutes chez les personnes âgées [6]

La chute peut avoir des significations diverses chez le sujet âgé : signe essentiel d'une étiologie unique (chute symptôme), élément sémiologique d'un syndrome (syndrome de régression psychomotrice), évènement dont la signification est sociale ou psychique (chute symbole), enfin événement brutal lié à un facteur intercurrent (chute fortuite). Il existe encore trop souvent une différence de comportement médical face à la chute selon qu'elle est ou non précédée de malaise ou perte de connaissance. La première, teintée d'une gravité, fait l'objet d'une exploration active. La seconde, "mécanique", est souvent négligée. Les chutes sont rarement fortuites chez le sujet âgé. Qu'elle soit la conséquence d'une perte progressive des mécanismes d'adaptation à l'équilibre, d'une affection aiguë, d'un effet indésirable d'un médicament ou la manifestation d'un appel à l'aide, chaque chute est susceptible de provoquer la hantise d'une nouvelle chute, le corollaire étant la crainte et la limitation spontanée de la marche et la spirale de la dépendance. Toute chute chez un sujet âgé, quelles qu'en soient les circonstances, doit ainsi être prise en charge non comme un simple accident, mais comme une affection potentiellement grave.

La plupart des chutes n'entraînent pas de traumatisme physique sérieux. En effet, 6 à 8 % seulement des chutes seraient responsables de fractures, dont 1 fois sur 3 de l'extrémité supérieure du fémur. Ainsi, la morbidité des chutes ne peut en aucun cas se résumer à leurs seules conséquences traumatiques. Les conséquences psychomotrices représentent les conséquences les plus fréquentes et les plus graves des chutes, conduisant, en l'absence de prise en charge rapide et adaptée à la dépendance lourde et à l'installation d'un état grabataire. En effet, la chute peut être à l'origine d'une sidération des automatismes acquis, entraînant une perte des réactions d'adaptation posturale, avec difficulté à se maintenir. Ce tableau est appelé "syndrome post-chute". C'est une urgence gériatrique car, en l'absence de prise en charge adaptée, il évolue vers un syndrome de régression psychomotrice. Le syndrome post-chute associe une composante motrice et une composante psychologique. La composante motrice se définit par :

- des troubles de la statique en position assise avec rétro-pulsion (projection du tronc en arrière) et impossibilité de passage en anté-pulsion (projection du tronc en avant) ;
- une position debout non fonctionnelle, se caractérisant par une projection du tronc en arrière (rétro-pulsion), un appui podal postérieur et un soulèvement des orteils ;
- la marche, lorsqu'elle est possible, se fait à petits pas, avec appui talonnier anti-physiologique, élargissement du polygone de sustentation (polygone délimité par les pieds), flexion des genoux, sans temps unipodal, ni déroulement du pied au sol.

La composante psychologique, lorsqu'elle existe, se traduit par une anxiété majeure avec une peur du vide antérieur qui gêne la marche et peut conduire à un refus de toute tentative de verticalisation (prendre et tenir la station debout). A côté de ce tableau dramatique de syndrome post-chute, les conséquences psychologiques peuvent être plus insidieuses mais avoir une réelle gravité. La chute est l'occasion pour le patient âgé de prendre conscience de la fragilité de son état ; le choc émotionnel qui en résulte entraîne une perte de confiance en soi, un sentiment d'insécurité et de dévalorisation, avec repli sur soi, démotivation et restriction des activités. Ce tableau d'inhibition psychomotrice doit évoquer un syndrome dépressif. Toute personne âgée qui est restée au sol sans pouvoir se relever pendant plus de 3 heures doit bénéficier d'une psychothérapie de soutien dès le premier jour d'hospitalisation. La perte d'autonomie qui en résulte est volontiers aggravée par l'entourage qui, inquiet d'une nouvelle chute, a tendance à surprotéger la personne et refuse souvent le retour au domicile. Une négociation entre l'équipe médicale, la personne âgée et sa famille est indispensable afin de trouver un compromis entre la sécurité du patient et son désir de rester à domicile.

1.6 Vieillissement de la population en France et conséquences

Entre 1950 et 2006, la population française a augmenté de 19 millions de personnes. Cette forte augmentation de la population totale est due à l'accroissement naturel, mais aussi à l'accroissement de la population âgée de 60 ans ou plus. En effet, cette dernière population a plus que doublé et comptait près de 12 millions en 2006 [11]. Ce vieillissement résulte de l'effet conjugué de la baisse de la natalité et de la baisse de la mortalité chez les sujets âgés de 60 ans ou plus. En fait, au début du XXe siècle, les sujets âgés de 60 ans ou plus représentaient un peu plus de 12 % de la population totale et ceux de moins de 20 ans plus d'un tiers de la population. En 2006, ces proportions étaient respectivement de 20 % et de 25 % seulement. Ainsi, la proportion des 60 ans ou plus a-t-elle presque doublé au cours du dernier siècle. Mais il devrait y

avoir aussi pratiquement doublement entre 1920 et 2020. Ainsi nous pouvons déduire que la forte augmentation de la population en France va de pair avec son vieillissement.

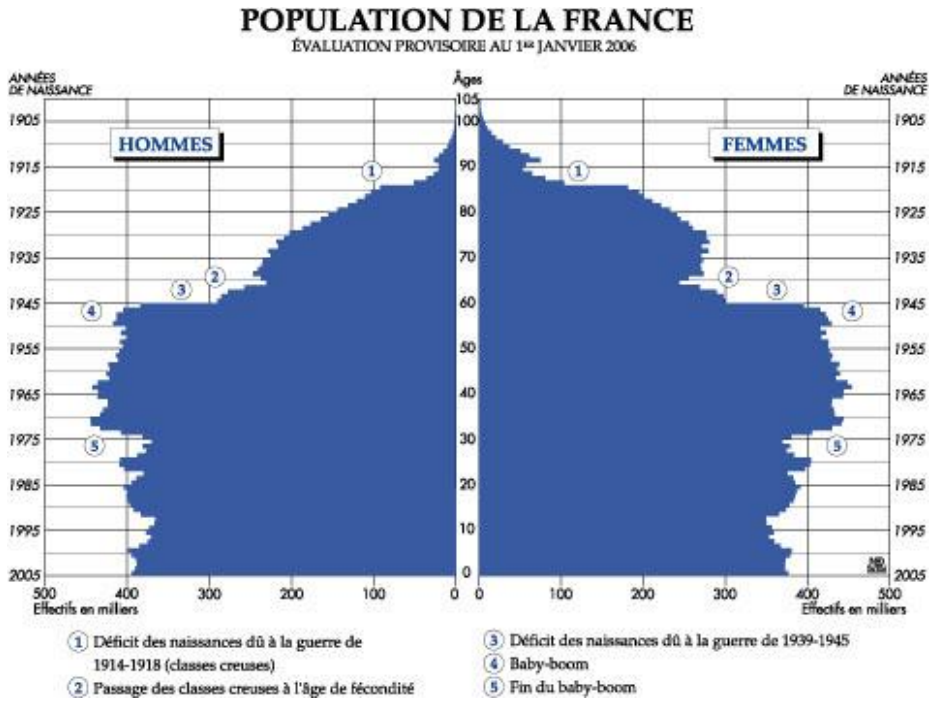


fig. 1.1 : Pyramide des âges pour la population en France (estimation 2006, source : INED)

Parmi cette population des plus de 60 ans, ce sont surtout les personnes de 85 ans et plus qui posent d'importants problèmes d'accueil ; de 200 000 personnes en 1950, l'effectif s'approchait de 1 250 000 en 2000 et devrait dépasser les 2 millions en 2020. D'autre part, le nombre de centenaires, de l'ordre de quelques centaines dans les années soixante, et de 17 267 en 2007 devrait dépasser les 30000 en 2030 [11].

D'ailleurs, ce vieillissement peut être expliqué par la forte progression de l'espérance de vie aux âges élevés. En effet, l'espérance de vie à 60 ans n'était au début du XXe siècle, que de 13 ans pour les hommes et d'un peu moins de 15 ans pour les femmes ; elle est en 2007 de plus de 21 ans pour les hommes et de près de 27 ans pour les femmes et en 2020 elle devrait être respectivement de l'ordre de 23 et de 28 ans. Alors même que la situation des françaises peut être qualifiée d'exceptionnelle, les progrès attendus continuent à être importants : environ six ans et demi d'espérance de vie à 85 ans actuellement et un an de plus en 2020, le progrès étant un peu moins net pour le sexe masculin (un peu plus de cinq ans et moins de six ans).

Le vieillissement de la population a des conséquences sociologiques importantes. En effet, six millions de personnes vivent seules en France : il s'agit majoritairement de personnes âgées.

L'isolement est principalement dépendant du veuvage pour les générations actuelles de personnes âgées. En 1990, 72,6 % des femmes âgées de plus de 80 ans étaient veuves, soit plus de 1 100 000 femmes. Le veuvage féminin résulte de deux grands mécanismes : la surmortalité masculine et la différence d'âge entre les époux. A l'inverse la probabilité pour un homme de se trouver isolé est faible. Une femme sur deux est seule à 70 ans et un homme sur deux l'est à 85 ans. Ceci a une influence considérable sur le maintien à domicile et majore le risque d'institutionnalisation pour les femmes. D'autre part, en 1990 environ 3 millions de personnes de plus de 60 ans vivent dans les communes rurales où ils représentent 23 % des habitants. Les autres résident dans des communes urbaines composant 19 % de la population citadine. Lorsque les personnes âgées vivent en ville, elles résident surtout dans le centre. Ceci est d'autant plus vrai que l'agglomération est importante. A l'opposé, dans les zones rurales, la population est dispersée et éloignée des services, des réseaux de soutien professionnels et des centres de soins bien équipés. L'organisation de la vie des personnes âgées en milieu rural est précaire contrairement aux stéréotypes, dès qu'il y a absence d'entourage propre et/ou survenue d'une dépendance, même réduite.

1.7 Conclusion

En conclusion nous pouvons constater un risque important de chute chez toute personne âgée à domicile. Le problème que pose une chute est lié à la gravité de la situation psychologique, médicale et sociologique qui en découle. Or, l'ampleur de ce problème devient plus grande avec l'accroissement de la population concernée. Les chiffres exposés ci dessus montrent un vieillissement progressif de la population en France et justifient la qualification du problème de chute comme problème de santé publique ; dans ce contexte, l'évaluation du risque de chute chez la personne âgée à domicile, dans le but d'éviter son occurrence ensuite, devient intéressante voire primordiale. Cette réflexion constitue la motivation principale du projet PARAChute et de ce travail de thèse.

Les contraintes que nous ne sommes fixées dans le cadre de PARAChute, constituent l'originalité de l'approche. Désormais dans le but de prévenir les chutes à domicile, l'équilibre dynamique de chaque personne devrait être évalué à partir de ces activités quotidiennes à domicile et non plus à partir d'exercices accomplis au cabinet médical ; d'autre part, le suivi de l'évolution du risque de chute sera journalier et non soumis aux dates de visites médicales. Ces deux bases du projet devraient rendre la prévention d'une chute plus efficace.

Bibliographie

[1] Buchner D, Cress M, De Lateur B, Esselman P et al., “The effect of strength and endurance training on gait, balance, fall risk, and health services use in community-living older adults”, Journal of Gerontology And Biologic Medical Sciences, vol. 52, pp 218-224, 1997.

[2] Wolf SL, Barnhart HX, Kutner NG, McNeely E et al., “Reducing frailty and falls in older persons : an investigation of Tai Chi and computerized balance training Atlanta FICSIT Group. Frailty and Injuries : Cooperative Studies of Intervention Techniques”, Journal of American Geriatric Society, vol. 44, pp 489-497, 1996.

[3] Campbell J, Robertson MC, Gardner M, Norton RN et Buchner DM, “Falls prevention over 2 years : a randomised controlled trial in women 80 years and older”, Journal of Age & Ageing, vol. 28, pp. 513-518, 1999.

[4] Robertson MC, Devlin N, Gardner M et Campbell J, “Effectiveness and economic evaluation of a nurse delivered home exercise programme to prevent falls. 1 : Randomised controlled trial”, Biologic Medical Journal, vol. 322, pp. 697-701, 2001.

[5] Myers AH, Young Y et Langlois JA, “Prevention of falls in the elderly”, Bone, vol. 18, pp. 87-101, 1996.

[6] Le collège national des enseignants en gériatrie, Corpus de Gériatrie, Tome 1, 2000.

[7] Meyer P, Boucher D, Martinet N et André J, “Troubles attentionnels et chutes du sujet âgé” , Annales Médicales de Nancy, vol. 37, pp 145-156, 1998.

[8] Lord S, “Instabilité des chutes chez la personne âgée : les troubles de la posture et les risques de chute” , Année Gérontologique, vol. supplémentaire, pp 65-67, 1995.

[9] Baudoin C, Fardellone P, Potard V et Sebert J, “Fractures of the proximal femur in Picardy”, Journal of Osteoporos, vol. 3, pp 43-49, 1993.

[10] Comité Français d’Education pour la Santé, “Prévention des chutes chez les personnes âgées : des conseils pour préserver l'autonomie et le bien-être de nos aînés”, Caisse Nationale de l'Assurance Maladie des Travailleurs Salariés, Dossier de presse, novembre 2000.

[11] Institut National d’Etudes Démographiques (INED) : http:// www.ined.fr

Chapitre 2 : Différentes approches pour l'évaluation des risques de chute chez les personnes âgées

Après avoir présenté les objectifs et les contraintes de ce travail, nous allons exposer dans ce chapitre les approches déjà utilisées pour détecter les risques de chute et pour évaluer l'équilibre dynamique. Chaque approche sera exposée selon deux aspects : l'aspect technologique de l'approche se résumant aux capteurs et aux méthodes utilisés et l'aspect médical représenté par les paramètres estimés et leur pertinence pour l'évaluation des risques de chute. Nous étudierons ces deux aspects, vis à vis de leur compatibilité avec les contraintes imposées et les objectifs souhaités (absence d'une tierce personne, objectivité etc..). Ainsi, nous noterons tous les paramètres pertinents pour l'évaluation de l'équilibre dynamique et les techniques compatibles avec les contraintes imposées.

Nous détaillerons les tests cliniques utilisés par les gériatres pour évaluer les risques de chute chez une personne âgée durant sa visite médicale régulière, tels que les tests de Berg, Tinetti, etc.. Ces tests validés par la communauté médicale, consistent à demander à la personne en question, d'effectuer un certain nombre d'exercices simulant des mouvements de la vie quotidienne ; par suite, la qualité de l'équilibre de chaque patient est évaluée par le gériatre, suivant son appréciation de la capacité de ce patient à accomplir ces exercices. Nous retiendrons de ces tests ceux qui sont objectifs, simples et qui, pour une personne âgée, ne présentent aucun danger à être effectués à domicile sans la présence d'un personnel médical ou paramédical.

Nous exposerons ensuite les méthodes qui utilisent des capteurs que la personne porte sur elle même ; nous décrirons les approches faisant intervenir les accéléromètres, les gyroscopes et les goniomètres. Seront aussi détaillées les méthodes d'évaluation d'équilibre par électromyographie ou par utilisation de semelles équipées de capteurs de pression. Intuitivement nous pouvons exclure l'utilisation de ces capteurs, étant en contradiction avec les exigences de l'étude. Par contre, nous nous intéressons aux paramètres décrits comme pertinents pour l'analyse de l'équilibre dynamique et qui sont fournis par ces capteurs.

Nous parlerons aussi des systèmes qu'on peut intégrer dans des parcours pour évaluer certains paramètres de la marche. Ainsi seront décrites les approches par plateformes de forces, télémètres, tapis actimétriques et systèmes de reconstruction 3D du mouvement humain. Une discussion sur les technologies utilisées et sur les paramètres évalués sera établie.

Après avoir présenté les différentes approches déjà appliquées, nous dresserons la liste des paramètres et des tests qu'on a identifiés comme pertinents et compatibles avec le projet PARAChute. Suivra une discussion sur les outils à utiliser ou développer pour recueillir ces données d'une manière passive et peu coûteuse.

2.1 Tests cliniques usuels

Plusieurs tests indicateurs ont déjà été développés pour évaluer l'équilibre des personnes âgées. Ces tests comprennent des manœuvres complexes à effectuer et nécessitent souvent la présence d'un médecin qui évaluera le score de chaque personne. Or, les médecins n'ont pas souvent le temps pour superviser de tels tests d'où le besoin d'outils plus simples, rapides et moins coûteux. Une fois que les personnes sont identifiées par ces tests comme ayant un grand risque de chute, le médecin peut les soumettre à un traitement pour réduire ce risque (par exemple des programmes de rééducation ou de renforcement musculaire [1]). La validité de ces tests a été prouvée auparavant par les experts en gériatrie à partir des observations et des suivis des patients. Le but du système à développer sera de reproduire certains de ces tests ou au moins aider à les accomplir en fournissant une estimation de quelques paramètres utilisés.

Les gériatres évaluent le risque de chute d'une personne en utilisant des questionnaires bien définis avec une subjectivité (de la part du médecin) dans l'évaluation des réponses. Ces questionnaires portent surtout sur l'historique des chutes, les antécédents médicaux, les traitements médicaux suivis et l'état physique actuel de la personne âgée. Ces facteurs peuvent augmenter la susceptibilité aux chutes mais ne peuvent pas être considérés comme des indicateurs fiables de risque de chute. En effet l'altération de l'équilibre dynamique durant la marche peut être causée par d'autres facteurs que ceux cités ci dessus, d'où le besoin de rendre ces tests plus globaux en évaluant l'aptitude des sujets à effectuer certaines tâches de l'activité quotidienne et certains mouvements. Dans cette partie nous exposerons certains de ces tests qui ont été validés par la communauté médicale.

2.1.1 Test de Berg

Ce test introduit par Berg et al. [2] porte sur l'observation de 14 mouvements de la vie quotidienne. Il aide à l'analyse de l'équilibre et de la sécurité en déplacement des personnes à mobilité réduite. Il concerne les personnes en perte d'autonomie peu marquée, mais à risque. Pour chaque aspect de ce questionnaire un score allant de 0 à 4 (0 = pas de difficultés, 1 = un peu difficile, 2 = assez difficile, 3 = très difficile, 4 = incapacité) est attribué par le gériatre. Les aspects à observer sont les suivants :

1. capacité à rester assis sans soutien;
2. capacité à effectuer un changement de position : passer de assis à debout;
3. changement de position : de debout à assis;
4. capacité à effectuer des transferts : vers les toilettes, vers un siège de voiture;
5. capacité à rester debout sans soutien;
6. capacité à se tenir debout les yeux fermés;
7. capacité à se tenir debout, les pieds collés l'un à l'autre;
8. capacité à se tenir debout, un pied devant l'autre (talon contre pointe);
9. tenir debout sur une seule jambe;
10. tourner le tronc en gardant les pieds fixes;
11. capacité à ramasser un objet au sol;
12. tourner de 180 degrés (tour complet, ou volte-face);
13. monter sur un tabouret de 40 cm de haut;
14. allonger les deux bras devant soi en restant debout immobile.

La cotation maximale de 56 points est catastrophique, le risque de chute devient sérieux à partir d'un index de 45 points additionnés. Plusieurs études [3,4] ont montré la validité de ce test à évaluer le risque de chute chez les personnes âgées. Par contre, ce test ne s'avère pas adéquat pour une personne vivant seule ou en perte d'autonomie et par conséquent sera incompatible avec les objectifs attendus.

2.1.2 L'indice 'DGI' (Dynamic Gait Index)

Ce test [5] est constitué de 8 exercices. Le score pour chacune des tâches élémentaires varie de 0 à 3 (0 = grave altération, 1 = altération modéré, 2 = faible altération, 3 = normal). Les mouvements à accomplir sont :

1. marcher avec une vitesse normale;

2. marcher avec changement de vitesse;
3. marcher avec des mouvements horizontaux de la tête;
4. marcher avec des mouvements verticaux de la tête;
5. marcher avec demi tour immédiat;
6. marcher sur obstacle (tapis, marche, etc..);
7. contourner des obstacles;
8. monter les escaliers.

Un score inférieur à 19/24 indique un risque important de chute. Dans une étude menée sur une population de 103 personnes (31 personnes ayant déclaré une chute dans les 6 mois précédents l'étude) avec des problèmes vestibulaires, Whitney et al.[6] montrent que parmi les personnes ayant déjà chuté, celles qui ont obtenu un score de 18 ou moins pour le DGI sont 2.7 fois plus nombreuses que les autres. Dans ce contexte, la sensibilité du test de DGI était estimée à 70% et sa spécificité à 51%. D'autre part, ce test, qui est simple à accomplir, présente toujours un risque pour la personne en question.

2.1.3 Le test de '8 foot Up and Go'

Ce test très simple à effectuer, consiste à mesurer le temps que met une personne pour se lever d'une chaise, parcourir une distance de 2.66 m (8 foot) et retourner au fauteuil, durant un temps minimal. Rose et al. [7], ont appliqué ce test pour classer une population formée de 71 chuteurs et 63 non chuteurs. Un test statistique d'ANOVA, montre que les deux groupes peuvent être différenciés par ce test. Pour optimiser la classification, un temps seuil de 8.5s (pour effectuer la manœuvre) était choisi. Avec ces données, 86% des personnes non chuteuses étaient bien classifiées et 78% des chuteurs l'étaient. La prédiction globale du test étant de 82%. Ce test est objectif, simple à accomplir pour une personne âgée vivant seule et pourrait être automatisée. En conséquence, ce test est bien compatible avec notre approche.

2.1.4 Le test de Tinetti

Ce test [8] qui peut servir de repère pour l'enchaînement des exercices d'équilibre, s'adresse aux personnes âgées et en perte d'autonomie. Il analyse l'équilibre au cours de diverses situations de la vie quotidienne et ne nécessite pas d'expérience particulière de l'examinateur, ni de matériel spécifique. Il a été validé par rapport à des épreuves de posturographie et présente une sensibilité et une spécificité satisfaisantes. Différentes études ont démontré que les sujets ayant un score médiocre au test de Tinetti, présentent en général un risque de chute plus élevé. Ce test doit être réalisé en suivant l'ordre des épreuves et en observant la marche et l'équilibre.

L'observation porte sur les deux aspects statique et dynamique de l'équilibre. Le sous-score total au test de Tinetti statique (13 épreuves) est calculé en comptant 0 pour une réponse normale, 1 point pour une réponse adaptée et 2 points pour une réponse anormale. Le sous-score total au test de Tinetti dynamique (9 épreuves) est calculé en comptant 0 pour une réponse normale et 1 point pour une réponse anormale. Les aspects à évaluer durant ce test sont les suivants :

- Equilibre statique :

  1. équilibre assis droit sur une chaise;
  2. lever d'une chaise (si possible sans l'aide des bras);
  3. équilibre debout, juste après s'être levé;
  4. équilibre debout, les yeux ouverts, les pieds joints;
  5. équilibre debout les yeux fermés, les pieds joints;
  6. équilibre après un tour complet;
  7. résistance à une poussée sternale;
  8. équilibre après rotation de la tête;
  9. équilibre en station unipodale;
  10. équilibre avec hyper extension de la tête en arrière;
  11. équilibre avec extension de la colonne vertébrale et élévation des membres supérieurs;
  12. équilibre penché en avant;
  13. équilibre en s'asseyant.

- Equilibre dynamique (à observer durant la marche)

  1. l'initiation de la marche;
  2. la hauteur des pas (droite, gauche);
  3. la longueur des pas (droite, gauche);
  4. la symétrie des pas (droite, gauche);
  5. la régularité de la marche;
  6. déviation du trajet;
  7. la stabilité du tronc;
  8. l'espacement des pieds durant la marche;
  9. demi tour pendant marche.

Les exercices portant sur l'équilibre statique présentent des risques pour une personne âgée voulant les accomplir d'une manière autonome et sont à exclure. D'autre part, l'évaluation de

chacune de ces manœuvres reste subjective et doit être faite par un médecin. Par contre, les paramètres utilisés pour évaluer l'équilibre dynamique sont objectifs et peuvent être extraits d'une manière automatique. Pour ces raisons, ces paramètres (hauteur, longueur des pas etc..) seront à retenir pour notre approche.

2.1.5 Tests divers

Murphy et al. [9], ont observé 50 personnes âgées pendant 14 mois, en les faisant passer des tests divers. Leur but était de trouver les tests qui discriminent le mieux les chuteurs des autres. Pendant cette période, 34 des 50 personnes observées ont subi au moins une chute. Les tests effectués sont :

1. 'Floor Transfer': Mesure le temps que prend une personne pour s'asseoir par terre, étant debout, et puis pour se relever ;
2. '5 step test' : Mesure le temps que prend la personne pour faire 5 pas, de 10 cm chacun. Chaque pas consiste à avancer et puis revenir à la position initiale ;
3. '5 chair stands' : Mesure le temps nécessaire pour se lever d'un fauteuil, sans l'aide des bras, à 5 reprises ;
4. 'Tandem stance' : Mesure le temps (10s max.) que passe la personne debout, les pieds en position amant-aval ;
5. 'Semi-Tandem stance' : Mesure le temps (10s max.) que passe la personne debout, les pieds en position demi amant-aval ;
6. 'Side by side stance' : Mesure le temps (10s max.) que passe la personne debout, les pieds cote à cote ;
7. 'CTSIB' : Mesure le temps maximal (30s max.) que passe la personne debout, les pieds cote à cote, avec yeux ouverts et puis fermés ;
8. 'POAM-B' : 9 tâches élémentaires (issu du test de Tinetti) ;
9. 'Fonctional Reach test' : Mesure la distance maximale, que peut atteindre la personne avec ses mains tout en restant fixe et sans perdre l'équilibre ;
10. 'PPT' : 7 taches incluant manger, écrire, porter les habits, poser un livre sur un étagère, récupérer une pièce de monnaie tombée par terre, tourner de 360 degrés, et marcher pour une distance de 15m ;
11. '5 min. walk' : Mesure la distance parcourue pendant 5 minutes de marche rapide.

Les résultats obtenus montrent que les résultats de neuf épreuves différaient d'une manière claire d'un groupe à l'autre. Ces épreuves sont : 'floor transfer', '5-step test', 'tandem stance', 'POAM-B', 'FR', 'penny pick up', '360 deg. Turn', '50ft. walk', et '5 min. walk'. Les tests '360 deg. Turn', '5-step test', 'POAM-B' et '5 min. walk' sont bien corrélés au test '50ft. walk'. Les valeurs seuils sont choisies de façon à optimiser la classification. Les meilleurs tests discriminateurs qui ressortent de cet étude sont :

- 'Floor transfer' : 67% des chuteurs n'étaient pas capables de l'accomplir, 100% des non chuteurs l'ont réussi ;
- '50 ft walk' : La moyenne pour les chuteurs est de 33.2 s, contre 15.7 s pour les non chuteurs.

En conclusion, ces deux manœuvres combinées classifient correctement les personnes observées dans cette étude, avec 95.5% de réussite (bonne identification de 81% des chuteurs et de 100% des non chuteurs). Le point faible de cette étude est sans doute le nombre réduit des personnes observées, ce qui impose une certaine réserve sur la validité des résultats. Le test 'floor transfer' parait un peu encombrant à accomplir par des personnes âgées, au contraire le '50 ft walk' reste simple et pratique.

2.1.6 Conclusion

L'avantage des tests classiques réside dans le fait qu'ils sont méthodiques, peu coûteux et fiables. Néanmoins, la plupart de ces tests doivent être supervisés par un médecin, un infirmier ou tout autre personnel hospitalier ou médical. D'ailleurs, l'évaluation qui en découle suit la décision du médecin et par la suite est subjective. Pour ces raisons, ces tests subjectifs ne peuvent pas servir comme référence dans le contexte de notre étude.

D'autre part, les tests de '8 foot up and go' et '50ft walk', se résumant à une mesure du temps de parcours, sont simples à accomplir, objectifs et peuvent être potentiellement automatisés. Les paramètres intervenant dans l'évaluation de l'équilibre dynamique du test de Tinetti sont recueillis à partir des observations visuelles et possèdent les caractéristiques citées ci dessus. En conséquence, nous pouvons considérer ces paramètres ainsi que la mesure de vitesse de la marche sur un parcours donné, comme étant des paramètres intéressants à évaluer dans le cadre de notre approche.

2.2 Approches ambulatoires par capteurs portables

Pour surmonter les inconvénients des tests cliniques, comme la subjectivité ou le manque de fiabilité de certains de ces tests, d'autres techniques ont été développées dans le but d'évaluer d'une manière objective l'équilibre dynamique et la stabilité durant la marche. En effet, avec le développement de la microélectronique, des microcontrôleurs de plus en plus puissants, des capteurs miniatures et des cartes mémoires, de nouvelles solutions ont été adoptées pour évaluer l'équilibre ambulatoire. Ces techniques portent sur l'utilisation de capteurs portables mesurant des positions durant la marche, des vitesses, l'activité électrique musculaire à partir de signaux EMG et les accélérations en utilisant des gyroscopes et des accéléromètres. Dans ce qui suit, nous allons exposer les différentes techniques déjà utilisées dans ce domaine. Notons cependant que les capteurs portatifs restent encombrants et ne sont pas adaptés pour une utilisation à domicile. Par contre, nous essaierons d'identifier les paramètres calculés à l'aide de ces capteurs et qui s'avèrent pertinents pour la détection d'une altération de l'équilibre dynamique.

2.2.1 Electromyographie

La technique d'électromyographie (EMG) (fig. 2.1) consiste à placer des électrodes en regard d'un muscle pour détecter et mesurer son activité électrique. Cette technique a été utilisée dans plusieurs études [10,11] pour montrer une altération de l'activité musculaire posturale chez les personnes avec un grand risque de chute. Brauer et al. [12] ont utilisé l'étude des signaux EMG pour classifier une population de personnes âgées entre 'chuteurs' et 'non-chuteurs'. Néanmoins, la précision des résultats obtenus était mauvaise pour cette classification (sensibilité 28% et spécificité 87,6%). D'autre part, l'étude d'un mouvement complexe nécessite l'utilisation de plusieurs électrodes et par conséquence le traitement d'un grand nombre de signaux. L'estimation de la force d'un muscle durant sa contraction pose aussi des difficultés.

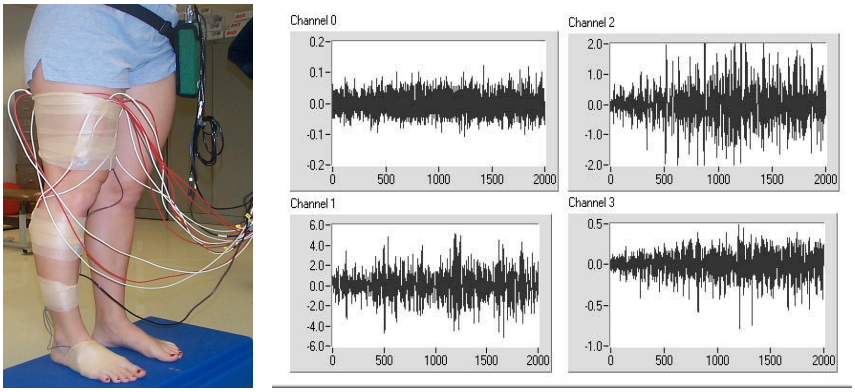


fig. 2.1 : Dispositif expérimental (à gauche) pour récupérer les signaux EMG avec les électrodes placées en regard du muscle étudié et les signaux obtenus (à droite)

En conclusion, cette technique peut nous fournir de l'information sur les réponses musculaires de la personne âgée mais pas sur le mouvement résultant (la marche), qui est fourni par plusieurs muscles contrôlant les articulations. Une étude de l'équilibre dynamique nécessite alors la mesure de l'activité électrique de plusieurs muscles et par conséquent l'utilisation de plusieurs électrodes et le traitement d'un nombre important de signaux. Or, ces électrodes sont encombrantes car elles doivent être fixées au corps et nécessitent un câblage spécifique les reliant à une centrale d'acquisition portable. Cette approche semble alors incompatible avec notre approche du problème, tenant compte des instruments utilisés.

2.2.2 Semelles instrumentées et capteurs FSR

Les systèmes ambulatoires utilisant des semelles équipées avec des capteurs de pression piézoélectriques [13] ou des capteurs FSR (Force Sensing Resistor) fixés sur la plante des pieds (fig. 2.2) aident à évaluer les paramètres temporels de la marche. Les capteurs FSR sont des résistances modulables suivant la pression qu'on y exerce. Leur structure est extrêmement simple : il s'agit de deux feuilles de polymères laminées ensemble (leur épaisseur ne dépassant pas 0,75 mm). Quand une force est appliquée, le contact entre les deux feuilles est plus grand et la conductivité plus importante. Surnommé "l'accélérateur électronique", le FSR est un capteur qui répond en faisant décroître sa résistance d'autant plus que l'on appuie plus fort sur le composant. Peu sensible aux bruits et aux vibrations, sa large plage d'impédance permet l'emploi d'une électronique d'interface simplifiée. Il est important de noter que le FSR ne fournit pas une mesure précise de la pression exercée (erreur de mesure entre 15% et 25%).

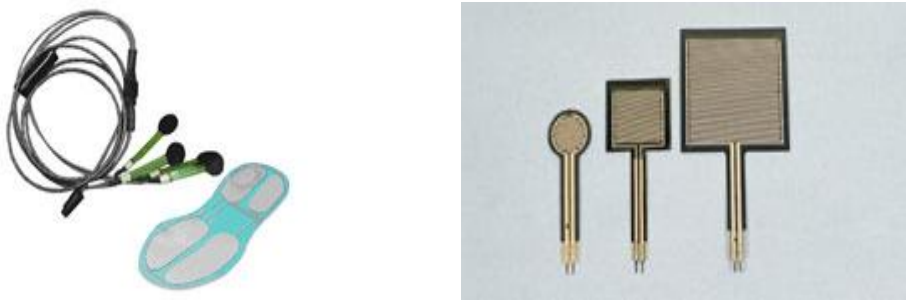


fig. 2.2 : Semelle instrumentée Noraxon® avec des capteurs de pression (à gauche) et différentes formes d'un capteur FSR (à droite)

Cette technique ambulatoire présente des résultats satisfaisants pour la détection des différentes phases du cycle de la marche normale (contact cheville-sol, contact orteil-sol, lever de la cheville etc..). Par contre, des problèmes comme le bon positionnement des capteurs ou leur

durée de vie réduite en raison des contraintes mécaniques, limitent l'application de ce type de technique. Hausdorff et al. [14] ont utilisé des capteurs FSR pour classifier une communauté de 53 personnes âgées (de plus de 70 ans) décrites comme 'chuteurs' ou non selon leur historique. L'étude portait sur la mesure de la variabilité des paramètres temporels de marche (temps d'appui, du double appui etc..) pour chaque personne durant 6 minutes de marche. Les sujets étiquetés comme 'chuteurs' ont montré une plus grande variabilité que celle déduite chez les personnes âgées non chuteuses ou chez les sujets jeunes. Par contre la vitesse moyenne de la marche a été identique chez les sujets âgés 'chuteurs' et 'non-chuteurs'.

Le capteur utilisé n'est pas facile à utiliser, possède une durée de vie réduite et par conséquent devient cher pour une utilisation à longue durée (suivi longitudinal dans le temps). En plus, il est encombrant, étant relié à une carte d'acquisition portable. En revanche, nous retiendrons de cette approche quelques paramètres temporels de marche qui nous semblent intéressants comme le temps d'appui et du double appui.

2.2.3 Goniométrie

Un goniomètre est utilisé dans l'analyse de la marche pour évaluer l'amplitude des flexions et des extensions des articulations (cheville, genou, hanche etc..) durant la marche. C'est un capteur électromécanique à deux parties qu'on attache aux muscles qui entourent les os formant l'articulation dont nous voulons suivre le mouvement. En effet un goniomètre fournit une tension proportionnelle à l'angle formé entre les deux surfaces où les extrémités du capteur sont fixées. Un goniomètre de base est formé d'un potentiomètre circulaire. Les goniomètres électriques, plus pratiques à utiliser, comprennent une partie mobile flexible entre deux extrémités qu'on fixe autour de l'articulation (fig. 2.3).



(a) (b) (c)

fig. 2.3 : (a) Goniomètre mécanique, (b) goniomètre électrique, (c) goniomètre électrique attaché autour d'un genou pour mesurer l'amplitude de sa rotation

Ce type de goniomètres permet de mesurer l'angle entre deux segments quand le centre de rotation bouge durant la rotation ou quand il n'est pas bien défini (comme dans le cas de la

marche). La nouvelle génération de goniomètres utilise une technique de fibres optiques et n'a pas encore prouvé sa validité. Les principaux inconvénients de ce capteur sont sa faible robustesse et la procédure encombrante pour le fixer efficacement sur l'articulation.

Nitz et al. [15] utilisent des goniomètres pour mesurer la flexion aux chevilles chez 380 femmes âgées de 40 à 80 ans. Les résultats montrent une corrélation importante entre l'étendue de la flexion de la cheville, le vieillissement et le nombre de chutes. L'étude montre aussi qu'une étendue plus réduite des flexions des chevilles est observée chez les femmes plus âgées et peut être à l'origine de leurs chutes. L'étude des flexions aux hanches et aux chevilles permet aussi d'évaluer l'équilibre postural de la personne et son aptitude à établir son équilibre durant la marche.

Un goniomètre est difficile à placer et à étalonner par une personne âgée vivant seule. D'autre part, il est encombrant à porter, d'où sa non-conformité aux contraintes de notre étude. Cependant, l'angle de flexion de la cheville est un paramètre à inclure dans notre étude.

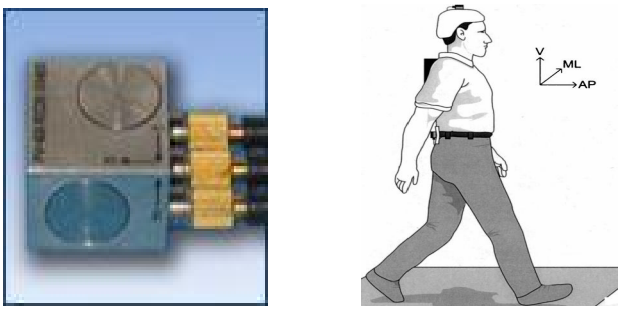
2.2.4 Utilisation d'accéléromètres

Les accéléromètres miniatures sont souvent utilisés pour l'analyse de la marche et de l'activité. Récemment, plusieurs capteurs à base d'accéléromètres ont été aussi proposés pour la détection de l'occurrence d'une chute chez les personnes âgées. Noury et al. [16] utilisent un capteur formé de 3 accéléromètres disposés orthogonalement et d'un microcontrôleur qui détermine l'inclinaison du corps, pour détecter d'une manière autonome la chute d'une personne âgée et émettre une alarme à partir d'un système intelligent à base de règles. Les accéléromètres tri axiaux (fig. 2.4) peuvent être fixés au niveau du bassin, de la cheville, ou de la tête. L'étude des signaux d'accélération fournit un ensemble de paramètres de la marche :

- la vitesse de marche en m/s ;
- la cadence de marche en pas/s ;
- la longueur moyenne de pas en m ;
- la valeur RMS (Root Mean Square) des signaux d'accélération suivant trois axes (medio latéral, vertical, antéropostérieur) ;
- l'analyse de l'harmonique des signaux d'accélération.

Menz et al. [17] utilisent deux accéléromètres tri axiaux attachés au niveau du bassin et de la tête, pour détecter les différences entre les personnes âgées et les sujets plus jeunes vis a vis des

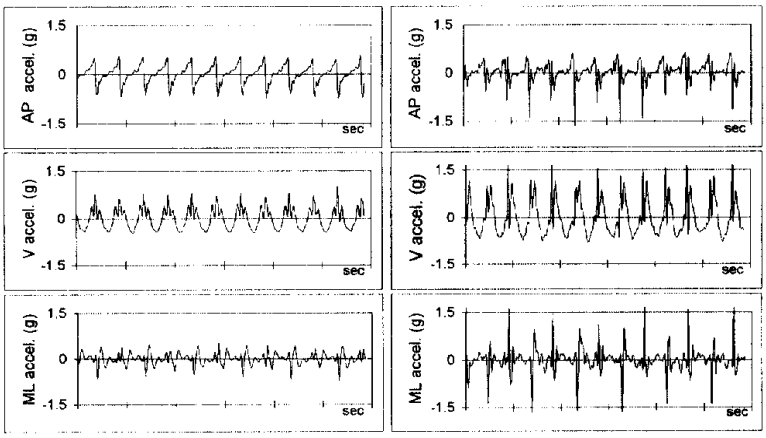
paramètres cités ci dessus. Une différence entre les groupes d'âge concernant la vitesse de marche et la longueur du pas, était bien observée. En effet les personnes âgées avaient une vitesse plus réduite et une plus petite longueur de pas en comparaison avec les sujets jeunes. De même pour les valeurs RMS des signaux d'accélération, qui avaient des plus petites valeurs chez les plus âgés. En conclusion, nous pouvons dire que les personnes âgées adoptent une stratégie plus prudente durant la marche, caractérisée par des vitesses, des longueurs de pas et des accélérations réduites, et les signaux d'accélération traduisent bien ces différences.



(a) (b)

fig. 2.4 : (a) Accéléromètre tri axial miniature et (b) son positionnement sur le corps durant la marche

Moe-Nilssen [18], décrit un système portable fixé au niveau du pelvis pour évaluer les accélérations (ML, V, AP). Il utilise ce système pour comparer les signaux d'accélération de sujets sains à ceux de personnes ayant des problèmes de marche. En se basant sur les résultats obtenus (fig. 2.5), il déduit que les amplitudes et les fréquences des accélérations chez les personnes ayant des problèmes de marche sont plus grandes que chez les sujets sains.



(a) (b)

fig. 2.5 : Profils d'accélération d'un sujet jeune (contrôle) (a) et d'un sujet ayant des problèmes d'équilibre (b)

En menant une étude sur 53 personnes âgées (20 'chuteurs'), Auvinet et al. [19] mesurent cinq variables caractérisant la marche sur un parcours de 40m, en utilisant un accéléromètre portable :

1. vitesse de marche en m/s ;
2. fréquence des pas en Hz;
3. longueur des pas en m ;
4. symétrie des pas;
5. régularité des pas.

Les variables vitesse, fréquence et longueur de pas prennent des valeurs plus réduites chez les 'chuteurs', mais ne sont pas considérées comme des indicateurs fiables pour la prédiction des chutes. Par contre, l'irrégularité des longueurs des pas est considérée comme un facteur pertinent pour la prédiction des chutes.

L'utilisation de tout capteur portable (les accéléromètres) est exclue de notre étude. Par contre, la fréquence des signaux d'accélération et les longueurs des pas seront des paramètres à exploiter.

2.2.5 Utilisation de gyroscopes

Un gyroscope (fig. 2.6) mesure la vitesse de rotation du segment sur lequel il est fixé. Le mouvement humain étant formé par un ensemble de rotations d'articulations, l'utilisation de gyroscopes s'avère intéressante. Contrairement aux accéléromètres le signal fourni par un gyroscope n'est pas perturbé par la gravité terrestre et du coup fournit une mesure fiable et précise. Les rotations des articulations sont déduites par une simple intégration du signal fourni par le gyroscope. Ainsi, des estimations des angles de flexion des genoux [20], de la rotation du tronc [21] et des cycles de marche [22] ont été réalisées. Najafi [23] utilise un gyroscope modifié pour surveiller l'activité et évaluer le risque de chute chez les personnes âgées avec succès.



fig. 2.6 : Gyroscope portable

Malgré sa fiabilité, l'utilisation d'un gyroscope semble peu pratique à cause de son prix élevé, sa consommation en énergie et sa sensibilité aux chocs. D'autre part un travail important de filtrage et de traitement de signal doit être réalisé pour éliminer les artefacts dans le signal brut fourni. Toutefois, l'angle de flexion extension du genou sera retenu parmi les paramètres exploitables.

2.3 Approches passives par parcours instrumentés

L'utilisation des capteurs portables pose un problème d'acceptabilité et d'encombrement chez les sujets concernés, d'où l'idée de développer des systèmes passifs externes (exo-capteurs). Parmi ces systèmes, nous allons présenter ci dessous l'approche par plate-forme de forces, par tapis actimétriques, par télémètre et par système optique de suivi de mouvement à base de marqueurs.

2.3.1 Plate-forme de force

Une plate-forme de force classique (fig. 2.7) est composée de quatre capteurs de force tri axiaux permettant, à partir de la mesure des forces de réaction exercées au passage de la personne sur la plate-forme, d'évaluer les 3 composantes orthogonales de la résultante des forces qui y sont appliquées, le couple vertical et les coordonnées (dans le plan antéropostérieur et médiolatéral) du centre de pression. Une étude temporelle des signaux provenant des capteurs de pression permet aussi de détecter les moments exacts du contact pied-sol et du lever du pied durant la marche. Dans le cadre de l'étude de l'équilibre dynamique, la plate-forme de forces est souvent intégrée dans un parcours de marche.

Des études [24] montrent, qu'en comparaison avec les jeunes, les personnes âgées en bonne santé ont une oscillation posturale spontanée plus importante et des réflexes posturaux plus lents pour rétablir leur équilibre à la suite d'une poussée. En effet, l'oscillation posturale reflète bien la stratégie d'équilibre de chaque personne et par conséquent est utilisée comme un facteur d'évaluation des risques de chutes chez les personnes âgées. La procédure la plus utilisée pour évaluer l'aptitude de chaque personne à maintenir une stabilité posturale se réduit aux mesures des positions de son centre de pression et des forces de réaction en utilisant une plate-forme de forces. Cheng et al. [25] mesurent la force de réaction sous chaque pied pour estimer le temps de transition entre une position assise et debout et utilisent ce paramètre pour différencier les 'chuteurs' des 'non-chuteurs'. Ce paramètre reste quand même subjectif étant déterminé par le médecin suivant son estimation de la durée de la phase de la préparation à se lever. La plate-

forme de forces était aussi utilisée par Duncan et al. [26] pour effectuer le test du 'Fonctionnal Reach' (décrit précédemment).

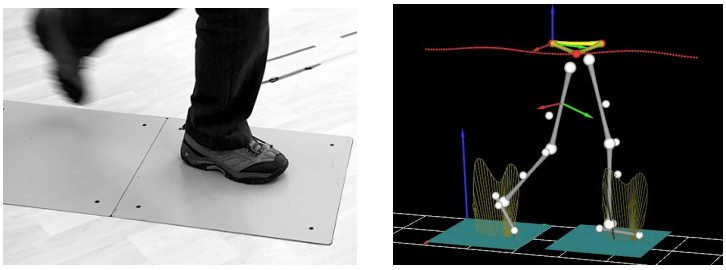


fig.2.7 : Plate-forme de forces et mesure des forces de réaction durant la marche

L'intégration d'une plate forme de force chez une personne âgée à domicile n'est pas pratique. En effet, ce capteur est coûteux et peut perturber le mouvement spontanée normale de la personne en question. L'oscillation posturale semble pourtant être un paramètre pertinent.

2.3.2 Tapis actimétriques

Plusieurs études montrent la différence des paramètres spatio-temporels entre une population de personnes âgées 'chuteurs' et une population de 'non chuteurs'. L'une des méthodes les plus pratiques pour estimer ces paramètres se base sur l'intégration dans un parcours de marche bien défini, d'une matrice de résistances modulables selon la force exercées [27] ou d'une matrice de capteurs de pression [28]. Le système GAITRite® [29] (fig. 2.8) commercialisé sous forme de tapis utilise ce type de technologies.



Tapis équipé avec des capteurs de pression pour détecter le contact pied-sol

fig. 2.8 : Le système GAITRite ® formé par un tapis de 366 cm de longueur et 61 cm de largeur et l'interface montrant les différents paramètres mesurés. Ce tapis contient une matrice de 48*288 capteurs situés à 1.27 cm l'un de l'autre.

Basser et al. [31], utilisent le système GAITRite pour prédire les chutes dans une population de 71 personnes âgées. Ce système calcule un index PFI (Probability of Fall Index) à partir de la mesure de neuf paramètres temporels et six paramètres spatiaux. Ces mesures sont recueillies lors du passage de la personne observée sur un tapis instrumenté de longueur 3.66m. Les personnes en question sont classées suivant leur index PFI obtenu. Les résultats des observations montrent que ce système n'identifie les 'chuteurs' correctement qu'avec un taux de réussite de 52%. Il s'agit donc d'un système non fiable pour déceler la déviation du comportement vers une chute.

Pour expliquer la contradiction entre le fait d'avoir des valeurs bien différentes des paramètres de marche chez les chuteurs et l'incapacité à prédire les chutes à partir de ces mêmes paramètres, Maki BE [32] mène une étude psychologique et physiologique sur un ensemble de 101 seniors. Les paramètres spatio-temporels sont calculés et comparés. La plus petite longueur du pas, le plus grand temps du double support et d'autres différences entre les variables chez les chuteurs, sont liés plutôt à la peur de rechuter et non à la chute elle-même. Par contre, la plus grande variance (entre les pas) de la longueur du pas, de la vitesse, et du temps du double support indiquent un risque de chute et ne résultent pas de la peur. Seule, une plus grande largeur du pas indique une peur et un risque de chute à la fois. Cependant l'augmentation de la variance de la vitesse est considérée comme le meilleur indicateur indépendant de risque de chutes.

En conclusion, nous pouvons dire que les variables (longueur de pas, vitesse, etc.) utilisées souvent pour prédire les chutes, peuvent en effet être des facteurs stabilisants liés à la peur de tomber. Une largeur du pas plus importante, n'améliore pas nécessairement la stabilité, mais peut indiquer une grande probabilité que la personne tombe. Seule la variance de certaines variables (longueur des pas, vitesse, temps du double support), peut être utile pour identifier les personnes ayant un risque de chute élevé. En effet, nous avons déjà identifié ces variables, extraites à l'aide d'autres techniques citées ci-dessus, comme étant des variables révélatrices d'un trouble d'équilibre. Par contre, l'usage d'un tapis actimétrique peut constituer un facteur extrinsèque poussant à la chute. Le prix exorbitant de ce type d'instrument constitue un inconvénient supplémentaire de cette technique.

2.3.3 Télémètre

Les télémètres à ultrasons (fig. 2.9) étaient aussi utilisés [30] pour mesurer la vitesse instantanée de la marche d'une personne et la fréquence de ses pas. Ce système comporte un transpondeur porté par la personne qu'on veut suivre et un émetteur infrarouge / récepteur

ultrason. La base émet des impulsions infrarouges puissantes qui atteignent le transpondeur et le poussent alors à émettre un signal ultrasonore qui sera reçu par la base. Le calcul du temps écoulé entre l'émission du signal infrarouge et la réception du signal ultrasonore permet alors de déterminer la vitesse du transpondeur et donc du sujet. Une autre technique consiste à installer le long du parcours des émetteurs récepteurs infrarouges dont le signal sera interrompu au passage de la personne, déclenchant ainsi un compteur. Cette dernière technique est simple, mais ne fournit que la vitesse moyenne de la marche et reste encombrante et coûteuse à installer. Par conséquent, cette technique ne sera pas utilisée pour évaluer les paramètres de la marche.

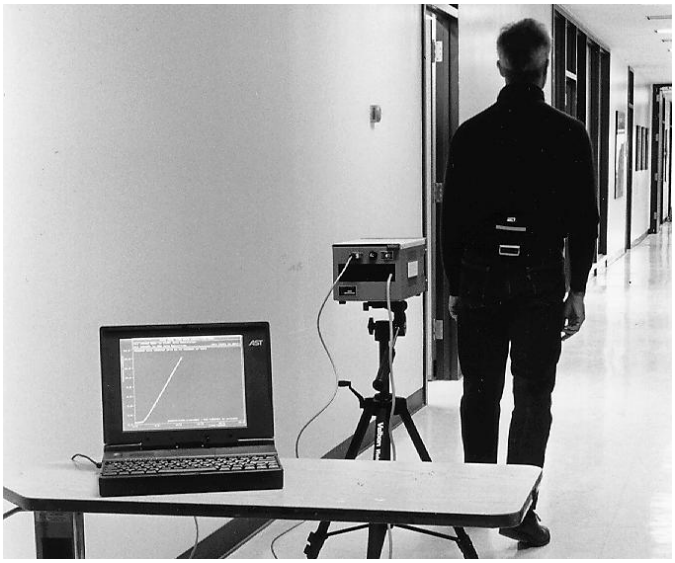


fig. 2.9 : Télémètre à ultrasons utilisé pour mesurer la vitesse instantanée et la fréquence de pas d'une personne.

2.3.4 Systèmes de suivi du mouvement 3D

Il existe deux grandes catégories de systèmes destinés à la mesure des mouvements du corps. La première utilise des dispositifs fournissant un enregistrement visuel des positions des segments du corps, alors que la seconde utilise des capteurs magnétiques pour déterminer la position et l'orientation des segments dans l'espace. Les dispositifs utilisant la capture d'image peuvent, à leur tour, être subdivisés en deux catégories en fonction du type de marqueurs mis en œuvre. Nous distinguons les marqueurs actifs et les marqueurs passifs. Les systèmes passifs utilisent des marqueurs lumino-réfléchissants (Vicon [33], Elite [34] etc..), c'est à dire qu'ils réfléchissent la lumière vers le capteur, alors que les systèmes actifs (Visualeyez [35]) utilisent des marqueurs qui comportent une source de lumière à base de LED (Light Emitting Diode).

Les systèmes de suivi du mouvement à base de marqueurs sont de plus en plus utilisés pour l'analyse du mouvement humain. Ces systèmes fournissent une modélisation cinématique du corps et aident à l'analyse de certains paramètres de la marche [36]. Le principe de ces systèmes repose sur l'utilisation de marqueurs attachés à certains points-clés du corps (fig. 2.10). Contrairement à un système qui utilise des marqueurs actifs, un système à marqueurs passifs est moins encombrant grâce aux marqueurs sans fil.



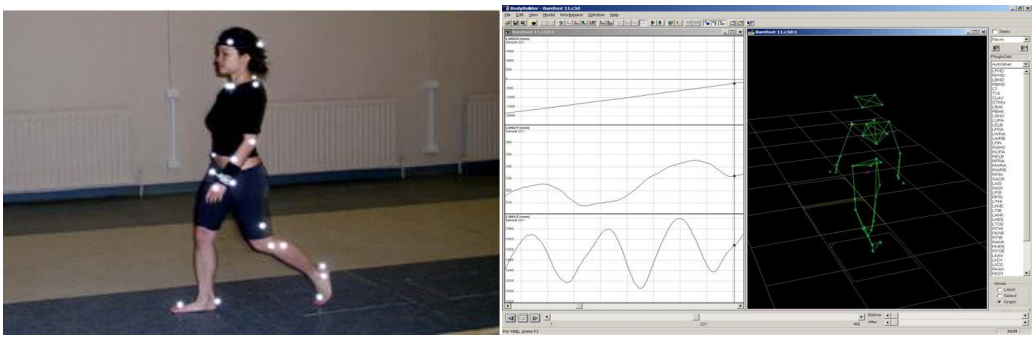
(a) (b) (c)

fig. 2.10 : (a) marqueurs lumino - réfléchissants utilisés dans les systèmes dits 'passifs' ; (b) Caméra Vicon® équipée d'un émetteur/récepteur infrarouge ; (c) marqueurs 'actifs' (LED) fixés sur une combinaison pour le système Visualeyez®

Tout système à base de marqueurs doit être installé dans un espace équipé de plusieurs caméras à émetteurs/récepteurs infrarouges. Chaque marqueur réfléchit le signal infrarouge émis par les caméras (pour les marqueurs 'passifs') ou émet un signal infrarouge (pour les marqueurs 'actifs'). Ce signal-réponse ne sera capté que par la caméra ayant une ligne de vue directe sur le marqueur, ce qui rend un certain nombre de marqueurs 'invisibles' et 'indétectables' pour une caméra donnée à un instant donné, d'où le besoin d'un nombre important de caméras (6 au minimum en général). Pour déterminer la position 3D de chaque marqueur dans le repère absolu de l'espace d'expérimentation, une méthode de triangulation est appliquée. Pour pouvoir localiser ce marqueur il faut qu'il soit visible par au moins deux caméras. De plus un calibrage des caméras s'avère nécessaire. Ehara et al. [37] établissent une description et une comparaison détaillée des performances de différents systèmes commerciaux de suivi 3D du mouvement.

Pour l'analyse de la marche, les marqueurs sont souvent placés sur les points-clés dont nous voulons suivre le mouvement. Ainsi, les positions 3D des chevilles, épaules, hanches, genoux etc. sont estimées avec une fréquence allant jusqu'à 100 Hz. Le modèle cinématique en fils de fer du corps en mouvement permet aux médecins spécialistes de détecter les anomalies au niveau des rotations des différentes articulations et du mouvement de certains segments du corps

pendant la marche ou tout autre mouvement. Ces anomalies (par exemple le balancement du corps aux chevilles ou le balancement du tronc) sont souvent considérées comme des facteurs prédicatifs de chute. Au sein des laboratoires d'analyse de mouvement, ces systèmes sont souvent combinés à une plate-forme de forces, ce qui rend l'évaluation du centre de masse du corps plus précise.



(a) (b)

fig. 2.11 : (a) Les marqueurs passifs sont fixés sur les points-clés du corps, (b) Capture d'écran du rendu d'un système Vicon®

A l'aide d'un système de suivi 3D du mouvement humain, nous pouvons détecter les positions 3D de tout point du corps humain. A partir de ces positions, la majorité des paramètres identifiés ci-dessus comme pertinents peuvent être estimés. Toutefois, il n'est pas envisageable d'installer ces systèmes à domicile, car ils sont chers et ils nécessitant un espace bien dégagé pour assurer le recouvrement des champs de vision.

2.4 Discussion et conclusion

Dans ce chapitre, nous avons présenté une analyse des différentes approches déjà utilisées pour l'évaluation de la qualité d'équilibre. Un récapitulatif de la conformité de ces approches par rapport aux exigences de l'étude PARAChute sera présenté dans le tableau 2.1.

	Tests cliniques	Approches ambulatoires	Parcours instrumentés	Systèmes de suivi 3D à base de marqueurs
Objectivité des tests utilisés	Non	Oui	Oui	Oui
Adaptabilité au domicile	Oui	Oui	Non	Non
Absence de capteurs encombrants	Oui	Non	Oui	Non
Absence d'une tierce personne	Non	Oui	Oui	Oui
Coût réduit	Oui	Non	Non	Non
Estimation de plusieurs paramètres	Non	Non	Non	Oui

Tableau 2.1 : Comparaison des différentes approches par rapport aux contraintes et objectifs de l'étude PARAChute.

Compte tenu des contraintes et objectifs de notre étude, aucune approche (technologie et paramètres utilisés) ne sera retenue seule. Par contre, nous avons identifié quelques paramètres objectifs qui ont été utilisés avec succès, pour différencier une population de personnes âgées 'non chuteuses', d'une autre composée de personnes ayant des problèmes d'équilibre durant la marche. L'idée sera alors de trouver une nouvelle technique (capteur) peu coûteuse qui puisse fournir tous ces paramètres, d'une manière passive, non encombrante.

2.4.1 Incompatibilité des méthodes et des technologies utilisées avec les contraintes du projet PARAChute

Les tests psychologiques et cliniques exposés ci-dessus ont été validés comme indicateurs fiables de risque de chute par la communauté des gériatres et font partie des protocoles médicaux pour établir un bilan gériatrique de la personne âgée durant sa visite médicale régulière. L'évaluation de ces tests est souvent subjective parce que le résultat dépend de la décision de chaque médecin. D'autre part, les exercices à effectuer sont souvent dangereux à effectuer pour une personne âgée sans la présence d'un personnel médical ou infirmier. Dans le cadre de notre étude, nous recherchons des indicateurs objectifs que nous pouvons estimer à partir de mouvements simples que la personne pourra effectuer d'une manière autonome à domicile, ce qui rend la majorité de ces tests incompatibles avec notre approche.

Les tests basés sur l'utilisation de capteurs portables fournissent un diagnostic fiable et objectif ; la présence d'une tierce personne pour accomplir ces tests n'est pas nécessaire. Toutefois, l'utilisation des capteurs portables ne présente pas une solution convenable puisque les personnes suivies, surtout les plus âgées, se changent souvent d'habits et risquent de ne pas remettre les capteurs par oubli ou par manque de volonté. Certaines personnes ne sont pas toujours suffisamment habillées pour porter les capteurs sur eux ; par exemple les malades d'Alzheimer se déshabillent souvent et ainsi abandonnent les capteurs [38]. De plus, les sujets âgés sont très sensibles à tout changement, même petit, dans leur environnement de vie [39] et se sentent gênés et encombrés par les câbles et les capteurs. Les paramètres estimés par ces capteurs sont intéressants pour notre étude mais il fallait développer un autre capteur 'passif' pour les extraire, la seule approche admissible étant celle d'exo-capteurs.

Un parcours équipé d'un système optique pour le suivi du mouvement, d'une plate-forme de forces ou d'un tapis actimétrique, permet une analyse complète de la marche, mais il doit être installé dans un endroit qui y est dédié (un laboratoire de mouvement) et ne peut pas s'intégrer facilement à un environnement de vie classique (maison). D'autre part, le sujet à suivre ne peut

bouger que dans un espace réduit, clos et qui doit être bien calibré. Par conséquent, l'analyse de la marche s'effectue sur un nombre réduit de pas en admettant que ces pas correspondent au mouvement habituel normal de la personne, ce qui n'est pas toujours valide. De plus, le coût d'un système optique ou d'un tapis équipé est exorbitant. L'installation d'un tel parcours peut aussi perturber l'environnement de vie quotidienne de la personne à suivre et par conséquent construire un facteur extrinsèque précipitant la chute.

Pour les systèmes optiques à base de marqueurs, il faut qu'un marqueur soit visible par deux caméras ou plus pour qu'on puisse suivre son mouvement. Donc, à un moment donné si un marqueur est caché (par un habit par exemple) le système sera incapable de déterminer sa position. La condition idéale pour l'utilisation d'un tel système est que la personne à suivre soit en habits moulants, ce qui constitue un facteur rédhibitoire dans le cadre de notre étude. D'autre part, ces systèmes sont complexes et exigent la présence d'un opérateur bien formé pour les utiliser. Les marqueurs doivent être attachés au corps et par la suite deviennent une source d'encombrement pour la personne suivie. Pour toutes ces raisons l'usage d'un système optique à marqueurs n'est pas envisageable dans le cadre de notre étude.

2.4.2 Récapitulation des paramètres identifiés comme pertinents pour l'évaluation de l'équilibre dynamique

Toutes les approches exposées dans ce chapitre sont incompatibles avec notre approche du point de vue technologique. Toutefois, ces approches se basent sur l'estimation d'un certain nombre de variables qui étaient identifiées dans la littérature comme étant pertinentes pour l'évaluation de l'équilibre dynamique. En effet, l'analyse des paramètres spatio-temporels, des accélérations et des rotations des différentes articulations durant la marche devra permettre d'établir un indicateur pour l'évaluation du risque de chute chez les personnes âgées. La liste des paramètres que nous avons identifiés, à partir des approches exposées, comme intéressants est la suivante :

- la vitesse moyenne de la personne durant la marche (valeur et variance) ;
- l'accélération du tronc durant la marche ;
- les longueurs des pas gauches et droits ;
- les largeurs des pas (espacement des pieds) ;
- la régularité des pas (variance des longueurs des pas) ;

- la symétrie des pas (différence des longueurs entre le pas gauche et le pas droit) ;
- les hauteurs des pas gauches et droits (élévations des pieds par rapport au sol) ;
- oscillation du tronc durant la marche ;
- angles de flexion aux chevilles ;
- angles de flexion aux genoux ;
- temps d'appui et de double appui ;
- déviation d'un trajet donné.

Ces paramètres sont à extraire durant la marche normale quotidienne de la personne. Pour que la variabilité (entre pas) de ces paramètres, exprimée à l'aide de la variance, soit significative, il faut que la personne effectue au moins six pas. Une description plus détaillée de ces paramètres et de la méthode pour estimer chacun sera détaillée dans le chapitre 5.

2.4.3 Quel type de capteurs adopter ?

Pour pouvoir estimer ces paramètres, il suffit de déterminer les positions 3D de quelques points-clés du corps avec une fréquence donnée durant la marche, comme dans le cas des systèmes à marqueurs, mais en adoptant une approche passive sans marqueurs ni capteurs portables. L'idée est alors de développer un système de suivi du mouvement 3D passif et peu coûteux à partir d'un flux vidéo correspondant au mouvement naturel de la personne suivie ; le capteur dans ce système sera une caméra vidéo classique bon marché (webcam, caméra numérique etc..). Pour que cette solution soit non intrusive, il faut s'assurer qu'aucune image de la personne suivie ne soit transmise hors domicile. Dans le chapitre 3 nous exposerons les différentes techniques utilisées pour la reconstruction des mouvements 3D sans l'utilisation de marqueurs. La solution retenue sera détaillée dans le chapitre 4.

Bibliographie

[1] Wolf S, Barnhart H et al., ‘‘Reducing Frailty and Falls in Older Persons : An Investigation of Tai Chi and Computerized Balance Training’’, Journal of the American Geriatrics Society, Vol.44, pp 489-497, 1996.

[2] Berg KO, Maki BE, Williams F et al., “Clinical and laboratory measures of postural balance in an elderly population”, Archives of physical and medical rehabilitation, Vol.73, pp 1073-1080, 1992.

[3] Andersson AG, Kamwendo k et al., “How to identify potential fallers in a stroke unit : validity indexes of 4 test methods”, Journal of rehabilitation medicine, Vol.38, pp 186-191, 2006.

[4] Chiu AY, Au-Yeung SS et Lo SK, “A comparison of four functional tests in discriminating fallers from non-fallers in older people”, Disability and rehabilitation, Vol. 25, pp 45-50, 2003.

[5] Shumway-Cook A et Wollacott M, “Motor Control : Theory and Practical Applications”, Williams and Wilkins, 1995.

[6] Whitney SL, Marchetti GF, Schade A et Wrisley DM, “The sensitivity and specificity of the Timed "Up & Go" and the Dynamic Gait Index for self-reported falls in persons with vestibular disorders”, Journal of vestibular research : equilibrium and orientation, Vol.14, pp 397-409, 2004.

[7] Rose D, Jones C.J et Lucchese N, “Predicting the probability of falls in community-residing older adults using the 8 foot up and go : anew measure of functional mobility”, Journal of Aging and Physical Activity, Vol. 10, pp 466-475,2002.

[8] Tinetti M.E, “Performance-oriented assessment of mobility problems in elderly patients”, Journal of American Geriatrics Society, Vol. 34, pp 119-126, 1986.

[9] Murphy M, Olson S.L, Protas E.J et Overby A.R, “Screening for falls in community-dwelling elderly”, Journal of Aging and Physical Activity ,Vol. 11, pp 66-80, 2003.

[10] Mankovskii N, Mints Y.A et Lysenyuk U.P, ‘‘Regulation of the preparatory period for complex voluntary movement in old and extremely old age’’, Journal of Human Physiology, Vol. 6 : pp 46-50, 1980.

[11] Woollacott M et Manchester D, ‘‘Anticipatory postural adjustments in older adults : Are changes in response characteristics due to changes in strategy?’’ Journal of Gerontology ,Vol. 48, no.2, pp 64-70, 1993.

[12] Brauer S.G, Burns Y.R et Galley P, ‘‘A prospective study of laboratory and clinical measures of postural stability to predict community-dwelling fallers’’, Journal of Gerontology, Vol. 55, no.8, pp 469-76, 2000.

[13] Noraxon®, http://www.noraxon.com

[14] Hausdorff JM, Edelberg HK, Mitchell SL et al, ''Increased gait unsteadiness in community-dwelling elderly fallers'', Archives of Phys. Medical Rehabilitation, vol.78, pp 278-283, 1997.

[15] Nitz J, Low Choy N, "The relationship between ankle dorsiflexion range, falls and activity level in women aged 40 to 80 years", New Zealand Journal of Physiotherapy, vol. 32(3), pp 121-125,2004.

[16] Noury N et al, "Un capteur intelligent pour la détection des chutes – Fusion multi-capteurs et décisions à base de règles" , 'Nouvelles méthodes d'instrumentation', Hermès-Elsevier, vol.2, pp 411-418,2004.

[17] Menz HB, Lord SR, Fitzpatrick RC, "Age-related differences in walking stability", Journal of Age and Ageing, vol. 32, pp 137-142, 2003.

[18] Moe-Nilssen R., "A new method for evaluating motor control in gait under real-life environmental conditions", Journal of Clinical Biomechanics, vol. 13, pp 320-335, 1998.

[19] Auvinet B, Berrut G, Touzard C, Moutel L et al, ''Gait abnormalities in elderly fallers'', Journal of Aging and Physical Activity, vol. 11, pp 40-52, 2003.

[20] Dejnabadi H, Aminian K, Najafi B et al, ''Joint and Segment Angles of Lower Limbs in Hip Ostheoarthritis and Total Hip Replaced Patients Measured Using Physilog System'', Gait & Posture, vol. 20, pp 102-107, 2004.

[21] Najafi B, Aminian K et al, ''An Ambulatory System for Physical Activity Monitoring in elderly'', Proceeding of the 1st International IEEE EMBS Special Topic Conference on Microtechnology in Medicine and Biology, Lyon, France, pp 562-566, oct.2000.

[22] Aminian K, Najafi B et al, ''Ambulatory gait analysis using gyroscopes'', Proceeding of the 25th Annual meeting of the. American Society of. Biomechanics (ASB), pp 309-310, 2001.

[23] Najafi B, "Physical activity monitoring and risk of falling evaluation in elderly people", thèse de l'école polytechnique fédérale de Lausanne, 2002.

[24] Maki BE et McIlroy WE., ''Postural control in the older adult'', Clin Geriatr Med, Vol.12, pp 635-58, 1996.

[25] Cheng PT, Liaw MY, Wong MK, Tang FT et al, ''The Sit-to-Stand Movement in Stroke Patients and Its Correlation with Falling'' Archives of Phys. Medical Rehabilitation, Vol. 79, pp 1043-1046, 1998.

[26] Duncan PW, Weiner DK, Chandler J et al, ''A new clinical measure of balance'', Journal of Gerontology : Medical Sciences, Vol.45, pp 192-197, 1990.

[27] Hirokawa S, Matsumura K, ''Gait analysis using a measuring walkway for temporal and distance factors'', Journal of Medical and Biological Engineering Computing, vol.25, no.5, pp577-582, 1987.

[28] Macellari V, Giacomozzi C, ''Multistep pressure platform as a stand-alone system for gait assessment'', Journal of Medical and Biological Engineering Computing, vol.34, no.4, pp 299-304, 1996.

[29] GaitRite®, http://www.gaitrite.com

[30] Weir RF, Childress D, ''A Portable, Real-Time, Clinical Gait Velocity Analysis System'', IEEE Transactions on Rehabilitation Engineering, vol. 5, no.4, pp 310-321, 1997.

[31] Besser MP, Selby-Silverstin L, ''Predicting fall risk in the elderly using temporal-spatial parameters of gait'', Proceedings of the Symposium of the International Society for Postural and Gait Research, Control of Posture and Gait, Maastricht, The Netherlands, pp 70-73.

[32] Maki BE, ''Gait Changes in older adults : predictors of falls or indicators of fear'', Report Centre for studies in aging, Sunnybrook Health Science Centre/ University of Toronto Canada, 2003.

[33] Vicon®, http://www.vicon.com

[34] Elite®, BTS Bioengineering http://www.bts.it

[35] Visualeyez®, Phoenix Technologies Incorporated http://www.ptiphoenix.com

[36] Simoneau GG, Krebs DE, ''Whole-body momentum during gait : A preliminary study of non-fallers and frequent fallers'', Journal of Applied Biomechanics, vol.16, no.1, pp 1-13, 2000.

[37] Ehara Y, Fujimoto H, Miyazaki S et al, ''Comparison of the performance of 3D camera systems'', Journal of Gait & Posture, vol. 5, no.3, pp 251-255, 1997.

[38] Burgio L, Scilley K, Hardin M et al, "Studying Disruptive Vocalization and Contextual Factors in the Nursing Home Using Computer-Assisted Real-Time Observation", Journal of Gerontology, vol. 49, pp 230-239, 1994.

[39] Burgio L, Scilley K, Hardin M et al, "Temporal patterns of disruptive vocalization in elderly nursing home residents", International Journal of Geriatric Psychiatry, vol. 16,pp 378-386, 2001.

Chapitre 3 : Suivi 3D du mouvement humain et filtrage particulaire

Dans le chapitre précédent nous avons détaillé plusieurs méthodes permettant de réaliser l'analyse de la marche. Les méthodes exposées présentent des inconvénients majeurs, ou sont mal adaptées aux contraintes et exigences de notre projet. Une analyse fine de la marche nécessite un suivi 3D complet du mouvement humain. Le capteur que nous avons choisi pour effectuer cette analyse est une (ou plusieurs) caméra(s) vidéo. Ce capteur a été choisi pour ses caractéristiques de coût, d'encombrement et aussi parce qu'il s'agit d'un capteur passif et non invasif.

Dans ce chapitre nous allons présenter les techniques utilisées pour le suivi 3D complet du mouvement humain à partir de flux vidéo. Les techniques de stéréovision combinées à l'utilisation de deux (ou plus) caméras bien calibrées, permettent de déterminer les coordonnées 3D de tout point situé dans les champs de vision des caméras. Ainsi, pour localiser un point donné il faut l'étiqueter dans l'image. Cette opération ne peut pas être réalisée automatiquement d'une manière fiable, d'où l'idée d'utiliser des modèles virtuels humanoïdes.

Les méthodes fondées sur l'utilisation de modèles 2D et de l'appariement 2D-3D ainsi que les approches utilisant les modèles de mouvement issus des déplacements des blocs de pixels dans l'image, seront exposées. Les résultats et les contraintes qu'imposent ces méthodes, n'étant pas adaptés à notre problématique, nous opterons pour le développement d'une nouvelle approche.

L'utilisation d'un modèle 3D humanoïde virtuel et paramétrable, simulant le mouvement humain, est une idée intéressante. En effet le suivi 3D peut ainsi être assimilé à la recherche de la configuration optimale des paramètres du modèle virtuel par rapport à l'observation qui est dans notre approche fournie par l'image.

Dans ce qui suit nous présenterons alors une formulation du problème auquel nous nous sommes confrontés. Après un bref rappel sur le problème de l'estimation dynamique (filtrage bayésien), nous parlerons de l'intérêt des filtres de Kalman et de l'approche particulaire pour effectuer cette estimation. Le filtrage particulaire étant l'outil le mieux adapté pour notre problème, des exemples d'application de cette approche pour le suivi 3D du mouvement humain seront aussi exposés et les bases de la solution que nous proposons seront établies.

3.1 Analyse globale du mouvement humain par stéréovision

Une approche globale du suivi 3D du mouvement humain peut être proposée en utilisant plusieurs méthodes développées pour la vision par ordinateur. En effet, il est nécessaire de comprendre le mécanisme de formation de l'image (voir Annexe A) et de calculer les paramètres de la caméra pour pouvoir effectuer la reconstruction 3D d'une scène.

Une reconstruction en 3D du mouvement d'un objet est possible à partir d'une seule caméra, connaissant d'autres informations extérieures à l'image telles les coordonnées du plan où se déplace cet objet, ses dimensions ou sa profondeur. Or, ce n'est pas facile de déterminer le plan du mouvement pour chaque scène ni les dimensions de l'objet. Par contre, l'estimation de la profondeur d'un objet, sans aucune information complémentaire, est possible. En effet, il faut utiliser deux vues différentes au moins (deux caméras, deux yeux etc..) pour calculer la profondeur par rapport à une caméra (distance dans la direction de l'axe de vue) d'un point de la scène.

Considérons un point P appartenant à une scène 3D observée par deux caméras de centres optiques O1 et O2. Le plan épipolaire est le plan comprenant les deux droites (O1P) et (O2P). Sa ligne de base est la droite (O1O2). Le point d'intersection de la ligne de base avec le plan Image1 est appelé l'épipôle e1 ; celui avec le plan Image2 sera l'épipôle e2. Nous noterons p1 et p2 les images de P sur les deux plans Image 1 et 2. La ligne épipolaire l1 associée au point p1 est l'intersection du plan Image1 avec le plan épipolaire ; de même, la ligne épipolaire l2 associée à p2 est formée par intersection du plan Image2 avec le plan épipolaire (fig. 3.1).

Si p1 et p2 sont les images du même point P, alors p1 doit appartenir à l1 et p2 à l2, c'est ce qu'on appelle la contrainte épipolaire. Comme $\overrightarrow{O_1P} = \overrightarrow{p1}$, $\overrightarrow{O_2P} = \overrightarrow{p2}$ et $\overrightarrow{O_1O_2} = \overrightarrow{t}$ sont coplanaires, nous pouvons écrire : $\overrightarrow{p1}.(\overrightarrow{t} \wedge \overrightarrow{p2}) = 0$ En exprimant cette relation dans le repère

de l'Image1 nous obtenons : $\overrightarrow{p1}.(\overrightarrow{t} \wedge R.\overrightarrow{p2}) = 0$ où $\overrightarrow{p1}$ et $\overrightarrow{p2}$ sont les vecteurs des coordonnées homogènes de p1 et p2 respectivement dans les repères Image1 et Image2, et *R* la matrice de rotation du repère Image2 par rapport au repère Image1. A partir de cette dernière relation nous pouvons déduire l'existence d'une Matrice *F* tel que : $\overrightarrow{p1}^T.F.\overrightarrow{p2} = 0$; *F* est appelée la matrice fondamentale du capteur stéréoscopique. Dans le cas où les paramètres intrinsèques et extrinsèques des caméras ne sont pas connus, il faut estimer la matrice fondamentale *F* à partir d'un ensemble d'appariements de points entre deux images issues des deux caméras. Ce processus est appelé le calibrage faible.

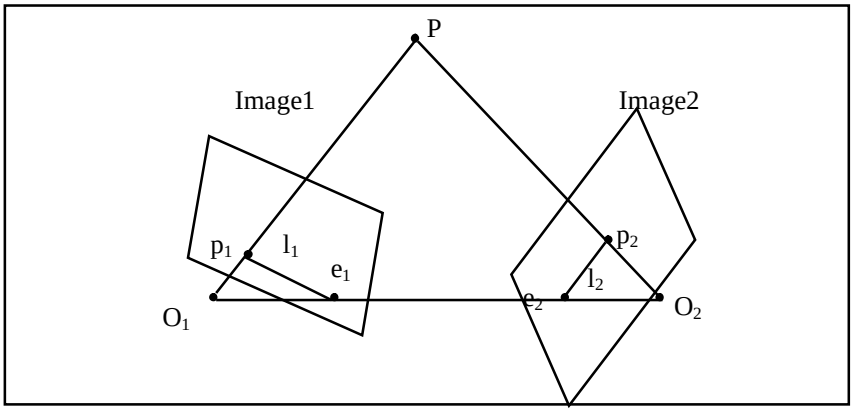


fig. 3.1 : Géométrie épipolaire

Connaissant les coordonnées des images de certains points dans l'image1 et celles des images des mêmes points dans l'image2, la matrice *F* peut être estimée. Le problème de l'appariement de deux points images est traité en deux étapes. En effet il faut d'abord choisir un point de l'image 1 et trouver un point 'candidat' lui correspondant dans l'image 2. Ce point 'candidat' doit respecter au moins la contrainte épipolaire. La deuxième étape correspond à tester la validité du point 'candidat'. Plusieurs algorithmes ont été proposés [1] pour effectuer l'appariement et pour estimer la matrice fondamentale d'un capteur stéréoscopique. Dans la pratique cette estimation présente beaucoup d'imprécisions dues à la présence de bruit dans la position des points détectés.

Connaissant la matrice fondamentale *F*, une correspondance peut être établie entre les images m_1 (par la caméra 1) et m_2 (par la caméra 2) d'un point M. En utilisant les matrices de projection établies pour chaque caméra à partir de ses paramètres extrinsèques et intrinsèques, les demi-droites $[O_1m_1)$ et $[O_2m_2)$ reliant les centres optiques aux images seront déterminées ; l'intersection de ces deux demi-droites sera le point 3D, M recherché. En présence de bruit dans l'image, cette intersection est souvent déterminée avec une certaine imprécision. En appliquant

la même démarche pour plusieurs correspondances, une reconstruction complète de la scène observée peut être établie. Ainsi pour déterminer la position 3D d'un point du corps humain, il sera nécessaire d'estimer sa position dans l'une des images.

Cette méthode de reconstruction nécessite l'étalonnage précis des caméras et l'estimation de la matrice fondamentale du capteur stéréoscopique, ce qui introduit une complexité de calcul. D'autre part les erreurs d'appariement d'images et les imprécisions introduites par le bruit rendent les résultats obtenus moins précis. Pour pouvoir reconstruire le mouvement humain, il faut à chaque instant étiqueter, dans les images fournies par les deux (ou plus) caméras, tous les points dont nous souhaitons suivre le mouvement. L'étiquetage des extrémités du corps (pieds, tête etc..) peut être effectuée automatiquement. Par conséquent, nous ne pouvons obtenir qu'une analyse globale du mouvement, n'ayant pas les positions des différents points dont nous avons besoin pour calculer tous les paramètres identifiés.

Cependant, l'étiquetage automatique des autres articulations du corps dans une image est une opération compliquée et peu fiable. L'étiquetage manuel est alors nécessaire. En revanche, dans PARAChute nous excluons toute intervention humaine dans le calcul des paramètres de la marche, d'où l'incompatibilité de cette méthode avec nos objectifs.

3.2 Analyse fine du mouvement humain par utilisation de modèles humanoïdes

L' approche décrite ci dessus fournit une estimation de la position d'un point dans un repère 3D à partir de sa position dans une image en utilisant plusieurs caméras (par triangulation), ou par application des équations de la géométrie perspective et de certaines contraintes. L'estimation des mouvements humains doit être déduite à partir des positions 3D de plusieurs points (articulations) du corps humain. Pour pouvoir suivre le mouvement 3D du corps à partir des flux vidéo, un étiquetage des positions des différentes articulations doit être effectué *a priori* dans les images. Cette dernière opération s'avère compliquée et peu précise et doit être faite d'une façon manuelle pour chaque image du flux vidéo. D'autre part, les paramètres intrinsèques et extrinsèques des caméras doivent être calculés et une calibration du couple stéréoscopique doit être appliqué en plus, dans le cas d'utilisation de plusieurs caméras. Pour toutes ces raisons, plusieurs méthodes utilisant des modèles visuels représentant le corps humain ont été proposées pour faire le suivi 3D du mouvement. Dans ce qui suit nous allons décrire quelques unes de ces méthodes.

3.2.1 Le système w4

Ce système a été créé par Haritaoglu [2] pour suivre le mouvement 2D et détecter la nature de l'activité humaine à partir de flux vidéos monochromatiques. Le système consiste à déterminer, dans le flux vidéo, une boite 2D englobante du corps en mouvement par l'utilisation d'algorithmes développés pour l'extraction de l'arrière plan. Des modèles d'apparence sont ensuite construits pour chaque personne pour pouvoir différencier plusieurs acteurs dans une même scène. D'autre part des modèles dynamiques sont élaborés pour décrire une succession de mouvements durant une action et ainsi pouvoir attribuer une action donnée à une suite de mouvements observés. L'étiquetage de chaque membre du corps humain est réalisé suivant une méthode heuristique. En effet, la hauteur de la personne suivie est estimée comme équivalente à la hauteur de la boite englobante. La tête, le torse et les jambes sont considérés comme enfermés dans des sous boites respectivement situées à 1/5, 1/2 et 1/2 de la hauteur de la boite globale (fig. 3.2). La position 2D de chaque membre dans sa sous boite est ensuite raffinée en utilisant les textures.

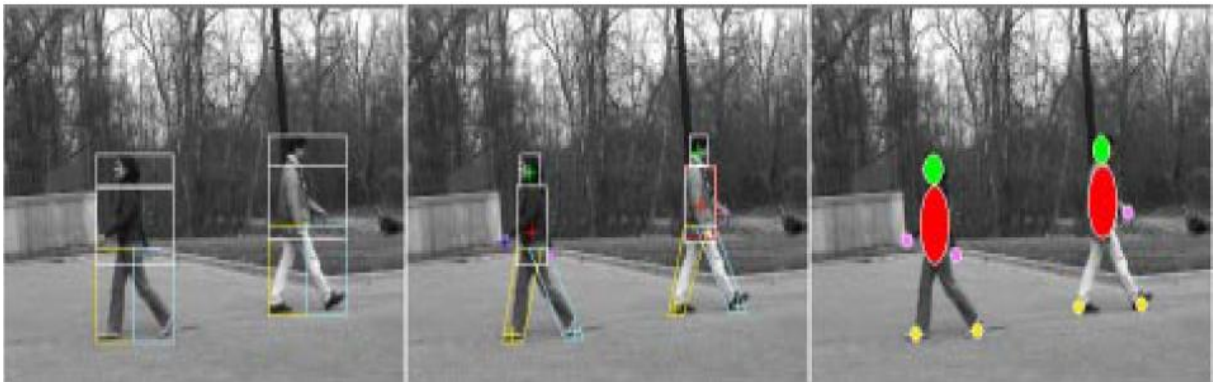


fig. 3.2 : Etiquetage des différents membres du corps humain par le système w4s [2]

En présence de bruit dans l'image (ombre) et d'occlusions, le système est incapable de bien suivre le mouvement d'où l'utilisation d'une deuxième caméra dans le système w4s [2]. L'utilisation d'une deuxième vue permet aussi de déterminer l'orientation du corps dans l'image et ainsi le suivi sera effectué en 2D ½. Toutefois, ce système utilise une méthode heuristique pour la détermination de la position de chaque membre et fournit ainsi une analyse du mouvement peu précise et globale. En effet, seules les positions 3D des points extrémités du corps peuvent être déterminées à l'aide du système w4s et en utilisant deux caméras calibrées. Une analyse plus fine du mouvement nécessite, en plus, l'estimation des mouvements 3D d'autres points ce qui rend ce système inadapté à notre approche.

3.2.2 Utilisation des modèles mathématiques de mouvement

Cette méthode a été utilisée par Bregler et al. [3] pour le suivi du mouvement d'une chaîne cinématique. En estimant que l'intensité du pixel représentant un point en mouvement dans une image reste la même entre deux images successives, ils essayent de trouver un modèle de mouvement $u\ (x, y, \phi)$ décrivant le déplacement d'un point en fonction de sa position *(x,y)* dans l'image et d'un vecteur de paramètres ϕ (ayant une dimension *K* équivalente au nombre de degrés de liberté de l'objet) ; ce modèle sera déduit à partir de l'équation : $I(x+u_x(x,y,\phi), y+u_y(x,y,\phi), t+1) = I(x,y,t)$ où $I(x,y,t)$ est l'intensité du pixel *(x,y)* à l'instant *t*. En utilisant le développement du premier ordre de cette équation, la formulation classique en gradients sera la suivante : $I_t(x,y) + [I_x(x,y), I_y(x,y)].u(x,y,\phi) = 0$ où $I_t(x,y)$ sera le gradient temporel de l'intensité du pixel et $[I_x(x,y), I_y(x,y)]$ celui suivant *x* et *y*. En observant le mouvement de *N* points (*N* > *K*) de l'objet rigide dans l'image, *N* équations pourront être établies, le vecteur ϕ sera calculé et ainsi le modèle *u* sera estimé.

En se basant sur l'observation que tout mouvement rigide peut être assimilé à une rotation autour d'un axe 3D et à une translation par rapport à cet axe (6 degrés de liberté), le modèle *u* peut être simplifié et une relation exponentielle peut être établie entre ce modèle et les valeurs des degrés de liberté de l'objet 3D en mouvement. Cette représentation est dite en 'twist' [3]. Ainsi en calculant le modèle *u* à partir de deux images successives et des degrés de liberté dans la première image, le mouvement effectué (les nouvelles valeurs des degrés de liberté) peut être estimé.

Pour le suivi du mouvement humain, chaque partie du corps sera considérée comme un objet et ainsi un calcul complexe du modèle de mouvement doit être effectué pour chacune [3]. Les points choisis pour effectuer ce calcul sont les projections dans l'image de points 3D appartenant à un ellipsoïde (représentant les parties du corps) reliant deux articulations. Cette représentation fausse le suivi du mouvement du fait que plusieurs points de l'environnement seront considérés comme des points du corps. Pour minimiser ce problème, des modèles statistiques ou des méthodes de recherche de type EM (Expectation Maximisation) peuvent être utilisés, ce qui introduit une complexité supplémentaire à l'algorithme du suivi. D'autre part le suivi se base sur l'étiquetage manuel des différentes positions des articulations du corps dans la première image. La posture 3D initiale du corps dans l'environnement doit être connue également. Ces données ne peuvent généralement pas être fournies avec précision ce qui constitue un inconvénient

majeur de la méthode. De plus, l'approche proposée utilise une méthode différentielle et par conséquent fournira de mauvais résultats dans le cas où le mouvement effectué entre deux images est rapide comme dans le cas du suivi d'une marche rapide ou dans le cas d'utilisation de vidéos prises à basse fréquence.

3.2.3 Méthode de Gavrila et Davis [4]

Gavrila et Davis ont proposé une nouvelle méthode [4] pour réaliser le suivi 3D du mouvement de plusieurs personnes dans une scène donnée. Elle se base sur la construction d'un modèle visuel individuel des personnes à suivre et la comparaison d'images de synthèse représentant plusieurs poses de ce modèle à l'image réelle. A partir de deux images (une de profil et l'autre de face) d'une personne, et en utilisant les contours du corps détectés dans ces images, un modèle 3D adapté à cette personne sera construit (fig. 3.3) ; il sera composé à partir d'une famille de formes géométriques dite la famille des *super-quadrics*. Pour simuler le mouvement humain, ce modèle a été doté de 22 degrés de liberté.

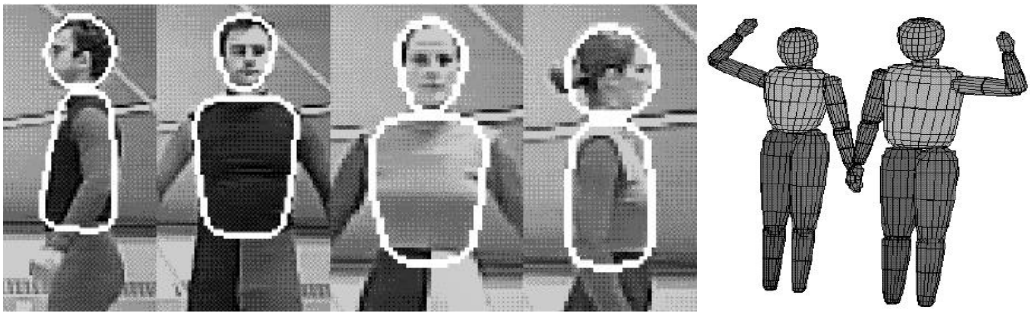


fig.3.3 : Construction d'un modèle 3D individuel par personne [4]

Le problème du suivi du mouvement sera réduit à l'estimation de la configuration du modèle qui correspond le mieux aux images. La stratégie adoptée est de créer un ensemble de configurations et les comparer à l'image pour en déduire la plus probable. Le critère de similarité utilisé est basé sur la comparaison des contours de la personne dans l'image réelle et de la configuration du modèle synthétique. La distance entre les contours est calculée en utilisant la fonction dite de *chamfer*. La distance *DD(S, R)* de *chamfer*, entre un ensemble de points *S* et un autre ensemble de points de référence *R*, est obtenue par la sommation des distances élémentaires entre chaque point de *S* et le point qui lui est le plus proche de *R* :

$$DD(S,R) = \sum_{t \in T} dd(s,R) = \sum_{t \in T} \min_{r \in R} \|s - r\|$$

La création de l'ensemble des configurations est accomplie d'une manière factorisée. En effet, à cause de la grande dimension du vecteur représentant la configuration du modèle,

l'espace de recherche des configurations est énorme. L'idée est de décomposer cet espace de recherche en sous espaces correspondant chacun à une partie du corps (torse, jambes etc..). Ainsi la recherche s'effectuera sur plusieurs positions du torse pour en trouver la meilleure avant de passer aux autres parties et ainsi de suite.

Ce choix de factorisation peut être justifié à deux conditions nécessaires. D'une part, la courbe de vraisemblance exprimée en fonction des positions d'une partie donnée du corps ne doit pas être multimodale, sinon la décision sur la bonne position de cette partie pourrait être inexacte. D'autre part, les différentes parties du corps doivent être indépendantes les unes des autres vis a vis de la fonction de vraisemblance. La position d'une partie donnée ne doit pas affecter l'estimation de vraisemblance d'une autre partie. Or, dans le cas du corps humain et de son mouvement complexe, des occlusions peuvent avoir lieu, surtout avec l'utilisation d'un nombre réduit de caméras. A cause de ces occlusions plusieurs configurations différentes du modèle peuvent avoir la même vraisemblance comme ayant un même contour. Enfin le mouvement de certaines parties du corps peut intervenir dans la formation du contour d'une autre partie et ainsi dans l'évaluation de la validité d'une configuration donnée, d'où l'absence de la notion d'indépendance. En conclusion la méthode utilisée pour créer les configurations paraît inadaptée ce qui rend l'approche correspondante inapplicable surtout en excluant l'usage d'un nombre important de caméras couvrant tous les angles de vue de la scène.

3.2.4 Suivi 2D et reconstruction 3D à l'aide de modèles appris

En se basant sur des profils de marche, appris *a priori*, Howe et al. [5] proposent une approche pour effectuer un suivi 3D du mouvement humain à partir d'une seule caméra. Leur méthode se décompose en deux étapes ; la première consiste à adapter un modèle 2D du corps humain, à l'image du corps suivi (2D). Ce modèle est formé par un ensemble de surfaces coplanaires représentant 14 segments du corps et reliant 20 points de ce corps. Les positions de ces surfaces sont contrôlées par un ensemble de 34 paramètres. Le suivi en 2D consiste à détecter le mouvement de blocs de pixels décrivant chacun une partie du corps. Les positions 2D des 20 points constituant le corps sont ensuite estimées à partir des positions estimées des différents surfaces. Pour que ce suivi soit possible, une étape d'initialisation, où un opérateur adapte le modèle 2D à la première image, doit être effectuée.

La deuxième étape de l'algorithme consiste en une reconstitution 3D des positions 2D des points du corps. Cette reconstitution est effectuée en établissant une correspondance entre un ensemble de vecteurs - positions 2D successifs et un ensemble de vecteurs - positions 3D. En fait

à partir d'un ensemble d'apprentissage, à chaque ensemble de 11 vecteurs - positions 3D successifs (positions dans 11 images successives), l'ensemble des positions 2D correspondantes dans l'image est associé. Une fois les vecteurs positions 2D estimés, une classification du type EM (Expectation Maximisation) est appliquée pour attribuer ces vecteurs à un ensemble de vecteurs 2D de l'ensemble d'apprentissage et par conséquence à l'ensemble de vecteurs positions 3D qui lui y est associé.

Le suivi 2D nécessite une initialisation manuelle mais peut être assez fiable. Cependant, la méthode de reconstruction 3D présente certaines restrictions. L'ensemble d'apprentissage reste limité et ne pourra pas regrouper toutes les configurations possibles du mouvement humain, ce qui peut fausser la reconstruction en associant un vecteur 3D (le plus probable par rapport au suivi 2D), sans aucune garantie de précision. D'autre part, le mouvement suivi doit être filmé dans les mêmes conditions que celles des vidéos de l'ensemble d'apprentissage (même distance focale de caméras, même distance entre la scène et la caméra etc..).

3.2.5 Discussion et conclusion

Nous avons détaillé plusieurs méthodes développées dans la littérature pour effectuer un suivi du mouvement humain. Les méthodes basées sur la géométrie 3D sont fiables mais nécessitent un travail considérable de calibration et ne fournissent qu'une analyse globale de la marche. Une analyse plus fine du mouvement humain s'avère nécessaire dans le cadre de notre étude. En effet l'estimation de certains paramètres de la marche nécessite le suivi 3D du mouvement de certains points du corps ce qui implique l'incompatibilité de l'analyse globale avec notre démarche.

L'utilisation de modèles humanoïdes représentant le corps humain est intéressante, du fait qu'on peut suivre le mouvement de plusieurs points du corps en estimant le mouvement du point du modèle virtuel qui lui correspond. Le suivi se réduit ainsi à l'estimation de la configuration du modèle virtuel qui correspond le mieux au mouvement détecté dans la vidéo. Toutefois, l'utilisation des modèles pour le suivi 2D ne fournit pas les données nécessaires pour calculer les paramètres que nous souhaitons estimer, d'où le besoin d'un modèle en 3D. D'autre part, la reconstitution d'une scène en 3D, à partir d'un suivi 2D et d'une base d'apprentissage, fausse l'estimation de tout mouvement n'apparaissant pas dans la base d'apprentissage. Le but de notre système étant de déceler ces mouvements et ces comportements inhabituels, toute méthode utilisant des modèles appris de la marche sera exclue dans notre approche.

En conclusion, la solution retenue pour extraire les paramètres de la marche se basera sur un modèle 3D humanoïde générique. L'estimation de sa configuration à partir des images vidéos sera accomplie grâce à un algorithme modifié basé sur une approche particulaire de l'estimation bayésienne dynamique.

3.3 Formulation du problème

Dans l'approche que nous avons adoptée, le problème du suivi se ramène à l'estimation de la configuration des paramètres du modèle 3D, de façon à ce que ce dernier représente au mieux la posture et l'apparence du corps réel dans l'image. La comparaison d'une configuration possible du modèle à l'image réelle passe par la comparaison de sa projection en image 2D synthétique à cette dernière. La solution banale de ce problème est de rechercher cette configuration optimale d'une façon exhaustive. En revanche, cette méthode possède une complexité de calcul exponentielle et en conséquence est inapplicable. En effet, l'espace de recherche des configurations possibles du modèle 3D virtuel est formé par la combinaison des valeurs possibles de chacun de ses degrés de liberté. La dimension du vecteur configuration du modèle étant importante, et les valeurs possibles pour chaque degré de liberté étant nombreuses, l'espace de recherche contient alors un nombre très grand de configurations, ce qui implique l'utilisation d'un nombre important d'images. Or la création/comparaison d'un grand nombre d'images est très coûteuse en temps de calcul et en terme d'occupation d'espace mémoire, d'où l'impossibilité technique d'appliquer ce type de recherche. Une autre solution est alors à trouver.

Grâce à la continuité dans le mouvement humain durant la marche, nous pouvons prédire la configuration future du modèle humanoïde à partir de son état actuel. D'autre part, le mouvement du modèle 3D, représentant la marche humaine, est un processus markovien de premier ordre. En effet, la configuration actuelle de la posture ne dépend que de sa configuration précédente ; la fluidité dans le mouvement humain et les contraintes biomécaniques justifient cette propriété. La modélisation de la dynamique dans le temps, du vecteur regroupant les paramètres de configuration du modèle 3D humanoïde, peut être réalisée par une représentation discrète d'état. Dans le cadre d'une représentation d'état, ce vecteur sera désormais appelé le vecteur d'état. Le problème consiste alors à estimer le vecteur d'état du système dynamique (le modèle 3D humanoïde en mouvement) au vu de l'observation (image) et du vecteur d'état du système estimé à l'instant précédent. D'une manière plus générale, nous cherchons à estimer, pour chaque observation, la valeur du vecteur d'état d'un processus aléatoire Markovien de

premier ordre. Dans ce qui suit, nous allons faire un petit rappel sur les méthodes permettant de réaliser ce type d'estimation dite dynamique.

3.4 La théorie de l'estimation dynamique [6]

Les bases de l'estimation dynamique reposent sur la construction de la probabilité conditionnelle à l'information disponible. D'un point de vue global, l'estimation dynamique consiste à évaluer à un instant $\boldsymbol{t}$, d'une manière récursive, un processus représenté par son vecteur d'état $\boldsymbol{X_t}$ à partir d'une observation $\boldsymbol{Y_t}$. Le caractère optimal de l'estimateur de $\boldsymbol{X}$ dépend du critère à optimiser (minimum de la variance d'état par exemple).

Nous rappelons ici le problème de filtrage à temps discret de processus markoviens. Soit $\boldsymbol{X}$ un processus stochastique Markovien à valeurs dans $\mathbb{R}^m$ ($m \geq 1$), défini sur un espace probabilisé par :

(3.1)
$$\begin{cases} \boldsymbol{X_t = f_t\,(X_{t-1}, \pi_t)} \quad \boldsymbol{(t \geq 1)} \\ \boldsymbol{X_0} : \text{Condition initiale} \end{cases}$$

où $\boldsymbol{f_t}$ désigne une fonction de transition mesurable de $\mathbb{R}^m * \mathbb{R}^n$ dans $\mathbb{R}^m$ et $\boldsymbol{\pi_t}$ un bruit (de dynamique) de loi connue à valeurs dans $\mathbb{R}^n$, et dont les réalisations temporelles sont indépendantes.

Le processus $\boldsymbol{X}$ est observé au travers d'un processus $\boldsymbol{Y}$ défini dans le même espace probabilisé, à valeurs dans $\mathbb{R}^p$ de la forme suivante :

(3.2)
$$\boldsymbol{Y_t = h_t\,(X_t, v_t)} \quad \boldsymbol{(t \geq 1)}$$

où $\boldsymbol{h_t}$ est une fonction d'observation mesurable de $\mathbb{R}^m * \mathbb{R}^p$ dans $\mathbb{R}^p$ et $\boldsymbol{v_t}$ est un bruit de mesure de loi connue à valeurs dans $\mathbb{R}^p$ et dont les réalisations temporelles sont indépendantes.

Nous noterons par la suite $\underline{\boldsymbol{Y_t}} = \boldsymbol{(Y0, Y_1, \ldots, Y_{t-1}, Y_t)}$ l'ensemble des observation de l'instant 0 jusqu'à l'instant $\boldsymbol{t}$, et $\underline{\boldsymbol{X_t}} = \boldsymbol{(X_0, X_1, \ldots, X_{t-1}, X_t)}$ l'historique des vecteurs d'état à l'instant $\boldsymbol{t}$. Nous parlons de filtrage lorsqu'on cherche à estimer $\boldsymbol{X_t}$ à partir des mesures $\underline{\boldsymbol{Y_t}}$ $\forall$ $\boldsymbol{t}$. La prédiction et le

lissage sont des extensions du filtrage, où l'on cherche à l'instant τ, à estimer X_τ respectivement à partir de mesures $\underline{Y}_t$ antérieures ($t < \tau$) ou postérieures ($t > \tau$).

Une loi marginale d'un vecteur aléatoire étant la loi de tout sous - vecteur de ce vecteur, le problème d'estimation (au sens du minimum de variance de l'erreur d'estimation) consiste à déterminer la mesure de probabilité conditionnelle (dite encore *a posteriori*) aux observations $p(\underline{X}_t \,|\underline{Y}_t)$, ou de manière particulière sa marginale $p(X_t \,|\underline{Y}_t)$. L'estimateur optimal de X_t sachant $\underline{Y}_t$, noté $\hat{X}_t$ s'écrit :

(3.3) $$\hat{X}_t = E\,(X_t \,|\underline{Y}_t) = \int_{X_t} X_t \;\; p(X_t \,|\underline{Y}_t) = \int_{\underline{X}_t} X_t \;\; p(\underline{X}_t \,|\underline{Y}_t)$$

Le problème de filtrage étant un problème dynamique, il est important de mettre en évidence les relations récursives permettant de calculer à chaque instant cette mesure de probabilité. Cette opération peut être décomposée en deux parties :

- Une partie prédictive fondée sur les équations de la représentation d'état du système, afin d'engendrer *a priori* la mesure de probabilité à l'instant t, connaissant les $t-1$ dernières observations. La construction de cette probabilité *a priori* $p\,(X_t \,|\underline{Y}_{t-1})$ se fait grâce à l'équation de Chapman Kolmogorov :

  (3.4) $$p(X_t \,|\underline{Y}_{t-1}) = \int_{X_{t-1}} p(X_t \,|X_{t-1})\; p(X_{t-1} \,|\underline{Y}_{t-1})$$

  où $p(X_t \,|X_{t-1})$ représente la probabilité de transition de l'état X_{t-1} à X_t. Comme X est un processus markovien de premier ordre nous pouvons écrire $p(X_t \,|X_{t-1}, \underline{Y}_{t-1}) = p(X_t \,|X_{t-1})$ dans (3.4).

- Une partie corrective qui repose sur la règle de Bayes et sur la propriété Markovienne du processus X :

  (3.5) $$p(X_t|\underline{Y}_t) = \frac{p(Y_t|X_t)}{p(Y_t|\underline{Y}_{t-1})}\, p(X_t|\underline{Y}_{t-1})$$

  où $p(Y_t \,|\underline{Y}_{t-1})$ est une constante de normalisation $p(Y_t \,|\underline{Y}_{t-1}) = \int_{Xt} p(Y_t \,|X_t)\; p(X_t \,|\underline{Y}_{t-1})$ qui dépend de la fonction de vraisemblance $p(Y_t \,|X_t)$ définie à partir du modèle d'observation

(3.2) et de la loi statistique que suit $\boldsymbol{v_t}$. L'observation $\boldsymbol{Y_t}$ est utilisée dans l'étape de correction pour modifier la densité *a priori* $\boldsymbol{p(X_t \mid \underline{Y}_{t-1})}$ et obtenir la densité *a posteriori* $\boldsymbol{p(X_t \mid \underline{Y}_t)}$ de l'état à l'instant $\boldsymbol{t}$. $\boldsymbol{p(\underline{X}_t \mid \underline{Y}_t)}$ peut alors être écrite sous la forme :

(3.6)
$$p(\underline{X}_t \mid \underline{Y}_t) = \frac{\prod_{\tau=0}^{t} p(Y_\tau \mid X_\tau) \prod_{\tau=1}^{t} p(X_\tau \mid X_{\tau-1}) p(X_0)}{\int_{\underline{X}_t} \prod_{\tau=0}^{t} p(Y_\tau \mid X_\tau) \prod_{\tau=1}^{t} p(X_\tau \mid X_{\tau-1}) p(X_0)}$$

En résumé, les processus markoviens mesurés sont entièrement définis par une loi de transition $\boldsymbol{p(X_t \mid X_{t-1})}$ et une loi d'observation $\boldsymbol{p(Y_t \mid X_t)}$. Les équations (3.4), (3.5) et (3.6) forment la base de la solution bayésienne optimale.

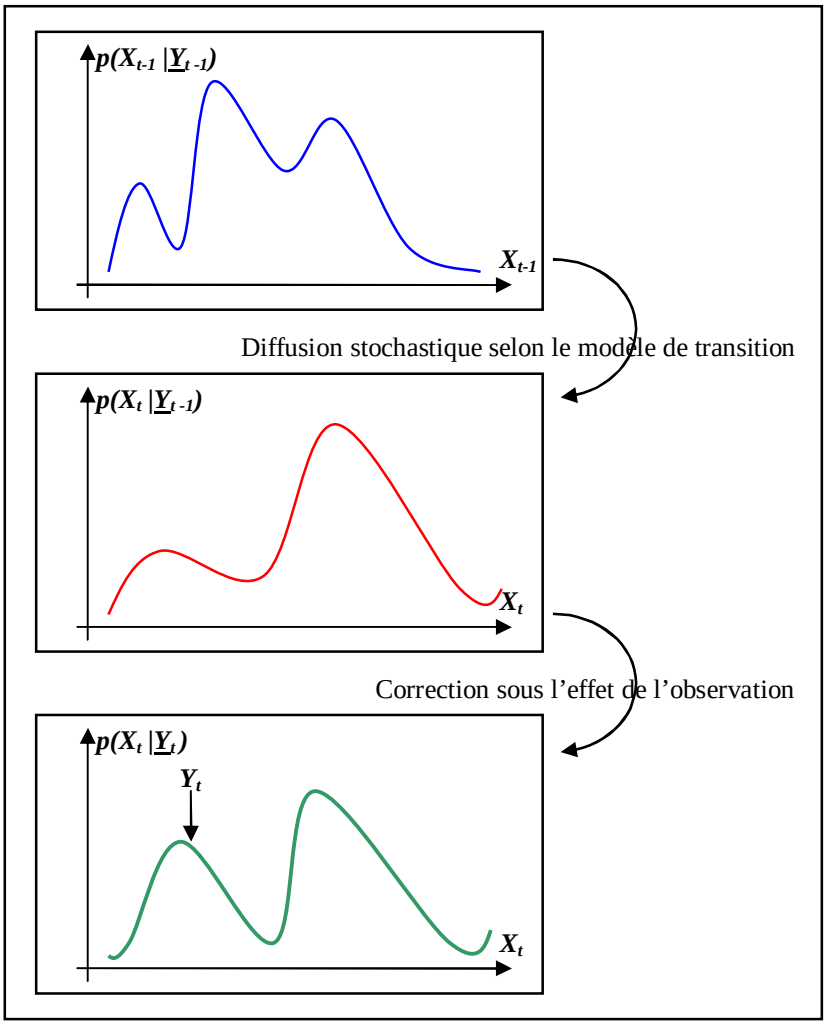


fig. 3.4 : Propagation dynamique de la densité de probabilité conditionnelle (*a posteriori*) d'un processus $\boldsymbol{X}$ sur un pas de temps $\boldsymbol{(t-1, t)}$. Cette propagation se fait en deux étapes : Une diffusion stochastique accomplie à travers le modèle de transition du processus (éq. 3.1) ; une étape corrective conditionnée par l'observation à l'instant $\boldsymbol{t}$, basée sur la formule de Bayes (éq. 3.2).

Cette propagation récursive de la densité de probabilité conditionnelle est une solution théorique qui ne peut pas être déterminée d'une manière analytique en général. Cependant, des solutions existent dans des cas particuliers, comme le filtre de Kalman et l'approche par maillage de l'espace d'état que nous allons exposer dans le paragraphe suivant. Dans le cas où la résolution analytique de ces équations est impossible, des solutions optimales approximatives basées sur des méthodes de Monte Carlo peuvent être utilisées (Algorithme de Condensation ou Filtrage Particulaire).

3.5 Filtres Bayésiens optimaux

3.5.1 Filtre de Kalman

Le but d'un estimateur bayésien dynamique est de calculer (de manière approchée), pour chaque instant d'observation, la fonction de densité de probabilité *a posteriori* $p(X_t | \underline{Y}_t)$ (i.e. la distribution de probabilité conditionnelle) de l'état à l'instant courant, sachant l'historique des observations. La suite des mesures de probabilité ainsi formée est dite filtre optimal. Elle évolue dans le temps par une alternance de prédictions et d'estimations (ou corrections). A chaque instant d'observation, la prédiction consiste à faire évoluer la densité de probabilité $p(X_t | \underline{Y}_t)$ à l'aide d'un modèle d'évolution (ou noyau de transition de l'état) entre les instants t et $t+1$. L'étape d'estimation met à jour la probabilité ainsi prédite en fonction de l'observation $\underline{Y}_{t+1}$, par simple application de la formule de Bayes.

En filtrage, nous cherchons un algorithme permettant d'effectuer le calcul du filtre optimal qui satisfasse en outre des contraintes de temps réel. Concrètement, le temps de calcul d'une estimée doit être plus petit que la période d'observation. Pour satisfaire cette contrainte (et accessoirement économiser la mémoire de l'ordinateur), un algorithme récursif est recherché de préférence. Ainsi, le calcul de l'estimée à l'instant courant ne doit dépendre que de l'observation courante et de l'estimée précédente : les observations passées ne sont pas réutilisées. Ces contraintes sont primordiales pour une bonne efficacité de l'algorithme.

En 1961, Kalman et Bucy [7,8] révolutionnent la théorie de l'estimation, en fournissant le premier algorithme de filtrage récursif. Le filtre de Kalman-Bucy permet en effet le calcul exact et rapide d'un filtre optimal lorsque les modèles d'état et d'observation ne font intervenir que des fonctions linéaires et des bruits additifs gaussiens. Kalman suppose que, dans ce cas, la loi conditionnelle recherchée est gaussienne, monomodale. Il suffit alors de propager sa moyenne et

sa variance et par conséquent, le problème est alors de dimension finie. En effet si $p(X_{t-1} | \underline{Y}_{t-1})$ suit une loi gaussienne, il a montré que $p(X_t | \underline{Y}_t)$ l'est aussi à condition que :

- Le bruit de mesure v_t et le bruit du modèle π_t suivent une loi gaussienne à paramètres connus.
- La fonction de transition $f_t (X_{t-1}, \pi_t)$ soit une fonction linéaire connue de X_{t-1} et π_t.
- La fonction d'observation $h_t (X_t, v_t)$ soit une fonction linéaire connue de X_t et v_t.

Les équations (3.1) et (3.2) peuvent alors être écrites de la façon suivante :

(3.7) $$X_t = F_t * X_{t-1} + \pi_t \quad (t \geq 1)$$

(3.8) $$Y_t = H_t * X_t + v_t \quad (t \geq 1)$$

où F_t et H_t sont des matrices connues représentant les fonctions linéaires f et h à l'instant t.

On appelle Q_t et R_t les matrices de covariance des vecteurs π_t et v_t respectivement. De plus, π_t et v_t sont considérés comme des bruits blancs gaussiens (moyenne nulle) indépendants. Les matrices F_t , H_t , Q_t et R_t sont variantes en fonction du temps. L'algorithme du filtre de Kalman dérivé à partir des équations (3.4) et (3.5) peut être alors présenté d'une manière récursive :

(3.9) $$p(X_{t-1} | \underline{Y}_{t-1}) = N (X_{t-1} ; m_{t-1|t-1} , P_{t-1|t-1})$$

(3.10) $$p(X_t | \underline{Y}_{t-1}) = N (X_t ; m_{t|t-1} , P_{t|t-1})$$

(3.11) $$p(X_t | \underline{Y}_t) = N (X_t ; m_{t|t} , P_{t|t})$$

(3.12) $$m_{t|t-1} = F_t * m_{t-1|t-1}$$

(3.13) $$P_{t|t-1} = Q_t + F_t * P_{t-1|t-1} * F_t^T$$

(3.14) $$m_{t|t} = m_{t|t-1} + K_t (Y_t - H_t * m_{t|t-1})$$

(3.15) $$P_{t|t} = P_{t|t-1} - K_t * H_t * P_{t|t-1}$$

où $N(x ; m, P)$ est une densité gaussienne avec un argument x, un vecteur moyenne m et une matrice covariance P et où K_t est dit la matrice gain de Kalman :

(3.16) $$S_t = H_t * P_{t|t-1} * H_t^T + R_t$$

(3.17) $$K_t = P_{t|t-1} * H_t^T * S_t^{-1}$$

Dans le cas général, le problème du filtrage non linéaire n'admet pas de solution de dimension finie. En pratique, des versions dérivées du filtre de Kalman-Bucy comme le filtre de Kalman étendu, qui linéarise le modèle autour de l'estimée courante, sont souvent utilisées. Bien

que cet algorithme soit, depuis une trentaine d'années, le filtre non linéaire le plus utilisé, il présente certains points faibles lorsque le système est non linéaire ou lorsque le filtre est mal initialisé.

3.5.2 Approche par maillage de l'espace d'état [9]

Cette méthode fournit la solution récursive optimale pour filtrer la densité conditionnelle *a posteriori* $p(X_t \mid \underline{Y}_t)$ si l'espace d'état est discret et est composé par un nombre fini d'états. Supposons que l'espace d'état à l'instant $t-1$ soit formé par les états discret X^i_{t-1}, $i=1..N_s$. Pour tout état X^i_{t-1} , $w^i_{t-1|t-1}$ est noté comme étant la probabilité conditionnelle de l'état étant donné l'historique des observations à $t-1$, $p(X_{t-1}= X^i_{t-1}| \underline{Y}_{t-1}) = w^i_{t-1|t-1}$. La probabilité conditionnelle *a posteriori* sera alors représentée par :

(3.18) $$p(X_{t-1}| \underline{Y}_{t-1}) = \sum_{i=1}^{N_s} w^i_{t-1|t-1}\, \delta(X_{t-1} - X^i_{t-1})$$

où $\delta(.)$ est une fonction de dirac. En remplaçant (3.18) dans les équations de prédiction (3.4) et de correction (3.5), ces dernières seront de la forme :

(3.19) $$p(X_t| \underline{Y}_{t-1}) = \sum_{i=1}^{N_s} w^i_{t|t-1}\, \delta(X_t - X^i_t)$$

(3.20) $$p(X_t| \underline{Y}_t) = \sum_{i=1}^{N_s} w^i_{t|t}\, \delta(X_t - X^i_t)$$

avec

(3.21) $$w^i_{t|t-1} = \sum_{j=1}^{N_s} w^j_{t-1|t-1}\, p(X^i_t \mid X^j_{t-1})$$

(3.22) $$w^i_{t|t} \cong \frac{w^i_{t|t-1}\, p(Y_t \mid X^i_t)}{\sum_{j=1}^{N_s} w^j_{t|t-1}\, p(Y_t \mid X^j_t)}$$

Cette solution est optimale dans le cas où $p(X^i_t \mid X^j_{t-1})$ et $p(Y_t| X^i_t)$ sont des mesures connues et où l'espace d'état est discret (formé par un ensemble fini d'états). Or souvent, nous avons affaire à des espaces continus qui nécessitent un très grand nombre d'états afin d'être discrétisés de manière efficace et précise. Cette approche reste alors difficile.

3.5.3 Conclusion

Le filtre de Kalman représente une solution optimale du problème d'estimation, dans le cadre strict où la densité de probabilité conditionnelle est gaussienne. D'autres approches plus

précises proposent de calculer la solution des équations du filtrage par des techniques de maillage de l'espace l'état. Cependant la complexité de ces méthodes numériques, augmentant fortement avec la dimension de l'espace, les rend inutilisables à partir d'une certaine dimension de l'espace d'état. Plus récemment, les méthodes de Monte Carlo (ou particulaires) ont été proposées comme une alternative attrayante [10,11,12]. Connues pour leur quasi-insensibilité à la dimension de l'espace et aux non linéarités du système, leur utilisation semble particulièrement adaptée au filtrage non linéaire. Les méthodes particulaires proposent d'approcher la loi conditionnelle recherchée par une mesure empirique changeant les poids d'un système de particules aléatoires évoluant avec le modèle et les observations.

3.6 Approches particulaires pour l'estimation dynamique

Dans le cadre d'une estimation dynamique et d'une observation additive gaussienne, le problème principal pour déterminer la loi de probabilité conditionnelle $\boldsymbol{P(\underline{X}_t \mid \underline{Y}_t)}$, réside dans la connaissance de la probabilité de transition $\boldsymbol{P(X_t \mid X_{t-1})}$. Dans le cas où le modèle d'évolution est non linéaire, nous ne savons généralement pas évaluer analytiquement cette probabilité, même si $\boldsymbol{P(X_0)}$ est initialement gaussienne. En effet il n'existe pas de solution optimale en dimension finie de ce problème d'estimation, mais uniquement des approximations convergentes, contrairement au cas linéaire gaussien par exemple (filtrage de Kalman). La méthode particulaire permet de réaliser une approximation convergente de la solution, par un nombre fini $\boldsymbol{N}$ de particules qui permettent une discrétisation de l'espace d'état. L'essentiel de l'approche particulaire réside dans une adaptation dynamique de la loi des grands nombres.

3.6.1 Principe et procédure de base (dite *'a priori'*) [6]

L'utilisation des tirages de Monte-Carlo, pour approcher des mesures de probabilité au sens faible, introduit de manière naturelle des mesures de Dirac. Soit $\boldsymbol{X}$ une variable aléatoire à valeurs dans $\mathbb{R}^m$ distribuée selon la densité de probabilité $\boldsymbol{p(X)}$, et $\boldsymbol{(X^i)_{i=1..N}}$ $\boldsymbol{N}$ variables aléatoires de même loi que $\boldsymbol{X}$. Pour toute fonction $\boldsymbol{\varphi(X)}$:

(3.23) $$\bar{\varphi} = \int \varphi(X) p(X) \qquad \int \left| \varphi(X) - \bar{\varphi} \right|^2 p(X) < \infty$$

La loi des grands nombres assure :

(3.24) $$\bar{\varphi}^N = \frac{1}{N} \sum_{i=1}^{N} \varphi(X^i) \xrightarrow{N \to +\infty} \bar{\varphi}$$

Donc, l'approximation $\boldsymbol{p^N}$ au sens faible de la mesure de probabilité $\boldsymbol{p}$ s'écrit :

(3.25) $$\bar{\varphi}^N = \int \varphi(X) p^N(X) \quad \Leftrightarrow \quad p^N = \frac{1}{N}\sum_{i=1}^{N} \delta_{X^i}(X)$$

où δ_{X^i} est une mesure de Dirac en $\boldsymbol{X^i}$. Cette opération revient donc à approcher la densité de probabilité par un peigne de Dirac équipondéré. Le principe de la méthode particulaire repose sur ce principe. Nous cherchons donc une approximation de $\boldsymbol{p(X_t|\, \underline{Y_t})}$ par $\boldsymbol{N}$ mesures de Dirac. Dans le cadre particulaire, toute variable $\boldsymbol{X^i}$ sera appelée 'particule'. Il découle de la loi des grands nombres que l'on construit récursivement, le peigne de Dirac pondéré à partir de la dynamique du système (éq. 3.4). L'évolution des particules est donnée par la loi de transition (flot du système) ; ces particules sont ensuite pondérées via la loi d'observation.

L'approche particulaire '*a priori*' peut être décomposée de la manière suivante :

1. **Initialisation :** Nous initialisons les particules $\boldsymbol{X^i_0}$ selon la loi *a priori* $\boldsymbol{p(X_0)}$, et les poids $\boldsymbol{p^i_0}$ à ***1/N.***

2. **Exploration :** A l'instant $\boldsymbol{t}$, chaque particule $\boldsymbol{(X^i_{t-1})_{i=1..N}}$ de l'espace d'état, évolue selon le flot stochastique du système en tirant $\boldsymbol{N}$ variables indépendantes $\boldsymbol{\pi^i_t}$ de bruit suivant la loi *a priori* (connue) $\boldsymbol{p(\pi_t)}$:

   (3.26) $$\boldsymbol{X^i_t = f_t\,(\,X^i_{t-1},\, \pi_t\,)}$$

3. **Pondération :** Nous construisons, de manière itérative, la pondération normalisée $\boldsymbol{p^i_t}$ associée à chaque particule, à partir de la règle de Bayes :

   (3.27) $$p_t^i = \frac{p(Y_t|X_t^i)}{\sum_{j=1}^{N} p(Y_t|X_t^j)}\, p_{t-1}^i$$

4. **Estimation :** L'estimateur particulaire s'écrit alors :

   (3.28) $$\hat{X}_t^N = \sum_{i=1}^{N} p_t^i X_t^i$$

Pour un traitement dynamique, nous itérons à chaque instant les parties 2, 3 et 4. L'algorithme particulaire '*a priori*' se décompose en deux parties : une partie évolution *a priori* de chaque particule, reposant sur la connaissance de la loi $\boldsymbol{p(X_t|\, X_{t-1})}$, et une autre qui associe, au

regard de l'observation, une pondération bayésienne à chaque particule. Une forte pondération associée correspond alors à une particule très vraisemblable.

Les tirages de Monte Carlo proposent une prise en compte globale des modèles non linéaires dont la convergence n'est limitée que par le nombre de particules utilisées. La procédure de base décrite ci dessus n'est satisfaisante qu'en temps fini. En effet, son principal problème vient du comportement asymptotiquement dégénérescent des poids des particules, qui est représentatif de la probabilité des trajectoires individuelles. Lorsque le temps croît, les différences relatives entre les poids des particules augmentent jusqu'à ce que la masse totale de probabilité soit concentrée sur une seule particule. Le comportement est alors loin de l'optimalité attendue, puisque les ***N-1*** particules restantes demeurent inutilisées. Ce problème se résout généralement en procédant à une régularisation du poids des particules. D'autre part, l'algorithme de base n'est pas adapté aux processus à variance non bornée. Dans ce cas, les particules suivent le flot divergent du processus *a priori* et la majorité d'entre elles acquiert un poids négligeable. Cette difficulté peut se résoudre par l'introduction d'une évolution conditionnelle (aux mesures) des particules et non plus *a priori*, ce qui contraint les particules à rester au voisinage du processus $\boldsymbol{X}$. Plusieurs algorithmes [12,13,14] tentent de contourner cette difficulté en apportant des modifications à cette méthode de base. Dans ce qui suit nous allons exposer l'un de ces algorithmes dit l'algorithme de Condensation qui a été souvent utilisé, soit sous sa forme initiale soit sous une forme modifiée, pour effectuer le suivi 3D du mouvement humain à partir de modèles humanoïdes.

3.6.2 Algorithme de Condensation [12]

L'algorithme de Condensation a été conçu pour effectuer la détection d'objets dans une image donnée. Cet algorithme est capable de gérer les densités de probabilité non gaussiennes et multimodales. D'autre part il permet de modéliser l'incertitude en transmettant les configurations peu probables à un instant $\boldsymbol{t}$ vers les étapes futures, ce qui leur donne une chance d'être choisies ultérieurement. Dans le cadre du filtrage particulaire, chaque configuration probable $(X^i_t)_{i=1..N}$ du vecteur d'état $\boldsymbol{X}$ à un instant $\boldsymbol{t}$ est représentée par une 'particule'. A chaque pas de temps, un 'poids' w^i_t représentant la vraisemblance de la configuration X^i_t en vue de l'observation Y_t, est calculé pour toute particule. La densité *a posteriori* $p(X_t|\ \underline{Y_t})$ est alors représentée d'une manière approximative par un ensemble de $\boldsymbol{N}$ particules X^i_t 'pondérées' (fig. 3.5).

L'étape d'initialisation de l'algorithme consiste à choisir $\boldsymbol{N}$ particules suivant une loi uniforme avec des poids équivalents à $\boldsymbol{1/N}$.

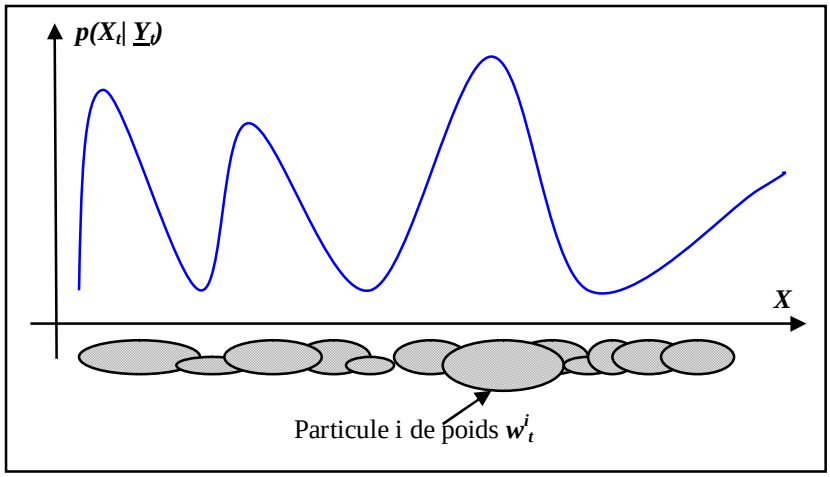


fig. 3.5 : Représentation de la probabilité conditionnelle *a posteriori* $p(X_t|\underline{Y_t})$ à l'instant t par l'ensemble des particules 'pondérées', dans le cas où le processus à estimer X est monodimensionnel. Chaque particule i est représentée par une ellipse dont la surface est proportionnelle à son 'poids' w^i_t.

Les trois étapes essentielles de l'algorithme qui sont itérées à chaque pas de temps seront alors :

- **Sélection (ou 'ré échantillonnage') :** En utilisant une méthode de tirage de Monte Carlo, N particules sont choisies parmi l'ensemble des N particules X^i_{t-1} créées à $t-1$. Ainsi toute particule pourrait être choisie une ou plusieurs fois ou pas retenue suivant sa vraisemblance à l'instant $t-1$. Ce choix se fait suivant la densité de probabilité *a priori* $p(X_t|\underline{Y_{t-1}})$. Cette probabilité est en général multimodale et ne peut pas être représentée analytiquement mais sera construite à partir des particules et leurs poids à $t-1$. En effet, la probabilité de choisir une particule est proportionnelle à son poids normalisé établi à $t-1$. En conséquence, les particules ayant un poids plus grand que les autres seront plus nombreuses dans l'ensemble des particules ainsi créé à la fin de cette étape. Dans le cas où le nombre N des particules est petit (relativement par rapport à la dimension du vecteur d'état) et où une particule a un poids normalisé très grand par rapport aux autres, nous risquons de choisir cette particule plusieurs fois et d'empêcher les autres particules (configurations) de survivre. La convergence de l'algorithme est donc garantie par l'utilisation d'un nombre important de particules.

- **Prédiction (ou diffusion) :** Toute particule i (parmi les N particules choisies à l'étape de sélection) évolue suivant le modèle d'évolution (loi de transition) du processus ; $X^i_t = F_t * X^i_{t-1} + \pi_t$ dans le cas d'un modèle linéaire avec un bruit blanc gaussien π_t. Dans un cadre plus général $X^i_t = f_t (X^i_{t-1}, \pi_t)$.

- **Mesure :** Etant donné l'observation Y_t (l'image à l'instant t), le poids de chaque particule X^i_t est calculé. Les poids w^i_t seront ensuite normalisés et l'ensemble de particules pondérées sera

transmis pour être utilisé à l'instant suivant ***t+1***. Le vecteur d'état $\boldsymbol{X_t}$ est estimé comme étant représenté par la particule la plus 'lourde', ou par la moyenne de l'ensemble des ***N*** particules pondérées. La densité de probabilité *a posteriori* $\boldsymbol{p(X_t|\ \underline{Y_t})}$ sera alors représentée par cet ensemble de particules pondérées.

Un nombre minimum de particules est nécessaire pour représenter la densité *a posteriori* d'une manière fidèle et pour avoir une estimation fiable du vecteur d'état. Dans un espace à grande dimension, ce nombre de particules devient important et augmente la complexité de l'algorithme. En effet le nombre de particules nécessaires pour avoir une même performance augmente avec l'augmentation de la dimension du vecteur d'état. MacCormick et al. [15] citent deux paramètres pour décrire l'efficacité d'un filtre de type particulaire : le 'survival diagnostic' *D* et le 'survival rate' α. *D* indique le nombre de particules qui vont survivre à l'étape de 'Sélection' et prend des valeurs comprises entre 1 (dans le cas où une particule à un poids normalisé = 1) et ***N*** (dans le cas où les ***N*** particules ont des poids normalisés égaux à 1/N). Avec un petit 'survival diagnostic' l'algorithme peut diverger de la bonne estimation recherchée. Le 'survival rate' α (α <<1) multiplié par le nombre ***N*** de particules est approximativement égal au nombre des particules survivant à la 'Sélection'. Le paramètre α donne une indication sur la qualité de l'estimation ; quand α est petit, l'estimation est souvent peu précise. Le nombre minimal de particules $\boldsymbol{N_{min}}$ nécessaires pour estimer un vecteur d'état de dimension *d* avec un *D* minimal et un α donné est [15] :

(3.29) $$N_{\min} \geq \frac{D_{\min}}{\alpha^{d}}$$

Lorsqu'on veut utiliser l'algorithme de Condensation pour estimer un vecteur d'état de grande dimension, le nombre minimal de particules devient alors très important. D'autre part, l'utilisation d'un grand nombre de particules augmente la complexité temporelle de l'algorithme de base et le rend pratiquement inapplicable, vu le fait que le poids de chaque particule est calculé à chaque pas de temps. Notre idée est alors de modifier cet algorithme et de l'adapter pour pouvoir l'utiliser dans l'estimation du mouvement humain tout en gardant une complexité temporelle acceptable.

3.6.3 Application du filtrage particulaire pour le suivi du mouvement humain

L'algorithme de condensation a souvent été utilisé pour le suivi 3D du mouvement humain. Sidenbladh et al. [16] décrivent un modèle 3D humanoïde et un modèle d'apparence et appliquent l'algorithme de condensation avec 10^4 particules (5 minutes de traitement et de calcul

par image) pour trouver la configuration du modèle 3D (les angles et les vitesses angulaires aux articulations), qui convient le mieux à l'image 2D. L'étape de prédiction du vecteur d'état, dans le cadre du filtrage particulaire, se fait suivant un modèle temporel qui était déjà établi à partir d'un ensemble d'apprentissage. Or, nous ne pouvons pas inclure tous les mouvements possibles dans un même modèle d'où le besoin d'établir un modèle par catégorie de mouvement (marche, course etc..). D'autre part pour une catégorie donnée nous ne pouvons pas inclure toutes les configurations possibles des angles. Par conséquent, toutes les particules produites dans l'étape de prédiction seront forcées de suivre le modèle appris avec le risque de perdre la configuration réelle qu'on veut estimer. Comme nous cherchons à trouver un algorithme générique pour le suivi de tous les types de mouvement, l'utilisation d'un modèle de mouvement est exclue.

Cohen et al. [17] proposent une méthode basée sur le filtrage particulaire et les Support Vector Machine (SVM), pour effectuer le suivi 3D du mouvement humain. En extrayant la silhouette de la personne suivie, dans les vidéos prises simultanément avec plusieurs vues sur la scène, les axes principaux du corps humain (torse, tête, bras etc..) sont étiquetés dans l'image à l'aide d'une analyse en composantes principales (ACP). Par exemple, en appliquant une ACP sur les points médians de la silhouette, le premier axe principal est considéré comme étant la colonne vertébrale du corps. L'algorithme de Condensation a été ensuite utilisé pour estimer la configuration des axes d'un modèle 3D en squelette. Dans ce cadre, la somme des distances orthogonales entre tout axe du modèle virtuel et l'axe étiqueté dans l'image qui lui correspond est considérée comme une fonction de vraisemblance. La configuration des axes une fois trouvée, une classification est appliquée pour trouver la posture d'un modèle 3D humanoïde qui correspond au squelette estimé. Cette classification implique un processus d'apprentissage supervisé à l'aide de SVM. Or, l'étiquetage des axes est une opération complexe et ne peut pas être exacte. D'autre part, l'opération de classification qui se fonde sur un ensemble d'apprentissage, rend la méthode restrictive. En effet, pour tout mouvement à détecter il faut construire un ensemble d'apprentissage. Ne pouvant pas cerner tous les mouvements humains possibles, le nombre des mouvements détectés est alors réduit. D'autre part, cette méthode permet de reconnaître les mouvements humains effectués dans les images mais ne fournit pas les positions 3D exactes des différentes articulations.

Pour réduire le nombre de particules utilisées dans l'algorithme de Condensation, Deutscher et al. [18] y intègrent une technique de recuit simulé. Entre deux pas de temps (deux images consécutives), une recherche multi –couches est appliquée dans le but de retrouver le maximum global de la fonction de vraisemblance et éviter de tomber dans les maxima locaux. En fait, toute

particule évolue plusieurs fois par pas de temps, par ajout d'une perturbation, pour atteindre le maximum de la fonction de vraisemblance. Le résultat de l'algorithme n'est plus alors un ensemble de particules représentant la probabilité *a posteriori*, mais plutôt un ensemble de particules concentrées autour d'un maximum global de la fonction de vraisemblance. A l'aide d'un modèle 3D virtuel du corps et d'une fonction de vraisemblance combinant la silhouette et le contour du corps dans la vidéo, l'algorithme modifié est appliqué pour le suivi complet du mouvement humain. Cette approche fournit de bons résultats mais privilégie l'exploitation du voisinage de la particule la plus 'lourde' à l'exploration de l'espace de recherche d'état (exploration des maxima locaux de la fonction de vraisemblance). En effet, la recherche multi couches à un instant ***t*** fournit un ensemble de particules peuplant le voisinage (multidimensionnel) du maximum global de la fonction de vraisemblance à cet instant. Or, cette fonction, qui est multimodale, reste sensible aux bruits (ombres, occlusions etc..). Par conséquent, la particule fournissant le maximum global de la fonction d'observation à un instant ***t*** donné, n'est pas forcément issue (par diffusion stochastique) du voisinage de celle fournissant le maximum global à l'instant ***t-1***, si l'observation était erronée à cet instant. Or, toutes les particules crées à ***t*** sont issues de ce voisinage qui ne contient plus le maximum global d'où la divergence de cet algorithme. L'utilisation de cet algorithme a été alors exclue.

3.7 Conclusion

Le filtrage particulaire fournit la solution la plus générale pour estimer un vecteur d'état compte tenu d'une observation. En effet cette technique permet de gérer les modèles non linéaires d'évolution et les densités de probabilité multimodales et non gaussiennes. D'autre part, l'algorithme de Condensation de base est pratiquement inapplicable dans le cas des vecteurs d'état à grande dimension. Les versions modifiées de cet algorithme semblent également inadaptées à notre approche. Dans le chapitre suivant nous allons exposer l'approche que nous avons proposée pour résoudre le problème du suivi 3D du mouvement humain, en utilisant un modèle virtuel du corps humain, une fonction de vraisemblance et une version modifiée et adaptée de l'algorithme de Condensation.

Bibliographie

[1] Armangué X et Salvi J, "Overall view regarding fundamental matrix estimation", Journal of Image and Vision Computing, Vol. 21, No. 2, pp. 205-220, 2003.

[2] Haritaoglu I, Harwood D et Davis L, "w 4s : A real-time system for detecting and tracking people in 2.5D", Proceedings of the 5th European Conf. Computer Vision 1998, vol. 1, pp. 877-892, 1998.

[3] Bregler C et Malik J, "Tracking people with twists and exponential maps", Proceedings of the Conference on Computer Vision and Pattern Recognition 1998.

[4] Gavrila D et Davis L, "3D model based tracking of humans in action a multi view approach", Proceedings of the Conference on Computer Vision and Pattern Recognition 1996, pp. 73-80.

[5] Howe N.R, Leventon M.E et Freeman W.T, "Bayesian reconstruction of 3D human motion from single camera video", Advances in Neural Information Processing Systems MIT Press 2000, vol. 12, pp. 820-826.

[6] Noyer J.C, ''Traitement non linéaire du signal radar par filtrage particulaire'', Mémoire de thèse de doctorat, Université Paul Sabatier Toulouse III, 1996.

[7] Kalman R.E, "A new approach to linear filtering and prediction problems", Transactions ASME, Journal of Basic Engineering, vol. 82, pp. 34-45, 1960.

[8] Kalman R.E. et Bucy R, "A new approach to linear filtering and prediction theory", Transactions ASME, Journal of Basic Engineering, vol. 83, pp.95-108, 1961.

[9] Arulampalam S, Maskell S, Gordon N. et Clapp T, "A tutorial on Particle Filters for Online Non-linear/Non-Gaussian Bayesian Tracking", QinetiQ Ltd, 2001.

[10] Bergman N, "Recursive Bayesian Estimation : Navigation and Tracking Applications", PhD Thesis, Linkoping University Sweden, 1999.

[11] Gustafsson F, Bergman N et Forsell U, "Particles filters for positioning, Navigation and Tracking", IEEE Transactions on Signal Processing, Special issue on Monte Carlo methods for statistical signal processing, 2002.

[12] Isard M.A et Blake A, "CONDENSATION- conditional density propagation for visual tracking", International Journal of Computer Vision, vol 29(1), pp. 5-28, 1998.

[13] Gordon N.J, Salmond D.J. et Smith A.F.M, "A novel approach to nonlinear/non-Gaussian Bayesian state estimation", IEE Proceedings on Radar and Signal Processing, vol 140, pp. 107-113, 1993.

[14] Doucet A, Godsill S et Andrieu C, "On sequential Monte Carlo sampling methods for Bayesian filtering", Journal of statistics and computing, vol. 10, no. 3, pp. 197–208, 2000.

[15] MacCormick J et Blake A, “Partitioned sampling, articulated objects and interface quality hand tracking”, European Conference on Computer Vision 2000.

[16] Sidenbladh H, Black M.J et Fleet D.J, “Stochastic tracking of 3D human figures using 2D image motion”, Proceedings of 6th European Conference on Computer Vision, Dublin 2000.

[17] Cohen I et Mun Wai L, “3D Body Reconstruction for Immersive Interaction”, Proceedings of the Second international Workshop on Articulated Motion and Deformable Objects, Palma de Mallorca Spain, pp. 119-130, 2002.

[18] Deutscher J, Blake A et Reid I, “Articulated body motion capture by annealed particle filtering”, Proceedings of the Conference Computer Vision and Pattern Recognition, vol. 2, pp. 1144-1149, 2000.

Chapitre 4 : Approche proposée

Dans le chapitre précédent nous avons montré comment nous pouvons réduire le problème du suivi 3D du mouvement humain à un problème d'estimation dynamique d'un vecteur d'état à partir d'une observation donnée. Le vecteur d'état à estimer est la configuration d'un modèle 3D humanoïde virtuel, simulant le mouvement humain, dont nous souhaitons trouver la configuration maximisant la vraisemblance par rapport à la vraie posture du corps.

Dans ce chapitre nous détaillerons le modèle 3D virtuel humanoïde que nous utilisons. Il est formé par les points clés du corps dont nous voulons suivre le mouvement et par des segments les reliant. Ce modèle simulant le mouvement humain est doté d'un ensemble de degrés de liberté dont la configuration déterminera la posture. En plus du modèle virtuel, il faut développer une fonction évaluant sa vraisemblance à la vraie posture du corps représentée dans l'image.

Cette vraisemblance est calculée à partir de la comparaison des projections 2D du modèle (image synthétique) et du corps (image vidéo). Pour pouvoir comparer les deux postures réelle et virtuelle, il faut extraire ce qu'on appellera l'image 'silhouette' représentant la vrai posture dans le flux vidéo. La comparaison des images portera alors sur l'estimation de la surface de recouvrement entre les deux images synthétique et 'silhouette'.

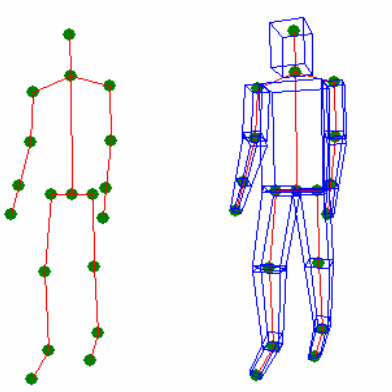
Un rappel sur la théorie de l'estimation bayésienne dynamique ainsi qu'une description des filtres – estimateurs optimaux (filtre de Kalman, maillage de l'espace d'état) ont été présentés. Or, ces filtres optimaux nécessitent des conditions strictes pour être appliqués, telle que la nécessité d'avoir une densité de probabilité conditionnelle gaussienne. Cette condition n'est pas vérifiée pour le problème que nous traitons, d'où la nécessité d'adopter des approches plus générales. La famille de méthodes dites de filtrage particulaire offre ce cadre plus global et permet d'estimer des vecteurs d'état dont la probabilité conditionnelle à l'observation est multi modale ou non gaussienne. Parmi ces méthodes, l'algorithme de Condensation qui a été déjà utilisé pour le suivi 3D complet du mouvement humain, a été détaillé. Cependant, cet algorithme présente un inconvénient majeur qui est sa complexité exponentielle par rapport à la dimension

du vecteur d'état. Dans le but de surmonter cette difficulté, nous avons modifié cet algorithme pour le rendre plus efficace avec un nombre de particules plus petit. L'algorithme modifié que nous avons développé et qui sera exposé est appelé 'Interval Particle Filtering' (IPF).

Voulant réduire la complexité de l'algorithme IPF et diminuer le nombre des particules utilisées, nous en présenterons une version factorisée en représentant le processus dynamique à l'aide d'un réseau bayésien dynamique. L'algorithme IPF est alors utilisé pour effectuer l'inférence dans ce réseau.

4.1 Le modèle 3D virtuel utilisé

Comme dans toutes les approches du suivi 3D du mouvement humain, nous utilisons un modèle 3D visuel articulé représentant le corps humain. Ce modèle est formé de 19 points ou articulations, représentant les points clés du corps dont nous voulons suivre le mouvement (tête, coudes, genoux, chevilles etc.). Ces 19 points seront ensuite reliés par 17 segments(fig. 4.1). Les longueurs de ces segments sont calculées suivant le modèle de l'homme de Vitruve de Leonard De Vinci.



(a) (b)

fig. 4.1 : Le modèle 3D virtuel articulé ; (a) Le modèle en fils de fer est composé de 19 points clés reliés par 17 segments formant ainsi 5 chaînes cinématiques. Chaque point est muni d'un certain nombre de degrés de liberté (rotations 3D). Du fait de la présence de certaines restrictions sur le mouvement des différentes parties du corps, le modèle sera représenté par un vecteur formé de 31 degrés de liberté. (b) Des volumes seront ensuite ajoutés au modèle initial en squelette pour le rendre « plus humain ».

Ce modèle en fils de fer doit représenter un squelette humain et simuler le mouvement des différents membres du corps. Par conséquent, chaque articulation est dotée d'un certain nombre de degrés de liberté (au maximum 3) représentant les angles de rotation d'Euler autour des axes ***(x, y, z)***. Grâce à des contraintes sur le mouvement humain, nous pouvons réduire le nombre des

degrés de liberté attribués au modèle squelette à 31 dimensions. Notre modèle en fils de fer sera alors configuré à partir d'un vecteur formé par 31 degrés de liberté. Une description détaillée du modèle et des degrés de liberté sera établie dans l'annexe B. Des parallélépipèdes, représentant les membres du corps, sont ensuite ajoutés à ce modèle pour le rendre plus proche de la réalité. Par ailleurs, des valeurs extrêmes, au delà desquelles tout mouvement sera impossible, sont imposées pour chaque degré de liberté. Ces contraintes peuvent être facilement modifiées selon la nature du mouvement qu'on veut détecter. Par exemple, pour une marche normale, l'angle de rotation aux hanches varie entre –30 degrés et 60 degrés. Vu la nature du corps humain, le modèle sera composé de chaînes cinématiques ; la partie du corps dont le mouvement entraîne le mouvement d'une autre partie formera une chaîne avec cette dernière.

4.2 La fonction de vraisemblance utilisée

En plus du développement d'un modèle 3D humanoïde, il faut développer une fonction pour évaluer le degré de ressemblance de la configuration établie à partir des valeurs des 31 degrés de liberté avec la vraie position du corps. Cette comparaison n'étant pas possible en 3D, nous allons comparer les projections 2D du modèle et du corps. Cette projection du corps est donnée à travers la vidéo. Pour que la comparaison soit valide il faut alors projeter le modèle 3D à l'aide d'une caméra virtuelle ayant les mêmes paramètres et la même orientation (par rapport au corps) que la caméra réelle.

Dans un contexte de filtrage particulaire, la fonction de ressemblance décrivant le degré de similitude entre une configuration du vecteur d'état et l'information fournie par l'observation (image), est dite 'poids'. Dans certaines approches [1] [2], des textures sont appliquées pour rendre les images du corps humain plus réalistes. Ces textures interviennent aussi dans l'évaluation du poids de chaque configuration. Malgré les avantages de l'utilisation des textures, la création de ce type d'images est spécifique pour chaque personne et dépend des conditions de l'acquisition vidéo (habits, éclairage, etc.), ce qui rend cette approche mal adaptée à notre problématique. La détection des contours et l'extraction de l'arrière plan dans les images, sont utilisées dans [3] pour produire une fonction 'poids' simple et générale. Cette approche fournit des résultats satisfaisants, d'où l'idée d'adopter les mêmes techniques. Ainsi, la fonction 'poids' que nous avons développée est basée sur une méthode simple d'extraction d'arrière plan. En effet une image 'silhouette', du corps dont nous voulons suivre le mouvement dans la vidéo est produite à partir de l'image réelle. Cette image 'silhouette' est ensuite comparée à une image synthétique (la projection 2D du modèle 3D du corps) représentant la configuration du modèle à

laquelle nous voulons assigner un 'poids' (fig. 4.2). Après soustraction de l'image synthétique à l'image 'silhouette', la fonction de ressemblance ***w*** sera calculée par :

(4.1)
$$w = \frac{N_c}{N_s + N_m}$$

où :

N_c = Nombre de pixels (non noirs) communs entre les deux images synthétique et 'silhouette' (les pixels en bleu).

N_s = Nombre de pixels (non noirs) formant l'image 'silhouette' mais n'appartenant pas à l'image synthétique (les pixels en blanc).

N_m= Nombre de pixels (non noirs) formant l'image synthétique du modèle mais n'appartenant pas à l'image 'silhouette' (les pixels en rouge).



(a) **(b)** **(c)** **(d)**

fig. 4.2 : Evaluation de la ressemblance ; l'image originale (a) est transformée en image 'silhouette' (b) par soustraction de l'arrière plan. (c) La configuration du modèle 3D est représentée par une image synthétique de sa projection en 2D. La comparaison de la configuration du modèle à l'image 'silhouette' est faite en (d) en soustrayant la première à la deuxième et par calcul du nombre N_c des pixels communs aux deux images, du nombre N_s de pixels de la silhouette n'étant pas dans l'image synthétique et du nombre N_m de pixels de l'image synthétique n'appartenant pas à l'image 'silhouette'.

Le choix de cette fonction est motivé par le fait qu'on cherche à trouver la configuration du modèle qui maximise la ressemblance de sa projection 2D à la vraie posture du corps dans l'image. Ceci peut être interprété par un nombre plus grand de pixels communs à la silhouette et l'image synthétique N_c, et un nombre plus petit de pixels n'appartenant pas à la zone de recouvrement (N_s+N_m). Dans le cas où nous utilisons plusieurs caméras, la fonction *w* sera donné par:

(4.2)
$$w = \frac{\sum_{i=1}^{c} w_i}{c}$$

où ***c*** est le nombre de caméras et w_i le poids assigné à partir de l'image fournie par la caméra ***i***.

Pour obtenir l'image 'silhouette', il suffit de soustraire, pixel par pixel, l'image courante contenant le corps en mouvement, à l'image d'arrière plan de référence dont nous disposons directement ou qu'on peut construire à fur et à mesure [4].

(4.3) $$M(x,y) = \left| I_t(x,y) - B(x,y) \right|$$

où ***M*** est l'image résultante, ***I*** est l'image courante à l'instant ***t*** et ***B*** est l'image de référence.

Cette méthode est simple, mais reste très sensible aux ombres et à la différence d'illumination, d'où la nécessité d'appliquer un certain nombre de filtres pour éliminer les défauts.

4.2.1 Filtre seuil

Si la différence (en valeur absolue) entre les intensités de deux pixels est plus petite qu'un seuil fixé, ces deux pixels seront considérés comme identiques et par la suite leur soustraction sera nulle. Si la différence est supérieure au seuil, les deux pixels seront considérés différents et nous donnerons une valeur fixée (1 = pixel blanc) à leur soustraction (fig. 4.3). La valeur du seuil doit être bien réglée de façon à ne pas inclure de l'ombre dans l'image (seuil très petit) ni couper une partie du corps (seuil très grand). Ce réglage est effectué manuellement suivant l'éclairage et la nature de la scène filmée.

(4.4) $$M_t(x,y) = \begin{cases} 0 \quad si \quad \left| I_t(x,y) - B(x,y) \right| < seuil \\ \\ 1 \quad si \quad \left| I_t(x,y) - B(x,y) \right| \geq seuil \end{cases}$$

4.2.2 Filtre Médian

Pour éliminer le bruit résultant des petites différences d'illumination dans l'image, nous appliquons un filtre médian à l'image produite après seuillage. Ce filtre aide à garder la netteté d'une image en supprimant tout pixel ayant une intensité différente de ses voisins. L'idée de ce filtre est de remplacer un point donné par la valeur médiane des intensités triées d'un certain nombre de ses voisins. La complexité de l'algorithme est exponentielle par rapport au nombre des pixels voisins impliqués, dit ordre du filtre. Pour cela, nous utiliserons un filtre médian d'ordre 3.



(a) (b) (c) (d)

fig. 4.3 : Construction de l'image 'silhouette' ; par soustraction de l'image de référence (b) à l'image courante (a) nous obtenons une image de différence (c) ; en appliquant un seuillage adapté et un filtre médian nous obtenons une image 'silhouette' nette (d).

4.2.3 Filtre HSV

Dans le cas de présence de fortes zones d'ombres, l'application des deux filtres présentés ci dessus s'avère inefficace. En revanche, Cucchiara et al. [5] proposent une méthode efficace pour l'élimination des ombres, en utilisant l'espace de couleur HSV (Hue, Saturation, Value ou Brightness), dit aussi TSV en français (Teinte Saturation, Valeur), à la place du système RGB classique. Cette méthode se fonde sur l'hypothèse que les ombres réduisent la Saturation et la Valeur d'une surface et maintiennent sa valeur de Teinte. En représentant l'image dans l'espace de couleur HSV et en appliquant quatre équations de seuillage, les ombres dans les images peuvent être réduites d'une façon efficace (fig. 4.4). Toutefois, le réglage de ce filtre, nécessitant la détermination des valeurs de quatre seuils différents, reste compliqué.



(a) (b)

fig. 4.4 : Exemple d'utilisation d'un filtre HSV ; (a) En utilisant un seuil d'une petite valeur pour la différenciation des images courante et de référence, nous pouvons introduire de l'ombre dans l'image 'silhouette' obtenue. (b) Par application d'un filtre de type HSV, nous remarquons que les ombres sont effacées.

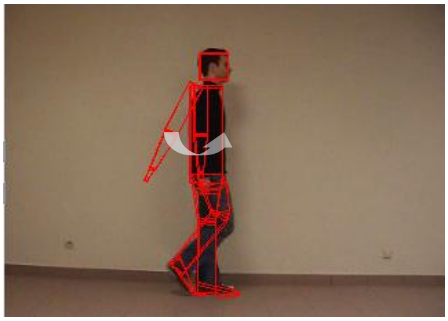
Pour obtenir l'image 'silhouette' nous effectuons la soustraction de l'arrière plan puis nous appliquons un filtre seuil suivi d'un filtre médian. Dans le cas de présence de grandes zones d'ombres, un filtre HSV doit être ensuite appliqué.

La fonction 'poids' qu'on a détaillée ci dessus est simple et générale. La sensibilité de cette fonction aux ombres peut être réduite en réglant les paramètres des filtres appliqués. Cette fonction servira à évaluer la vraisemblance d'une configuration donnée du modèle 3D vis a vis d'une image donnée. La configuration ayant le plus grand degré de vraisemblance, sera considérée comme étant l'estimation la plus probable de la vraie posture du corps suivi dans la vidéo.

4.3 Quel estimateur bayésien utiliser ?

Le choix de l'estimateur bayésien à utiliser dans notre approche dépend essentiellement de la nature de l'observation (fonction de vraisemblance) dont nous disposons et de la nature de la densité de probabilité conditionnelle *a posteriori*. La fonction de vraisemblance que nous utilisons pour attribuer un 'poids' (la probabilité $p(Y_t/X_t^i)$ de l'observation sachant un état donné X^i), à une configuration donnée du vecteur d'état, se fonde sur la comparaison d'une image en 2D représentant ce vecteur d'état à l'image vidéo réelle, constituant l'observation. Par conséquent nous ne pouvons pas construire un modèle d'observation pour notre système. D'autre part, l'image illustrant le vecteur d'état n'est qu'une projection en 2D (suivant la position de la caméra vidéo) d'une configuration 3D des positions des membres du corps. Plusieurs configurations différentes du vecteur d'état (présence d'occlusions par exemple) peuvent ainsi être représentées par la même image et d'ailleurs avoir le même 'poids', d'où la multi modalité de la fonction de vraisemblance et par conséquent de la probabilité conditionnelle *a posteriori*.

D'autre part les variables d'état sont dépendantes entre elles par rapport à l'observation dans des cas particuliers. En effet la probabilité *a posteriori* $p(x_i|Y)$ d'une variable d'état donnée x_i peut suivre une répartition gaussienne ou non gaussienne suivant la valeur d'une autre variable d'état x_j dans certains cas. Par exemple, prenons le cas particulier où l'on suit le mouvement de rotation du bras droit (fig. 4.5), le bras gauche étant aligné le long du tronc. Nous remarquons que la probabilité *a posteriori* de l'angle de rotation du bras droit suit une répartition gaussienne si nous observons le mouvement à partir d'images prises de profil. Dans le cas où le bras gauche (de notre modèle) est levé vers le haut (fig. 4.6) nous remarquons que cette probabilité n'est plus gaussienne mais devient bi-modale.



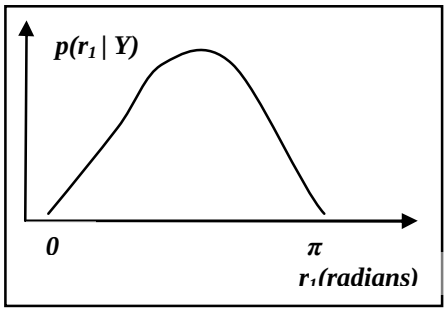
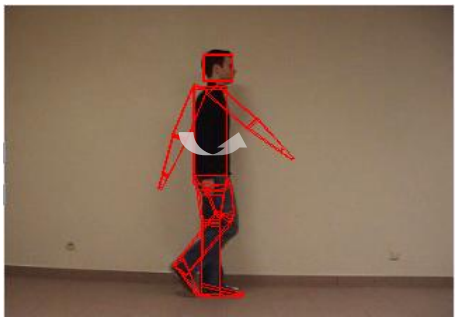


fig. 4.5 : Approximation de la densité de probabilité *a posteriori* de l'angle ***r1*** de rotation du bras droit dans les cas où le bras gauche est caché par le tronc. Nous remarquons que la répartition est quasi gaussienne.



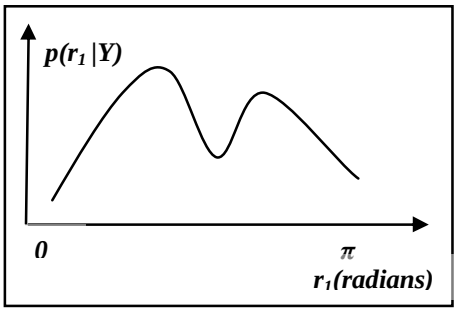


fig. 4.6 : Approximation de la densité de probabilité *a posteriori* de l'angle ***r1*** de rotation du bras droit dans les cas où le bras gauche est levé. Nous remarquons que la répartition devient bimodale.

Par ailleurs, les variables de notre vecteur d'état représentent des rotations qui sont contraintes par des limites biomécaniques. Par exemple, l'angle de flexion - extension du genou ne peut pas dépasser 180 degrés. D'ailleurs, sa probabilité *a posteriori* ne peut être gaussienne. Pour toutes ces raisons, l'utilisation d'un filtre de Kalman pour effectuer le suivi du mouvement était exclue à cause des densités non gaussiennes impliquées.

Deutscher et al. [6] ont également démontré la nature non gaussienne de la probabilité conditionnelle, dans le cas du suivi du mouvement humain 3D, et ont montré l'incapacité du filtre de Kalman à suivre le mouvement dans certains cas particuliers. L'adoption d'une approche particulaire s'avère utile malgré la grande dimension du vecteur d'état à estimer. En effet, l'algorithme de Condensation a déjà été utilisé pour le suivi du mouvement humain pour sa capacité de gérer les densités multimodales. Cependant, une modification doit y être introduite, pour le rendre plus efficace et pour réduire sa complexité (nombre de particules utilisées). Cette modification est apportée à travers l'algorithme 'Interval Particle Filtering' que nous avons conçu et que nous allons détailler dans le paragraphe suivant.

4.4 L'algorithme 'Interval Particle Filtering' (IPF)

Tenant compte de la nature multimodale et non gaussienne des densités de probabilité impliquées, l'utilisation d'un algorithme de filtrage particulaire nous semble judicieuse. D'autre part, la grande dimension de notre vecteur d'état (31 degrés de liberté) nous oblige à modifier l'algorithme de Condensation. L'algorithme ainsi modifié est appelé 'Interval Particle Filtering' (IPF).

4.4.1 Différences avec l'algorithme de Condensation

Le but de tous les algorithmes de filtrage particulaire modifiés est de réduire le nombre de particules nécessaires pour avoir une estimation fiable du vecteur d'état recherché, surtout dans le cas où ce vecteur est de grande dimension et par conséquent, où la complexité de l'algorithme de base le rend pratiquement inapplicable. Dans notre approche, le filtrage particulaire sera interprété comme étant une recherche de la meilleure particule dans un espace de particules bien défini à chaque pas de temps. L'estimation de la probabilité *a posteriori* n'est alors plus appliquée.

L'algorithme de Condensation tend à 'peupler' la zone autour d'une particule, relativement à son poids à l'instant précédent. En effet, plus la particule est 'lourde', plus son voisinage sera 'peuplé'. Le problème de l'algorithme de Condensation réside dans le fait que les particules les plus 'lourdes' monopolisent la création des nouvelles particules. Ce problème apparaît souvent dans le cas où le nombre de particules utilisées est relativement petit par rapport à la dimension du vecteur d'état. En effet dans ce cas, les nouvelles particules seront créées dans le voisinage d'un nombre très réduit de particules, empêchant ainsi la création de nouvelles particules à partir d'autres moins probables à l'instant précédent. Or, la fonction de vraisemblance utilisée n'est pas optimale et présente des faiblesses, dans le cas d'occlusions par exemple, surtout dans une configuration avec une seule caméra. Par conséquent, la particule $\boldsymbol{p_i}$ la plus probable à l'instant $\boldsymbol{t}$ n'est pas forcément issue (par évolution) de la particule la plus 'lourde' à $\boldsymbol{t-1}$, mais elle est probablement dérivée d'une particule $\boldsymbol{p_j}$ appartenant à l'ensemble des particules les plus lourdes à $\boldsymbol{t-1}$. Si $\boldsymbol{p_j}$ possède un poids $\boldsymbol{w^j_{t-1}}$ qui ne lui permet pas de survivre à la 'Sélection' à $\boldsymbol{t}$, la probabilité de retrouver la particule $\boldsymbol{p_i}$ dans l'ensemble issu de la 'Sélection' devient petite et l'algorithme peut diverger de la bonne solution (fig. 4.7).

D'autre part, le fait que $\boldsymbol{p_j}$ survive à la 'Sélection' n'implique pas la création de $\boldsymbol{p_i}$. En effet, si le poids $\boldsymbol{w^j_{t-1}}$ est petit, ou si le nombre des particules survivant à la 'Sélection' est relativement

grand par rapport au nombre total de particules utilisées, le nombre de particules dérivées de $\boldsymbol{p_j}$ dans son voisinage sera réduit et la probabilité de retrouver $\boldsymbol{p_i}$ dans l'ensemble créé par évolution durant l'étape de 'prédiction' sera réduite. Le 'survival diagnostic' *D*, détaillé précédemment, est un indicateur qui décrit bien ces deux problèmes. En effet, un indicateur *D* de petite valeur peut aboutir à la perte de la bonne particule $\boldsymbol{p_j}$ après l'étape de sélection. En outre, un grand *D* augmente la probabilité de conserver $\boldsymbol{p_j}$ mais réduit celle de créer $\boldsymbol{p_i}$ par évolution de $\boldsymbol{p_j}$. En conclusion, pour avoir la bonne estimée à l'instant courant il faut utiliser un grand nombre de particules pour assurer la présence d'au moins une particule au voisinage de cette bonne estimée. Néanmoins, dans le cas d'un vecteur d'état de dimension 31, ce nombre devient très grand et rend la complexité de la méthode très importante. Pour cela cette méthode devient pratiquement non utilisable.

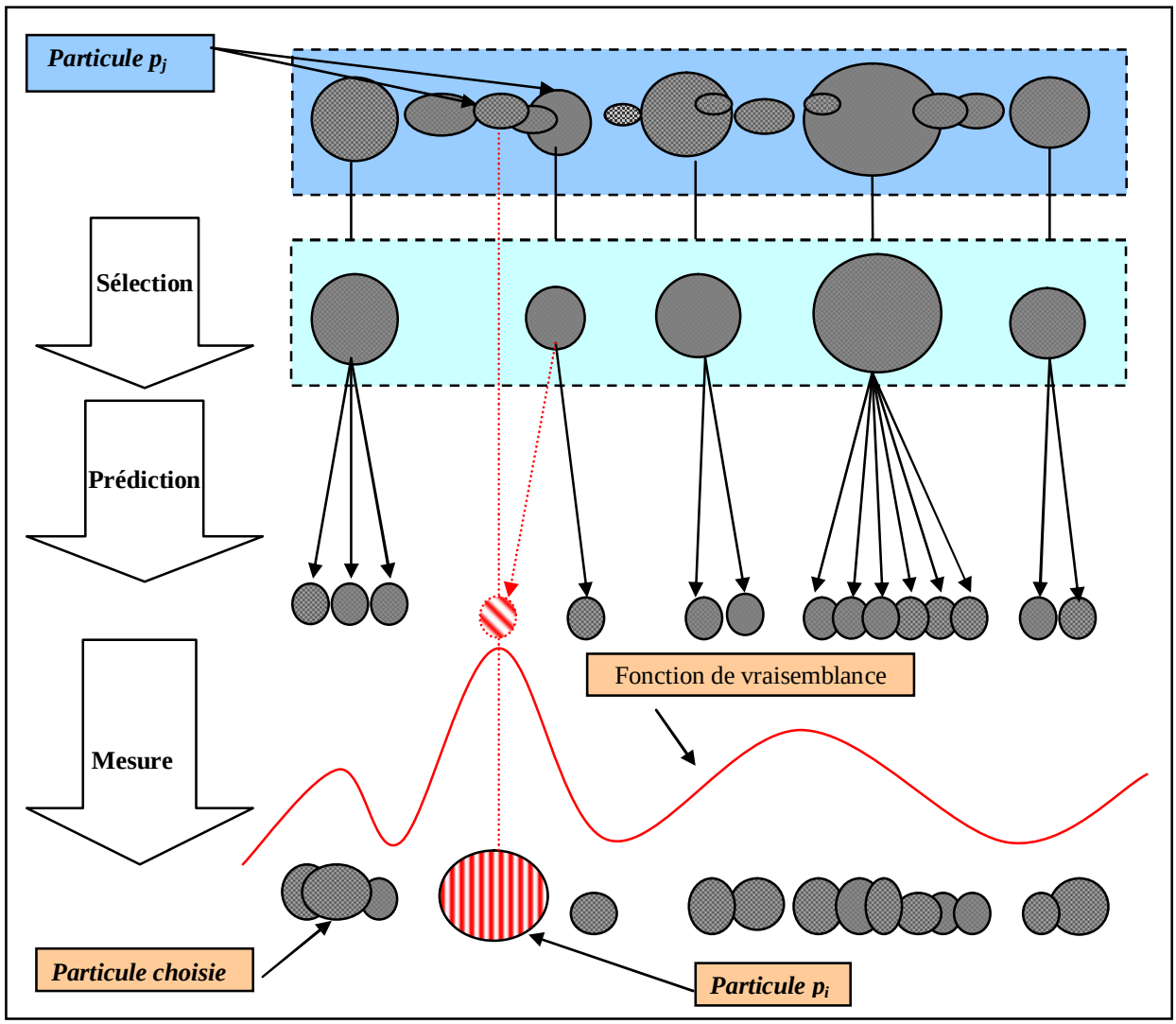


fig. 4.7 : Structure générale des trois étapes itératives de l'algorithme de Condensation à un instant ***t*** donné en considérant un vecteur d'état monodimensionnel. La divergence de l'algorithme résulte de l'absence de la particule $\boldsymbol{p_i}$ (en rayé) dans l'ensemble de particules crées

après l'étape de prédiction. Cet absence est due au petit poids des parents potentiels $\boldsymbol{p_j}$ de $\boldsymbol{p_i}$. En effet, ce poids ne laisse survivre aucun de ces parents ou réduit le nombre des particules 'fils' qui en sont issues par évolution.

L'idée de l'algorithme IPF que nous proposons, est de surmonter les deux points faibles de l'algorithme de Condensation cités ci-dessus, dans le cas où il est utilisé avec un nombre réduit de particules. L'algorithme IPF tend aussi à reconfigurer l'espace de recherche des particules d'une manière optimale, en fonction de l'application demandée (analyse de la marche) et de la fonction de vraisemblance utilisée. Les particules étant réparties de manière optimale, nous pouvons réduire ainsi le nombre nécessaire de ces particules pour avoir une bonne estimation. Le problème d'obtenir un petit 'survival diagnostic' à un instant donné est résolu en imposant une valeur constante de ce paramètre à tout instant. Cette valeur de *D* doit être grande mais en même temps doit rester relativement petite par rapport au nombre total de particules utilisées. D'autre part, pour donner une chance égale à toute particule survivant à la 'Sélection', le nombre de particules créées à partir de chacune sera constant lui aussi. Par conséquent, dans le cas d'une occlusion (observation erronée), une particule représentant le bon vecteur d'état mais ayant acquis, suite à l'observation, un poids moins important que les autres, donnera naissance au même nombre de particules que les autres particules durant l'étape de prédiction.

4.4.2 Evolution dynamique des particules dans IPF

Aucun modèle d'évolution (de transition) ne sera utilisé dans notre approche. Le but de notre système est de détecter les défauts de marche et doit s'adapter souvent à des personnes présentant des problèmes d'équilibre dynamique. Or, l'utilisation d'un modèle de mouvement appris, ou de modèles de marche pour robot bipède de type Denavit-Hartenburg, va obliger les particules à évoluer de façon à reproduire le mouvement appris ou désiré. Ce mouvement appris est souvent parfait et par conséquent, non représentatif du mouvement qu'on souhaite détecter, d'où le risque de fausser le suivi. Néanmoins, le mouvement humain est contraint par des limites et des capacités biomécaniques. En effet, les vitesses angulaires aux articulations sont bornées. L'évolution des angles aux articulations est alors limitée et prévisible. Dans le cas où l'on n'utilise pas de modèle de transition dans l'étape de prédiction de l'algorithme de Condensation, chaque particule évolue par ajout d'un bruit blanc gaussien. Or, comme il y a des limites sur l'évolution de tout degré de liberté du modèle, nous pouvons remplacer le bruit blanc gaussien par une loi uniforme bornée par les valeurs limites de l'évolution de ce degré de liberté, réduisant ainsi l'effet aléatoire du bruit gaussien. En résumé, l'algorithme IPF aide à 'peupler' avec le même nombre de particules les zones les plus 'probables' de l'espace d'état, d'une façon déterministe régulière.

4.4.3 Structure générale de l'algorithme IPF

Dans l'algorithme IPF nous maintenons la structure itérative des trois étapes de tout algorithme de type particulaire. La première étape est une étape de sélection, la deuxième étape concerne l'évolution des particules, la troisième est l'étape d'évaluation et de choix de la particule la plus probable. En utilisant ***N*** particules, la partie itérative de l'algorithme à l'instant ***t*** sera représentée par le schéma 4.8. Les trois étapes de notre algorithme sont les suivantes :

4.4.3.1 L'étape de 'Sélection'

L'ensemble des ***N*** particules crées à l'instant ***t-1*** est trié par ordre décroissant des 'poids'. Dans cet ensemble trié, ***M*** particules distinctes ayant les plus grands 'poids', sont choisies. Ce nombre ***M*** représentera le 'survival diagnostic' de l'algorithme. Ce paramètre qui peut prendre des valeurs allant de 1 jusqu'à N, est choisi d'une façon à être ni grand ni petit relativement au nombre total ***N*** de particules utilisées. La possibilité de gérer les densités de probabilité multimodales est ainsi préservée. D'autre part, comme les particules sont retenues grâce à leur poids à ***t-1***, l'historique des poids (et des observations) est introduit d'une façon implicite.

4.4.3.2 L'étape de 'Diffusion'

Aucun modèle temporel de transition ne sera utilisé pour faire évoluer les particules. Par contre, chaque articulation du corps humain possède une vitesse angulaire maximale et par conséquent l'évolution des angles à chaque articulation est bornée. Grâce à la présence de ces restrictions sur l'évolution des variables d'états (angles), nous pouvons déduire que la valeur de toute variable d'état x^k à l'instant ***t***, x^k_t (*k=1..31)* sera incluse dans un intervalle E^k_t de valeurs centrées sur la valeur x^k_{t-1} et délimité par les valeurs limites possibles de x^k définies en fonction du pas de temps donné de l'observation Δt (temps entre deux images successives) et de la vitesse angulaire maximale v^k de l'articulation correspondante à x^k :

$$x_t^k \in [x_{t-1}^k - v^k * \Delta t, x_{t-1}^k + v^k * \Delta t] \tag{4.5}$$

En adoptant le même raisonnement, nous pouvons dire que le vecteur d'état à l'instant ***t*** fera partie de l'intervalle multidimensionnel formé par la combinaison de tous les intervalles E^k_t (*k=1..31)*.

D'autre part, tout intervalle E^k_t est continu et doit être discrétisé à l'aide de ***q*** valeurs qui y appartiennent et par conséquent, toute variable x^k_t aura une valeur parmi les ***q*** valeurs possibles. Le paramètre ***q*** dépend de la précision recherchée. En effet, pour avoir plus de précision il faut discrétiser l'intervalle des valeurs probables par plus de valeurs. Dans le cas où le vecteur d'état

est de dimension 31, l'espace des vecteurs d'état possibles sera alors formé par q^{31} vecteurs possibles. Cet espace est alors très grand et nous retrouvons de nouveau le problème de complexité de l'algorithme. Notre idée est alors d'imposer à un nombre réduit n de variables d'état d'évoluer d'une façon déterministe et de laisser les autres variables évoluer par ajout d'un bruit blanc gaussien. Ainsi, le nombre des vecteurs d'état possibles est réduit de q^{31} vecteurs à q^n vecteurs.

Pour alléger les dimensions de cet espace de recherche, nous décomposons tout vecteur d'état X en deux vecteurs d'états L et R. Le premier vecteur L est formé par n degrés de liberté (variables d'état) qu'on va appeler les variables 'intéressantes' et le second R couvrira le reste des degrés de liberté. Chaque particule représente maintenant deux vecteurs d'état au lieu d'un. A tout instant t, L_{t-1} est mis à jour et remplacé par un intervalle multidimensionnel (espace) formé par la combinaison des intervalles déterministes construits, de la manière décrite ci dessus, autour des variables d'état appartenant à L. Cet espace multidimensionnel, une fois discrétisé, sera formé par q^n vecteurs couvrant une grille de vecteurs entourant le vecteur initial L_{t-1}. Le vecteur R_{t-1} évoluera par ajout d'un bruit blanc gaussien comme dans le cas de l'algorithme de Condensation. Chaque particule X^j_{t-1} parmi les M particules retenues sera alors remplacée par un ensemble de $I=q^n$ particules 'voisines', de façon à couvrir l'espace des configurations possibles (par rapport à l'état à $t-1$) des variables 'intéressantes'. A l'instant t, toute particule X^m_t, issue de X^j_{t-1} $(j=1..M)$ sera formée par un vecteur L^m_t appartenant à la grille de I vecteurs entourant L^j_{t-1} et un vecteur R^m_t obtenu par ajout d'un bruit blanc gaussien à R^j_{t-1} (fig. 4.9).

Pour que la décomposition du vecteur d'état soit bénéfique au niveau complexité, il faut que la dimension n du vecteur L soit petite par rapport à la dimension de R. Le choix des variables d'état formant L est alors délicat. Par contre, l'estimation des variables d'état formant L sera plus précise et moins aléatoire. En effet, les valeurs probables de ces variables sont fixées en fonction des contraintes locomotrices du corps humain. D'autre part, l'intervalle contenant ces valeurs probables est toujours représenté par des particules présentes dans l'espace de recherche d'états. Cet aspect déterministe garantit ainsi une convergence de l'algorithme pour le suivi de ces variables. En outre, les variables relatives aux membres du corps les plus volumineux ont plus d'influence sur la fonction de vraisemblance (une plus grande surface). Le fait d'avoir plus de précision sur l'estimation de ces variables rend alors la fonction de vraisemblance plus fiable. Il sera judicieux, en conséquence, d'inclure dans L les variables pour lesquelles nous souhaitons avoir plus de précision (paramètres servant à décrire la qualité de la marche) et qui peuvent avoir plus d'influence dans la fonction de vraisemblance utilisée.

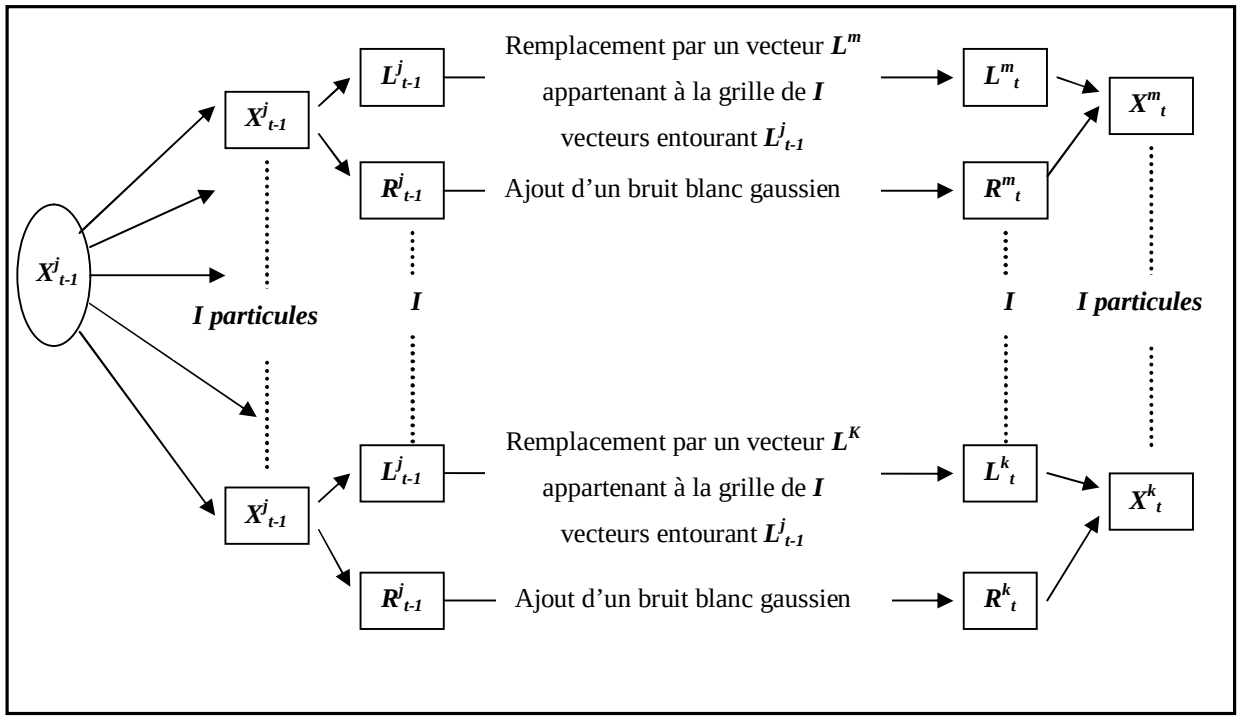


fig. 4.9 : L'étape de prédiction : toute particule est formée par deux vecteur ***L*** et ***R*** ; l'évolution de ***L*** est effectuée d'une manière déterministe, en le remplacement par un vecteur probable qui lui est 'voisin' ; ***R*** évolue par addition d'un bruit blanc gaussien. Ainsi toute particule donne 'naissance' à ***I*** particules couvrant un espace d'état probable qui l'entoure.

Dans le cas d'une succession d'observations erronées et où le paramètre ***M*** n'est pas grand, le risque de divergence de l'algorithme devient possible. Pour donner la possibilité à l'algorithme d'être robuste malgré une série de fausses estimations/observations, nous ajoutons à l'ensemble de particules, créées par évolution, un ensemble de ***S*** particules statiques (qui n'évoluent pas avec le temps) réparties de façon à couvrir l'espace de recherche d'état. De cette façon, toutes les zones de l'espace d'état sont 'peuplées' à tout instant ***t*** et le risque de divergence devient très faible. L'ensemble formé à la fin de cette étape contiendra ***N=M*I + S*** particules. Comme toute particule est remplacée par un intervalle multidimensionnel de particules, nous avons décidé d'appeler l'algorithme de filtrage particulaire que nous avons modifié, « Interval Particle Filtering ».

4.4.3.2 L'étape de 'Mesure'

Cette étape reste inchangée par rapport à l'algorithme de Condensation. Etant donné l'observation Y_t (l'image à ***t***), le poids de chaque particule créée durant l'étape de diffusion est calculé et le nouvel ensemble de particules pondérées est propagé pour être utilisé à l'étape ***t+1***. L'estimation du vecteur d'état (la posture du corps) à ***t*** est ainsi représentée par la particule (la configuration du modèle) ayant le poids le plus important dans l'ensemble.

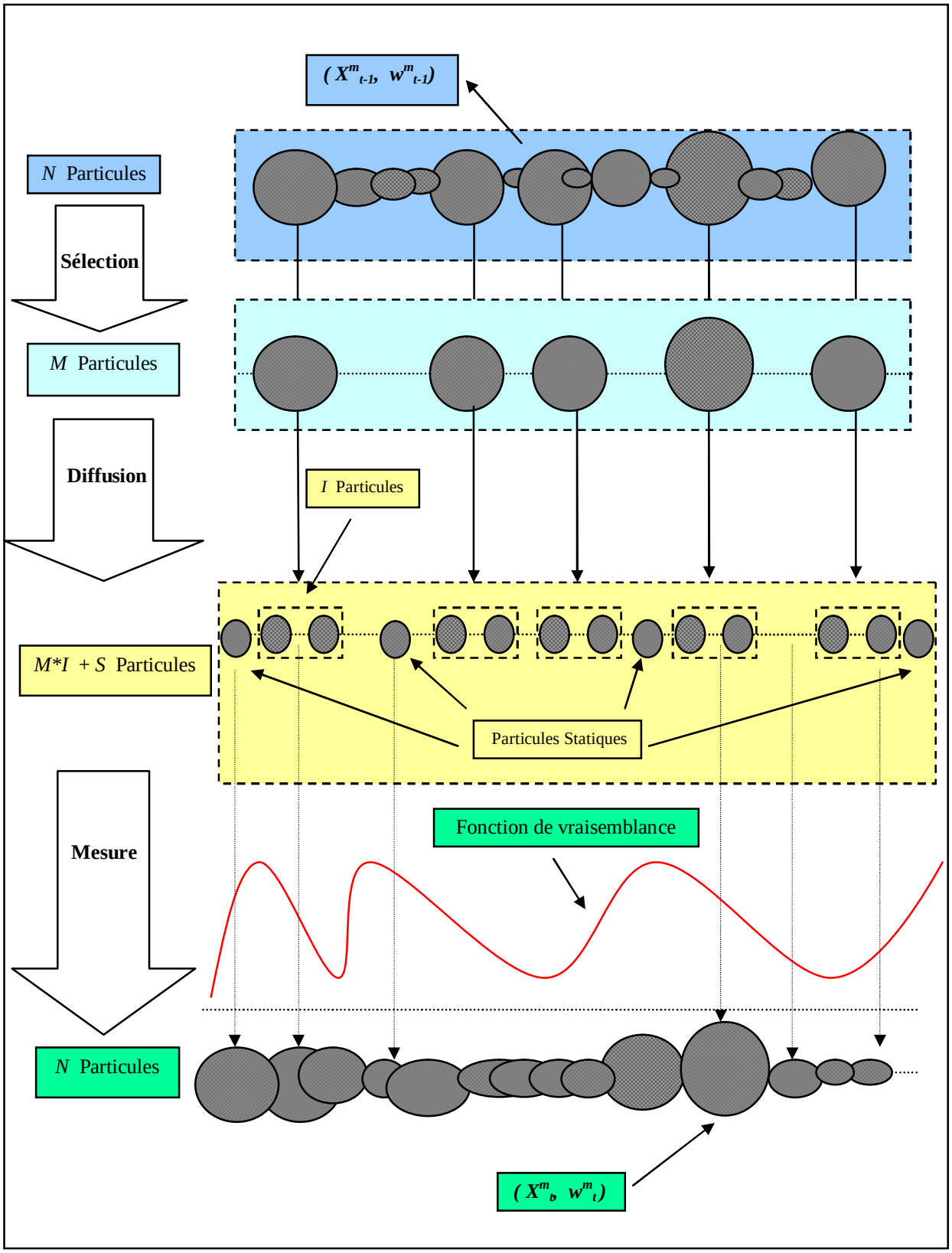


fig. 4.8 : Structure générale des trois étapes itératives de l'algorithme IPF développé à un instant ***t*** donné ; les ***M*** particules ayant les poids les plus importants seront remplacées chacune par un intervalle multidimensionnel (intervalle de dimension 1 dans la figure) discrétisé à l'aide de ***I*** particules couvrant son 'voisinage'. ***S*** particules statiques sont ajoutées à l'ensemble créé. La particule ayant le 'poids' le plus grand parmi ses paires vis à vis l'(les) image(s) sera considérée comme l'estimée à ***t***.

4.4.4 Discussion

Grâce à la diffusion déterministe et au maintien d'un nombre constant de particules, l'algorithme IPF maintient l'équilibre entre l'exploitation et l'exploration de l'espace de recherche des configurations possibles. L'ensemble des particules créées par IPF couvrent d'une manière efficace tout l'espace de recherche limitant ainsi le risque de divergence de l'algorithme. Cette garantie de convergence constitue la différence majeure avec une approche de type recuit simulé [3] utilisant le même nombre de particules. Cette dernière approche privilégie l'exploitation de l'espace de recherche à son exploration, fournissant ainsi des résultats plus précis mais au dépend d'un risque très important de divergence.

Dans le but d'obtenir une estimation plus précise utilisant IPF il faudra comme dans toute approche particulaire, utiliser un nombre plus important de particules et ainsi peupler plus fortement l'espace de recherche des configurations possibles du modèle 3D virtuel. Or, utiliser plus de particules implique un nombre plus important de créations/comparaisons d'images ce qui augmente la complexité de l'algorithme. Notre idée est alors de développer une version modifiée de IPF dans le but de réduire sa complexité exponentielle et ainsi obtenir des résultats plus précis en utilisant le même nombre de particules, ou garder la même échelle de précision en utilisant un nombre moins important de particules. Cette amélioration est apportée à travers l'algorithme que nous avons appelé 'Factored Interval Particle Filtering' (FIPF).

4.5 L'algorithme 'Factored Interval Particle Filtering' (FIPF)

La grande dimension du vecteur d'état utilisé est à l'origine de la complexité temporelle (temps d'exécution) des algorithmes de type particulaire. Pour casser cette complexité nous pourrions estimer les variables d'état chacune à part. Or, ces variables ne sont pas indépendantes vis à vis de l'observation dans notre cas (voir paragraphe 4.3) et ne peuvent pas êtres évaluées indépendamment. Cependant, le mouvement du corps peut être modélisé par une chaîne cinématique. Notre idée est alors de profiter de cette représentation pour factoriser le vecteur d'état et l'observation (la fonction 'poids') qu'il engendre, en utilisant le formalisme des réseaux bayésiens dynamiques qui sont de plus en plus utilisés pour factoriser les problèmes de calcul de grande dimension. La version factorisée de l'algorithme IPF appelée 'Factored Interval Particle Filtering' (FIPF) sera présentée dans ce paragraphe.

4.5.1 Réseaux bayésiens

Les réseaux baysésiens fournissent un cadre naturel pour la représentation de données incertaines et complexes. Ils sont issus du mariage de deux champs différents des mathématiques :

- la théorie des graphes qui fournit le cadre nécessaire pour une modélisation qualitative des connaissances ;
- la théorie des probabilités qui permet d'introduire une information quantitative dans ces connaissances.

En effet, les réseaux bayésiens sont des modèles graphiques qui encodent des relations de dépendance probabiliste entre des variables et qui donnent lieu à la représentation sous forme factorisée d'une distribution de probabilité jointe. Ils se présentent comme des graphes directionnels acycliques dont les nœuds sont associés à des variables aléatoires (discrètes ou continues avec une distribution particulière) et dont les liens encodent les relations de causalité entre les variables. L'absence de lien entre deux variables est liée à la notion d'indépendance conditionnelle qui peut exister entre ces deux variables.

En fait, la probabilité conditionnelle de deux variables (ou ensembles de variables aléatoires) ***X*** et ***Y*** s'écrit de la forme suivante :

(4.6) $$p(X = x|Y = y) = p(X|Y) = \frac{p(X,Y)}{p(Y)}$$

La probabilité jointe des deux variables ***X*** et ***Y*** s'écrit alors :

(4.7) $$p(X,Y) = p(X|Y).p(Y) = p(Y|X).p(X)$$

D'autre part la variable ***X*** est dite indépendante de la variable ***Y*** (et réciproquement) si et seulement si :

(4.8) $$p(X|Y) = p(X) \text{ et par conséquent } p(Y|X) = p(Y)$$

La probabilité jointe des deux variables indépendantes ***X*** et ***Y*** est alors égale au produit de leur probabilité marginale. Cette probabilité s'écrit alors :

(4.9) $$p(X,Y) = p(X).p(Y)$$

La probabilité jointe de N variables aléatoires $\{X^1, X^2, .., X^N\}$ peut être écrite sous la forme de produit de N probabilités conditionnelles. Ce produit est le résultat d'une application répétée de la formule (4.7) dans n'importe quel ordre :

$$p(X^1, X^2, .., X^N) = p(X^N \mid X^{N-1}, ..., X^2, X^1).p(X^2 \mid X^1).p(X^1) \tag{4.10}$$

En utilisant la propriété d'indépendance entre les variables et en notant l'ensemble des variables dépendant d'une variable X^j donnée par $PAR(X^j)$, nous pouvons écrire :

$$p(X^j \mid X^{j-1}, ..., X^2, X^1) = p(X^j \mid PAR(X^j)) \tag{4.11}$$

Ainsi la formule (4.10) s'écrira :

$$p(X^1, X^2, .., X^N) = \prod_{j=1}^{N} p(X^j \mid PAR(X^j)) \tag{4.12}$$

Cette propriété permet de simplifier les calculs de probabilités. En effet, pour calculer la probabilité jointe il suffira alors de calculer les probabilités conditionnelles de chaque variable par rapport aux variables qui en sont dépendantes uniquement. Dans un réseau bayésien un arc représente généralement une relation de causalité entre deux variables et son absence indique l'indépendance de ces deux variables. Ainsi, un arc orienté allant d'une variable X vers une variable Y indique que X est une cause possible de Y ou a une influence causale directe sur Y. En général cet arc est orienté dans le sens 'cause' vers 'effet'. Par conséquent, la probabilité jointe de l'ensemble des variables constituant le réseau bayésien est calculée à partir des probabilités conditionnelles de chaque variable 'nœud' par rapport à ses parents (dans le graphe).

4.5.2 Réseaux bayésiens dynamiques

Un réseau bayésien dynamique [7,8] est une représentation séquentielle d'un réseau bayésien statique B, utilisée pour modéliser les distributions de probabilités sur un ensemble semi infini d'ensembles de variables aléatoires Z_1, Z_2 .. où Z_t est la réalisation temporelle à l'instant t d'un processus Z modélisé par le réseau B. Le terme 'dynamique' est utilisé pour indiquer l'évolution dans le temps du processus Z et non pas pour indiquer un changement dans la structure du réseau B qui reste invariable.

D'une manière générale Z se décompose en trois sous ensembles ; le sous ensemble X des variables d'états non observables (cachées), le sous ensemble Y des variables observées et le

sous ensemble ***U*** des variables d'entrées (ou de contrôle). Ainsi, nous pouvons décrire ce modèle comme étant une généralisation des modèles de Markov cachés (HMM, POMDP). Un réseau bayésien dynamique est utilisé pour représenter des processus à temps discret. Chaque ajout d'un nouveau réseau statique qui est désormais appelé 'tranche' du réseau dynamique, correspond à un pas de temps supplémentaire de l'observation et représente l'évolution du processus stochastique ***Z***. L'ajout d'un réseau statique implique aussi l'ajout des arcs le reliant aux réseaux précédents.

Un réseau bayésien dynamique est défini par un couple (B_1, $B\rightarrow$), où B_1 est un réseau bayésien définissant l'état *a priori* $p(Z_1)$, et $B\rightarrow$ un réseau bayésien temporel représentant deux pas de temps (deux reproductions du réseau avec les arcs qui y correspondent), qui modélise $p(Z_t | Z_{t-1})$ à l'aide d'un graphe acyclique dirigé :

$$p(Z_t | Z_{t-1}) = \prod_{i=1}^{N} p(Z_t^i | PAR(Z_t^i)) \tag{4.13}$$

où Z_t^i est le i^{ème} nœud du réseau à un instant ***t*** donné et qui peut être membre de X_t, Y_t ou U_t et $PAR(Z_t^i)$ est l'ensembles des nœuds 'parents' de Z_t^i dans le graphe. Les nœuds dans le premier pas de temps du réseau temporel n'ont pas de paramètres associés, mais chaque nœud (variable) possède une distribution de probabilité conditionnelle à partir du deuxième pas de temps, qui définit $p(Z_t | Z_{t-1})$ pour tout instant $t > 1$. Ces distributions ont une forme arbitraire (table de probabilités, gaussiennes, etc.).

Les parents d'un nœud Z_t^i peuvent aussi bien appartenir au réseau représentant le même instant ***t*** ou celui représentant l'instant précédent ***t-1*** (comme il s'agit d'un processus markovien de premier ordre). Les arcs reliant les nœuds entre les différents pas de temps (tranches), vont de gauche à droite, représentant ainsi le déroulement séquentiel du temps. Tout arc reliant un nœud Z_t^i à un autre nœud Z_{t-1}^i est dit 'persistant'. Les arcs reliant deux tranches du réseau dynamique décrivent l'évolution temporelle 'persistante' des variables, alors que les arcs au sein d'une même tranche (même pas de temps) décrivent les relations de causalité immédiate entre les variables. Les paramètres des distributions de probabilité sont considérés comme invariantes avec le temps. Si ces paramètres sont amenés à changer avec le temps, il faut alors les inclure dans les réseaux comme des variables aléatoires.

La sémantique du réseau bayésien dynamique à un instant ***T*** donné peut être définie en déroulant le réseau ***B***→ plusieurs fois jusqu'à atteindre ***T*** (fig. 4.10). La probabilité jointe résultante sera calculée par :

(4.14) $$p(\underline{Z_t}) = \prod_{t=1}^{T} \prod_{i=1}^{N} p(Z_t^i \mid PAR(Z_t^i))$$

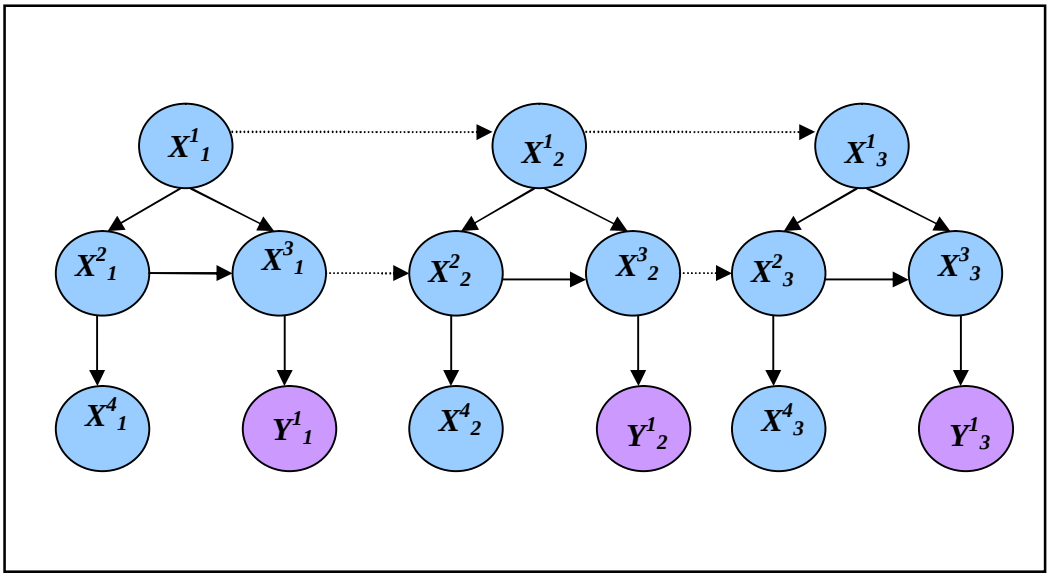


fig. 4.10 : Exemple d'un réseau bayésien dynamique sur trois pas de temps (***t***=1..3) formé par quatre variables cachées ***X*** et une variable d'observation ***Y***. Les arcs en pointillé représentent les phénomènes temporels 'persistants' et les arcs pleins représentent les liens de causalité immédiate.

En général les variables observations dans une tranche donnée sont indépendantes des variables formant la tranche précédente (réalisations temporelles indépendantes). Dans le cas d'absence de variables ***U*** de contrôle, un réseau bayésien dynamique est considéré comme un modèle probabiliste temporel basé sur une représentation factorisée des densités de probabilité de transition $p(X_t|X_{t-1})$ et d'observation $p(Y_t|X_t)$ où X_t et Y_t peuvent être décomposées en plusieurs variables : $X_t = (X^1_t, X^2_t .., X^N_t)$ et $Y_t = (Y^1_t, Y^2_t .., Y^M_t)$. La représentation factorisée de $p(X_t, Y_t \mid X_{t-1})$ est donnée par un graphe acyclique où tout nœud est associé à une variable X^i_t ou Y^j_t :

(4.15) $$p(X_t, Y_t \mid X_{t-1}) = \prod_i p(X_t^i \mid PAR(X_t^i)).\prod_j p(Y_t^j \mid PAR(Y_t^j))$$

où $PAR(X^i_t) \in (X_{t-1}, X_t)$, $PAR(Y^j_t) \in (X_t)$ sont les parents de X^i_t et Y^j_t dans le graphe acyclique dirigé.

4.5.3 Inférence dans les réseaux bayésiens et utilisation du filtrage particulaire

L'inférence dans un réseau bayésien consiste à calculer les probabilités des variables inconnues (cachées) en fonction des variables connues (observées). Ce calcul peut se faire facilement en utilisant la formule de probabilité de Bayes. Il existe dans la littérature plusieurs algorithmes d'inférence exacte dans les réseaux à structure arborescente (où chaque nœud ne peut avoir qu'un seul parent). Parmi ces algorithmes nous pouvons citer les méthodes variationnelles [9], l'algorithme de l'arbre de jonction [10] et l'algorithme de Pearl [11].

Cependant, quand la distribution de probabilité conditionnelle est aléatoire et ne suit pas une loi bien déterminée (gaussienne par exemple) ou quand les tables de probabilités conditionnelles sont absentes, toute inférence exacte est impossible. Pour résoudre ce problème, une estimation approximative stochastique est alors à appliquer pour effectuer l'inférence dans le réseau dynamique. En fait, l'approximation de la densité de probabilité peut être estimée à l'aide du filtrage particulaire. Ng et al. [12] utilisent l'algorithme de condensation pour estimer les densités de probabilité de certaines variables dans un réseau bayésien dynamique utilisé pour la fusion de données. Murphy [7] applique le filtrage particulaire pour effectuer l'inférence approximative dans un réseau bayésien modélisant un filtre de Kalman et dans un autre réseau de type SLAM (Simultaneous Localisation And Mapping). En effet ces approches consistent à créer un ensemble de particules représentant les variables cachées du réseau à chaque pas de temps (suivant l'algorithme de Condensation) et à les 'mesurer' par rapport aux observations et ainsi estimer les probabilités des variables cachées par rapport aux variables observées.

Ces méthodes d'inférence sont fondées sur l'algorithme de Condensation. Dans le but de conserver les avantages de IPF par rapport à Condensation, nous avons décidé d'adopter une nouvelle méthode d'inférence dans les réseaux bayésiens, fondée sur l'utilisation de IPF. Cette méthode sera présentée dans le paragraphe 4.5.6.

4.5.4 Structure du réseau bayésien utilisé

Notre approche consiste à modéliser le processus de suivi du mouvement humain représenté par le vecteur de configuration des degrés de liberté, à l'aide d'un réseau bayésien dynamique. Ainsi, toute variable d'état parmi 31 sera représentée par un nœud caché ayant plusieurs parents, et l'observation représentée par un nœud ayant 31 parents. Un arbre d'une taille pareille et avec des dépendances complexes, engendre des grandes complexités d'ordre algorithmique pour le calcul de la probabilité jointe du vecteur d'état. Cette représentation sera alors sans aucun intérêt par rapport à IPF.

Pour que la représentation à l'aide des réseaux bayésiens soit bénéfique au niveau de la complexité, nous avons essayé de réduire le nombre des nœuds (variables) et des arcs dans le graphe, en factorisant le vecteur d'état en plusieurs sous vecteurs, représenté chacun par une variable cachée dans le réseau. La réduction de la complexité passe aussi par la factorisation de la variable observation en plusieurs variables dépendant chacune d'une variable cachée unique. Cette factorisation est effectuée en utilisant une modélisation du mouvement en chaîne cinématique. En effet, comme nous traitons le mouvement humain, il est alors naturel de décrire le mouvement de chaque membre du corps par une seule variable dans le réseau et d'observer ce mouvement à l'aide d'une seule variable.

4.5.4.1 Factorisation du vecteur d'état à l'aide d'une chaîne cinématique

Une chaîne cinématique est définie par un ensemble d'articulations reliant N segments $S^i_{i=1..N}$. Les positions des différents éléments de la chaîne sont déterminées à partir de l'ensemble R des positions relatives (l'une par rapport à l'autre) des N segments voisins formant la chaîne $R= (R^1, R2,\ldots R^M)$. La dimension d'un élément R^i décrivant la position d'un segment S^i par rapport à un autre segment S^j dépend des degrés de liberté attribués à l'articulation reliant ces deux segments.

Dans une chaîne cinématique à boucle ouverte, tout segment est relié à un autre segment par un chemin unique ; la chaîne prend alors la forme d'un arbre. Ainsi, si la chaîne comprend N segments nous aurons $N-1$ positions relatives R^i. Nous numérotons les positions relatives R^i dans un ordre topologique en spécifiant l'un des segments comme étant l'origine de la structure. De même les positions absolues des segments sont notées dans un ordre topologique $X=(X^1,.., X^N)$. Dans une telle configuration, la position absolue X^i d'un segment S^i ne dépendra que de la position absolue $PAR(X^i)$ de son parent (unique) et de sa position relative R^{i-1} par rapport à ce parent. Par conséquent, la position absolue X^i d'un segment, sachant la position absolue $PAR(X^i)$ de son segment parent est indépendante des positions absolues des autres segments de X. Cette dernière propriété nous permet de déduire une indépendance conditionnelle entre toute variable (position absolue) et toute autre variable représentant un segment de la chaîne cinématique à l'exception de son segment parent. Les positions absolues des segments d'une chaîne cinématique peuvent alors être représentées à l'aide d'un réseau bayésien possédant la même structure arborescente de la chaîne. Le modèle 3D humanoïde virtuel que nous utilisons peut être représenté par une chaîne cinématique à boucle ouverte formée de 14 segments représentant les membres du corps auxquels nous avons attribué des degrés de liberté (fig. 4.11).

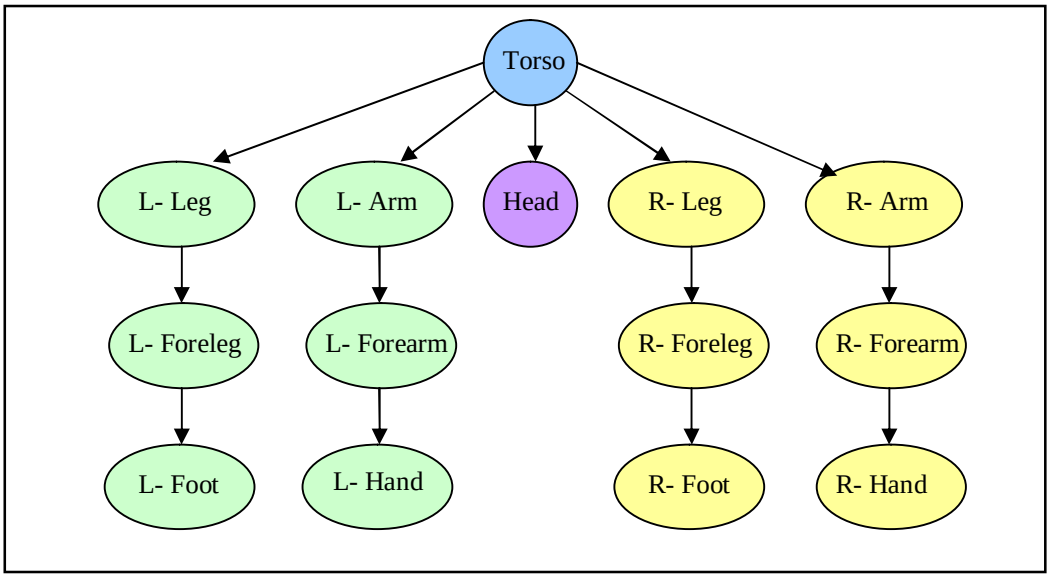


fig. 4.11 : Représentation du modèle humanoïde avec une chaîne cinématique formée de 14 segments ; le tronc a été choisi comme étant la racine de cette chaîne cinématique. Par exemple, la position absolue de la main ne dépend que de sa position relative par rapport à l'avant bras (le parent dans la chaîne) et de la position absolue de cet avant bras.

Dans le vecteur d'état qu'on souhaite estimer, (31 degrés de liberté) nous regroupons les variables d'état relatives au mouvement d'un même segment dans une seule variable. Ainsi, nous pouvons factoriser le vecteur d'état $\boldsymbol{X}$ à l'aide d'un réseau bayésien formé par 14 variables $\boldsymbol{X^i}$. La structure de ce réseau sera identique à la structure de la chaîne cinématique représentant le modèle virtuel (fig. 4.10). Le vecteur d'état étant structuré en arbre, sa probabilité (jointe) est alors le produit des probabilités locales de chaque variable de cet arbre. Nous pouvons alors écrire :

(4.16)
$$p(X) = \prod_{i=1}^{14} p(X^i \mid PAR(X^i))$$

Ainsi, à l'aide de la factorisation, nous avons réduit le nombre des distributions de probabilité impliquées à 13. (une par variable parmi 13 ayant chacun un seul parent).

4.5.4.2 Factorisation de la fonction d'observation ('poids')

Dans le cadre du suivi du mouvement à partir d'images vidéo, les segments formant le corps ne sont pas des variables observables directement. En effet, la fonction 'poids' (d'observation) dont nous disposons, décrit la configuration du corps entier (l'ensemble des segments). Cependant, cette fonction peut être factorisée conformément à la factorisation du vecteur d'état. Ne pouvant pas évaluer la vraisemblance de chaque segment du corps d'une manière indépendante, nous allons effectuer cette évaluation d'une manière hiérarchique, segment après

segment. Ainsi, nous évaluons d'abord la vraisemblance du segment S^1 (et non pas la vraisemblance de tout le corps), en utilisant la fonction 'poids' que nous avons développée. Les pixels communs à la projection de la configuration de S^1 et à l'image 'silhouette' seront masqués et n'interviennent plus dans le calcul du 'poids' d'un autre segment enfant (par rapport à la chaîne d'évaluation) de S^1. L'image résultante sera utilisée comme image de référence pour le calcul du 'poids' du segment fils direct (par rapport à l'évaluation) de S^1. Cette évaluation en chaîne du vecteur d'état nous permet alors de factoriser la fonction 'poids' de l'observation Y en 14 fonctions et d'écrire la fonction d'observation sous la forme suivante :

$$p(Y|X) = \prod_{i=1}^{14} p(Y^i | X^i, PRE(Y^i)) \quad (4.17)$$

où $PRE(Y^i)$ est la valeur de la fonction 'poids' du segment évalué juste avant S^i.

4.5.4.3 Représentation de la dynamique des particules

L'évolution des particules est introduite par les liens temporels qu'on ajoute au réseau bayésien, le transformant ainsi en un réseau bayésien dynamique. Pour effectuer le suivi 3D du mouvement humain, aucun modèle dynamique de la marche n'a été utilisé. Par conséquent, la position du segment S^i à l'instant t ne dépend que de sa position à l'instant $t-1$:

$$p(X_t^i | X_{t-1}) = p(X_t^i | X_{t-1}^i) \quad (4.18)$$

En effet par application de l'algorithme IPF, l'évolution d'une variable d'état à l'instant t est effectuée soit par le remplacement dans une particule de sa valeur donnée à $t-1$, par un intervalle de valeurs possibles construit d'une façon déterministe suivant les contraintes sur l'évolution de cette variable, soit par ajout d'un bruit blanc gaussien à cette valeur à $t-1$. En ajoutant les liens temporels et les observations à la structure en arbre des variables, nous obtenons un réseau bayésien dynamique (fig. 4.12) qu'on utilisera ensuite pour réduire la complexité de l'algorithme IPF.

4.5.5 Intérêt de la factorisation

Le problème des algorithmes de filtrage particulaire est le grand nombre de particules nécessaires pour bien estimer un vecteur d'état de grande dimension. En revanche, nous remarquons qu'un nombre important de ces particules créées durant les étapes de 'ré échantillonnage' et de 'diffusion' à un instant t, présentent une faible vraisemblance par rapport à l'observation Y_t et par conséquent, ne survivent pas au 'ré échantillonnage' à l'instant $t+1$.

L'idée essentielle est alors de réduire le nombre de ces particules 'inutiles' en utilisant la factorisation du vecteur d'état et de l'observation, pour appliquer un ré échantillonnage hiérarchique des composantes de ce vecteur d'état ; dans le cadre d'une approche particulaire, les particules à ***t*** survivent ou non à cette étape, suivant leur 'poids' calculé par rapport à l'observation à ***t-1***. Or, il est plus intéressant de prendre en compte l'observation à ***t***, quand elle est disponible (comme dans le cas de notre approche), et non pas celle à ***t-1***, pour la création des particules durant l'étape de 'ré échantillonnage' à ***t***. Ainsi, en utilisant la factorisation de l'observation et de l'état, nous introduisons cette observation dans le processus de création de particules. Ce principe a été inspiré par la procédure de mesure de vraisemblance utilisée dans [13,14] où les éléments du vecteur d'état sont échantillonnés dans un ordre topologique. En utilisant la factorisation, nous pouvons réduire le nombre de particules utilisées à tout instant en les éliminant dès leur création au cas où elles semblent peu probables vu l'observation à venir.

D'autre part, l'algorithme IPF que nous avons introduit crée un espace de recherche de particules plus efficace et mieux adapté à l'application demandée (le mouvement humain) et utilise un nombre plus réduit de particules que dans un approche classique (algorithme de Condensation). Dans l'étape de 'diffusion' de cet algorithme, chaque particule est remplacée par un intervalle de particules. Ainsi, en utilisant le principe de factorisation dans la création des éléments de cet intervalle, nous pouvons réduire leur nombre et par conséquent la complexité de l'algorithme. Nous appellons le nouvel algorithme conçu en se basant sur le principe de factorisation et sur IPF, 'Factored Interval Particle Filtering' (FIPF).

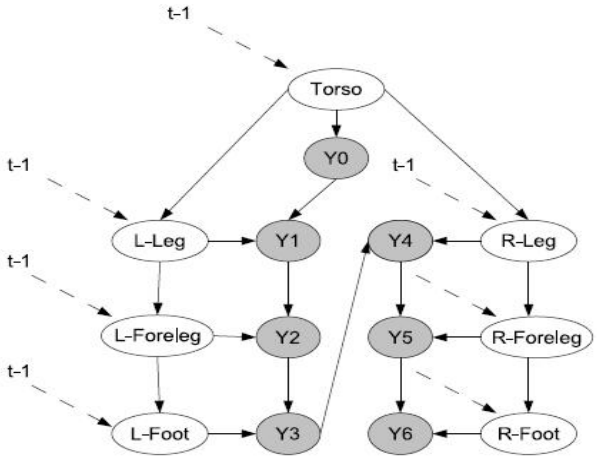


fig. 4.12 : Représentation d'une partie du réseau bayésien dynamique, construit à partir de la chaîne cinématique du modèle 3D virtuel, dans le but d'alléger la complexité de IPF. Les nœuds ($\boldsymbol{Y}^i$) en gris représentent les observations et les nœuds en blanc les variables d'état relatives à chaque segment.

4.5.6 Algorithme d'inférence utilisé dans FIPF

Dans une approche particulaire, toute particule choisie durant l'étape de 'Sélection' (***M*** particules dans le cas d'IPF) donne naissance à un ensemble de particules. Chacun de ces ensembles est construit de la façon suivante à partir de la particule 'parent' : pour effectuer une construction hiérarchique du vecteur d'état à l'instant ***t***, et en profiter pour réduire le nombre de particules utilisées, nous proposons d'effectuer une étape de ré échantillonnage chaque fois que nous rencontrons un nœud observé, en parcourant, d'une manière topologique, le réseau bayésien construit (fig. 4.12). L'algorithme d'inférence appliqué pour chaque nœud est le suivant :

- Si le nœud correspond à une variable d'état X^n_t, toute particule z^i_t donne naissance à un nouveau sous ensemble de particules dont le vecteur d'état sera complété par des échantillons de la probabilité $p(X^n_t | PAR(X^n_t)=u_j)$ (où u_j est la valeur du parent de X^n_t dans la particule z^i_t), ou par un intervalle de valeurs possibles (dans le cas de FIPF). En effet, z^i_t est toute particule créée durant le passage par le nœud 'variable parent' et conservée après passage par le nœud d'observation précédent, ou représente la particule 'parent' initiale si X^n_t est le premier nœud d'état visité. Ainsi, l'ensemble des particules est agrandi à chaque passage par un nœud d'état, ce qui correspond à l'étape de 'diffusion' de Condensation.
- Si le nœud correspond à un nœud d'observation Y^K, le poids w^i_t d'une particule z^i_t sera mis à jour suivant l'équation suivante : $w^i_t = w^i_{t-1} * p(Y^K = y^k_t | PAR(Y^K_t)=u_i)$ où u_i est la valeur du parent de Y^K_t dans la particule z^i_t. Dans cette étape, un nombre réduit de particules (les plus 'lourdes') sera conservé. Ainsi, à chaque passage par un nœud d'observation, l'ensemble des particules sera réduit.

La différence entre cette méthode d'inférence et les méthodes d'inférence [7,12] utilisant le filtrage particulaire déjà utilisées dans la littérature, est que la création des particules se fait au fur et à mesure qu'on parcourt le graphe et non pas *a priori* avant le parcours. D'autre part, l'identification des particules 'peu probables' est effectuée à chaque passage par une variable d'observation ce qui nous permet de réduire le nombre des particules utilisées. Ainsi, l'observation intervient dans la formation de l'ensemble des particules en gardant exclusivement celles qui sont les plus probables à la fin du parcours. La création de l'ensemble des particules est effectuée suivant l'algorithme IPF (diffusion déterministe) et non pas suivant l'algorithme de Condensation, ce qui nous permet de créer un ensemble plus adapté au problème et garantir la convergence de l'algorithme utilisé.

Cet algorithme a été développé mais n'a pas encore été testé et réglé. En effet, il sera nécessaire de spécifier à chaque ré échantillonnage le nombre de particules créées ou éliminées. D'un point de vue théorique, la factorisation présentée ci-dessus doit réduire d'une manière significative le nombre de particules utilisées. Ainsi, nous aurons la possibilité d'augmenter la précision du suivi sans augmenter pour autant la complexité. En effet, dans IPF le nombre de particules nécessaires augmente d'une manière exponentielle avec l'augmentation de la dimension du vecteur à estimer. Cependant, dans FIPF ce nombre de particules augmente d'une façon polynomiale avec l'accroissement de la dimension du vecteur d'état et exponentiellement avec l'augmentation de la dimension de chaque variable du réseau.

4.6 Conclusion

Dans ce chapitre nous avons détaillé l'approche que nous avons utilisée pour effectuer un suivi 3D du mouvement humain. Notre méthode se fonde sur l'utilisation d'un modèle 3D virtuel humanoïde et d'une fonction 'poids' développée pour évaluer la vraisemblance d'une configuration donnée du modèle par rapport à la vraie posture du corps dans les images.

Pour trouver la configuration du modèle qui maximise la fonction de vraisemblance, nous avons adopté une approche particulaire modifiée de l'algorithme de filtrage particulaire de base. En effet, l'algorithme de Condensation (ou de filtrage particulaire) nécessite un nombre important de particules pour donner une estimation fiable du vecteur d'état recherché. L'algorithme IPF, que nous avons développé et détaillé dans ce chapitre, crée une ensemble de particules mieux reparties dans l'espace de recherche et fournit une garantie de la convergence de l'algorithme. Ainsi, en utilisant un nombre plus réduit de particules qu'avec Condensation, nous pouvons couvrir d'une manière plus efficace l'espace de recherche de configuration et obtenir une bonne estimation.

IPF présente une complexité exponentielle par rapport à la dimension du vecteur d'état à estimer. Pour remédier à ce défaut, nous avons proposé une version modifiée de l'algorithme que nous avons appelée FIPF. Cette version factorisée de l'algorithme IPF utilise une représentation factorisée du vecteur d'état à l'aide d'un réseau bayésien dynamique. Pour calculer la probabilité d'une configuration donnée, une nouvelle méthode d'inférence dans les réseaux bayésiens a été développée. Cette méthode d'inférence utilise l'algorithme IPF et fait intervenir l'observation courante dans la formation des particules, ce qui nous permet d'éliminer celles qui sont pas probables dès leur création et ainsi réduire le nombre final de particules utilisées. Dans le chapitre suivant nous décrirons la mise en œuvre pratique de l'algorithme IPF et le dispositif

expérimental utilisé pour valider les résultats obtenus. Nous exposerons également les résultats de quelques reconstructions 3D.

Bibliographie

[1] Sidenbladh H, Black M.J et Fleet D.J, "Stochastic tracking of 3D human figures using 2D image motion", Proceedings of 6th European Conference on Computer Vision, Dublin 2000.

[2] Cohen I et Mun Wai L, "3D Body Reconstruction for Immersive Interaction", Proceedings of the Second international Workshop on Articulated Motion and Deformable Objects, Palma de Mallorca Spain, pp. 119-130, 2002.

[3] Deutscher J, Blake A et Reid I, "Articulated body motion capture by annealed particle filtering", Proceedings of the Conference Computer Vision and Pattern Recognition, vol. 2, pp. 1144-1149, 2000.

[4] Wenstop O, "Motion detection for image information", Proceedings of the 3rd Scandinavian conference on image analysis, Norway, pp. 381-386, 1983.

[5] Cucchiara R, Grana C, Piccardi M, Prati A et Sirotti S, "Improving Shadow Suppression in Moving Object Detection with HSV Color Information", Proceedings of Intelligent Transportation Systems Conference, pp. 334-339, 2001.

[6] Deutscher J, Blake A, North B et Bascle B, "Tracking through singularities and discontinuities by random sampling", Proceedings of the 7th International Conference on Computer Vision, 1999.

[7] Murphy K.P, "Dynamic Bayesian Networks : Representation, Inference and Learning", PhD thesis, University of California Berkeley, 2002.

[8] Smaili C, "Utilisation des réseaux bayésiens dans un système d'aide à la décision", Mémoire de DRT, Université Nancy 2, 2005.

[9] Castillo E, Gutiérrez J.M et Hadi A, "Expert Systems and Probabilistic Network Models", Springer-Verlag New York, Inc, 1997.

[10] Lauritzen S, Cowell R, Dawid A.P et Spiegelhalter D, "Probabilistic Networks and Expert Systems", Springer-Verlag New York, Inc, 1999.

[11] Pearl J, "Probabilistic reasoning in intelligent systems : Networks of plausible inference", Morgan Kaufman Publishers, Inc., San Mateo, CA, 2nd edition, 1988.

[12] Ng B, Pfeffer A, "Factored Particle Filtering for Data Fusion and Situation Assessment in Urban Environments", Proceedings of 7th IEEE international Conference on Information Fusion, 2005.

[13] Fung R.M et Chang K.C, "Weighing and integrating evidence for stochastic simulation in Bayesian networks", UAI '89 : Proceedings of the Fifth Annual Conference on Uncertainty in Artificial Intelligence, pp. 209–220, Amsterdam, 1989.

[14] Shachter R et Peot M, “Simulation approaches to general probabilistic inference on belief networks”, UAI ’89 : Proceedings of the Fifth Annual Conference on Uncertainty in Artificial Intelligence, pp. 221–234, Amsterdam, 1989.

Chapitre 5 : Mise en oeuvre et résultats

Dans le chapitre précédent, nous avons détaillé l'approche que nous avons adoptée pour effectuer le suivi 3D du mouvement humain. Ainsi ont été détaillés l'algorithme IPF que nous avons développé et qui se fonde sur une approche particulaire, le modèle virtuel utilisé, ainsi que la fonction de vraisemblance utilisée pour évaluer la ressemblance d'une configuration de ce modèle à l'image réelle. Une version modifiée de l'algorithme IPF a été aussi détaillée.

Dans ce chapitre, nous allons présenter le dispositif expérimental et le matériel utilisés pour tester et valider les algorithmes développés. La façon dont l'algorithme IPF a été appliqué sera aussi détaillée ; nous exposerons la méthode utilisée pour déterminer les paramètres de calibration de la caméra virtuelle que nous utilisons pour représenter les différents configurations du modèle humanoïde. Ensuite, nous expliquerons le choix des paramètres de l'algorithme IPF que nous avons appliqué. L'étape d'initialisation de cet algorithme sera également détaillée.

L'approche que nous avons utilisée pour calculer, à partir des positions 3D estimées des articulations, les paramètres de la marche identifiés au chapitre 2, sera ensuite expliquée. Les résultats du suivi, obtenus à partir de notre système, seront comparés aux résultats du suivi, pour les mêmes sujets et les mêmes mouvements effectués, obtenus à l'aide d'un système à base de marqueurs du type VICON. En effet cette comparaison porte sur les paramètres de la marche calculés à partir des données fournies par IPF et par VICON. Nous ferons également une comparaison des reconstructions 3D des mouvements de quelques points clé du corps durant la marche normale, par les deux systèmes.

Pour montrer la capacité de notre système à bien suivre le mouvement 3D du corps, nous montrerons les reconstructions 3D de quelques mouvements complexes (lever de chaise, occlusion etc..). Nous discuterons aussi de la complexité temporelle de notre approche et des améliorations que nous pouvons y apporter, au niveau des technologies utilisées, pour réduire le temps de calcul.

5.1 Dispositif expérimental

Dans le but de tester l'approche proposée, nous avons utilisé le système que nous avons développé pour effectuer le suivi 3D du mouvement d'un ensemble de sujets. Pour valider les résultats obtenus, nous avons effectué simultanément le suivi du mouvement de ces personnes à l'aide d'un système à base de marqueurs passifs de marque VICON.

5.1.1 Description du matériel utilisé

Le système VICON a été choisi comme référence pour la précision et la fiabilité du suivi qu'il fournit. Le modèle VICON que nous avons utilisé est formé de six caméras émettrices infrarouges reliées à une centrale de calcul et de visualisation ; les marqueurs qui sont fixées sur le corps en mouvement réfléchissent les signaux infrarouges émis par les caméras ce qui permet de déterminer les droites reliant chaque caméra aux différents marqueurs. La position 3D d'un marqueur est alors calculée par triangulation (voir paragraphe 3.1) à condition qu'il puisse réfléchir les signaux d'au moins deux caméras différentes. Ainsi, le système VICON calcule les positions 3D de tous les marqueurs et fournit les positions 3D des différentes articulations correspondantes grâce à un modèle d'interpolation. La fréquence d'émission des signaux infra rouges utilisés par ce modèle étant de 50 Hz, les positions 3D sont calculées à cette même fréquence. Notre système calcule les positions 3D de 19 articulations du corps humain ; pour comparer ces positions 3D aux positions fournies par VICON, il a fallu attacher les marqueurs réfléchissants à ces 19 articulations.

Etant contraint par l'utilisation du système VICON, l'expérimentation s'est déroulée dans une salle qui y est dédiée, au laboratoire de mouvement de l'IRR (Institut Régional de Rééducation de Lorraine). La zone couverte par les champs de vision des différentes caméras infra rouges du VICON constitue un parcours de 3m de largeur et de 6m de longueur . En ce qui concerne notre système, il était composé de deux caméscopes numériques commerciaux (Sony et Panasonic). Une caméra était placée perpendiculairement au parcours (fig.5.1) (de profil) et l'autre avec un angle de 45 degrés par rapport à ce parcours (en trois quarts). Les vidéos étaient compressées sous format DV avec une résolution de 720*576 pixels et une fréquence de 25 images/sec. Ces vidéos étaient ensuite synchronisées avec les acquisitions VICON manuellement, et les images obtenues étaient utilisées comme observation pour l'algorithme IPF détaillé auparavant. Les algorithmes ont été développés en C++ Builder et les calculs ont été effectués à l'aide d'une machine P4 3Ghz. Pour chaque image, nous pouvons estimer, à partir de l'algorithme IPF et avec une fréquence de 25Hz, la configuration des différents degrés de liberté

(angles de rotations aux articulations) du corps et les positions en 3D dans le repère caméra des 19 points suivis. Comme dans VICON, ces données sont fournies sous forme d'un fichier Excel. Les positions 3D étant fournies dans le repère de notre caméra (pour notre système) et dans un repère absolu (par rapport à la salle d'expérimentation pour VICON), une translation et une rotation ont été appliquées à notre repère pour que les positions soient calculées dans un même repère pour les deux systèmes.

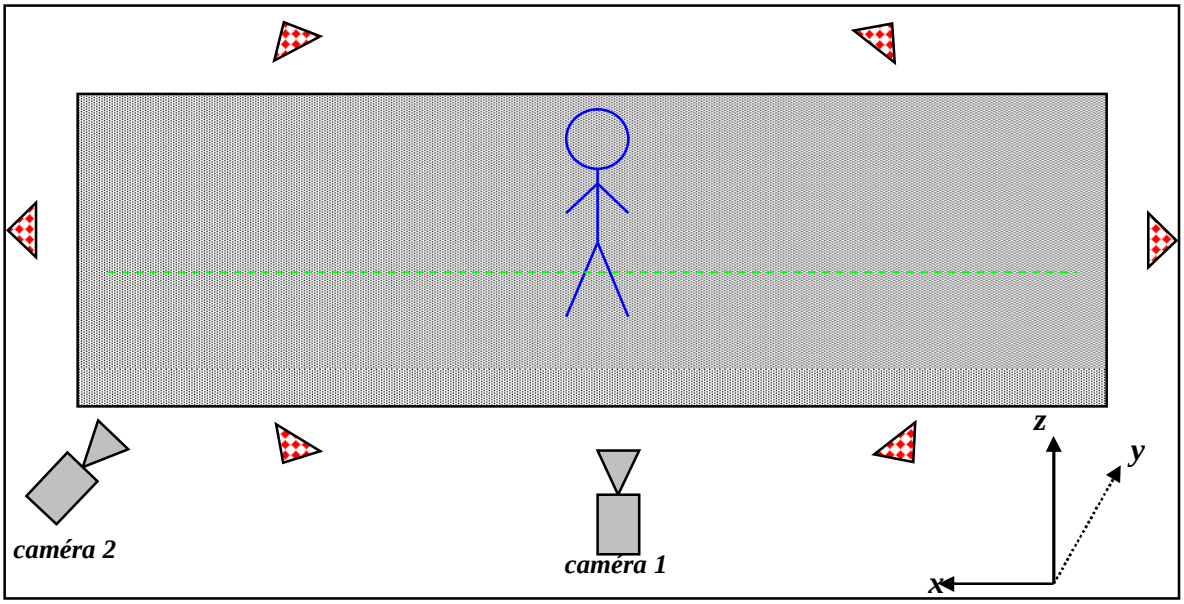


fig. 5.1 : Dispositif expérimental utilisé pour le test du système développé ; les emplacements des caméras infrarouges sont spécifiés par les petits triangles pointillés (rouges). Notre caméra 1 est placée perpendiculairement au parcours (en pointillé) ; la caméra 2 est positionnée avec un angle de 45 degrés par rapport au parcours.

5.1.2 Description de l'ensemble des sujets tests et du protocole adopté

Quinze sujets (3 femmes et 12 hommes) ont participé à l'expérimentation. Ces sujets sont jeunes (moyenne d'âge de 29.4 ans) et ne souffrent pas de problèmes particuliers de marche. Ce choix d'une population jeune nous était imposé à cause des contraintes administratives (assurances, passage par un comité d'éthique etc..) et logistiques, qui rendaient l'application de l'expérimentation sur une population âgée impossible. D'autre part, le but de cet expérimentation étant de comparer la qualité de reconstruction 3D de notre système par rapport à un système de référence (VICON), la présence de sujet âgés n'était pas nécessaire.

Nous avons demandé à chaque sujet de traverser le parcours d'une manière naturelle, pieds nus. L'essai est répété 3 fois avec 2 minutes de repos entre les essais. Etant contraints par la zone couverte par le système VICON, à chaque traversée, seuls 3 cycles complets de marche s'étalant

sur environ 3 sec, ont été enregistrés. Ainsi nous avons pour chaque essai 3 pas gauches et 3 pas droits.

5.2 Paramètres et réglages des algorithmes utilisées

Les vidéos recueillies à l'aide des caméras ont été décompressées et décomposées en images (25 images /sec) en utilisant la librairie « vfw » de C++. Dans chaque image, le corps en mouvement a été détecté et l'image a été transformée en image 'silhouette'. Cette image est ensuite comparée aux projections 2D par une caméra virtuelle, des différentes configurations du modèle 3D virtuel, créées à l'aide de l'algorithme IPF. La configuration dont la projection possède le plus grand degré de ressemblance à l'image 'silhouette' est choisie comme étant l'estimée. L'algorithme IPF nécessite une étape d'initialisation pour trouver la configuration initiale du modèle dans la première image. D'autre part, pour bien estimer la configuration des degrés de liberté du corps, il faut estimer les coordonnées 3D du point 'origine' du modèle. Toutes ces étapes de prétraitement sont nécessaires et seront exposées dans ce qui suit.

5.2.1 Création de l'image 'silhouette'

Pour toute caméra, une image de l'arrière plan est extraite. Nous soustrayons cet arrière plan de toute image montrant le corps en mouvement. L'image résultante ne contient alors que l'image de ce corps. Or, l'expérience étant menée dans une salle éclairée artificiellement, l'image résultante est fortement bruitée par les ombres. Pour réduire ces ombres, l'utilisation d'un filtre de seuillage et d'un filtre médian s'est avérée efficace et suffisante pour éliminer tous les bruits dans les images. Aucun filtre HSV n'a alors été appliqué pour le prétraitement des images résultant de cette expérimentation.

5.2.2 Réglage et calibration de la caméra virtuelle

Toute configuration du modèle 3D virtuel fournit les coordonnées des points formant ce modèle dans un repère virtuel. Or pour donner un 'poids' à chacune de ces configurations, il faut comparer sa projection 2D à la projection de la vraie posture du corps dans l'image (image 'silhouette'). Pour que cette comparaison soit valide, il est nécessaire d'établir la correspondance entre tout point du corps $\boldsymbol{P(x,y,z)}$ représenté à travers la caméra vidéo dans l'image 'silhouette' par le point $\boldsymbol{P_i(x_i,y_i)}$, et le point $\boldsymbol{P_v(x_v,y_v,z_v)}$ le modélisant dans le monde virtuel, représenté à travers une caméra virtuelle dans l'image synthétique par le point $\boldsymbol{P_{iv}(x_{iv},y_{iv})}$. En effet, il faut établir les translations et les rotations permettant de réaliser le changement de repère entre les deux mondes (réel et virtuel) et déterminer les paramètres intrinsèques et extrinsèques (voir

Annexe A) des caméras réelles et virtuelles intervenant dans la formation des images réelles et synthétiques respectivement.

Or, c'est nous qui créons le repère virtuel et la caméra virtuelle qui fournit l'image synthétique. Ainsi, nous maîtrisons les paramètres de ce monde virtuel. Pour simplifier le problème de correspondance entre les mondes virtuel et réel, nous supposons que les repères des deux mondes (réel et virtuel) sont superposés. Il est alors judicieux de représenter tout point ***P(x,y,z)*** du squelette du corps réel par un point $\boldsymbol{P_v(x_v,y_v,z_v)}$ du repère 3D virtuel ayant les mêmes coordonnées. Le modèle 3D en fils de fer ainsi construit, aura les mêmes dimensions (dans le repère virtuel) qu'un squelette 3D réel. Les parallélépipèdes, ayant chacun une forme et un volume approchés de ceux de la partie du corps qu'il représente, seront ensuite ajoutés à ce modèle pour lui donner un aspect humain.

Après avoir créé une équivalence entre les deux repères réel et virtuel, nous supposons utiliser une caméra virtuelle ayant les mêmes paramètres extrinsèques et intrinsèques que la caméra réelle utilisée. Ainsi, nous supposons que la caméra virtuelle est placée dans l'espace au même endroit (connu) que la caméra réelle avec les mêmes orientations. Il reste alors à déterminer les paramètres de calibration intervenant dans la formation de l'image par la caméra réelle, et de les imposer comme paramètres de calibration à la caméra virtuelle.

Une méthode de calcul de ces paramètres est expliquée dans l'annexe A. Cependant, l'opération de calibration doit être effectuée pour chaque caméra utilisée et à chaque changement de la position de la caméra. Notre idée est alors de trouver une méthode alternative plus simple pour déterminer ces paramètres.

En effet, nous pouvons fixer quelques paramètres intrinsèques de la caméra virtuelle. Nous imposons que l'image produite soit sans distorsion et que le centre optique soit placé au centre de l'image. De plus, nous supposons que le rapport pixel/cm et la distance entre les pixels dans l'image formée soient tous les deux égaux à 1. Compte tenu des valeurs de ces paramètres, l'équation de projection $\boldsymbol{(x_{iv},y_{iv})}$ dans le plan image de la caméra virtuelle, d'un point $\boldsymbol{(x_v,y_v,z_v)}$ du repère virtuel peut alors s'écrire :

$$x_{iv} = f_v.\frac{x_v}{z_v}$$
$$y_{iv} = f_v.\frac{y_v}{z_v} \qquad (5.1)$$

où f_v est une variable inconnue représentant la distance focale de la caméra virtuelle.

Or, l'équivalence entre les images synthétiques et réelles (équivalence entre les deux caméras) implique que la projection (x_i,y_i) dans le plan image réelle de n'importe quel point $P(x,y,z)$ réel donné, soit égale à la projection (x_{iv},y_{iv}) par la caméra virtuelle, du point $P_v(x_v,y_v,z_v)$ représentant P dans le repère virtuel :

(5.2) $$(x_i,y_i)= (x_{iv},y_{iv})$$

Or, $(x,y,z) = (x_v,y_v,z_v)$; en se basant sur les équations (5.1) et (5.2) nous pouvons écrire :

(5.3) $$x_i = x_{iv} = f_v.\frac{x}{z}$$
$$y_i = y_{iv} = f_v.\frac{y}{z}$$

à partir de (5.3) nous pouvons écrire :

(5.4) $$h_r = f_v.\frac{h}{z}$$

où h est la hauteur réelle (en cm) d'une personne suivie et h_r la hauteur de la silhouette extraite de cette personne.

La hauteur réelle h de toute personne suivie est connue. Si à un instant t donné, nous disposons d'une image représentant la personne debout, la hauteur h_r (en pixels) peut être calculée facilement à partir de l'image 'silhouette' correspondante. Si de plus nous connaissons la distance z séparant la personne de la caméra à cet instant t, nous pouvons alors estimer f_v. Connaissant cette distance focale f_v, tous les paramètres nécessaires pour la projection du modèle à l'aide de la caméra virtuelle sont alors déterminés.

Cette méthode reste heuristique et moins précise que la calibration exacte de la caméra réelle mais plus simple, moins encombrante et générique. Cependant, pour qu'elle puisse être appliquée, il faut disposer d'une image (l'image initiale par exemple) où la personne suivie est debout, et connaître la distance entre la personne et la caméra à ce moment précis. De plus, il faut connaître la position et l'orientation de la caméra réelle dans un repère euclidien absolu. Elle peut être effectuée automatiquement au début du traitement de chaque séquence vidéo et ne présente pas de complexité temporelle (<0.5s). Si nous ne disposons pas de toutes les données

précises nécessaires pour appliquer cette méthode, le calcul de la distance focale virtuelle sera faussé, ce qui va entraîner une imprécision dans les résultats obtenus pour le suivi.

La calibration exacte complète d'une caméra est effectuée à l'aide d'images d'une mire dite de calibration. Cette opération de calibration fournit les paramètres extrinsèques et intrinsèques inconnus de la caméra. Elle doit être effectuée pour toute caméra et à chaque fois qu'on change sa position et son orientation. La calibration doit être réalisée au préalable pour pouvoir injecter les paramètres à la caméra virtuelle. Cette opération est plus lourde que la méthode décrite ci dessus mais elle est précise et effectuée une seule fois pour toute installation.

Notre choix de l'une ou de l'autre des deux méthodes de calibration a été souvent imposé par les conditions pratiques des expérimentations menées.

5.2.3 Choix des paramètres de l'algorithme

Pour pouvoir appliquer l'algorithme IPF que nous avons développé, il faut déterminer certains paramètres (paragraphe 4.5.4). Tout d'abord il faut spécifier le 'survival diagnostic' de l'algorithme qui représente le nombre de particules distinctes choisies après l'étape de sélection à chaque itération. Une petite valeur de ce paramètre peut entraîner une divergence de l'algorithme et une valeur proche du nombre total des particules $\boldsymbol{N}$ indiquera une mauvaise précision. L'idée est alors de fixer ce paramètre. En choisissant de garder à $\boldsymbol{t}$ les $\boldsymbol{M}$ particules les plus lourdes à $\boldsymbol{t-1}$, le 'survival diagnostic' de notre algorithme est de $\boldsymbol{M}$. Or, nous introduisons à chaque étape de prédiction un nombre $\boldsymbol{S}$ de particules statiques qui sont alors supposées survivre au 'ré échantillonnage'. Ainsi, le 'survival diagnostic' sera égal à $\boldsymbol{M+S}$. Après plusieurs essais nous avons choisi de fixer $\boldsymbol{M}$ à 25.

D'autre part chacune de ces $\boldsymbol{M}$ particules sera remplacée par un intervalle multidimensionnel discrétisé par $\boldsymbol{I}$ particules. Cet intervalle couvre d'une façon déterministe l'espace de valeurs des variables 'intéressantes'. Il faut alors déterminer ces variables. En effet il faut choisir un nombre réduit de ces variables pour ne pas augmenter la complexité de l'algorithme. Par conséquent il faut choisir les variables ayant plus d'influence sur la fonction 'poids' et sur lesquelles nous voulons avoir la plus grande précision. Compte tenu du fait que la fonction 'poids' se base sur des images prises de profil (ou en trois quarts) et qu'on s'intéresse surtout au mouvement des membres inférieurs du corps, nous allons étiqueter 5 variables comme 'intéressantes' : les angles de flexion aux deux hanches et aux deux genoux et l'angle d'oscillation du tronc dans le plan

médio latéral. En effet, ces membres du corps ont la plus grande surface dans l'image 'silhouette' et représentent les degrés de liberté intervenant directement dans la marche humaine.

Comme nous traitons des images prises à 25 Hz, le pas de temps Δt entre deux images est de 40 ms. La vitesse angulaire maximale des angles de flexion aux hanches et aux genoux est de 100 degrés/s pour tout type de mouvement (marche course etc..) normal ou anormal. Cette vitesse est en fait liée aux contraintes biomécaniques. L'intervalle E^k_t des valeurs possibles à t d'une variable x^k (parmi ces angles de flexion citées ci dessus) sera alors :

$$x_t^k \in [x_{t-1}^k - 4°, x_{t-1}^k + 4°]$$

Si nous souhaitons avoir une précision de 2 degrés pour chacune de ces variables nous pouvons discrétiser cet intervalle (de largeur 8 degrés) par 3 valeurs { $x^k_{t-1} - 4$; x^k_{t-1} ; $x^k_{t-1} + 4$ }. En effet la différence entre deux valeurs représentant l'intervalle étant de 4 degrés, le choix entre ces deux valeurs sera alors effectué avec une erreur maximale égale à la moitié de leur différence (2 degrés). Pour l'oscillation médio latérale x^0, la vitesse angulaire maximale est de 50 degrés/s ; l'intervalle E^0_t des valeurs possibles sera alors :

$$x_t^0 \in [x_{t-1}^0 - 2°, x_{t-1}^0 + 2°]$$

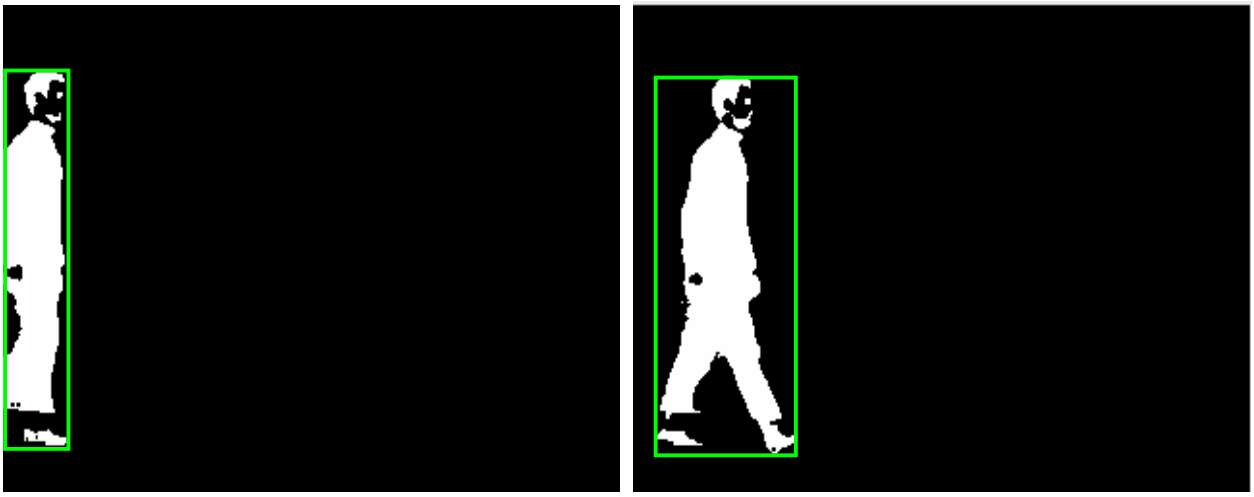
Nous allons représenter cet intervalle également par 3 valeurs. L'intervalle multidimensionnel devra alors contenir 3^5=243 particules ; or, certaines de ces configurations sont impossibles durant la marche (les deux jambes en avant par exemple). Par conséquent, nous avons pu réduire ce nombre à 189 particules. Ainsi toute particule parmi les M choisies sera remplacée par I=189 particules.

Les S particules statiques sont réparties sur une grille couvrant l'espace de recherche des états ; ces particules aident à garantir la convergence de l'algorithme. En effet nous utilisons 1125 particules statiques à chaque étape de prédiction qu'on ajoute au $M*I$ particules déjà créées. Le nombre total N de particules utilisées sera alors 5850 particules. Le 'survival diagnostic' de l'algorithme est de 1150. Le temps de traitement pour estimer, pour toute image, la positon du point 'origine' et la configuration du modèle est de 6 sec.

5.2.4 Initialisation de l'algorithme 'IPF'

Dans le cas où la personne suivie entre progressivement dans le champ de vision de la caméra et en ressort progressivement, nous considérons que l'instant $t=0$ initial de notre

algorithme correspond à l'instant où la personne suivie entre complètement dans ce champ de vision (silhouette complète dans l'image). En effet, à partir de la première image dans le flux, nous calculons (par balayage des lignes et des colonnes) pour toute image de la séquence vidéo, un rectangle encadrant la silhouette humaine. Tant que les points formant ce rectangle sont situés aux bordures de l'image (pixels de la première ou de la dernière, ligne ou colonne de la matrice formant l'image), nous pouvons déduire que la personne n'est pas encore complètement entrée dans le champ (fig.5.2). Après l'identification de l'instant initial, dès que les points du cadre atteignent les bordures nous déduisons par symétrie que la personne commence à sortir du champ de vison et le traitement est alors arrêté. Durant l'expérimentation menée, toute personne partait d'une position couverte par la caméra utilisée. Par conséquent, la première image extraite de toute séquence a été considérée comme l'image à l'instant initial ***t=0*** de l'algorithme.



(a) (b)

fig. 5.2 : Pour chaque image nous calculons le rectangle (vert) qui encadre la silhouette ; si des points de ce rectangle se situent à l'une des bordures de l'image nous constaterons que la personne n'est pas entièrement dans le champ de vision de la caméra (a) et l'image ne sera pas traitée. En revanche dès que tous les points du cadre sont à l'intérieur de l'image (b) nous lançons l'algorithme de suivi.

A tout instant ***t***, l'algorithme IPF utilise les ***N*** particules créées à l'instant ***t-1***, pour en créer de nouvelles et trouver celle qui représente la configuration la plus probable des degrés de liberté du corps à l'instant ***t***. Or, à l'instant ***t=0*** (première image), aucun ensemble de particules n'est encore créé. L'algorithme IPF n'est pas alors applicable à cet instant. Une étape d'initialisation s'avère donc obligatoire. Comme nous n'avons aucune indication sur la vraie configuration des degrés de liberté à cet instant, l'idée est de rechercher dans un espace formé par la combinaison des ensembles de valeurs possibles pour chaque variable, la configuration ayant la valeur de la fonction 'poids' la plus grande par rapport à la première image. Or, si nous voulons représenter

par exemple, chacun de ces ensembles par 6 valeurs (ce qui ne constitue d'ailleurs pas une bonne précision), notre espace de recherche sera alors formé par 6^{31} particules. Ce nombre de particules est très grand et rend cette recherche longue et lourde. L'idée est alors d'alléger cette recherche, en adoptant la même logique utilisée pour l'étape de prédiction dans IPF. Ainsi, l'espace de recherche à l'instant ***t=0*** sera formé par des particules recouvrant l'espace des valeurs possibles des variables 'intéressantes' d'une manière exhaustive et déterministe tandis que les autres variables prendront des valeurs aléatoires pour toute particule. Les variables identifiées comme 'intéressantes' sont les cinq variables citées dans le paragraphe précédent. En choisissant de représenter les valeurs possibles de ces variables par 6 valeurs, l'espace de recherche sera formé par 6^5 = 7776 particules. Nous attribuerons à chacune de ces particules le 'poids' qui lui correspond en utilisant la fonction 'poids' développée. La particule la plus 'lourde' sera choisie comme étant la configuration initiale la plus probable. Les ***M*** particules les plus lourdes seront ensuite transmises, pour être utilisées à l'instant ***t=1*** dans l'étape de prédiction l'algorithme IPF.

5.2.5 Détermination des coordonnées du point 'origine' et de l'orientation du corps

Le modèle 3D humanoïde utilisé est supposé être représenté en 2D à partir d'une caméra virtuelle. Il est caractérisé par le point particulier représentant le sacrum (le point le plus bas de la colonne vertébrale), qu'on appelle 'origine'. En effet, les coordonnées des points du modèle sont spécifiées dans un repère 3D euclidien ayant comme origine ce point particulier et qu'on appelle repère corps. Les positions des 19 points du modèle dans le repère caméra 3D d'origine la caméra (virtuelle ou réelle, les deux repères étant superposés) sont alors déterminées à partir de la position 3D (O_x, O_y, O_z) de son point 'origine' dans ce repère, de l'orientation du repère corps par rapport au repère caméra (R_x, R_y, R_z) et de la configuration de ses degrés de liberté (fig. 5.3).

L'image synthétique d'une configuration donnée du modèle est produite par la caméra virtuelle à partir des positions 3D dans le repère caméra de ces points. Pour bien estimer la configuration des degrés de liberté, il faut que la comparaison des images 2D 'silhouette' et synthétique soit valide et par conséquent, il faut que le repère corps virtuel soit superposé au repère corps réel. Pour cela, à tout instant ***t***, il faut trouver les coordonnées du point 'origine' du modèle et son orientation par rapport au repère caméra de façon à ce que les deux repères corps soit superposés. Ces paramètres étant estimés, nous pouvons ensuite utiliser l'algorithme IPF pour trouver la configuration du modèle humanoïde qui maximise la vraisemblance par rapport à l'image 'silhouette'.

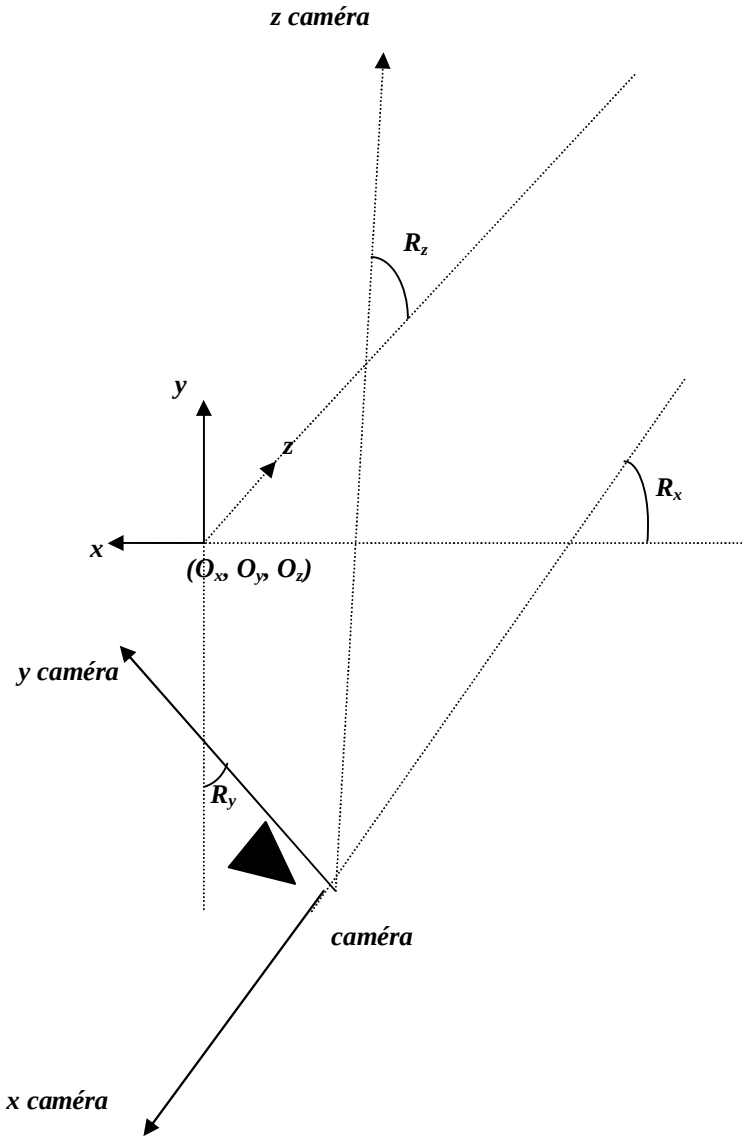


fig. 5.3 : Représentation du modèle 3D virtuel dans le repère caméra virtuelle

L'image 'silhouette' à ***t*** ainsi que l'estimation à ***t-1*** de la configuration du modèle 3D sont utilisées pour calculer ces paramètres. Pour déterminer l'orientation du corps à ***t***, nous recherchons dans le voisinage de l'angle d'orientation à ***t-1***, la valeur permettant au modèle d'avoir le poids le plus grand par rapport à l'image 'silhouette' à ***t***. D'autre part, pour trouver les coordonnées (O_x^t, O_y^t, O_z^t) du point origine à ***t***, nous effectuons une recherche hiérarchique (***z*** , puis ***x***, puis ***y***) dans le voisinage du point à ***t-1*** $(O_x^{t-1}, O_y^{t-1}, O_z^{t-1})$, de façon à trouver la meilleure correspondance entre la partie haute du tronc du corps réel dans l'image 'silhouette' (fig. 5.4) et du modèle virtuel dans l'image synthétique du modèle 3D ayant comme 'origine' le point candidat et comme configuration la particule choisie à ***t-1***. Cette partie du corps (le haut du tronc) a été choisie comme étant une partie facile à étiqueter et pour ayant une forme (représentation dans la 'silhouette') qui varie peu avec le mouvement (ne subit pas d'occlusions,

ne se détache pas etc..). La recherche de la position initiale se fait selon le même critère mais en balayant une plage de valeurs plus grandes.

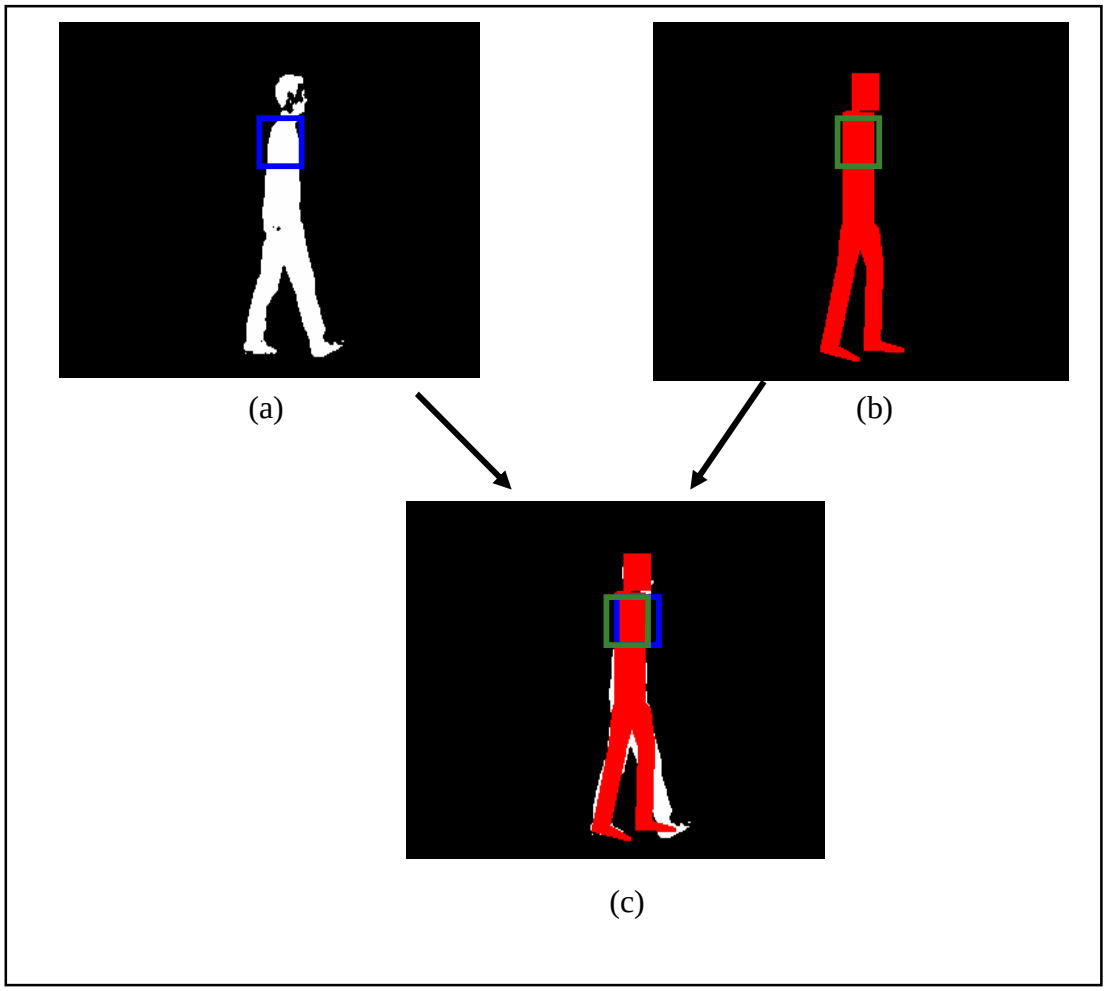


fig. 5.4 : Recherche de la position (O_x^t, O_y^t, O_z^t) du point 'origine' à t : nous recherchons les coordonnées du point 'origine' de façon à superposer (c) le cadre (bleu) englobant la partie haute du tronc dans l'image 'silhouette' à t (a) et le cadre (vert) englobant la partie haute du tronc dans l'image synthétique du modèle à $t-1$ (b).

5.2.6 Schéma général de la méthode

A partir de l'image 'silhouette' extraite à $t=0$, nous recherchons la distance focale de la caméra virtuelle (si elle n'est pas calculée par calibration exacte), la configuration initiale du modèle (à partir de l'ensemble d'initialisation) et nous recherchons les coordonnées (O_x^0, O_y^0, O_z^0) du point origine. A partir de l'instant $t=1$ et pour toute image à un instant donné t, nous recherchons les coordonnées du point origine (O_x^t, O_y^t, O_z^t) à partir de $(O_x^{t-1}, O_y^{t-1}, O_z^{t-1})$ et de la configuration du modèle à $t-1$, X_{t-1}, selon la méthode décrite ci dessus. Nous appliquons l'algorithme IPF ensuite avec l'ensemble de N particules créées à t-1, pour estimer la configuration du modèle à t, X_t (fig. 5.5).

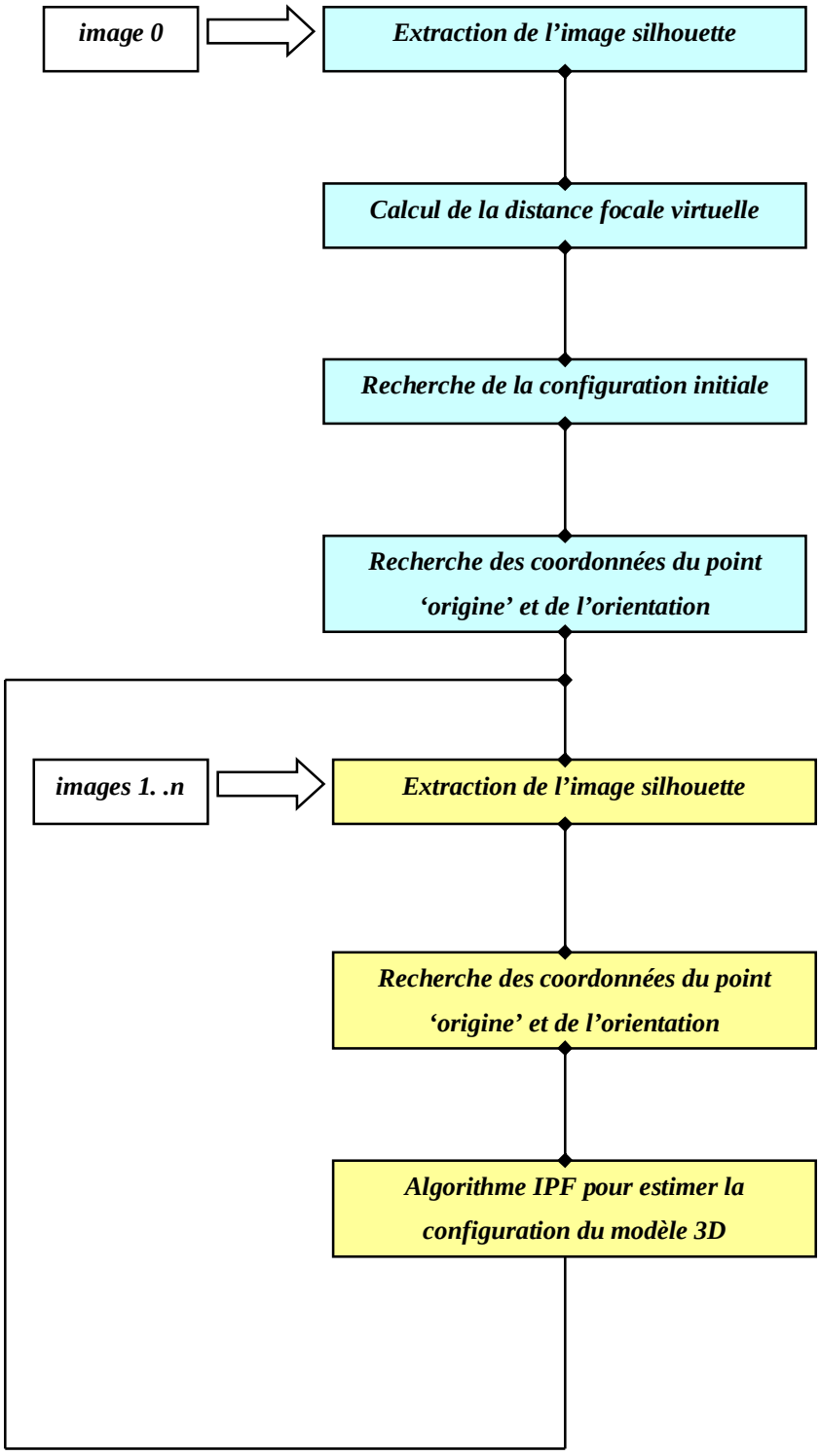


fig. 5.5 : Schéma général de l'application utilisée

5.3 Extraction des paramètres

A partir des images du corps en mouvement et en utilisant les algorithmes développés, nous recouvrons les angles de rotation aux articulations et les positions 3D de 19 points du corps humain pour toute image ***i (i=0..f)***. En effet, les positions 3D de ces points sont calculées dans le repère corps, connaissant les longueurs des segments formant le modèle (estimées suivant la

hauteur de toute personne) et les angles de rotation aux articulations joignant ces segments (degrés de liberté formant le vecteur d'état). Ces positions sont ensuite calculées dans le repère de la salle d'expérimentation par deux changements de repère : un changement de repère corps – caméra connaissant les coordonnées du point 'origine' dans le repère caméra (qu'on calcule suivant la méthode décrite ci dessus), et puis un changement de repère caméra – salle connaissant la position de la caméra dans la salle. Le système VICON fournit également les coordonnées 3D dans le repère salle, des points du corps sur lesquels les marqueurs étaient placés. A partir de ces positions 3D nous devons déterminer les paramètres, listés dans le paragraphe 2.5.2 et qui seront étudiés ultérieurement dans le but d'évaluer le risque de chute. Dans ce paragraphe nous allons détailler chacun de ces paramètres et la méthode pour l'extraire.

5.3.1 Vitesse et accélération

Pour calculer la vitesse moyenne de marche (suivant un axe donné), il suffit de diviser la distance parcourue durant un passage par le temps écoulé. Cette distance parcourue correspond à la différence entre la coordonnée (***x*** dans notre cas) du point 'origine'(sacrum) à l'instant ***0*** (image ***0***) et à l'instant final (image ***f***). Le temps écoulé correspond au nombre total des images-1 ***f***, multiplié par la différence temporelle ***Δt*** entre deux images successives. La fréquence d'acquisition de nos vidéos étant de 25 Hz, la différence temporelle ***Δt*** entre deux images successives est de 40ms. Par contre la fréquence d'acquisition du système VICON est de 50 Hz ; par conséquent la différence entre deux pas de temps sera de ***Δt*** 20ms. La vitesse moyenne en cm/s peut alors s'écrire :

(5.5) $$v_m = \frac{\left| x_f(sacrum) - x_0(sacrum) \right|}{f.\Delta t}$$

La vitesse instantanée en cm/s à un instant ***t*** donné peut être alors écrite :

(5.6) $$v_t = \frac{\left| x_t(sacrum) - x_{t-1}(sacrum) \right|}{\Delta t}$$

Pour calculer l'accélération instantanée du tronc à tout instant ***t***, il suffit de dériver la vitesse instantanée au point sacrum :

(5.7) $$\alpha_t = \frac{\left| v_t(sacrum) - v_{t-1}(sacrum) \right|}{\Delta t}$$

5.3.2 Longueurs et largeur des pas

Par définition, les longueurs des pas gauche et droit sont mesurés lorsque le déplacement est effectué suivant un axe donné. La longueur du pas gauche est la distance, suivant cet axe, entre la cheville gauche et l'orteil droit au moment dit 'd'attaque' de la cheville gauche, où cette dernière touche le sol (fig. 5.6). Cette longueur sera le maximum de la distance orientée dans le sens de la marche, entre la cheville gauche et l'orteil droit, durant un cycle locomoteur (un pas gauche suivi d'un pas droit). Par symétrie, la longueur du pas droit est la distance, suivant l'axe de déplacement, entre la cheville droite et l'orteil gauche au moment dit 'd'attaque' de la cheville droite, où cette dernière touche le sol. Cette longueur sera le maximum de la distance orientée dans le sens de la marche, entre la cheville droite et l'orteil gauche, durant un cycle de marche.

La largeur du pas est un paramètre qui mesure l'espacement des jambes durant la marche et qui décrit la manière dont la personne essaie de maintenir son équilibre. En effet, la largeur du pas est la distance suivant l'axe perpendiculaire au déplacement, entre les chevilles gauche et droite (fig. 5.6) aux moments 'd'attaque', où l'une des deux chevilles touche le sol (l'autre cheville étant déjà par terre). La mesure de cette largeur doit se faire au même moment que l'estimation des longueurs des pas gauche et droit.

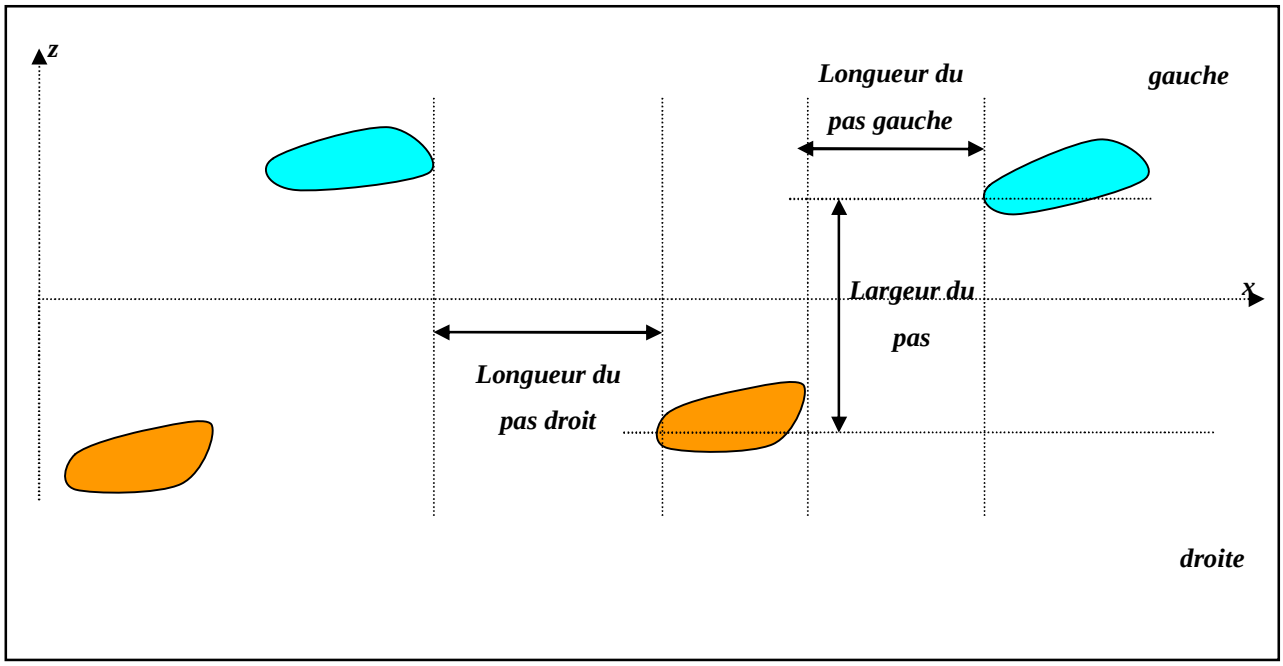


fig. 5.6 : Calcul des longueurs des pas gauche et droit et de la largeur du pas : le calcul des longueurs des pas se fait suivant l'axe du déplacement du corps (***x*** sur la figure), tandis que la calcul de la largeur se fait suivant l'axe perpendiculaire au déplacement (***z*** sur la figure).

Disposant des positions suivant les axes ***x*** et ***z*** des chevilles et des orteils, fournies par IPF et VICON pour chaque image, nous pouvons calculer les longueurs et les largeurs des pas

facilement pour toute image correspondant à un moment d'attaque d'une des deux chevilles. Nous identifions le moment 'd'attaque' d'une cheville comme étant le moment où cette cheville touche le sol. En effet nous disposons des coordonnées des chevilles suivant l'axe vertical d'un repère 3D euclidien absolu de la salle (***y*** dans notre repère). Or, la salle dédiée à l'expérimentation présente une inclinaison du plan du sol. Par conséquent, un point au niveau sol n'a pas une coordonnée nulle suivant ***y*** (fig.5.8). Le niveau du sol étant variable par rapport au repère absolu, nous n'étions pas capables de déterminer directement à partir des coordonnées suivant ***y*** de la cheville, le moment où elle touche le sol. Nous étions obligé de trouver une solution alternative.

En effet, après le moment d'attaque la cheville est immobilisée, et par conséquent garde une élévation constante suivant l'axe vertical. En traçant la courbe représentant la vitesse (obtenue par dérivation) du déplacement vertical de la cheville (fig. 5.7), nous pouvons alors identifier ce moment d'attaque comme étant le moment où cette vitesse devient nulle. Cette solution est générique et peut être appliquée pour les déplacements sur les plans d'inclinaison variable. En effet elle nous évitera d'estimer l'inclinaison du plan de mouvement pour toute séquence et de recalculer les coordonnées en tenant compte de cette inclinaison.

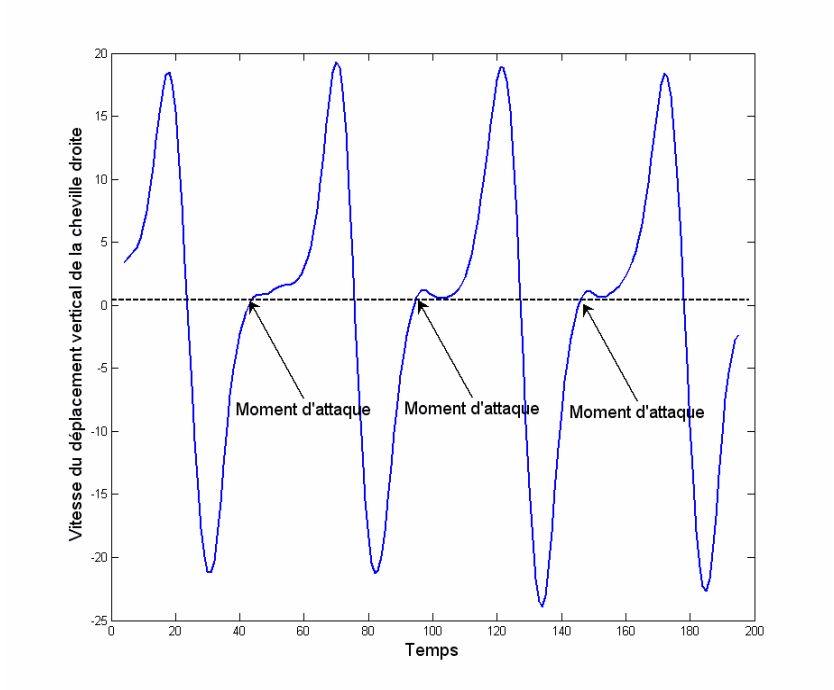


fig. 5.7 : Estimation du moment d'attaque de la cheville droite comme étant le moment où l'élévation verticale de cette cheville se stabilise (vitesse du déplacement vertical nulle).

5.3.3 Hauteurs des pas gauche et droit

La phase de la marche exécutée entre deux moments d'attaque de chevilles gauches (ou droites) est appelée cycle locomoteur gauche (ou droit). La hauteur d'un pas gauche (ou droit) correspond à la distance maximale suivant un axe vertical, séparant la cheville gauche (ou droite) du sol durant un cycle locomoteur gauche (ou droit) respectivement. En effet nous disposons des coordonnées des chevilles suivant l'axe vertical ***y*** de notre repère, mais ces données sont inexploitables directement à cause de l'inclinaison du plan sol ; pour remédier à ce défaut nous avons trouvé une solution générique qui peut être appliquée pour tout plan incliné ; en effet nous avons identifié les moments d'attaque des chevilles (voir paragraphe 5.3.2) et par conséquent le début et la fin de chaque cycle locomoteur. Nous avons alors considéré, pour chaque cycle locomoteur, que la coordonnée absolue suivant l'axe ***y*** du niveau sol est égale à la valeur minimale qu'atteint la cheville durant ce cycle. Ainsi, la hauteur du pas droit (gauche) est calculée comme étant la différence entre la valeur maximale et la valeur minimale qu'atteint la coordonnée suivant ***y*** de la cheville droite (gauche) durant un cycle locomoteur (fig. 5.8).

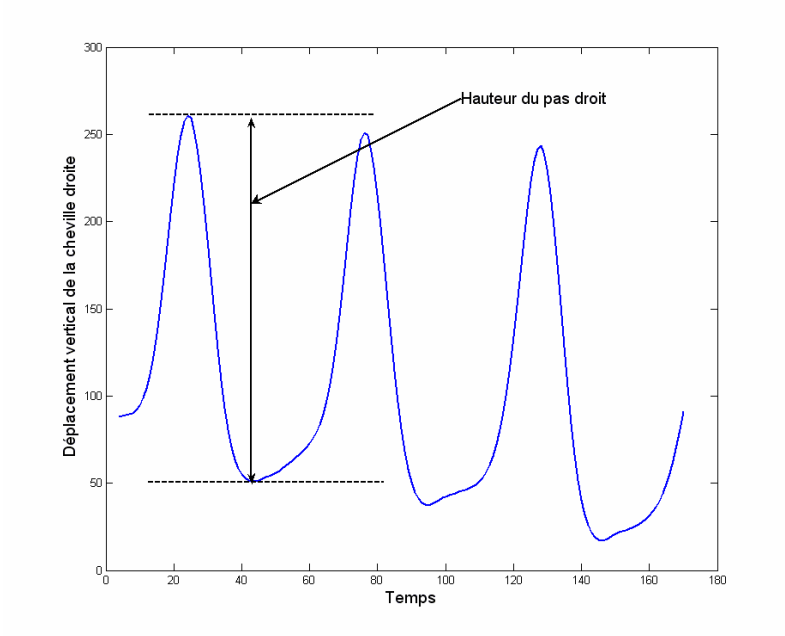


fig. 5.8 : Calcul de la hauteur du pas droit : à cause de l'inclinaison du sol, nous sommes obligés d'estimer pour chaque cycle locomoteur, la hauteur du pas comme étant la différence entre les valeurs maximales et minimales du déplacement vertical de la cheville durant ce cycle.

5.3.4 Régularité et symétrie des pas

Pour évaluer ces deux paramètres (la régularité et la symétrie des pas) il faut calculer les longueurs des pas gauches et droits. En effet, pour faire un constat sur la régularité des pas gauches et celle des pas droits, il suffit de calculer les longueurs de ces pas et de calculer les

variances pour chacune des deux longueurs. Une grande variance (sur au moins six pas) de ces longueurs indiquerait une irrégularité des pas. La symétrie des pas se résume à la comparaison des valeurs moyennes des longueurs des pas gauche et droits. Une grande différence des moyennes entre les pas gauches et droits signifierait une asymétrie des pas.

5.3.5 Oscillation du tronc dans le plan medio-lateral, angles de flexion aux genoux et chevilles

A partir des résultats obtenus par l'algorithme IPF, nous accèdons directement aux angles de flexion aux genoux et aux chevilles et à l'angle d'oscillation du tronc dans le plan medio latéral. En effet ces variables sont des degrés de liberté du modèle virtuel. Pour les résultats fournis par VICON, il faut estimer l'angle d'oscillation du tronc à partir des positions du sacrum et du cou, les angles de flexion aux genoux à partir des positions des genoux et des chevilles, et les angles de flexion aux chevilles à partir des positions des chevilles et des orteils (fig. 5.9). Cette estimation est possible connaissant les longueurs des segments du squelette (sacrum - cou, genou – cheville, cheville –orteil). Or, n'ayant pas accès à ces données pour le modèle adopté dans VICON, nous n'avons pas pu estimer ces paramètres pour ce système.

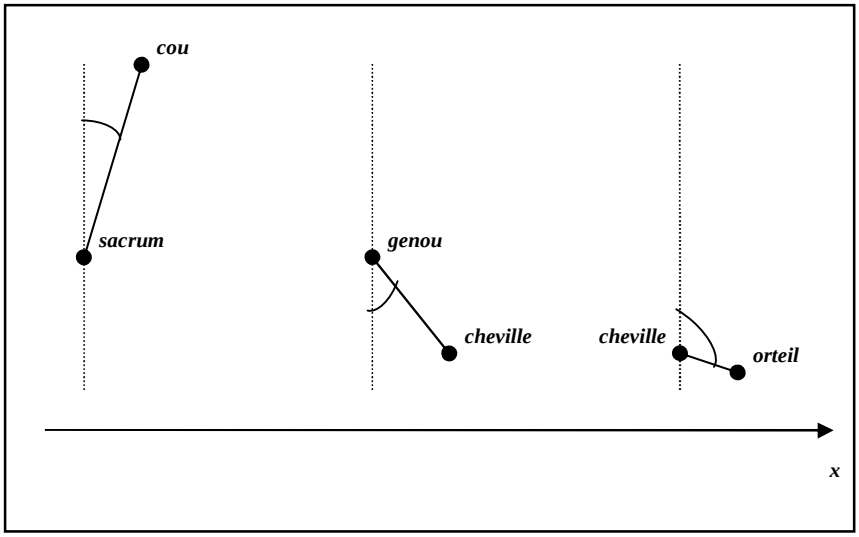


fig. 5.9 : Calcul des angles d'oscillation medio latérale du tronc, de flexion aux genoux et de flexion aux chevilles.

5.3.6 Temps d'appui et du double appui

Durant un cycle de locomotion gauche (ou droit) (fig. 5.10), le temps d'appui gauche (ou droit) est le temps qui s'écoule entre le moment d'attaque (contact avec le sol) de la cheville gauche (droite) et le moment du 'lever' (perte de contact avec le sol) de l'orteil gauche (droit) (fig.5.11).

Le moment d'attaque est calculé suivant la méthode décrite au paragraphe 5.3.2. Le moment du lever de l'orteil qui suit directement sa phase d'appui est estimé comme étant le moment après lequel la coordonnée suivant l'axe vertical de cet orteil n'est plus constante. Ainsi, ce moment de lever correspond à l'instant où la vitesse de déplacement vertical de l'orteil n'est plus nulle (fig. 5.12). Ayant les indices des images (pas du temps) où nous avons identifié les moments d'attaque de la cheville gauche (droite) et du lever de l'orteil gauche (droit), il suffit de calculer la différence temporelle entre ces images (Δt . différence des indices) qui sera égale au temps d'appui gauche (droit).

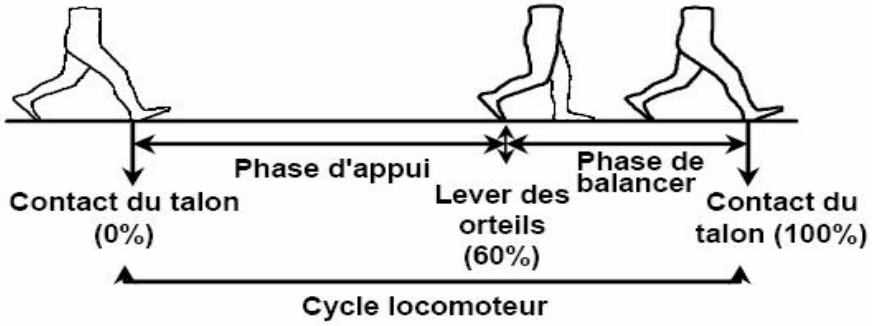


fig. 5.10 : Le cycle de marche par pied est composé de deux phases : une phase d'appui et une phase de balancer

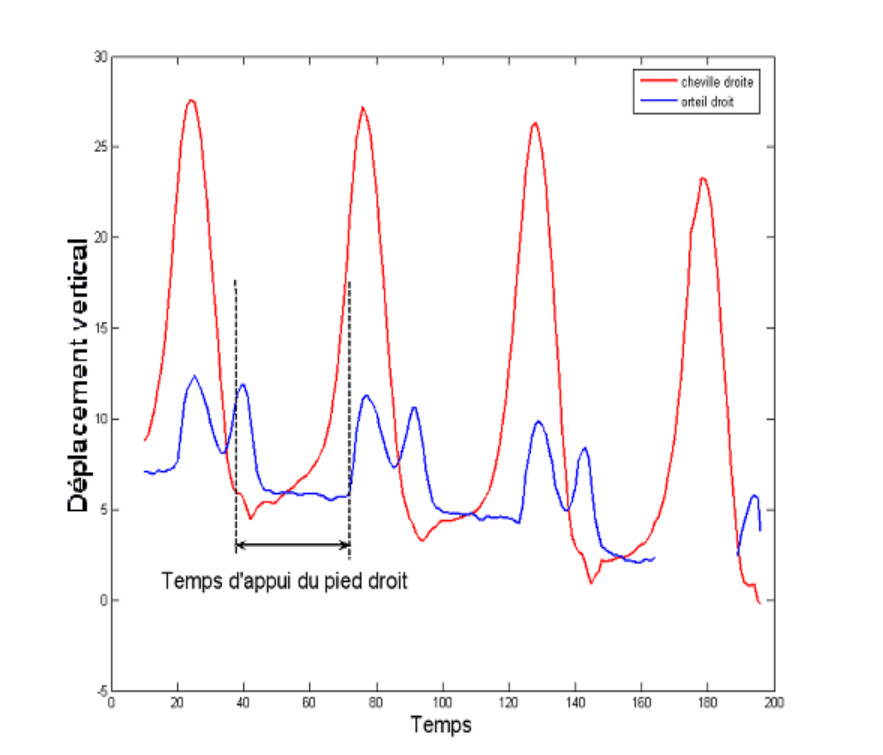


fig. 5.11 : Le temps d'appui d'un pied est le temps qui s'écoule entre le moment d'attaque de la cheville et le moment du lever de l'orteil du même pied.

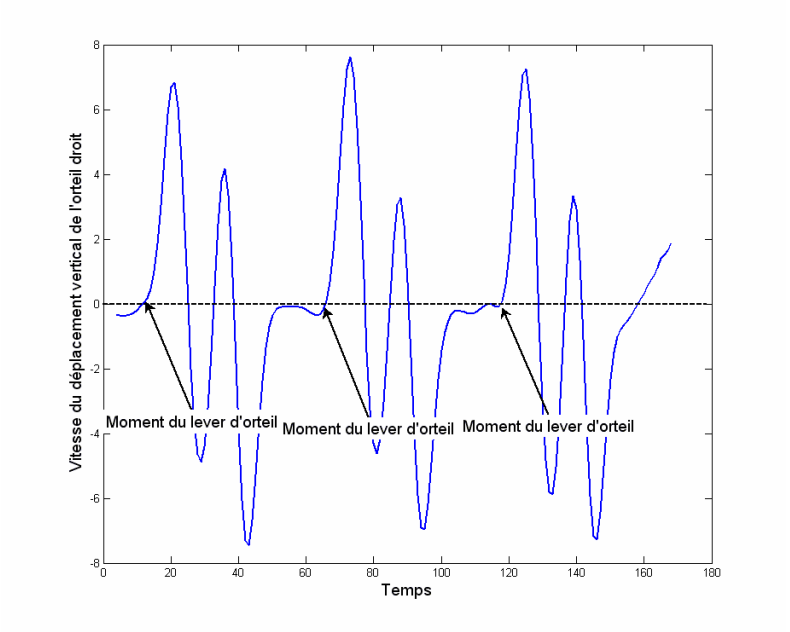


fig. 5.12 : Estimation du moment du lever de l'orteil comme étant le moment qui suit la période durant laquelle l'élévation verticale de cet orteil était stable ; c'est le moment où la vitesse du déplacement vertical de l'orteil commence à croître après avoir été à 0.

Le temps de double appui est la durée durant laquelle les deux pieds gauche et droit sont en phase d'appui. Cette phase de double appui qui est généralement très courte, commence alors au moment du contact de la cheville d'un pied avec le sol (moment d'attaque) et se termine au moment où l'orteil de l'autre pied s'élève du sol (moment du lever) (fig. 5.13). Disposant des moments d'attaque et du lever des deux chevilles et des deux orteils, le calcul de ce temps est alors possible.

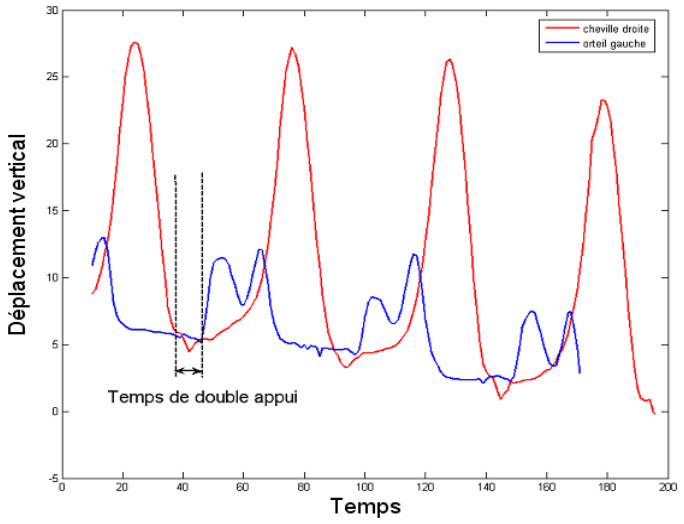


fig. 5.13 : Le temps du double appui est le temps qui s'écoule entre le moment d'attaque de la cheville d'un pied et le moment du lever de l'orteil de l'autre pied.

5.3.7 Déviation du trajet

Pour observer une déviation d'une ligne droite durant la marche, il suffit d'observer le déplacement du point sacrum suivant l'axe perpendiculaire au déplacement (***z*** dans notre repère) (fig. 5.14) dont nous disposons directement. Les critères pour évaluer cette déviation ne sont pas encore fixés et sont soumis à l'appréciation du médecin spécialiste.

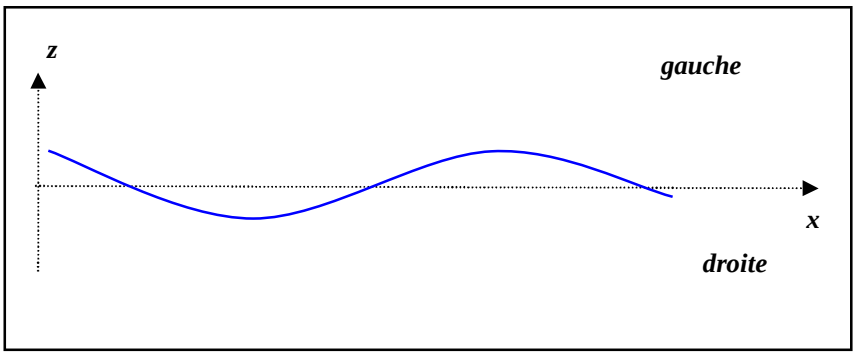


fig. 5.14 : Pour détecter une déviation du trajet, il suffit de tracer la courbe représentant les coordonnées du point sacrum suivant l'axe (***z*** sur la figure) perpendiculaire au déplacement (effectué suivant l'axe ***x***).

5.4 Résultats et comparaisons des paramètres

Nous allons comparer les paramètres de marche auxquels nous nous intéressons, estimés à partir de VICON (système référence) et de notre système. Pour chacun des paramètres nous comparons les moyennes estimées des variables par sujet pour 3 pas et nous calculons l'erreur d'estimation relative (en considérant les résultats de VICON comme la référence) par sujet. Une boite à moustaches décrivant cette erreur d'estimation sur l'ensemble des sujets sera aussi présentée.

5.4.1 Longueur du pas gauche

La longueur du pas gauche est estimée comme étant la distance suivant l'axe de la marche entre la cheville gauche et l'orteil droit au moment du contact cheville gauche sol. Cette longueur a été calculée une fois par cycle locomoteur gauche. Un récapitulatif des valeurs moyennes et des erreurs relatives d'estimation par sujet pour ce paramètre par VICON et par IPF est présenté au tableau 5.1.

L'erreur d'estimation obtenue varie entre -5.34% et 9.67%. Pour décrire la dispersion de l'erreur relative d'estimation, nous utilisons les boites à moustaches (fig. 5.15). Nous remarquons que le premier quartile se situe à -2.98% et le deuxième à 5.11%. La médiane se situe à -0.3% et la moyenne des valeurs absolues des erreurs est de 4.66%.

	Valeur moyenne de la longueur du pas gauche par VICON en cm	Valeur moyenne de la longueur du pas gauche par IPF en cm	Erreur sur l'estimation en %
Sujet 1	38.24	39.99	+4.5764
Sujet 2	36.5	34.6	-5.2055
Sujet 3	37.2	40.8	+9.6774
Sujet 4	41.3	39.2	-5.0847
Sujet 5	34.5	37.69	+9.2464
Sujet 6	28.7	27.89	-2.8223
Sujet 7	41.5	39.28	-5.3494
Sujet 8	39.01	42.26	+8.3312
Sujet 9	35.90	35.79	-0.3064
Sujet 10	34.76	36.66	+5.2934
Sujet 11	41.39	40.31	-2.6093
Sujet 12	39.45	40.35	+2.2814
Sujet 13	50	48.48	-3.0400
Sujet 14	33.5	32.87	-1.8806
Sujet 15	41.58	43.35	+4.2569
Moyenne	38.2353	38.6307	1.15

Tableau 5.1 : Estimation de la valeur moyenne de la longueur du pas gauche par VICON et par IPF, et de l'erreur sur cette estimation en prenant le système VICON comme référence.

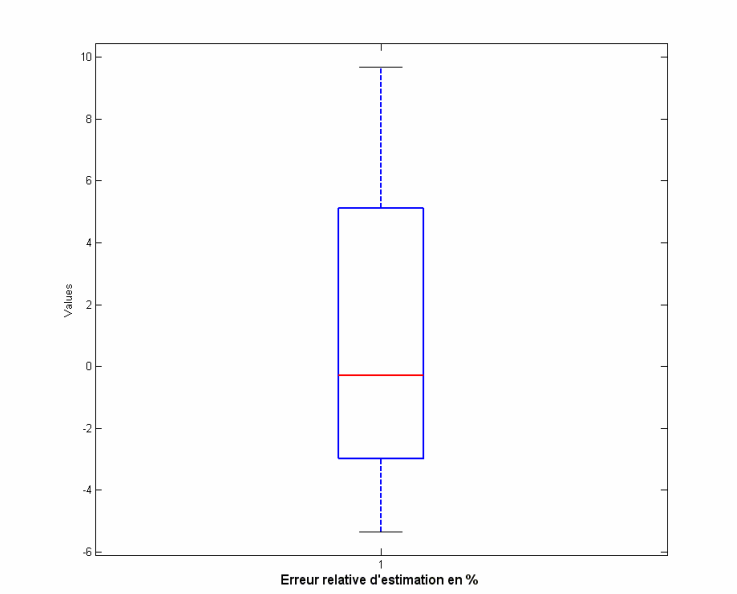


fig. 5.15 : Boite à moustaches représentant l'erreur relative d'estimation de la longueur du pas gauche.

5.4.2 Longueur du pas droit

La longueur du pas droit est estimée comme étant la distance suivant l'axe de la marche entre la cheville droite et l'orteil gauche au moment du contact de la cheville droite avec le sol ; cette longueur a été calculée une fois par cycle locomoteur droit. Un récapitulatif des valeurs moyennes et des erreurs relatives d'estimation par sujet pour ce paramètre par VICON et par IPF est présenté au tableau 5.2.

	Valeur moyenne de la longueur du pas droit par VICON en cm	Valeur moyenne de la longueur du pas droit par IPF en cm	Erreur sur l'estimation en %
Sujet 1	35.47	33.87	-4.51
Sujet 2	35.06	37.54	+7.07
Sujet 3	37.6	41.8	+11.17
Sujet 4	41.5	42.99	+3.59
Sujet 5	31.9	31.81	-0.282
Sujet 6	27.1	27.48	+1.40
Sujet 7	38.1	41.1	+7.874
Sujet 8	41.4	44.8	+8.21
Sujet 9	35.7	35.93	+0.644
Sujet 10	38.62	37.46	-3.00
Sujet 11	39.79	43.51	+9.35
Sujet 12	38.8	38.5	-0.7732
Sujet 13	51.4	47.94	-6.7315
Sujet 14	31.7	30.76	-2.9653
Sujet 15	36.2	34.4	-4.9724
Moyenne	37.35	37.99	1.7385

Tableau 5.2 : Estimation de la valeur moyenne de la longueur du pas droit par VICON et par IPF, et de l'erreur sur cette estimation en prenant le système VICON comme référence.

L'erreur d'estimation obtenue varie entre -6.73% et 11.17%. Pour décrire la dispersion de l'erreur relative d'estimation, nous utilisons les boites à moustaches (fig. 5.16). Nous remarquons que le premier quartile se situe à -2.99% et le deuxième à 7.67% ; la médiane se situe à 0.64% et la moyenne des valeurs absolues des erreurs est de 4.83%.

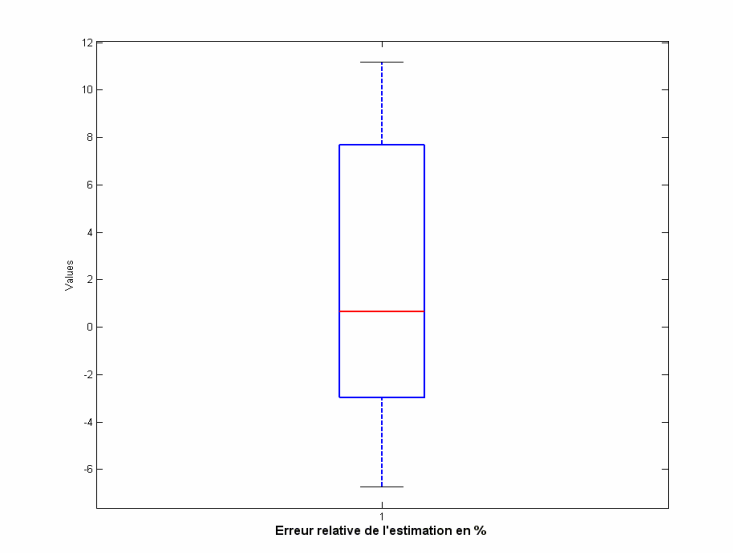


fig. 5.16 : Boite à moustaches représentant l'erreur relative d'estimation de la longueur du pas droit.

5.4.3 Largeur du pas

La largeur du pas du pas est la distance suivant l'axe perpendiculaire à l'axe de la marche entre les chevilles (droite et gauche) au moment du contact de l'une d'elle avec le sol. Cette

largeur a été calculée deux fois par cycle locomoteur. Un récapitulatif des valeurs moyennes et des erreurs relatives d'estimation par sujet pour ce paramètre par VICON et par IPF est présenté au tableau 5.3.

	Valeur moyenne de la largeur du pas par VICON en cm	Valeur moyenne de la largeur du pas par IPF en cm	Erreur sur l'estimation en %
Sujet 1	22.3	21.3	-4.48
Sujet 2	20.01	20.63	+3.09
Sujet 3	21.97	20.36	-7.32
Sujet 4	16.38	15.01	-8.36
Sujet 5	21.46	23.19	+8.06
Sujet 6	22.75	21.81	-4.13
Sujet 7	16.96	17.86	+5.3
Sujet 8	19.30	18.62	-3.52
Sujet 9	16.1	15.06	-6.4596
Sujet 10	26.20	25.51	-2.6336
Sujet 11	19.67	21.64	+10.015
Sujet 12	18.07	17.52	-3.0437
Sujet 13	18.99	21.02	+10.69
Sujet 14	16.2	15.4	-4.93
Sujet 15	20.62	21.25	+3.05
Moyenne	19.79	19.74	-0.312

Tableau 5.3 : Estimation de la valeur moyenne de la largeur du pas par VICON et par IPF, et de l'erreur sur cette estimation en prenant le système VICON comme référence.

L'erreur d'estimation obtenue varie entre -8.36% et 10.69 %. Pour décrire la dispersion de l'erreur relative d'estimation, nous utilisons les boites à moustaches (fig. 5.17). Nous remarquons que le premier quartile se situe à -4.82% et le deuxième à 4.75%. La médiane se situe à -3 % et la moyenne des valeurs absolues des erreurs est de 5.67%.

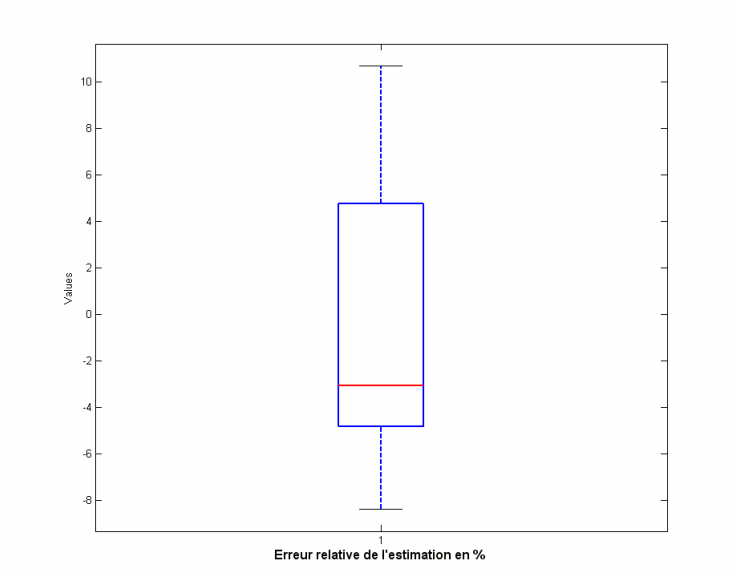


fig. 5.17 : Boite à moustaches représentant l'erreur relative d'estimation de la largeur du pas.

5.4.4 Vitesse instantanée

A chaque instant ***t*** nous calculons la vitesse du déplacement du tronc. Un récapitulatif des valeurs moyennes et des erreurs relatives d'estimation par sujet pour ce paramètre par VICON et par IPF est présenté au tableau 5.4.

	Valeur moyenne de la vitesse instantanée par VICON en cm/s	Valeur moyenne de la vitesse instantanée par IPF en cm/s	Erreur sur l'estimation en %
Sujet 1	129.5	129.46	-0.03
Sujet 2	124.61	124.04	-0.45
Sujet 3	129.76	124.05	-4.40
Sujet 4	137.32	132.70	-3.36
Sujet 5	123.98	116.65	-5.91
Sujet 6	107.53	98.21	-8.66
Sujet 7	138.159	130.90	-5.25
Sujet 8	125.01	112.10	-10.32
Sujet 9	115.62	112.58	-2.61
Sujet 10	114.23	105.64	-7.5
Sujet 11	137.7	132.4	-3.8
Sujet 12	129.83	131.43	+1.23
Sujet 13	154.73	158.17	+2.22
Sujet 14	116	108	-6.89
Sujet 15	114.55	104.90	-8.42
Moyenne	126.5686	121.4153	-4.28

Tableau 5.4 : Estimation de la valeur moyenne de la vitesse instantanée par VICON et par IPF, et de l'erreur relative sur cette estimation en prenant le système VICON comme référence.

L'erreur d'estimation obtenue varie entre -10.32% et 2.22 %. Pour décrire la dispersion de l'erreur relative d'estimation, nous utilisons les boites à moustaches (fig. 5.18). Nous remarquons que le premier quartile se situe à -7.36% et le deuxième à -1%. La médiane se situe à -4.4 % et la moyenne des valeurs absolues des erreurs est de 4.2%.

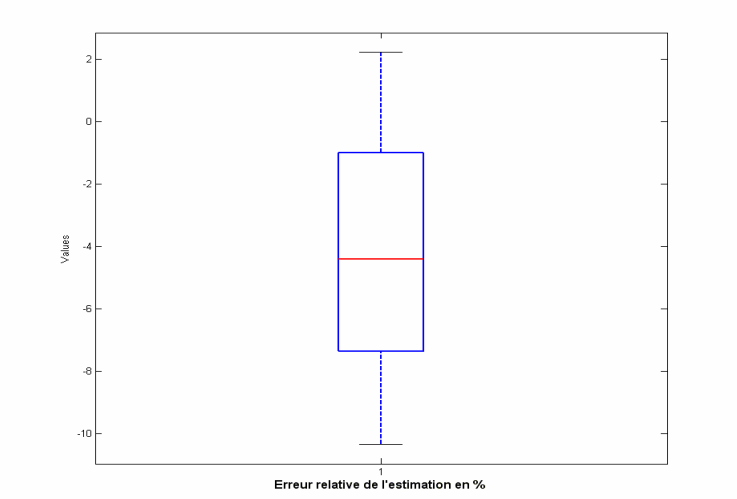


fig. 5.18 : Boite à moustaches représentant l'erreur relative d'estimation de la vitesse instantanée du tronc durant la marche.

5.4.5 Temps d'appui gauche, droit et de double appui

Les mouvements de contact des chevilles avec le sol et du lever des orteils en début de phase de balancement sont des mouvements rapides. Pour les détecter nous nous basons sur les résultats du suivi 3D de ces articulations. Or, la fréquence des images captées est de 25 images/sec pour notre système et de 50 Hz pour VICON. Par conséquent il y aura toujours une marge constante d'erreur dans l'estimation des temps d'appui allant jusqu'à 80 ms pour notre approche et de 40ms pour VICON. D'autre part, nous risquons de ne pas détecter la phase de double appui, ni par IPF ni par VICON, celle-ci étant en général très brève. Vu l'imprécision sur les données de référence, aucune erreur relative ne sera calculée. Dans le tableau 5.5 nous présenterons les résultats obtenus par IPF et VICON pour l'estimation des moyennes de ces paramètres.

	Temps d'appui par VICON en ms			Temps d'appui par IPF en ms			Erreur d'estimation en ms		
	Gauche	Droit	Double	Gauche	Droit	Double	Gauche	Droit	Double
Sujet 1	520	520	20	480	440	40	-40	-80	+20
Sujet 2	520	480	20	440	520	10	-80	+40	-10
Sujet 3	540	560	60	520	480	40	-20	-80	-20
Sujet 4	460	400	20	480	360	0	+20	-40	-20
Sujet 5	540	520	80	480	480	40	-60	-40	-40
Sujet 6	480	540	20	380	520	40	-100	-20	+20
Sujet 7	500	380	0	400	360	0	-100	-20	0
Sujet 8	560	320	40	480	280	20	-80	-40	-20
Sujet 9	600	580	60	520	480	80	-80	-100	+20
Sujet 10	560	540	20	520	640	0	-40	+100	-20
Sujet 11	480	400	40	400	400	40	-80	0	0
Sujet 12	500	540	0	420	480	0	-80	-60	0
Sujet 13	440	280	60	400	360	40	-40	+80	-20
Sujet 14	560	500	80	480	440	80	-80	-60	0
Sujet 15	680	700	20	640	600	40	-40	-100	-20

Tableau 5.5 : Estimation des valeurs moyennes des temps d'appui gauche, droit et double. A cause de la nature rapide des mouvements en question, les valeurs de référence (VICON) seront déterminées avec une précision de 40 ms. De même les valeurs fournies par IPF sont données avec une précision de 80ms.

5.4.6 Comparaison des mouvements de quelques points clés

En plus des comparaisons sur les paramètres de la marche, nous allons présenter la reconstitution des mouvements de certains points du corps par IPF en comparaison avec VICON durant une marche aléatoire (ne suit pas un axe donné) (fig.5.19). Les courbes représentant les positions des points suivis par IPF et VICON prennent des formes identiques, ce qui indique une bonne reconstruction par notre approche.

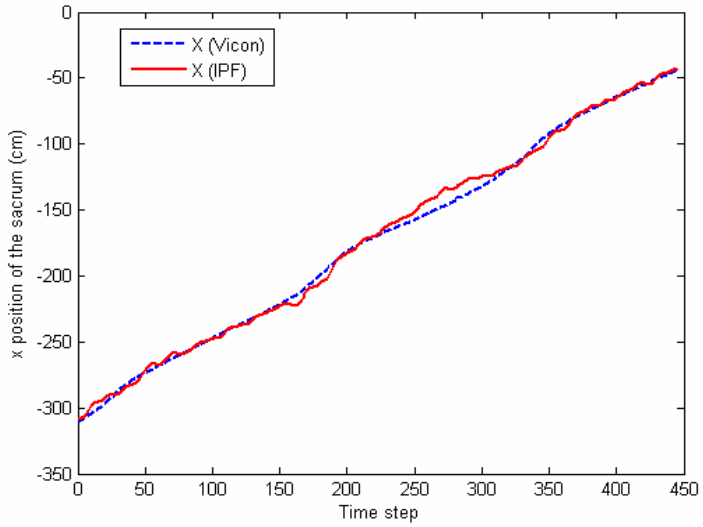


fig. 5.19 (a) : Comparaison du déplacement suivant un axe transversal x du point sacrum du corps. Les deux courbes présentent la même forme. Par contre, la courbe en pointillé représentant les résultats de VICON est plus lisse que la courbe des positions par IPF. Ceci est dû à la fréquence moins élevée de notre système (25 Hz contre 50 Hz) et au prétraitement des données par VICON.

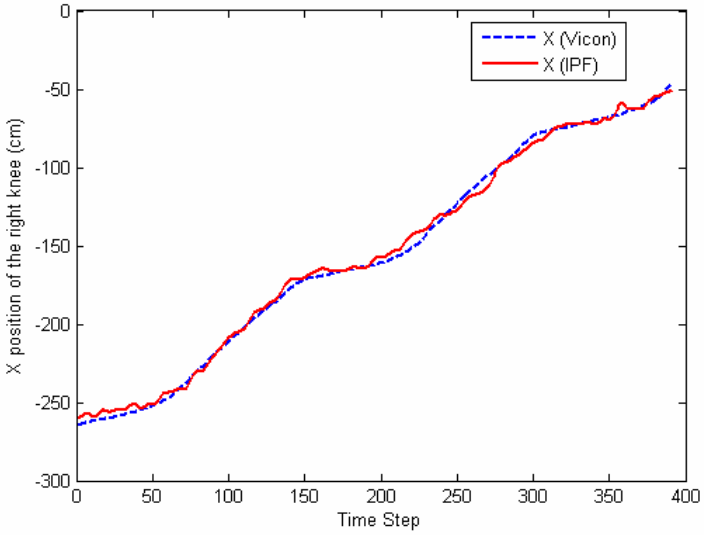


fig. 5.19 (b) : Déplacement transversal du genou droit obtenu par VICON (en pointillé) et IPF (en ligne pleine).

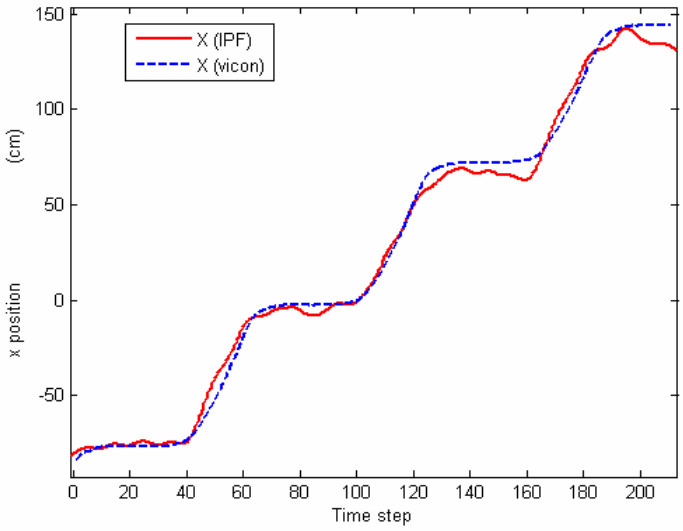


fig. 5.19 (c) : Comparaison du déplacement par VICON (en pointillé) et IPF (en ligne pleine) suivant un axe transversal x de la cheville droite.

5.4.7 Discussion

Pour les longueurs et les largeurs des pas nous remarquons que l'erreur d'estimation ne dépasse pas les 10 %. Cette erreur peut être expliquée par la fréquence d'acquisition des images utilisées par notre système. En effet, les calculs de ces paramètres se font à partir des positions des chevilles, estimées aux moments de leur contact avec le sol. Par conséquent, l'erreur dans l'estimation des moments de contact peut entraîner une erreur dans le calcul des longueurs et de la largeur des pas. Or, l'estimation de ces instants de contact est plus précise pour les données de VICON, ayant une fréquence d'acquisition deux fois plus grande que celle de notre système. Pour cela, le calcul des paramètres est plus précis en utilisant les données de VICON.

D'autre part, n'ayant pas accès au modèle virtuel utilisé par VICON pour déterminer les positions 3D fournies à partir des marqueurs, nous avons tenté de l'estimer de façon à suivre les mêmes points par les deux systèmes. Ce facteur peut influencer les résultats et introduire une marge d'erreur. En général, en utilisant plusieurs caméras avec une vitesse d'acquisition plus rapide les résultats seront améliorés et l'erreur diminuée.

La vitesse instantanée est calculée à partir des positions du point sacrum. La position de ce point constituant le point 'origine' de notre modèle est calculée à partir de l'image 'silhouette' à l'aide d'une recherche hiérarchique. L'erreur obtenue étant de 4% en moyenne, nous pouvons dire que cette méthode donne des résultats satisfaisants, malgré sa simplicité.

Les temps d'appui sont estimés par VICON avec une précision de 40 ms (durée de 2 images). Les temps d'appui fournis par IPF seront forcement moins précis, la différence

temporelle entre deux images successives étant de 40ms. L'erreur sur l'estimation du moment de contact de la cheville ou du moment de relève de l'orteil peut atteindre la durée d'une image, et par conséquent, en estimant mal les deux instants nous pouvons atteindre 80 ms d'erreur. Les résultats obtenus (tableau 5) montrent que la différence entre les temps estimés est en moyenne de 40 à 60 ms ce qui constitue déjà l'erreur d'estimation du VICON. En conclusion, nous pouvons déduire que les temps de référence estimés par VICON ne sont pas très précis et restent proches des valeurs estimées par IPF (erreur moyenne de 12%). La comparaison du suivi de quelques points clés du corps (sacrum, cheville, genou etc..) montre une similitude des positions fournies par les deux systèmes, avec une précision plus importante dans le système VICON.

5.5 Reconstruction 3D

En plus des comparaisons au système VICON, effectuées par rapport aux paramètres, nous avons appliqué les algorithmes développés dans le but d'une reconstruction 3D du mouvement. En effet, nous avons choisi un certain nombre de manœuvres complexes à titre d'illustration, pour tester la capacité de notre système à suivre le mouvement effectué en utilisant un seul caméscope numérique classique (fréquence de 25 images/sec). N'ayant pas de données de référence pour ces manœuvres, la qualité de la reconstruction est à observer à partir des images résultantes.

5.5.1 Marche rapide avec occlusion

L'occlusion d'une partie du corps est fréquente durant la marche, surtout si la caméra est placée perpendiculairement au sens du déplacement. L'algorithme IPF permet de surmonter cette difficulté en laissant survivre un nombre constant de configurations possibles et ainsi, après une occlusion la bonne configuration est toujours présente dans l'ensemble de particules candidates et le suivi reste correct (fig. 5.20).



t = 0 ***t = 0.5s*** ***t = 1s*** ***t = 1.5s***

fig. 5.20 : Reconstitution du mouvement 3D durant une marche rapide avec des occlusions temporaires de la jambe gauche et du bras gauche. Nous remarquons que l'algorithme développé arrive à bien suivre le mouvement des bras et jambes malgré les occlusions.

5.5.2 Levée d'une chaise

L'une des manœuvres intéressantes qui a aussi été étudiée, est la levée d'une chaise. En effet, ce mouvement donne une indication sur l'état d'équilibre de la personne. Au niveau algorithmique, la difficulté pour le suivi réside dans la détection du mouvement du point 'origine' du corps dont les coordonnées 3D changent rapidement et simultanément durant cette phase. D'autre part, l'orientation du corps, elle aussi, varie rapidement, ainsi que les mouvements du tronc et des membres inférieurs (fig. 5.21).



t = 0 ***t = 1s*** ***t = 2s*** ***t = 3s***

fig. 5.21 : Reconstitution du mouvement 3D d'une personne se levant d'un chaise pour changer d'orientation et marcher. A l'instant 0 (la première image) nous remarquons que l'algorithme utilisé réussit à trouver la bonne configuration initiale du corps. Le mouvement du tronc et des membres inférieurs est bien suivi et estimé durant la levée et dans la phase d'initialisation de la marche.

5.5.3 Attraper un objet par terre

Pour cet essai, nous avons demandé à la personne d'attraper un objet par terre durant la marche normale. Cet essai permet de montrer la capacité de l'algorithme à bien localiser le tronc et les bras durant le mouvement. D'autre part, lorsque la personne se baisse pour atteindre le sol, le corps prend une forme compacte et par conséquent une grande partie du corps sera cachée. Le suivi dans ces images là devient alors moins exact. Par contre, malgré la longue période d'occlusion, l'algorithme IPF permet de bien récupérer le mouvement lorsque la personne se lève (fig. 5.22).



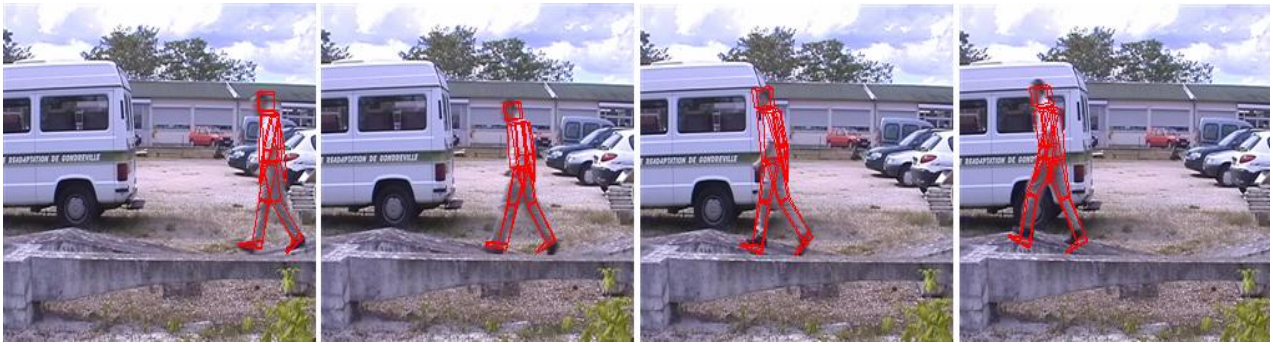
t = 0 ***t = 1s*** ***t = 2s*** ***t = 3s*** ***t = 4s***

fig. 5.22 : Suivi d'une personne rattrapant un crayon par terre durant sa marche ; le mouvement durant la descente est bien représenté. Par contre, le corps prend une forme compacte lorsque la personne se baisse, ce qui rend l'estimation des mouvements des jambes peu exacte (à ***t = 2s***).

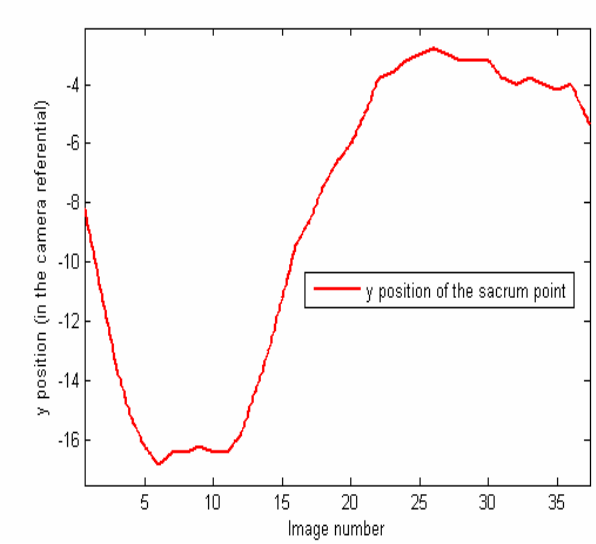
Malgré cette période d'occlusion, l'algorithme arrive toujours à trouver la bonne configuration quand le corps commence à reprendre sa forme normale.

5.5.4 Suivi en plein air

Contrairement au système VICON qui nécessite un étalonnage et un environnement qui lui y est dédié, notre système composé d'une seule caméra peut être utilisé pour le suivi du mouvement effectué en plein air (fig. 5.23 et 5.24). En effet cette propriété est intéressante vu qu'à l'extérieur nous ne sommes pas contraints par un trajet fixe et que l'espace permettant le mouvement est plus important, ce qui donne une liberté de mouvement et un aspect plus naturel à la marche que dans un espace contraint 'contrôlé'.



t = 0 ***t = 0.5s*** ***t = 1s*** ***t = 1.5s***



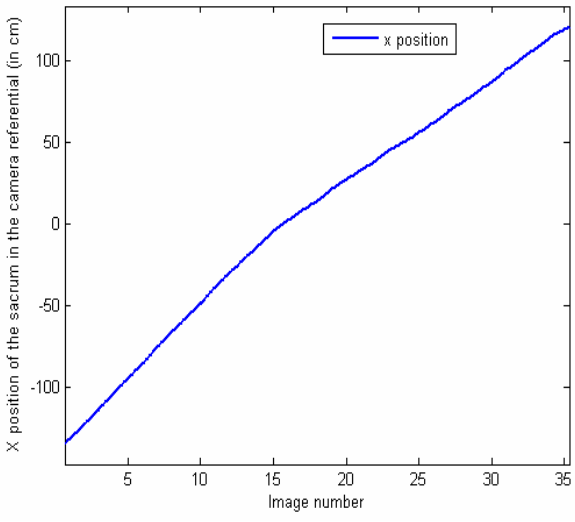
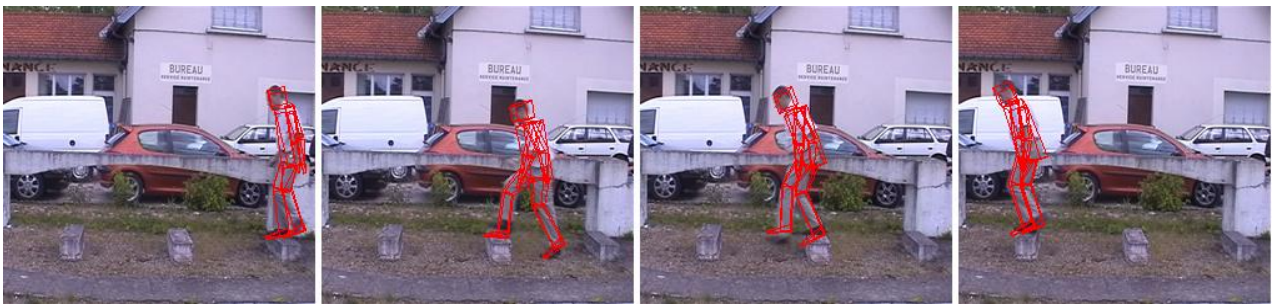


fig. 5.23 : Mouvement en plein air sur un chemin à multi niveaux ; l'algorithme IPF montre des résultats satisfaisants même avec des images prises dans un environnement extérieur. En observant la courbe du déplacement en hauteur (suivant l'axe ***y*** dans notre repère) du point sacrum du corps nous pouvons déduire que la personne bouge vers un niveau plus bas puis vers un niveau plus haut etc. Par conséquent, cette courbe combinée à la courbe de déplacement longitudinal (suivant ***x***) nous donne une indication sur la nature de l'activité que la personne effectue.



t = 0 ***t = 0.5s*** ***t = 1s*** ***t = 1.5s***

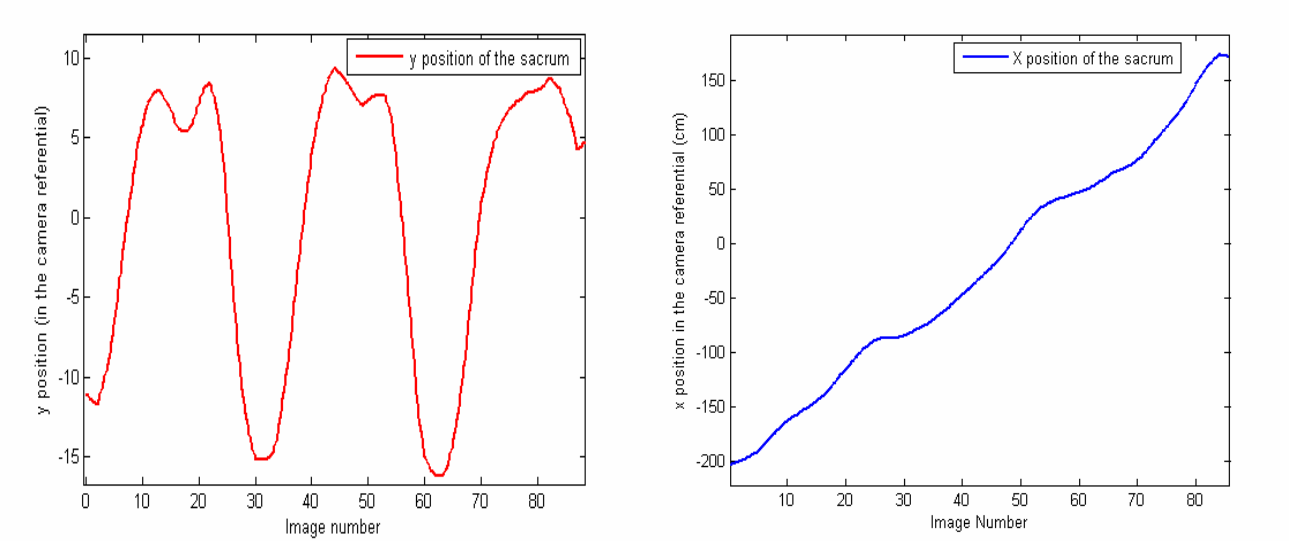


fig. 5. 24 : Mouvement en plein air ; la personne monte et descend des marches. Ce mouvement peut être détecté à l'aide des courbes du déplacement en hauteur (suivant ***y***) et longitudinal (suivant ***x***) présentées ci dessus.

5.5.5 Simulation d'une amputation

Le système développé a été appliqué pour détecter une anormalité de marche chez une personne simulant une amputation de jambe droite. En comparant les angles de flexion – extension des deux genoux gauche et droit (fig. 5.25) nous remarquons que les amplitudes des deux angles sont différentes et que le genou droit est bloqué, d'où la déduction d'une amputation du jambe droit.



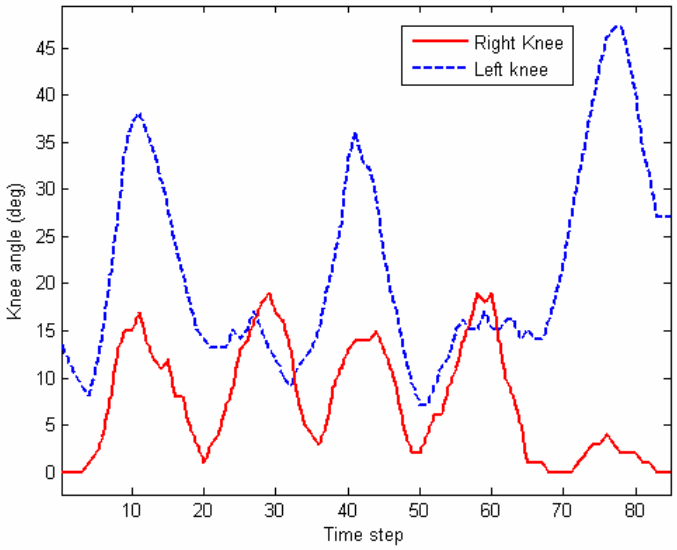


fig. 5.25 : Simulation d'une amputation : en comparant les courbes représentant les angles de flexions du genou gauche (en pointillé) et du genou droit (en plein) nous remarquons que l'amplitude maximale de la flexion du genou droit est petite par rapport à l'amplitude normale d'un genou en flexion. Par conséquent nous pouvons déduire une anormalité dans le mouvement de la jambe droite.

5.6 Complexité de l'algorithme et temps de calcul

Les résultats présentés prouvent que nous pouvons déterminer les paramètres de marche que nous recherchons, avec une erreur moyenne de 5%. La reconstitution 3D, à partir d'une seule caméra, de quelques mouvements complexes différents de la marche normale, a montré la capacité de notre algorithme et son efficacité.

En revanche, le nombre de particules utilisées (5850) implique à chaque instant ***t*** la création d'un grand nombre d'images et la comparaison de chacune à l'image 'silhouette'. Par conséquent, le temps nécessaire pour trouver une bonne estimation de la configuration des 31 degrés de liberté du corps est de 6s par image. La fréquence d'acquisition d'image étant de 25 Hz, le temps de calcul et de traitement pour une seconde de marche est de 150s. Or, le traitement ne doit pas être forcément effectué en temps réel. Par conséquent le temps de traitement reste raisonnable. En revanche, si nous souhaitons diminuer ce temps de calcul dans le but de créer un système de suivi 3D du mouvement en temps réel, il devient nécessaire de diminuer le temps de calcul par image. Ce système temps réel pourrait alors être utilisé pour d'autres applications, dans le domaine des interactions homme – machine par exemple.

La solution la plus simple pour réduire le temps de calcul est de réduire le nombre d'images impliquées en réduisant le nombre de particules utilisées. Or, comme pour tout algorithme

particulaire, la précision du suivi est proportionnelle au nombre des particules impliquées. Par conséquent, l'utilisation d'un nombre plus petit de particules entraîne des résultats moins précis (fig. 5.26) surtout que nous travaillons avec un vecteur d'état de grande dimension (31). Cette solution a été alors rejetée. La solution adoptée aborde la diminution du temps de calcul sur deux plans : sur le plan des technologies et outils utilisés et sur le plan de la conception de l'algorithme d'IPF.

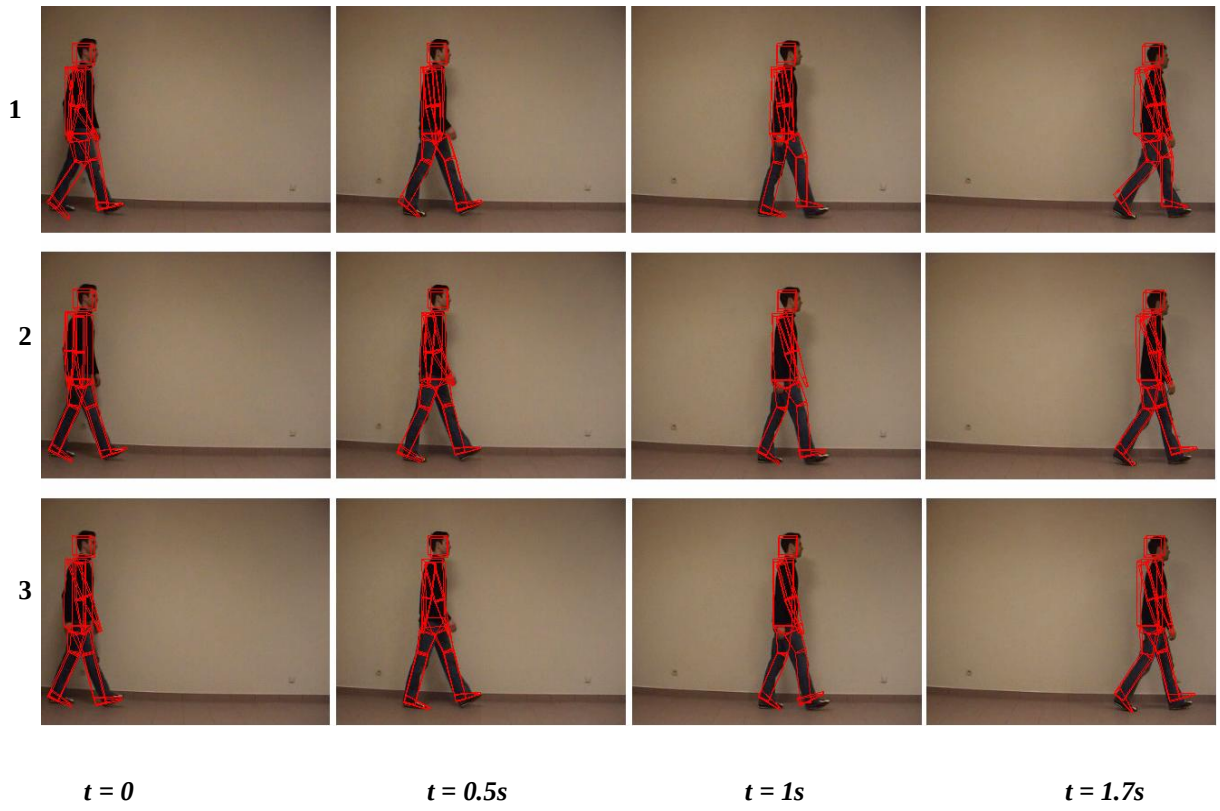


fig. 5.26 : Comparaison du suivi du mouvement effectué avec IPF pour la même séquence vidéo, en utilisant 220 particules (ligne 1), 1080 particules (ligne 2) et 4536 particules (ligne 3). Cet exemple confirme que l'utilisation d'un nombre plus important de particules fournit un résultat plus précis, ce qui est la propriété de tout algorithme particulaire.

5.6.1 Amélioration au niveau des technologies

On remarque que l'opération de projection du modèle 3D humanoïde par la caméra virtuelle sur un plan 2D occupe une grande partie de la durée de traitement et de calcul. Or, ce genre de projection peut être effectué d'une manière optimisée en utilisant les cartes graphiques qui, en principe, sont construites pour ce but. L'idée est alors d'utiliser la librairie OpenGL (Open Graphics Lab) pour effectuer les projections 3D/2D à l'aide du processeur de la carte graphique (le GPU) et non plus à l'aide du processeur de l'unité centrale (CPU). En effet, les premiers tests comparatifs ont permis d'entrevoir de réelles avancées en termes de performances. Le temps nécessaire pour projeter le modèle 3D à l'aide du CPU étant de 14 ms, nous avons réussi à

réduire ce temps à 3ms en utilisant le processeur d'une carte graphique de base. L'utilisation d'une carte graphique plus puissante et plus récente nous permettrait éventuellement de réduire encore ce temps.

D'autre part, en utilisant la carte graphique pour effectuer les calculs des projections, nous libèrons le CPU qui devient disponible pour effectuer d'autres calculs et traitements. Cette propriété impose la réorganisation de l'architecture des programmes développés. En effet, pour profiter de toute la puissance de calcul il serait judicieux de faire fonctionner les deux processeurs graphique et central en parallèle. Dans cette perspective, une architecture de programmation multi taches (Multi Thread) s'impose et sera à implémenter dans le futur.

5.6.2 Amélioration au niveau de l'algorithme développé

La réduction de la complexité peut aussi être apportée à travers l'algorithme FIPF que nous avons détaillé dans le chapitre précédent. En effet, par comparaison à IPF, cet algorithme devrait fournir des résultats de même précision mais avec un nombre plus petit de particules ce qui signifie un temps de calcul plus réduit. En effet, les factorisations du vecteur d'état et de la fonction d'observation permet d'introduire l'observation dans le processus de création de particules et ainsi éliminer les moins probables.

5.7 Conclusion

Dans ce chapitre nous avons présenté les résultats obtenus pour le suivi 3D du mouvement humain à partir l'algorithme IPF. Après un descriptif du matériel utilisé et de l'environnement dans lequel était effectué ce suivi, nous avons détaillé les méthodes utilisées pour extraire les paramètres de marche à partir des positions 3D des points suivis. Une description de la phase de prétraitement des données et de l'initialisation de l'algorithme IPF a aussi été présentée. Un choix des paramètres de réglage de cet algorithme a été fait et justifié également.

Une comparaison sur les moyennes des paramètres calculés à partir du suivi par VICON et par IPF montre que l'erreur relative sur l'estimation des paramètres ne dépasse pas les 10%. Ces résultats sont satisfaisants et encourageants. D'autre part l'algorithme développé a été testé avec succès pour le suivi du mouvement durant des manœuvres complexes (se lever d'une chaise, attraper un objet par terre etc.). Par ailleurs, en utilisant des séquences filmées en plein air, nous avons montré que le système que nous avons développé est capable de bien reconstruire le mouvement 3D dans n'importe quel environnement à l'opposé des systèmes à marqueurs qui nécessitent un étalonnage et un espace fermé qui leur est dédié.

L'erreur relative de mesure des paramètres par IPF est de 5% en moyenne. Pour avoir plus de précision dans les résultats obtenus, plusieurs solutions sont à envisager. L'utilisation d'un nombre plus grand de caméras améliore les résultats mais augmente le coût du système et implique une difficulté pratique reliée à la synchronisation des images obtenues par les deux caméras. Une meilleure estimation peut être apportée par l'utilisation de caméras fonctionnant à des fréquences plus importantes. Or, ces caméras sont chères et ne sont pas considérées comme des produits grand public, ce qui est en contradiction avec les exigences du projet PARAChute.

En conséquence, une amélioration de la précision du système passe forcément par une augmentation du nombre des particules utilisées. Or, l'utilisation de plus de particules introduit plus de complexité à l'algorithme IPF. L'algorithme FIPF que nous avons détaillé dans le chapitre précédent fournit la solution pour ce problème. En effet il doit fournir des résultats de même précision que IPF mais en utilisant un nombre plus petit de particules. Le temps de calcul peut lui aussi être réduit, en profitant de la puissance de calcul des cartes graphiques pour effectuer les projections 3D.

Chapitre 6 : Conclusions générales et perspectives

6.1 Motivations, objectifs et contraintes

Les chutes des personnes âgées à leur domicile constituent un problème majeur de santé publique à cause de leur fréquence et de leurs conséquences médicales, sociales et financières. En effet, le nombre de décès liés à ces chutes est estimé à 12000 décès par an en France. D'autre part, les chutes entraînent les sujets âgés dans la dépendance et ont un coût de traitement médical élevé. Cependant, ce problème devient encore plus important avec le vieillissement observé de la population française.

Le risque de chute peut être réduit chez une personne âgée grâce à l'intervention d'un acteur de santé. Motivé par ce constat, le projet PARAChute qui constitue le cadre général de ce travail de thèse, s'inscrit dans la logique d'une prévention primaire de chute chez les personnes âgées à domicile. Les approches déjà adoptées dans ce domaine se basent sur des tests cliniques effectués à l'hôpital où les patients sont dans des conditions modifiées par rapport aux conditions habituelles de leur vie quotidienne. L'objectif principal du projet PARAChute est alors de proposer un système permettant de détecter chez la personne âgée à domicile une détérioration du comportement locomoteur signifiant un risque de chute.

Le système que nous voulions concevoir se devait d'être simple et adapté aux conditions de vie quotidienne à domicile. Pour cela, plusieurs exigences devaient être respectées :

- ne pas perturber l'environnement de vie usuel de la personne suivie ;
- l'évaluation du risque de chute doit s'appuyer sur l'observation des activités quotidiennes de la personne ;
- la présence d'une tierce personne pour accomplir le protocole d'évaluation ne devait pas être obligatoire ;

- l'information fournie par le système devrait être réduite au strict nécessaire, de façon à préserver l'intimité de la personne, quelques soient les données recueillies.

L'objectif principal de cette thèse était la conception et le développement d'un système capable d'évaluer un ensemble de paramètres permettant d'analyser l'équilibre dynamique (durant la marche) d'une personne à partir de ses activités quotidiennes à domicile. Ce système fera partie du dispositif imaginé dans le cadre du projet PARAChute pour évaluer le risque de chute chez une personne âgée. Pour atteindre cet objectif plusieurs sous objectifs initiaux étaient alors à réaliser :

- identifier les paramètres et les tests utilisés pour évaluer le risque de perte de l'équilibre chez une personne âgée ;
- évaluer la compatibilité de chacun de ces paramètres et méthodes avec les exigences du projet PARAChute (simplicité du test, objectivité, absence d'une tierce personne pour l'effectuer) ;
- trouver les technologies et les capteurs permettant d'évaluer ces paramètres en respectant les contraintes de l'étude.

6.2 Synopsis de la thèse

Pour identifier les paramètres permettant de déceler une dérive vers la chute, nous avons mené avec l'aide des médecins gériatres une étude bibliographique sur les approches déjà utilisées dans ce domaine. En effet plusieurs techniques ont été développées pour évaluer l'équilibre dynamique d'une personne. Chacune de ces méthodes était étudiée à deux niveaux : Au niveau technologique et au niveau 'clinique'. Ces approches peuvent être regroupées en trois grandes familles : les tests cliniques, les approches ambulatoires et les parcours instrumentés.

Durant la visite médicale, les gériatres soumettent les patients à un bilan gériatrique complet comprenant une évaluation de l'équilibre. Celle ci est accomplie à travers des questionnaires et des tests bien définis. Les questionnaires portent sur l'historique des chutes, les traitements suivis, les antécédents médicaux et l'état physique actuel de la personne. Cependant, ces facteurs augmentent la susceptibilité aux chutes mais ne donnent pas une indication fiable sur le risque de chute. Les tests cliniques (Berg, Tinetti, DGI etc..) auxquels sont soumis les patients, évaluent leur capacité à effectuer certaines activités ou mouvements particuliers à un instant donné et dans un contexte particulier. Or cette évaluation de l'aptitude de chaque personne est à déterminer par un médecin et par conséquent elle est subjective. D'autre part, certaines des manœuvres incluses

dans ces tests sont dangereuses à faire effectuer par une personne vivant seule, sans l'aide d'un personnel médical ou paramédical. Pour toutes ces raisons, la majorité des tests cliniques était exclue dans notre approche. Seules les mesures objectives et simples (vitesse, longueur des pas etc..) ont été retenues en tant que paramètres pertinents.

Avec les avancées de la microélectronique, les capteurs sont devenus de plus en plus petits et portables d'où le développement de méthodes fondées sur l'utilisation de capteurs portables pour évaluer l'équilibre dynamique. Ces méthodes, au contraire des tests cliniques, utilisent des mesures objectives et fiables. Parmi les capteurs utilisés dans ces approches nous trouvons les accéléromètres, les goniomètres, les gyroscopes et les semelles équipées de capteurs de pression. Tous ces capteurs portables nécessitent un câblage spécial pour les relier à une centrale d'acquisition portable et par conséquent deviennent encombrants à porter pour la personne qu'on veut suivre. D'autre part, la présentation de ces capteurs sous forme intégrée dans des habits intelligents (pour les accéléromètres par exemple) ne résout pas ce problème. En effet, les sujets à surveiller sont souvent très âgées, changent souvent de vêtements et par conséquent risquent de ne pas remettre le capteur soit par oubli soit par manque de volonté. D'autre part, certains de ces capteurs sont chers et possèdent une durée de vie réduite. En conclusion, nous pouvons dire que l'intégration d'un capteur portable dans notre approche n'est pas envisageable. Par contre, les paramètres estimés par ces capteurs et validés par différentes études comme étant pertinents pour la détection d'une altération de l'équilibre dynamique, seront considérés comme paramètres 'intéressants' à estimer par notre système.

L'utilisation d'un tapis actimétrique s'avère utile pour évaluer un grand nombre de paramètres spatio - temporels de la marche (longueurs et largeur des pas etc..). Ces paramètres (et leur variabilité) constituent un indicateur fiable d'une évolution vers une chute. En revanche, ces tapis sont vendus à des prix exorbitants et peuvent perturber l'environnement de vie naturelle de la personne âgée et par conséquent n'ont pas été retenus dans l'étude.

Après avoir constitué la liste des paramètres que nous cherchons à calculer durant la marche, il fallait trouver le bon outil adapté aux contraintes fixées et capable de nous fournir les données nécessaires pour effectuer ce calcul. Or, nous avons remarqué que les paramètres listés peuvent être calculés à partir des positions 3D de quelques points clé du corps humain durant la marche. Un système de suivi 3D complet du mouvement humain pourrait alors être utilisé pour évaluer tous ces paramètres.

Les systèmes de suivi 3D du mouvement les plus performants sont les systèmes à base de caméras infrarouges et de marqueurs (VICON etc..). Dans un tel système, des marqueurs (émetteurs ou réfléchissants) sont fixés sur les points clés du corps dont nous voulons suivre le déplacement. Les coordonnées 3D dans l'espace d'un marqueur donné sont calculées par triangulation. Par conséquent le marqueur doit être dans le champ d'émission infrarouge de deux caméras au minimum. En effet, ce type de système doit être installé dans un espace bien calibré et qui lui est dédié et ne peut pas être intégré à n'importe quel environnement. D'autre part, il faut que les marqueurs soient toujours visibles par les caméras pour que l'estimation soit possible et par conséquent il est conseillé que les personnes suivies soient en habits 'légers' ou moulants. Enfin, le matériel utilisé par ces systèmes est un matériel sophistiqué et cher à acquérir. C'est pourquoi le recours à un tel système est complètement exclu. Par la suite, notre réflexion a été orientée vers l'utilisation de caméras vidéo commerciales et vers l'extraction des positions 3D à partir de leurs flux vidéo.

Cet objectif nous a alors poussés à étudier les méthodes de vision par ordinateur, utilisées pour les reconstructions 3D. Or, les techniques à base de stéréo vison nécessitent que tout point à suivre soit bien étiqueté dans chaque image. Cette opération est difficile et peu précise. L'analyse automatique de la marche à l'aide de la stéréo vision peut alors être qualifiée de globale. Ayant besoin d'une analyse fine du mouvement, différents systèmes étaient alors étudiés. Les systèmes réalisant le suivi en 2D ½ sont incapables de nous fournir tous les paramètres souhaités et ont été exclus. Pour d'autres approches, l'utilisation de modèles temporels de la marche ou de modèles de projection 2D/3D, appris auparavant, rend le suivi faussé puisque le mouvement détecté sera approximé à un mouvement appris ; or, le mouvement à détecter n'est pas forcément inclus dans l'ensemble d'apprentissage, surtout dans le cas où l'on cherche à estimer des mouvements anormaux comme dans notre cas. La solution retenue pour ce problème de suivi 3D a donc été d'utiliser un modèle 3D humanoïde virtuel simulant le mouvement humain et de trouver, pour chaque image, la configuration qui maximise sa vraisemblance à cette image. La configuration ainsi trouvée sera alors considérée comme étant la configuration des différents membres du corps.

Un modèle 3D virtuel humanoïde a été alors conçu et développé. Ce modèle est configuré à partir des 31 degrés de liberté que nous lui avons attribués. Ces degrés de liberté représentent des rotations d'Euler autour de 19 points clé du corps humain. En plus du modèle, il fallait trouver une fonction décrivant la vraisemblance d'une configuration donnée de ce modèle par rapport à l'image. L'idée a été alors de comparer l'image réelle à une image 2D synthétique représentant la

configuration du modèle 3D. Cette image synthétique n'est que la projection 2D par une caméra virtuelle, ayant les mêmes paramètres que la caméra réelle utilisée, du modèle 3D configuré. En effet, la comparaison des deux images porte sur la surface du recouvrement des pixels représentant les deux corps réel et virtuel. Une fois cette fonction développée, il reste à trouver l'algorithme permettant d'estimer la valeur du vecteur configuration vis à vis de l'observation fournie par l'image réelle.

Ainsi, le problème du suivi 3D se réduit à un problème d'estimation bayésienne dynamique d'un vecteur d'état, sachant une observation donnée. L'estimateur bayésien le plus connu et le plus utilisé est le filtre de Kalman. Or, pour que ce filtre soit applicable, il faut que les densités de probabilité conditionnelle *a posteriori* et de la probabilité d'observation soient gaussiennes et uni modales. Or, la probabilité *a posteriori* et celle de l'observation sont multimodales et non gaussiennes d'où le besoin d'appliquer une autre approche pour l'estimation. L'approche dite 'particulaire' (algorithme de Condensation) fournit ce cadre global. En effet cet algorithme permet de modéliser l'incertitude en transmettant les configurations peu probables à l'instant courant vers les étapes futures, ce qui leur donne une chance d'être choisies ultérieurement. Or, dans le but d'avoir une bonne estimation d'un vecteur d'état de grande dimension (comme dans le cas de notre application), il faut utiliser un nombre important de particules. Ce nombre de particules fait croître la complexité de l'algorithme de base et le rend pratiquement inapplicable. Pour réduire cette complexité et ce nombre nous avons modifié l'algorithme d'origine. L'algorithme ainsi développé a été appelé 'Interval Particle Filtering' (IPF). En considérant le filtrage particulaire comme étant une recherche de la meilleure configuration du vecteur d'état dans un espace de recherche donné, nous avons réorganisé cet espace de façon à le rendre plus compatible et efficace par rapport au suivi du mouvement humain.

Cet algorithme a été testé et validé par rapport aux paramètres recherchés, en le comparant aux résultats calculés à partir des postions 3D fournis par un système VICON pour les mêmes séquences de mouvement. L'erreur relative constatée sur un ensemble de paramètres est de 5% en moyenne. Ce résultat est assez satisfaisant par rapport à l'étude menée. Or, pour réduire le temps de traitement et d'estimation pour une image par IPF (6 sec), nous avons réduit le nombre de particules utilisées en factorisant le processus de leur création. D'autres améliorations d'ordre technique ont aussi été proposées.

6.3 Contributions essentielles

Les contributions et les apports essentiels de ce travail de thèse se résument dans les quatre points qui suivent :

6.3.1 Conception d'un nouveau système de suivi 3D du mouvement humain

Le résultat principal de cette thèse est la conception et le développement d'un système bon marché et sans marqueurs pour le suivi 3D du mouvement humain. Les systèmes de suivi commercialisés (VICON etc..), utilisent des capteurs assez coûteux et assez sophistiqués. Cependant, le système que nous avons développé fait appel à du matériel peu cher qu'on peut trouver dans le grand commerce (caméscope + pc de bureau). Avec le développement des nouvelles webcams qui deviennent de plus en plus performantes, le caméscope pourrait être remplacé par une webcam, diminuant ainsi le coût sans diminuer la précision du suivi. De plus, ce système peut être utilisé dans n'importe quel espace (fermé ou en plein air) et ne nécessite pas un environnement bien fermé et conditionné à son utilisation comme dans le cas des systèmes à marqueurs. En terme de reconstruction 3D, notre approche présente des résultats acceptables par comparaison au système VICON. Le système développé ne fournit pas un suivi 3D aussi précis qu'un système commercial spécifique de ce type mais le coût du suivi s'avère beaucoup plus réduit. Il pourrait bien servir dans d'autres applications que l'analyse de la marche comme dans le cadre des interactions homme machine.

6.3.2 Conception d'une méthode passive pour l'analyse fine de la marche humaine

La méthode d'analyse de la marche que nous avons proposée dans ce travail, se fondant sur le système du suivi 3D décrit ci dessus, est passive, non encombrante et bon marché, au contraire des méthodes déjà utilisées. Ces approches, utilisant souvent des capteurs fixés au corps et reliés à une centrale portable, peuvent être encombrantes et difficilement acceptables par les personnes concernées. Par contre notre système ne nécessite aucun capteur portable ni l'intervention de la personne. En effet il se base sur l'utilisation d'un caméscope numérique commercial (ou une webcam) qui peut s'intégrer parfaitement au domicile de la personne âgée et ne nécessite aucun dispositif spécial.

La personne suivie évoluera dans les conditions normales de sa vie quotidienne et ne se sentira pas contrainte par un capteur. En effet, il suffira d'installer la caméra (ou les caméras) dans une pièce du domicile, où la personne suivie passe régulièrement tous les jours. Pour éviter les problèmes de changement d'illumination et d'occlusions notre idée est alors d'installer ces

caméras dans un couloir puisque cette pièce est souvent illuminée artificiellement et par conséquent le niveau de lumière y reste constant. Ainsi, la silhouette de la personne en mouvement peut être extraite sans difficultés en réglant au préalable, lors de l'installation des caméras, les paramètres des filtres appliquées. D'autre part, pour traverser un couloir il est souvent nécessaire de faire un certain nombre de pas complets. Pour cela, nous aurons la possibilité de calculer tous les paramètres de marche dont nous avons besoin pour évaluer la qualité de l'équilibre. Le suivi de ces paramètres étant longitudinal, il suffira d'estimer leurs valeurs à partir d'un nombre réduit de pas par jour (10 par exemple). L'analyse est alors effectuée sur les courtes séquences de traversée de ce couloir. Par souci d'intimité, aucune image ne sortira du système. Les seules données transférables hors domicile seront les paramètres journaliers de la marche, calculés sur place à l'aide d'un PC de bureau. En conclusion, nous pouvons dire que notre dispositif ne nécessitera aucune intervention humaine pour récupérer les paramètres de la marche.

La comparaison du système développé avec le système VICON nous a menés à extraire des paramètres spatio-temporels durant le mouvement. Ainsi, nous avons développé les méthodes nécessaires pour calculer ces paramètres à partir des coordonnées 3D des articulations du corps, qu'elles soient fournies par VICON ou IPF. D'autre part, notre système peut être qualifié de complet. En effet, à l'aide d'un seul capteur (caméra), nous avons pu extraire des données qui d'habitude sont calculées à partir de plusieurs capteurs de technologies différentes. De plus, en disposant de toutes les positions des articulations du corps, nous pouvons extraire et tester la pertinence d'autres paramètres non encore exploités dans la littérature.

6.3.3 Développement d'un nouvel algorithme de filtrage particulaire IPF

L'algorithme IPF que nous avons introduit auparavant est une version améliorée de l'algorithme de Condensation. En effet, pour remédier au problème de dégénérescence des poids des particules, pouvant entraîner une divergence de l'algorithme de Condensation dans le cas d'occurrence d'une succession d'observation erronées, nous avons introduit à chaque itération un ensemble de particules statiques bien réparties dans l'espace des états et nous avons fixé le 'survival diagnostic' d'IPF en laissant un nombre fixe de particules distinctes (les plus lourdes à l'instant ***t-1***), survivre à l'étape de 'ré échantillonnage' à l'instant ***t***. Ainsi, la convergence de l'algorithme est garantie par l'existence de quelques particules bien réparties sur tout l'espace des états à tout instant.

D'autre part, n'utilisant aucun modèle dynamique de marche, aucun modèle de transition ne sera appliqué pour faire évoluer les particules. Ayant remarqué l'existence de contraintes sur l'évolution des variables d'état (angles dans le cas du suivi 3D), nous avons pu estimer, pour chaque variable, l'intervalle auquel appartiendra sa valeur à l'instant présent, à partir de sa valeur à l'instant précédent. Cet intervalle est discrétisé par un ensemble de valeurs. Ainsi, pour chaque variable, nous avons défini un ensemble de valeurs possibles et chaque particule survivant à l'étape de 'Sélection' sera remplacée par un intervalle multidimensionnel couvrant d'une manière déterministe son voisinage formé par la combinaison de l'ensemble de valeurs possibles de chaque variable d'état. IPF peuple d'une manière déterministe et identique le voisinage de toute particule survivante. En revanche, l'algorithme de Condensation peuple le voisinage de toute particule survivante suivant son poids à l'instant précédent. Ainsi, le voisinage des particules 'lourdes' sera mieux exploité que celui des particules moins 'lourdes' qui sera sous exploité.

L'algorithme IPF adopte une stratégie équilibrée à tout instant entre l'exploration et l'exploitation de l'espace des états. La stratégie de l'algorithme de Condensation penche plus vers l'exploitation des zones entourant les particules lourdes que vers l'exploration de l'espace des états. Or, quand l'exploitation d'une zone devient importante (impliquant la majorité des particules utilisées) l'estimation devient plus précise mais en même temps le risque de divergence devient important lui aussi, surtout si l'observation est erronée. Or, dans le cas où surviennent des occlusions durant le mouvement, ce qui est fréquent durant la marche humaine, la fonction de vraisemblance devient moins fiable. Il sera alors plus judicieux dans ce contexte, d'équilibrer l'exploitation et l'exploration. Ainsi, nous pouvons dire que IPF est mieux adapté et plus efficace pour le suivi du 3D du mouvement humain que l'algorithme de Condensation.

6.3.4 Conception d'une approche factorisée du filtrage particulaire

Dans le but de réduire la complexité des algorithmes de filtrage particulaire, nous avons présenté une version factorisée des étapes de 'Sélection' et de 'Diffusion'. Cette factorisation est possible seulement dans le cas où le vecteur d'état peut être représenté par une chaîne cinématique à boucle ouverte et où la fonction d'observation peut, elle aussi, être factorisée. Cette factorisation est effectuée sous le formalisme des réseaux bayésiens dynamiques. L'avantage offert par la construction factorisée hiérarchique (en groupes de variables d'état) des particules utilisées à un instant $\boldsymbol{t}$, est la possibilité d'y faire intervenir graduellement l'observation à $\boldsymbol{t}$. Ainsi, les particules présentant les 'poids' moins importants à chaque étape de

construction, sont éliminées dès leur création, réduisant en conséquent le nombre total de particules soumises à la comparaison avec l'image 'silhouette'.

Le mouvement du modèle virtuel étant assimilé à une chaîne cinématique à boucle ouverte, la factorisation a été appliquée au vecteur d'état que nous voulons estimer et à la fonction d'observation que nous utilisons. L'estimation du mouvement est alors représentée par un réseau bayésien dynamique. Pour effectuer l'inférence dans ce réseau nous avons utilisé une nouvelle méthode fondée sur l'algorithme IPF. L'algorithme ainsi conçu est appelé 'Factored Interval Particle Filtering' (FIPF).

En principe, en comparaison à un algorithme non factorisé, cette version factorisée devrait permettre de réduire le nombre de particules nécessaires pour avoir une bonne estimation, ou d'obtenir une estimation plus précise utilisant un nombre identique de particules.

6.4 Perspectives

Les travaux futurs pourraient porter sur la réduction de la complexité des algorithmes utilisés, l'amélioration de la précision du suivi, l'évaluation du système d'analyse de la marche et la conception d'un système intelligent pour la prise de décision concernant le risque de chute.

6.4.1 Vers un système de suivi 3D en temps réel ?

Pour arriver à transformer le système développé en un système fonctionnant en temps réel, des améliorations sur les technologies utilisées et les algorithmes appliqués seront à apporter.

Pour réduire la complexité temporelle de notre approche, il suffira de réduire le temps de calcul et de traitement pour toute comparaison image réelle - modèle 3D. L'opération la plus coûteuse en temps, durant cette comparaison, est la projection 3D/2D du modèle virtuel. En utilisant la librairie OpenGL, et moyennant une carte graphique de base, nous avons réussi à diviser par 5 le temps que nécessite cette projection. L'utilisation des processeurs des cartes graphiques (GPU) pour effectuer certaines tâches à la place du processeur central (CPU) sera alors une piste à exploiter davantage. Dans ce but, il faudra étudier et comparer les temps d'exécution de chaque tâche élémentaire. Il sera ensuite intéressant de confier au GPU toutes les opérations qu'il peut effectuer d'une façon optimale. Pour optimiser le temps de calcul, il sera aussi nécessaire d'établir une architecture 'multi-thread', permettant de faire fonctionner les deux processeurs (graphique et central) en parallèle. L'utilisation d'une carte graphique plus récente et

plus puissante que celle que nous avons testée devrait éventuellement réduire significativement les temps de traitement.

Au niveau algorithmique, la factorisation devrait apporter la solution au problème du nombre important de particules impliquées. L'algorithme FIPF, conçu dans cet esprit, constitue une avancée importante vers la réduction du nombre des particules. Cet algorithme a été développé, mais doit être testé ultérieurement. En effet, sa mise en œuvre nécessite quelques réglages à effectuer comme la détermination du nombre de particules à créer pour chaque variable d'état et du nombre de particules à garder après chaque observation partielle.

L'amélioration de la précision des estimations passe par l'augmentation du nombre des particules utilisées. Or, dans la perspective d'une réduction de la complexité temporelle d'IPF à travers FIPF, cette augmentation serait possible tout en gardant un temps de calcul raisonnable.

6.4.2 Vers un indicateur fiable de risque de chute chez les personnes âgées

Dans le but de tester le système de suivi développé, de valider le choix des paramètres de marche identifiés comme pertinents et en identifier d'autres, il est nécessaire de mener une expérimentation impliquant des personnes âgées et se déroulant sur une longue période de façon à permettre d'effectuer un suivi longitudinal, dans le temps, des paramètres pour chaque personne.

Effectivement, cette expérimentation a été mise en œuvre. En revanche, en raison des contraintes et des démarches administratives, elle n'a débuté que tardivement et par conséquent nous n'avons pas eu assez de données à montrer dans ce manuscrit ou à exploiter durant le déroulement de la thèse.

L'exercice consiste à construire pour toute personne âgée un bilan gériatrique complet qui décrit son état physique, son mode de vie, ses médicaments, son état de santé actuel et passé et l'évaluation de son équilibre effectuée par un médecin à partir de tests cliniques (Tinetti, Berg etc..). L'expérimentation se déroule en présence d'un gériatre, dans une salle au sein du service de gériatrie du CHRU de Nancy. Après avoir répondu à un questionnaire et avoir effectué une montée sur une pèse personne instrumentée (tâche 1 du projet PARAChute : évaluation de l'équilibre), toute personne répétera 3 fois (avec 5 minutes de repos entre les essais) la même manœuvre suivante : se lever d'une chaise, parcourir 6m, faire un demi tour et revenir pour s'asseoir de nouveau sur la chaise (fig. 6.1). Une sonnerie de téléphone faisant fonction de

distracteur est déclenchée pendant l'un des passages dans le but de détecter l'aptitude de la personne à assimiler une double tâche.

Pour filmer la scène, quatre caméras CMOS (1394 AVT Marlin F046) synchronisées ont été placées dans la salle. Le logiciel d'acquisition des quatre flux vidéo ainsi que le programme gérant la base de données ont été développés durant ce travail de thèse également.

L'étude des données obtenues portera sur le suivi longitudinal et la variabilité de quelques paramètres de marche de façon à identifier ceux qui permettent de déceler une dérive vers la chute. En plus des paramètres déjà cités auparavant, un nouvel ensemble de paramètres sera aussi mis en test pour décider ensuite de sa pertinence. Il sera également nécessaire d'étudier les corrélations qui peuvent exister entre certains paramètres et certains éléments du dossier médical de chaque personne. Une fois les paramètres fixés, un système expert intelligent devra être conçu dans le but de calculer un indicateur fiable du risque. Ce système se basera sur les modèles de décision markoviens (HMM etc..) ou sur les réseaux bayésiens dynamiques.

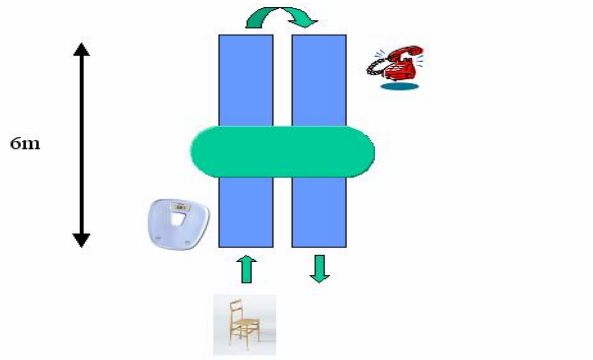


fig.6.1 : Parcours expérimental à effectuer par une personne âgée à trois reprises : la personne se lève de sa chaise parcourt 6 m, fait un demi tour et retourne sur sa chaise.

Annexe A : Formation de l'image et calibration d'une caméra

Modéliser le dispositif d'acquisition revient à modéliser la manière dont se forme l'image sur le capteur CCD de la caméra, et donc à caractériser la transformation entre la scène (espace à trois dimensions) et l'image de la scène (espace à deux dimensions). Plusieurs modèles sont utilisés pour représenter une caméra, dont le plus connu est le modèle dit à sténopé (par référence à l'appareil de sténopé). Dans ce modèle, la caméra est considérée comme un trou d'épingle qui constitue le centre optique. L'image acquise est alors formée par l'intersection entre la matrice CCD et les rayons provenant de la scène filmée, lesquels passent par le centre optique. La formation de cette image est donc le résultat d'une transformation rigide (rotation, translation et projection). Pour obtenir les coordonnées dans le plan image (l'espace à deux dimensions) d'un point de l'espace (à trois dimensions) de la scène il suffit de comprendre le processus de formation des images, dont nous allons décrire les différentes étapes, et d'appliquer les équations qui y sont associées. Il est donc nécessaire de déterminer les paramètres intervenant dans ce processus. Nous distinguons deux types de paramètres à déterminer :

- les paramètres intrinsèques, qui sont des paramètres internes à la caméra ;
- les paramètres extrinsèques à la caméra, qui sont les paramètres de la transformation rigide (rotation et translation).

Les paramètres intrinsèques dont nous avons besoin sont :

- $\boldsymbol{f}$: la distance focale effective de la caméra (à calculer) ;
- $\boldsymbol{N_{cx}}$: le nombre d'éléments de la matrice CCD dans la direction $\boldsymbol{x}$ (donné par le constructeur) ;
- $\boldsymbol{N_{fx}}$: le nombre de pixels dans une ligne de l'image numérique (donné par le constructeur) ;

- d_x, d_y : les distances en mm entre les éléments de la matrice dans les directions x et y, respectivement (données par le constructeur). Ces distances ne dépendent que de la taille du capteur CCD et de la résolution de l'image ;
- C_x, C_y : Les coordonnées du centre de distorsion radiale de la lentille souvent assimilé au centre de l'image sauf dans le cas d'assemblages optiques de mauvaise qualité ;
- ***K1, K2*** : coefficients de distorsion radiale de la caméra (à calculer) ;
- S_x : Facteur d'échelle (mm/pixel) de la caméra (donné par le constructeur) ;

Les paramètres extrinsèques dont nous avons besoin sont :

- la matrice de rotation R établie à partir des angles (R_x, R_y, R_z) d'Euler de rotation autour des axes (x,y,z) (à calculer) ;
- le vecteur T de translation (repère caméra – repère scène), formé par les paramètres de translation (T_x, T_y, T_z) (à calculer).

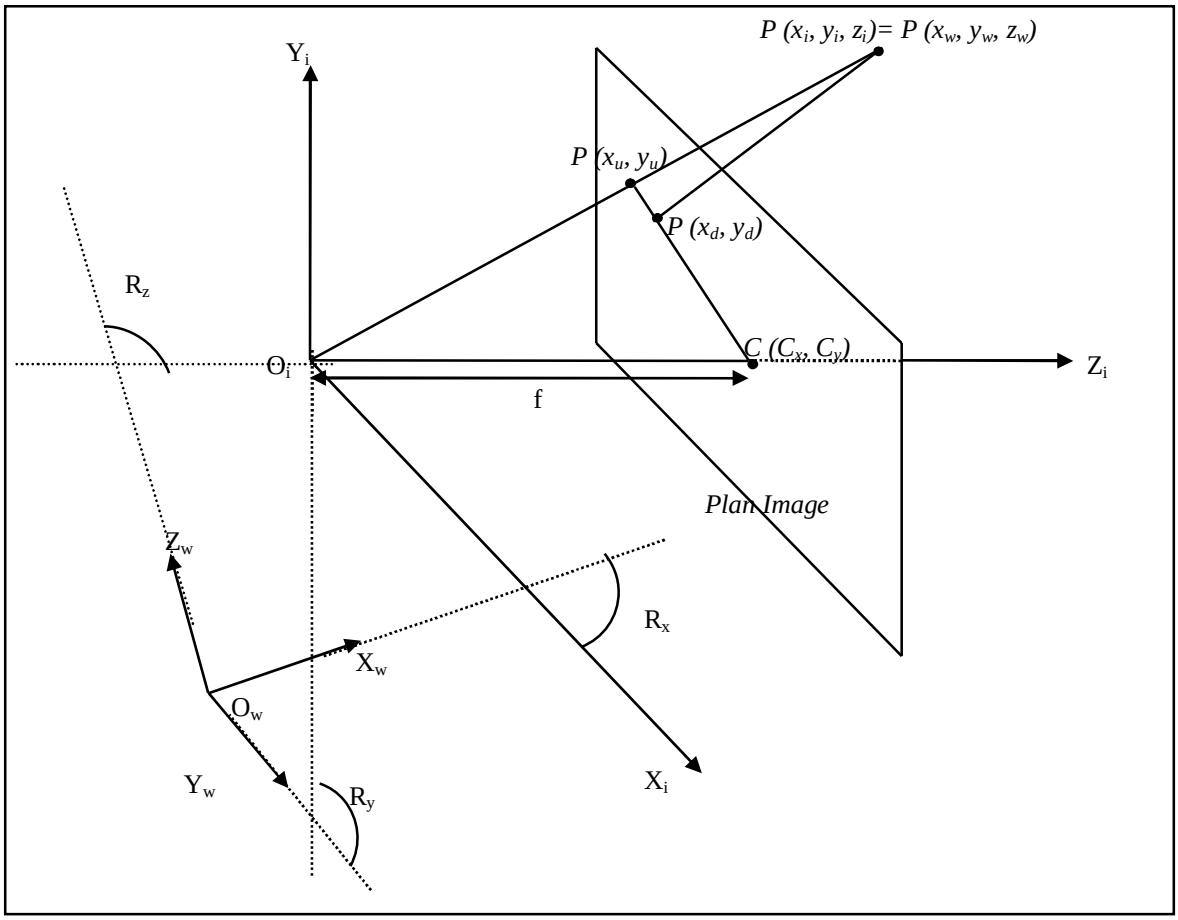


fig. A.1 : Le point P de coordonnées (x_w,y_w,z_w) dans le repère de la scène (O_w, X_w, Y_w, Z_w) est projeté dans le repère caméra (O_i, X_i, Y_i, Z_i). Son image sera formée à l'intersection du rayon le reliant à la caméra avec le plan image.

Nous allons présenter les différentes étapes du processus de formation de l'image. Considérons un point P de coordonnées (x_w, y_w, z_w) dans le monde réel (3D) dont nous cherchons les coordonnées (x_f, y_f) (2d) dans le plan image (fig. A.1).

La première étape consiste à effectuer la transformation rigide permettant de passer du repère scène vers le repère de la caméra. Cette transformation du point (x_w,y_w,z_w) en (x_i,y_i,z_i) est effectuée en utilisant l'équation suivante :

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = R \cdot \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + T$$

où $$R = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \quad \text{et} \quad T = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}$$

où
$$\begin{aligned}
r_1 &= \cos(R_y)\cos(R_z) \\
r_2 &= \cos(R_z)\sin(R_x)\sin(R_y) - \cos(R_x)\sin(R_z) \\
r_3 &= \sin(R_x)\sin(R_z) + \cos(R_x)\cos(R_z)\sin(R_y) \\
r_4 &= \cos(R_y)\sin(R_z) \\
r_5 &= \sin(R_x)\sin(R_y)\sin(R_z) + \cos(R_x)\cos(R_z) \\
r_6 &= \cos(R_x)\sin(R_y)\sin(R_z) - \cos(R_z)\sin(R_x) \\
r_7 &= -\sin(R_y) \\
r_8 &= \cos(R_y)\sin(R_x) \\
r_9 &= \cos(R_x)\cos(R_y)
\end{aligned}$$

L'impression de profondeur d'une image peut provenir des lois de la perspective. En partant du modèle à sténopé de la caméra nous allons établir ensuite les différentes équations de la projection perspective du point P dans le plan image de la caméra servant à récupérer les coordonnées non distordues (x_u, y_u). Pour cela, nous avons besoin de la focale effective f de la caméra :

$$\begin{cases} x_u = f \cdot \dfrac{x_i}{z_i} \\ y_u = f \cdot \dfrac{y_i}{z_i} \end{cases}$$

La prochaine étape consiste à prendre en compte la distorsion radiale due à la lentille. En effet la plupart des caméras analogiques et numériques souffrent de distorsions dues à leur lentille. Il existe deux types de distorsions observables :

- la distorsion en pelote d'épingle (pincushion) qui provoque la sphérisation des images depuis leur centre ;
- la distorsion en tonneau (barrel) qui provoque un pincement des images depuis leur centre.

Pour corriger et calculer ces distorsions, les formulations de distorsion de Brown [1] sont souvent utilisées : $\begin{cases} x_u = x_d + \Delta r_x + \Delta p_x \\ y_u = y_d + \Delta r_y + \Delta p_y \end{cases}$

avec x_u et y_u les coordonnées corrigées, x_d et y_d les coordonnées distordues avant correction, Δr_x et Δr_y les distorsions radiales pour les composantes x et y, et enfin Δp_x et Δp_y les distorsions de décentrage pour les composantes x et y. En fait, deux types de distorsions contribuent à altérer l'image : une distorsion radiale et une distorsion de décentrage. Cependant, la contribution de cette dernière étant très faible par rapport à la distorsion radiale, nous pouvons la négliger. En négligeant la distorsion de décentrage, et en supposant la distorsion radiale uniforme, nous obtenons les formules suivantes : $\begin{cases} x_u = x_d \cdot (1 + \Delta r) \\ y_u = y_d \cdot (1 + \Delta r) \end{cases}$

où $\Delta r = \mathrm{K}_1 \cdot r^2 + \mathrm{K}_2 \cdot r^4$ et $r^2 = {x_d}^2 + {y_d}^2$

Les coefficients K_1 et K_2 sont les paramètres de la distorsion radiale qu'il nous faut estimer. Ce calcul se fait à l'aide d'une mire composée d'une grille de points qui permet de juger de la distorsion.

Enfin, la dernière étape utilise les paramètres intrinsèques de la caméra, donnés par le constructeur, pour accomplir l'acquisition effective par la matrice CCD suivant les formules suivantes :

$$\begin{cases} x_f = \dfrac{x_d}{d_x} \cdot S_x \cdot \dfrac{N_{fx}}{N_{cx}} + C_x \\ y_f = \dfrac{y_d}{d_y} + C_y \end{cases}$$

où (x_f, y_f) sont les coordonnées finales du point P dans le plan image, (x_d,y_d) les coordonnées distordues du point P issues de l'étape précédente, S_x le coefficient d'incertitude, (C_x, C_y) les coordonnées du centre optique de projection.

Il nous faut maintenant calculer les paramètres de la caméra (modèle) dont nous ne disposons pas, pour pouvoir appliquer les équations de projection. Ce calcul peut être fait en utilisant la méthode de calibration de Tsai [2]. Cette méthode comporte deux étapes :

La première étape a pour but de calculer les paramètres extrinsèques de la caméra. Pour cela, il faut utiliser une mire étalon. A partir d'une image de cette mire, nous déterminons un ensemble de N points qui sera utilisé pour effectuer la mise en correspondance. Les coordonnées de ces points sont notées (x_{fi},y_{fi}) avec i compris entre 1 et N. Le pixel central de la mémoire image est pris comme origine du plan image. Ses coordonnées sont notées (C_x,C_y). Il faut maintenant calculer les coordonnées des points dans l'image réelle distordue avec les formules suivantes :

$$\begin{cases} x_{di} = \dfrac{d_x \cdot (x_{fi} - C_x)}{s_x} \cdot \dfrac{N_{cx}}{N_{fx}} \\ y_{di} = d_y (y_{fi} - C_y) \end{cases}$$

Posons ensuite : $$L = \left[\frac{r_1}{T_y} \quad \frac{r_2}{T_y} \quad \frac{T_x}{T_y} \quad \frac{r_4}{T_y} \quad \frac{r_5}{T_y} \right] = \left[a_1 \quad a_2 \quad a_3 \quad a_4 \quad a_5 \right]$$

Il faut alors utiliser la relation suivante : $\left[y_{di} \cdot x_{wi} \quad y_{di} \cdot y_{wi} \quad y_{di} \quad -x_{di} \cdot x_{wi} \quad -x_{di} \cdot y_{wi} \right] \cdot L^T = x_{di}$ pour déterminer les cinq inconnues de L. Une fois L calculé, il faut calculer la matrice R et les paramètres T_x et T_y de la translation. Dans un premier temps, il faut calculer $|T_y|$. Pour cela, nous définissons la matrice suivante :

$$C = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$$

Cette matrice étant orthogonale, nous obtenons alors :

$$T_y^2 = \frac{\left(S_r - \sqrt{S_r^2 - 4 \cdot (a_1 \cdot a_5 - a_4 \cdot a_2)^2} \right)}{2 \cdot (a_1 \cdot a_5 - a_4 \cdot a_2)^2}$$

où $S_r = \left(a_1^2 + a_2^2 + a_4^2 + a_5^2\right)$. Ce calcul permet d'obtenir $\boldsymbol{T_y}$ au signe près. En supposant que $\boldsymbol{T_y}$ est positif, il est possible de calculer $\boldsymbol{r_1, r_2, r_4, r_5}$ et $\boldsymbol{T_x}$. Il faut ensuite calculer les deux expressions suivantes :

$$\begin{cases} x = r_1 \cdot x_w + r_2 \cdot y_w + T_x \\ y = r_4 \cdot x_w + r_5 \cdot y_w + T_y \end{cases}$$

Si $\boldsymbol{x}$ et $\boldsymbol{y}$ sont de même signe, alors $\boldsymbol{T_y}$ est positif, sinon $\boldsymbol{T_y}$ est négatif. En fonction du signe, nous pouvons recalculer $\boldsymbol{r_1, r_2, r_4, r_5}$ et $\boldsymbol{T_x}$. La matrice de rotation sera alors la suivante :

$$R = \begin{bmatrix} r_1 & r_2 & s_2 \cdot \sqrt{1 - r_1^2 - r_2^2} \\ r_4 & r_5 & s_1 \cdot s_2 \sqrt{1 - r_4^2 - r_5^2} \\ s_2 \cdot r_7 & s_2 \cdot r_8 & r_9 \end{bmatrix}$$

où $$s_1 = \frac{r_1 \cdot r_4 + r_2 \cdot r_5}{\left|r_1 \cdot r_4 + r_2 \cdot r_5\right|} \qquad \text{et} \qquad s_2 = \frac{f}{|f|}$$

La seconde étape sert à estimer $\boldsymbol{T_z}$ et $\boldsymbol{f}$. Pour cela, nous utilisons le système surdéterminé suivant :

$$\begin{bmatrix} Y_i & -d_y \cdot y_i \end{bmatrix} \cdot \begin{bmatrix} f \\ T_z \end{bmatrix} = \left[r_7 \cdot x_{wi} + r_8 \cdot y_{wi}\right] \cdot d_y \cdot y_i$$

où $$Y_i = r_4 \cdot x_{wi} + r_5 \cdot y_{wi} + T_y$$

Ces équations nous fournissent une estimation de $\boldsymbol{T_x}$ et de $\boldsymbol{f}$ (et particulièrement le signe de $\boldsymbol{f}$ qui est utilisé dans l'étape précédente). Les solutions exactes de $\boldsymbol{T_x}$ et de $\boldsymbol{f}$ sont obtenues en résolvant l'équation suivante :

$$d_y \cdot Y + d_y \cdot Y \cdot \left(\kappa_1 \cdot r^2 + \kappa_2 \cdot r^4\right) = f \cdot \frac{r_4 \cdot x_w + r_5 \cdot y_w + r_6 \cdot z_w + T_y}{r_7 \cdot x_w + r_8 \cdot y_w + r_9 \cdot z_w + T_z}$$

où $$r = \sqrt{\left(\frac{d'_x}{s_x} \cdot X\right)^2 + \left(d_y \cdot Y\right)^2} \qquad \text{et} \qquad X = x_f - C_x ; Y = y_f - C_y$$

Ayant calculé tous ces paramètres, il suffira d'appliquer les équations de projection pour déterminer la position dans l'image de tout point de l'espace vu par la caméra.

La reconstruction 3D d'une scène passe par une inversion du processus de formation de l'image et des équations impliquées. En estimant tous les paramètres de la caméra, il est alors possible de calculer, pour un point donné de l'image, le rayon de vue contenant ce point, et par la suite de le localiser dans le repère de la scène moyennant une information supplémentaire (le plan sur lequel il se déplace, sa profondeur etc.).

Bibliographie

[1] Brown, DC, "Close Range Camera Calibration", Journal of Photogrametric Engineering, vol.37, No. 8, pp.855-866, 1971.

[2] Tsai RY, "An efficient and accurate camera calibration technique for 3D machine vision", IEEE Journal of Robotics and automation, RA-3, pp.364-374, 1986.

Annexe B: Description du modèle 3D virtuel utilisé

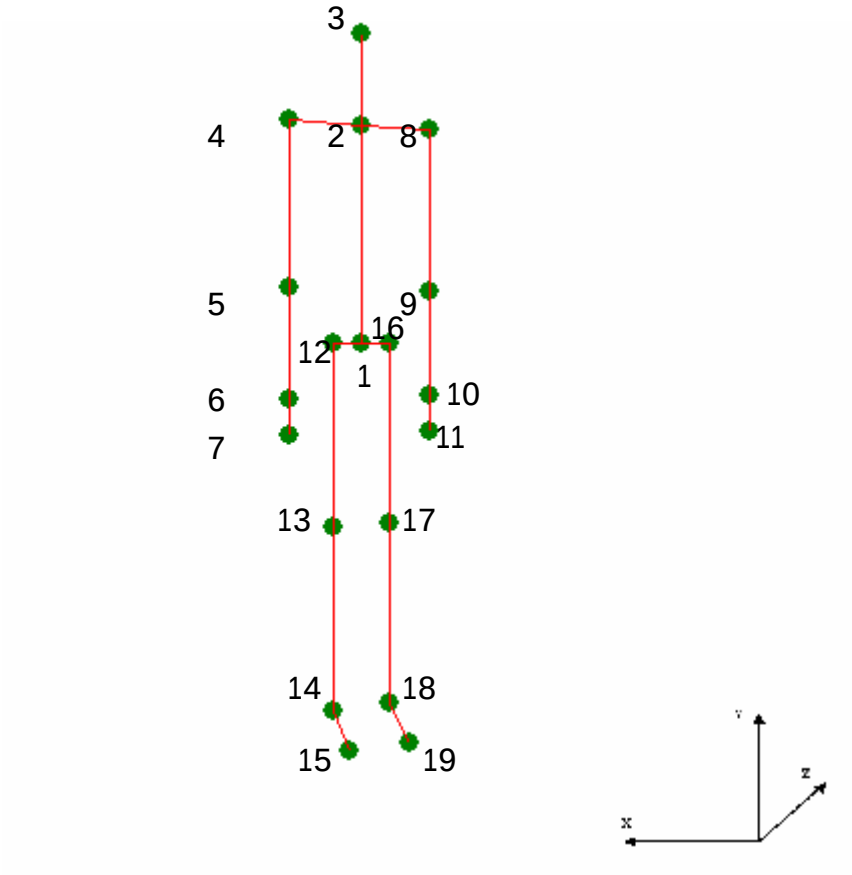


fig. B.1 : Modèle humanoïde en fils de fer formé par 19 points

Le modèle virtuel que nous utilisons est formé par 19 points et 17 segments les reliant (fig. B.1). Ces segments sont répartis en 5 chaînes cinématiques. Les 5 chaînes cinématiques contiennent les articulations suivantes (énumérées à partir de l'extrémité libre) :

- Bras droit : 7, 6, 5, 4
- Bras gauche : 11, 10, 9, 8
- Tronc : 11, 10, 9, 8 + 7, 6, 5, 4 + 3, 2, 1

- Jambe droite : 15, 14, 13, 12
- Jambe gauche : 19, 18, 17, 16

Les degrés de liberté attribués pour chacune des articulations de notre modèle (les mouvements permis) sont les suivants:

Notations :

$\boldsymbol{R_x}$ = Rotation par rapport à un axe parallèle à $\boldsymbol{x}$
$\boldsymbol{R_y}$ = Rotation par rapport à un axe parallèle à $\boldsymbol{y}$
$\boldsymbol{R_z}$ = Rotation par rapport à un axe parallèle à $\boldsymbol{z}$

mvt-ind = mouvement indépendant du point (n'entraîne pas le mouvement d'autres points du modèle)

Point 1 (sacrum) :

Fixe dans le référentiel mobile du modèle

Point 2 (cou) :

$\boldsymbol{R_x}$ de centre le point 1 ;
$\boldsymbol{R_z}$ de centre le point 1 ;

Point 3 (tête) :

$\boldsymbol{R_x}$ de centre le point 1 ;
$\boldsymbol{R_z}$ de centre le point 1 ;

$\boldsymbol{R_x}$ de centre le point 2 ; mvt-ind
$\boldsymbol{R_y}$ de centre le point 2 ; mvt-ind
$\boldsymbol{R_z}$ de centre le point 2 ; mvt-ind

Point 4 (épaule droit) :

$\boldsymbol{R_x}$ de centre le point 1 ;

$\boldsymbol{R_y}$ de centre le point 1 ;

$\boldsymbol{R_z}$ de centre le point 1 ;

$\boldsymbol{R_y}$ de centre le point 2 ; mvt-ind

$\boldsymbol{R_z}$ de centre le point 2 ; mvt-ind

Point 5 (coude droit) :

$\boldsymbol{R_x}$ de centre le point 1 ;

$\boldsymbol{R_y}$ de centre le point 1 ;

$\boldsymbol{R_z}$ de centre le point 1 ;

$\boldsymbol{R_y}$ de centre le point 2 ;

$\boldsymbol{R_z}$ de centre le point 2 ;

$\boldsymbol{R_x}$ de centre le point 4 ; mvt-ind

$\boldsymbol{R_z}$ de centre le point 4 ; mvt-ind

Point 6 (poignet droit) :

$\boldsymbol{R_x}$ de centre le point 1 ;

$\boldsymbol{R_y}$ de centre le point 1 ;

$\boldsymbol{R_z}$ de centre le point 1 ;

$\boldsymbol{R_y}$ de centre le point 2 ;

$\boldsymbol{R_z}$ de centre le point 2 ;

$\boldsymbol{R_x}$ de centre le point 4 ;

$\boldsymbol{R_z}$ de centre le point 4 ;

$\boldsymbol{R_x}$ de centre le point 5 ; mvt-ind

$\boldsymbol{R_y}$ de centre le point 5 ; mvt-ind

Point 7 (main droite) :

$\boldsymbol{R_x}$ de centre le point 1 ;

$\boldsymbol{R_y}$ de centre le point 1 ;

$\boldsymbol{R_z}$ de centre le point 1 ;

$\boldsymbol{R_y}$ de centre le point 2 ;

$\boldsymbol{R_z}$ de centre le point 2 ;

$\boldsymbol{R_x}$ de centre le point 4 ;

$\boldsymbol{R_z}$ de centre le point 4 ;

$\boldsymbol{R_x}$ de centre le point 5 ;

$\boldsymbol{R_y}$ de centre le point 5 ;

Point 8 (épaule gauche) :

$\boldsymbol{R_x}$ de centre le point 1 ;

$\boldsymbol{R_y}$ de centre le point 1 ;

$\boldsymbol{R_z}$ de centre le point 1 ;

$\boldsymbol{R_y}$ de centre le point 2 ; mvt-ind

$\boldsymbol{R_z}$ de centre le point 2 ; mvt-ind

Point 9 (coude gauche) :

$\boldsymbol{R_x}$ de centre le point 1 ;

$\boldsymbol{R_y}$ de centre le point 1 ;

$\boldsymbol{R_z}$ de centre le point 1 ;

$\boldsymbol{R_y}$ de centre le point 2 ;

$\boldsymbol{R_z}$ de centre le point 2 ;

$\boldsymbol{R_x}$ de centre le point 8 ; mvt-ind

$\boldsymbol{R_z}$ de centre le point 8 ; mvt-ind

Point 10 (poignet droit) :

$\boldsymbol{R_x}$ de centre le point 1 ;
$\boldsymbol{R_y}$ de centre le point 1 ;
$\boldsymbol{R_z}$ de centre le point 1 ;

$\boldsymbol{R_y}$ de centre le point 2 ;
$\boldsymbol{R_z}$ de centre le point 2 ;

$\boldsymbol{R_x}$ de centre le point 8 ;
$\boldsymbol{R_z}$ de centre le point 8 ;

$\boldsymbol{R_y}$ de centre le point 9 ; mvt-ind
$\boldsymbol{R_y}$ de centre le point 9 ; mvt-ind

Point 11 (main droite) :

$\boldsymbol{R_x}$ de centre le point 1 ;
$\boldsymbol{R_y}$ de centre le point 1 ;
$\boldsymbol{R_z}$ de centre le point 1 ;

$\boldsymbol{R_y}$ de centre le point 2 ;
$\boldsymbol{R_z}$ de centre le point 2 ;

$\boldsymbol{R_x}$ de centre le point 8 ;
$\boldsymbol{R_z}$ de centre le point 8 ;

$\boldsymbol{R_x}$ de centre le point 9 ;
$\boldsymbol{R_y}$ de centre le point 9 ;

Point 12 (hanche droite) :

Supposé fixe

Point 13 (genou droit) :

$\boldsymbol{R_x}$ de centre le point 12 ; mvt-ind
$\boldsymbol{R_y}$ de centre le point 12 ; mvt-ind
$\boldsymbol{R_z}$ de centre le point 12 ; mvt-ind

Point 14 (cheville droite) :

$\boldsymbol{R_x}$ de centre le point 12 ;
$\boldsymbol{R_y}$ de centre le point 12 ;
$\boldsymbol{R_z}$ de centre le point 12 ;

$\boldsymbol{R_x}$ de centre le point 13 ; mvt-ind

Point 15 (orteil droit) :

$\boldsymbol{R_x}$ de centre le point 12 ;
$\boldsymbol{R_y}$ de centre le point 12 ;
$\boldsymbol{R_z}$ de centre le point 12 ;

$\boldsymbol{R_x}$ de centre le point 13 ;

$\boldsymbol{R_x}$ de centre le point 14 ; mvt-ind
$\boldsymbol{R_y}$ de centre le point 14 ; mvt-ind
$\boldsymbol{R_z}$ de centre le point 14 ; mvt-ind

Point 16 (hanche gauche) :

Supposé fixe

Point 17 (genou gauche) :

$\boldsymbol{R_x}$ de centre le point 16 ; mvt-ind
$\boldsymbol{R_y}$ de centre le point 16 ; mvt-ind

$\boldsymbol{R_z}$ de centre le point 16 ; mvt-ind

Point 18 (cheville gauche) :

$\boldsymbol{R_x}$ de centre le point 16 ;
$\boldsymbol{R_y}$ de centre le point 16 ;
$\boldsymbol{R_z}$ de centre le point 16 ;

$\boldsymbol{R_x}$ de centre le point 17 ; mvt-ind

Point 19 (orteil gauche) :

$\boldsymbol{R_x}$ de centre le point 16 ;
$\boldsymbol{R_y}$ de centre le point 16 ;
$\boldsymbol{R_z}$ de centre le point 16 ;

$\boldsymbol{R_x}$ de centre le point 17 ;
$\boldsymbol{R_x}$ de centre le point 18 ; mvt-ind
$\boldsymbol{R_y}$ de centre le point 18 ; mvt-ind
$\boldsymbol{R_z}$ de centre le point 18 ; mvt-ind

Pour transformer ce modèle de squelette humain en un modèle humanoïde complet nous lui ajoutons des volumes représentant les différentes parties du corps (fig. B.2).

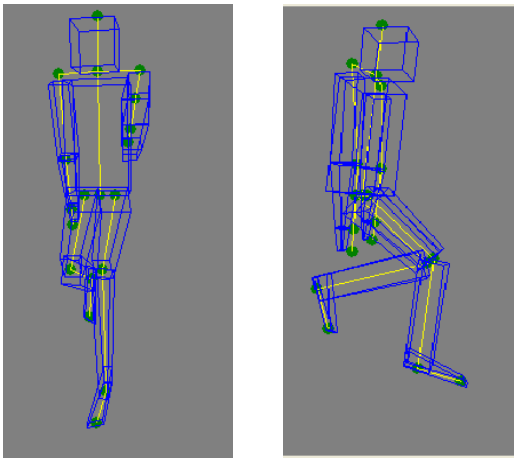


fig. B.2 : Modèle 3D virtuel après ajouts des volumes

Publications

Revues internationales avec comité de lecture

[2007a] David Hewson, Jacques Duchêne, François Charpillet, Jamal Saboune, Valérie Michel, Hassan Amoud, Michel Doussot, Jean Paysant, Anne Boyer et Jean Yves Horgel, *The PARAChute Project : Remote Monitoring of Posture and Gait for Fall Prevention*, EURASIP - European Journal of Applied Signal Processing, vol. 2007.

[2007b] Jamal Saboune et François Charpillet, *Marker less human motion tracking from a single camera using interval particle filtering*, IJAIT - International Journal on Artificial Intelligence Tools, vol. 16, n.4, pp. 593-609.

Revues nationales avec comité de lecture

[2008]Anne Boyer, François Charpillet, Jean Paysant et Jamal Saboune, *Vers une analyse écologique de la marche pour la prévention des chutes chez la personne âgée*, TSI – Technique et Science Informatiques : Numéro spécial 'Apport des STIC au maintien à domicile' (A paraître).

Conférences internationales avec actes et comité de lecture

[2005a] Jamal Saboune et François Charpillet, *Using Interval Particle Filtering for Marker less 3D Human Motion Capture*, ICTAI'20005 - 17th IEEE International Conference on Tools with Artificial Intelligence, pp. 621-627, Hong Kong Chine 2005.

[2005b] Jamal Saboune et François Charpillet, *Marker less human motion capture for gait analysis*, EMBEC05 - 3rd European Medical and Biological Engineering Conference, Prague République Tchèque 2005.

[2007c] Jamal Saboune, Cedric Rose et François Charpillet, Factored *interval particle filtering for gait analysis*, 29th Annual International Conference of the IEEE EMBS, pp. 3232-3235, Lyon France 2007.