



Analyse d'information tridimensionnelle issue de systèmes multi-caméras pour la détection de la chute et l'analyse de la marche

Edouard Auvinet

► To cite this version:

Edouard Auvinet. Analyse d'information tridimensionnelle issue de systèmes multi-caméras pour la détection de la chute et l'analyse de la marche. Education. Université Rennes 2; Université de Montréal. Faculté de médecine, 2012. Français. NNT : 2012REN20067 . tel-00946188

HAL Id: tel-00946188

<https://theses.hal.science/tel-00946188>

Submitted on 13 Feb 2014

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 Montréal
Institut de génie biomédical

Université Européenne de Bretagne
Laboratoire Mouvement Sport Santé - EA 1274

Analyse d'information tridimensionnelle issue de systèmes multi-caméras pour la détection de la chute et l'analyse de la marche

Par

Edouard Auvinet

Pour obtenir le grade de

Docteur de l'Université de Montréal
Discipline génie biomédical

Docteur de l'Université Rennes 2
Discipline Sciences et Techniques
des Activités Physiques et Sportives
(STAPS)

Juin 2012

© Copyright Edouard Auvinet 2012

Université de Montréal
Institut de génie biomédical

Université Européenne de Bretagne
Laboratoire Mouvement Sport Santé - EA 1274

Cette thèse intitulée :

**Analyse d'information tridimensionnelle issue de
systèmes multi-caméras pour la détection de la chute
et l'analyse de la marche**

Présentée par

Edouard Auvinet

a été évaluée par un jury composé des personnes suivantes :

Maxime Raison	Président-rapporteur
Jean Meunier	Co-directeur (Québec)
Franck Multon	Co-directeur (France)
Monique Thonnat	Examinatrice externe (France)
Robert Bergevin	Examineur externe (Québec)
Josette Noël	Représentante de la doyenne

Je dédie cette thèse :

À ma Mère,

à mon Père et mes Frères...

$$\sqrt{\text{Истина}} = \text{Любовъ}$$

"La racine carrée de la Vérité, c'est l'Amour".

Andréï Dmitriévitch SAKHAROV

Remerciements

Une thèse de doctorat est un travail personnel ... qui se fait en équipe.

Je souhaite remercier chaleureusement toutes les personnes qui de près ou de loin m'ont aidé à réaliser cette expérience scientifique, humaine et personnelle.

Le fruit de ces 5 années et demie que représente ce manuscrit n'aurait jamais abouti et mûrit sans l'aide, les conseils et le soutien de mes deux directeurs de thèse Jean Meunier et Franck Multon. Je leur en suis infiniment reconnaissant.

Je remercie le jury pour leur relecture et leurs commentaires très intéressants.

Je souhaite remercier Paul Delamarche de m'avoir accueilli au laboratoire du M2S. Je remercie les membres du laboratoire de Rennes 2 pour leur accueil, en particulier l'équipe de biomécanique.

Je remercie l'équipe du laboratoire de traitement d'image et de vision 3D du département d'informatique et de recherche opérationnelle pour leur accueil.

Je remercie le support informatique du DIRO, pour leur joie et leurs rigolades même lorsque l'alarme sonne...

Je remercie toute l'équipe d'Axper avec qui j'ai partagé une expérience extra-ordinaire.

Je remercie les organismes ANRT et MITACS pour leur soutien financier et les entreprises partenaires Axper et les ateliers de la marche.

Il en va de même pour toutes les équipes des deux universités qui ont permis de régler les questions administratives. En effet, faire une cotutelle avec une bourse CIFRE n'est pas très simple, grâce à leur aide et leur compétence nous avons réussi.

Je remercie tous mes amis qui m'ont supporté pendant toutes ces années et de m'avoir apporté la lumière dans les périodes les plus sombres.

Merci Monia pour ces deux fabuleuses années.

Je remercie Pierre Constantin pour tous ses précieux conseils.

Большое спасибо Наталии за Вашу поддержку и помощь

Je remercie ma Mère, mon Père et mes Frères de tout mon coeur.

Résumé

Cette thèse s'intéresse à définir de nouvelles méthodes cliniques d'investigation permettant de juger de l'impact de l'avance en âge sur la motricité. En particulier, cette thèse se focalise sur deux principales perturbations possibles lors de l'avance en âge : la chute et l'altération de la marche.

Ces deux perturbations motrices restent encore mal connues et leur analyse en clinique pose de véritables défis technologiques et scientifiques. Dans cette thèse, nous proposons des méthodes originales de détection qui peuvent être utilisées dans la vie courante ou en clinique, avec un minimum de contraintes techniques.

Dans une première partie, nous abordons le problème de la détection de la chute à domicile, qui a été largement traité dans les années précédentes. En particulier, nous proposons une approche permettant d'exploiter le volume du sujet, reconstruit à partir de plusieurs caméras calibrées. Ces méthodes sont généralement très sensibles aux occultations qui interviennent inévitablement dans le domicile et nous proposons donc une approche originale beaucoup plus robuste à ces occultations. L'efficacité et le fonctionnement en temps réel ont été validés sur plus d'une vingtaine de vidéos de chutes et de leurres, avec des résultats approchant les 100% de sensibilité et de spécificité en utilisant 4 caméras ou plus.

Dans une deuxième partie, nous allons un peu plus loin dans l'exploitation des volumes reconstruits d'une personne, lors d'une tâche motrice particulière : la marche sur tapis roulant, dans un cadre de diagnostic clinique. Dans cette partie, nous analysons plus particulièrement la qualité de la marche. Pour cela nous développons le concept d'utilisation de caméras de profondeur pour la quantification de l'asymétrie spatiale au cours du mouvement des membres inférieurs pendant la marche. Après avoir détecté chaque pas dans le temps, cette méthode réalise une comparaison de surfaces de chaque jambe avec sa correspondante symétrique du pas opposé. La validation effectuée sur une cohorte de 20 sujets montre la viabilité de la démarche.

Mots clés

Analyse de la marche, détection de la chute, multi-caméras, caméra de profondeur

Abstract

This thesis is concerned with defining new clinical investigation method to assess the impact of ageing on motricity. In particular, this thesis focuses on two main possible disturbance during ageing : the fall and walk impairment.

This two motricity disturbances still remain unclear and their clinical analysis presents real scientist and technological challenges. In this thesis, we propose novel measuring methods usable in everyday life or in the walking clinic, with a minimum of technical constraints.

In the first part, we address the problem of fall detection at home, which was widely discussed in previous years. In particular, we propose an approach to exploit the subject's volume, reconstructed from multiple calibrated cameras. These methods are generally very sensitive to occlusions that inevitably occur in the home and we therefore propose an original approach much more robust to these occultations. The efficiency and real-time operation has been validated on more than two dozen videos of falls and lures, with results approaching 100 % sensitivity and specificity with at least four or more cameras.

In the second part, we go a little further in the exploitation of reconstructed volumes of a person at a particular motor task : the treadmill, in a clinical diagnostic. In this section we analyze more specifically the quality of walking. For this we develop the concept of using depth camera for the quantification of the spatial and temporal asymmetry of lower limb movement during walking. After detecting each step in time, this method makes a comparison of surfaces of each leg with its corresponding symmetric leg in the opposite step. The validation performed on a cohort of 20 subjects showed the viability of the approach.

Key words

Gait analysis, fall detection, multiple cameras, depth camera

Table des matières

1	Introduction	17
1.1	Systèmes mis en jeu dans le mouvement de la marche	21
1.1.1	Le système musculo-squelettique	21
1.1.2	Le système nerveux	22
1.2	Le mouvement de la marche	23
1.3	L'impact de la pathologie sur la marche	25
1.4	Relation entre la marche, l'âge et la chute	26
1.5	Objectifs des travaux de la thèse	27
I	Étude de la chute	29
2	Détection de la chute	31
2.1	Relation entre l'âge et la chute	33
2.2	Quelques chiffres	34
2.3	État de l'art	34
2.3.1	Appareil d'alerte des secours en cas de chute	34
2.3.2	Dispositif de détection de chute automatique	35
2.4	Problématique	40
2.5	Méthode	40
2.5.1	Mode de représentation du volume	42
2.5.2	Discussion du choix de représentation du volume	45
2.5.3	Le calibrage	45
2.5.4	Mise en œuvre	48
2.6	Conclusions	49

3	Détection de la chute avec un système multi-caméras résistant aux occultations	51
3.1	abstract	54
3.2	Introduction	54
3.3	Method	57
3.3.1	Camera level	57
3.3.2	Data fusion level	61
3.3.3	Fall detection level	62
3.4	Materials and experiments	63
3.4.1	Experimental set-up	64
3.4.2	Fall scenarios	64
3.4.3	Data analysis	65
3.4.4	Statistical analysis	67
3.5	Results	68
3.5.1	VVDR behavior	68
3.5.2	Real-time implementation	69
3.6	Discussion	69
3.7	Conclusions	74
4	Discussion sur la chute	77
4.1	Commentaires sur la résistance à l'occultation	79
4.2	Propositions d'amélioration	79
4.2.1	Amélioration de l'étape de calibrage	80
4.2.2	Amélioration de l'algorithme de segmentation	80
4.2.3	Utilisation de caméras de profondeur	82
4.3	Conclusion	82
II	Étude de la marche	85
5	Analyse de la marche à partir d'une carte de profondeur	87
5.1	Introduction	89
5.2	Paramètres cliniques de la marche	91
5.3	La mesure quantitative de la marche	93
5.3.1	Mesure du mouvement monopoint	94
5.3.2	Mesure du mouvement multipoints	96
5.3.3	Bilan sur les méthodes de mesure	102

5.4	Mesure de la marche à l'aide de caméras	103
5.4.1	Système mono-caméra	104
5.4.2	Système multi-caméras	105
5.5	Système utilisant une caméra de profondeur	107
5.5.1	Système stéréoscopique	107
5.5.2	Structure temporelle	108
5.5.3	Structure spatiale	109
5.6	Utilisation des caméras de profondeur pour l'étude de la marche	112
5.7	Conclusions	114
6	Mesure de l'asymétrie des membres inférieurs avec une caméra de profondeur	115
6.1	abstract	118
6.2	Introduction	118
6.3	Previous works	120
6.4	Method	121
6.4.1	Depth map pretreatment	122
6.4.2	Key depth maps	123
6.4.3	Key depth map analysis for asymmetry measurement	124
6.5	Experimentation and validation	126
6.5.1	Experimentation	126
6.5.2	Validation	128
6.6	Results	128
6.7	Discussion	130
6.8	Conclusion	133
7	Discussion sur la mesure des paramètres de la marche	135
7.1	Mesure d'un index de la marche sans modèle biomécanique	137
7.2	Problème dû au capteur	139
7.2.1	Erreur due au codage de la profondeur	139
7.2.2	Erreur due au principe de mesure de la profondeur	141
7.3	Amélioration du protocole expérimental	142
7.4	En route vers la reconstruction 3D	142
7.5	Conclusions	144

8	Discussion et Conclusion	147
8.1	Contribution pour la communauté scientifique	149
8.2	Contribution pour la communauté médicale	150
8.3	Contribution pour la société	151

Chapitre

1

Introduction

Sommaire

1.1	Systèmes mis en jeu dans le mouvement de la marche	21
1.1.1	Le système musculo-squelettique	21
1.1.2	Le système nerveux	22
1.2	Le mouvement de la marche	23
1.3	L'impact de la pathologie sur la marche	25
1.4	Relation entre la marche, l'âge et la chute	26
1.5	Objectifs des travaux de la thèse	27

Le corps humain est capable de prouesses motrices très variées et complexes. Pour y parvenir, il est nécessaire de mobiliser près de 600 muscles, 200 os et une centaine d'articulations. Pour certains mouvements comme la marche, il en résulte une automatisation lentement acquise [39]. Cette automatisation s'effectue grâce au système nerveux central (SNC) qui intègre les informations provenant des capteurs sensoriels pour coordonner les commandes envoyées au système musculo-squelettique. Cette coordination est primordiale et dépend du bon fonctionnement général du SNC. Dès lors, si l'un de ces systèmes vient à faillir, le mouvement, en tant que résultante, est affecté. Les défaillances peuvent être le fait de causes très multiples, dont un grand nombre sont plus ou moins directement dues à l'avance en âge.

Ces dernières décennies, du fait des améliorations de la médecine, la part de la population des seniors dans les pays développés augmente de manière significative. L'un des effets associés au vieillissement est une régression du SNC (avec les pathologies pouvant être associées, comme Parkinson ou Alzheimer) et du système musculo-squelettique (avec les pathologies associées : arthrose, ostéoporose ...). Il en résulte alors une perte de la qualité du mouvement, que l'on observe tout particulièrement pour la marche et qui peut conduire, dans certains cas, jusqu'à la chute. Ainsi, la mesure de la qualité de la marche permet de qualifier le bon fonctionnement des différents systèmes nécessaires au contrôle du mouvement, en général. Ainsi par une mesure quantitative des paramètres de la marche, il peut être possible de diagnostiquer certaines pathologies ou encore de suivre leur évolution.

Globalement, pour effectuer ces diagnostics cliniques, on peut regrouper les approches en deux grandes familles :

- la mesure d'un paramètre en particulier, en un point de mesure, qui permet d'avoir une variable unique représentant l'état global du système. C'est le cas des mesures de type accélérométrique qui ont été utilisées pour quantifier l'irrégularité de la marche [16] ou bien encore pour détecter la chute d'une personne [28].
- les mesures en de multiples points, comme les systèmes opto-électroniques impliquant la pose d'un groupe de marqueurs sur des repères anatomiques. Dans ce cas, on fait l'hypothèse que l'information pertinente peut être répartie à différents lieux géographiques du corps, conduisant à des index globaux combinant plusieurs critères locaux [41]. Généralement, la localisation de ces points de mesure est liée à un modèle sous-jacent, qui a été défini en fonction des hypothèses initiales sur les paramètres à observer (généralement les principales articulations mises en jeu).

De ces deux approches sont nés un grand nombre de systèmes de mesures et d'index quantifiés qui reposent donc sur un échantillonnage plus ou moins fin du corps humain. D'un côté, les systèmes mono-point permettent de manipuler un paramètre unique, mais risquent de passer à côté d'informations locales très pertinentes. D'un autre côté, les systèmes multipoints tiennent compte d'un grand nombre de paramètres (déterminés à l'avance) mais rendent difficile l'analyse en raison de la taille du vecteur d'état manipulé.

Dans le domaine biomédical, l'un des grands défis est de définir le jeu minimal de paramètres (et la méthode de mesure associée) permettant d'effectuer un diagnostic pertinent du système de contrôle de la motricité. Dans cette thèse en génie biomédical, nous proposons de manipuler un échantillonnage beaucoup plus fin du corps humain (aller au-delà des quelques repères anatomiques utilisés dans les systèmes multipoints), multipliant les points de mesure, tout en ne retournant que l'information pertinente pour l'analyse clinique de la motricité. Pour cela, nous mettons en œuvre un nouveau type de mesure fondé sur l'enveloppe tridimensionnelle du sujet que nous illustrons dans deux applications complémentaires : la détection de la chute et la mesure des troubles de la marche.

En plus des aspects fondamentaux, il est important de rappeler que cette thèse s'est effectuée dans le cadre d'une convention Cifre qui implique un partenaire industriel et, donc, des résultats pratiques pouvant être facilement transférables. Dans le cas de cette thèse, l'industriel est une clinique, les "Ateliers de la marche", qui demande donc une portée clinique à ces recherches. Les méthodes abordées dans cette thèse doivent donc tenir compte de ces contraintes afin de proposer des méthodes facilement utilisables en clinique.

La suite du mémoire se focalise sur le cas particulier de la marche qui est l'un des mouvements les plus étudiés d'un point de vue clinique. Dans un premier temps, nous rappelons les principaux systèmes mis en jeu dans le contrôle de la motricité, et en particulier la marche. Nous présentons ensuite la marche d'un point de vue général, et comment les principales pathologies liées à l'avance en âge peuvent l'affecter. De là, nous donnons les principaux objectifs de la thèse et présentons plus précisément le plan du mémoire.

1.1 Systèmes mis en jeu dans le mouvement de la marche

La marche est un mouvement quotidien, que l'on acquiert dès le plus jeune âge, mais qui fait intervenir de nombreux processus de contrôle. Dans cette section, nous décrivons les deux principaux systèmes mis en jeu dans le contrôle de la marche : le système musculo-squelettique qui fait référence à des capacités motrices, et le système nerveux qui le pilote.

1.1.1 Le système musculo-squelettique

Le corps humain est doté d'une structure solide grâce aux os et mobile grâce aux articulations qui les relient. Cette ossature peut être décomposée en un ensemble de segments tels que le pied, la jambe, la cuisse, le bassin, le tronc, le bras, l'avant-bras, la main et la tête. Les articulations agissent comme des liaisons mécaniques (généralement des rotules) laissant un certain degré de liberté de mouvement.

Les muscles agissent sur le squelette en lui appliquant des forces aux différents points d'insertion. Ils ne peuvent agir qu'en contraction le long de la ligne reliant deux points d'insertion, ou des points de passage sur le squelette [45]. Le principe de fonctionnement consiste à resserrer le chevauchement de fibres d'actine et de myosine situées au cœur des cellules musculaires, comme illustré sur la figure 1.1.

Pour pouvoir effectuer un mouvement, deux muscles antagonistes sont placés de part et d'autre de l'articulation pour, alternativement exercer un moment de force sur celle-ci, favorisant ou luttant contre le mouvement. La mise en mouvement est due au résultat non nul du bilan des moments de force au centre de rotation. Au-delà des simples muscles agonistes et antagonistes, chaque articulation est associée à un ensemble de muscles, dont certains sont mono-articulaires, et d'autres agissent conjointement sur plusieurs articulations. Le système est dit redondant, car il existe une multitude de coordinations musculaires entraînant le même effet sur l'articulation. C'est la variation du nombre de petits groupes musculaires redondants recrutés qui permet de faire varier l'intensité de la force exercée. La nécessité de contrôler efficacement la contraction de chaque muscle est très importante pour permettre d'effectuer le mouvement désiré, c'est le rôle du système nerveux.

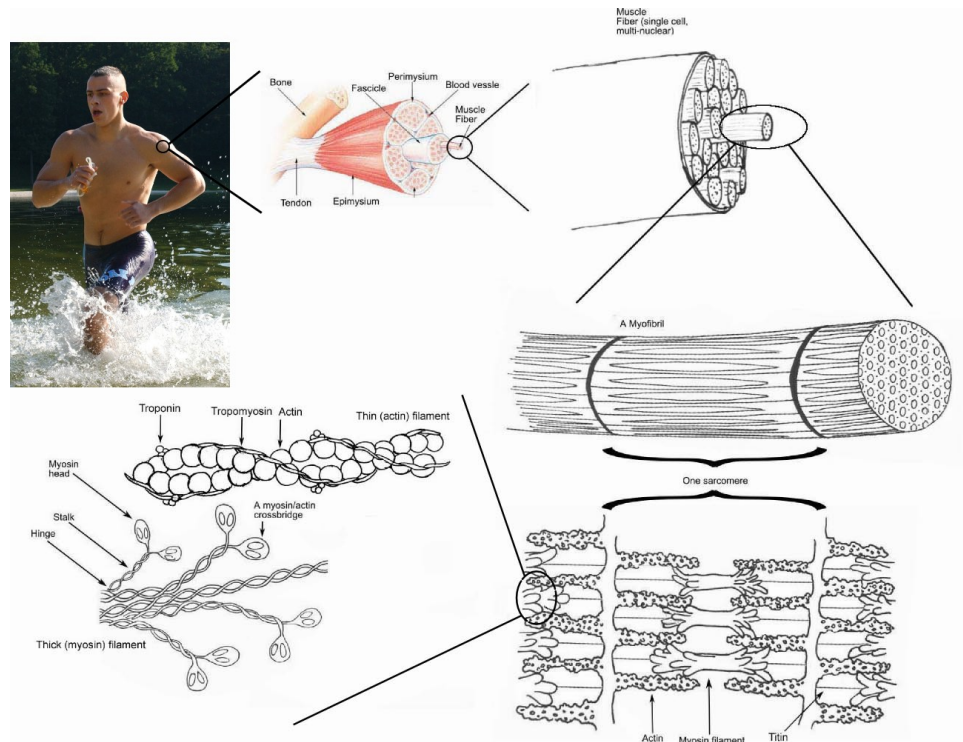


FIGURE 1.1 – Schéma de description du muscle. Source Wikipedia

1.1.2 Le système nerveux

Le système nerveux a pour objectif de mesurer, véhiculer et intégrer l'information. Pour cela il se décompose en deux parties, le système nerveux périphérique (SNP) et le système nerveux central (SNC). Le premier amène les informations des capteurs sensoriels vers le système nerveux central et de celui-ci aux muscles. Le système nerveux central intègre les données afférentes, les analyse puis élabore des commandes qui sont ensuite envoyées au système nerveux périphérique pour être relayées aux muscles.

Pour contrôler efficacement un mouvement, le système nerveux central se fonde sur des informations internes et externes. Les informations internes proviennent des capteurs d'extension qui prennent place directement au niveau des muscles et des ligaments dans le but d'indiquer leur état de tension. On parle alors de proprioception. L'oreille interne, située derrière l'oreille, contient le système vestibulaire qui donne une information sur la direction de la gravité.

Les informations externes proviennent principalement de la vue et des capteurs de pression situés sur la peau et plus particulièrement sous les pieds. Les autres sens peuvent aussi être impliqués, généralement dans une moindre mesure. Le système ner-

veux central traite ces informations pour définir les commandes musculaires nécessaires pour effectuer le mouvement voulu. Dans ces processus de traitement, il existe des boucles de rétroaction qui adaptent la commande musculaire en fonction du déroulement du mouvement retourné par le SNP et de la commande.

Un grand nombre de pathologies peuvent affecter la prise d'information, ou une ou plusieurs parties du système SNC ou du SNP. L'effet de ces pathologies est donc visible sur le mouvement, ce qui a conduit à définir de nombreux tests cliniques fondés sur une analyse quantifiée de ces mouvements. C'est ce que nous abordons dans les sections suivantes, en particulier pour le mouvement de marche.

1.2 Le mouvement de la marche

Dans cette thèse, nous abordons le mouvement spécifique de la marche puisqu'il est l'un des plus étudiés en clinique, avec la possibilité de diagnostiquer de nombreuses pathologies des systèmes nerveux et musculo-squelettiques. C'est un mouvement destiné à déplacer le corps qui est quasi-périodique. Un cycle de marche est ainsi défini par un intervalle de temps séparant deux événements similaires, comme les deux poses successives du talon gauche. Un pas est alors défini par l'intervalle de temps qui sépare un événement d'un côté du corps avec son homologue du côté opposé, comme l'intervalle entre la pose du talon gauche et du droit. Le cycle de marche se caractérise de trois manières différentes : spatiale, temporelle et fonctionnelle. La description spatiale revient à énumérer les différents événements qui se succèdent durant un cycle de marche (en nomenclature standard) d'après [165] et [73] (Figure 1.2 et 1.3) :

- Le contact initial du talon droit, début par convention.
- Le décollement des orteils opposés
- Le décollement du talon
- Le contact initial opposé
- Le décollement des orteils
- Les pieds adjacents
- Le tibia vertical

La description temporelle suit le même schéma et se fonde sur les durées de la phase de double appui (définie par le contact simultané des deux pieds sur le sol, qui intervient entre deux appuis monopodaux), de la phase d'appui monopodal (le pied du côté étudié est en contact avec le sol) et de balancement (où le pied du côté étudié n'est pas en contact

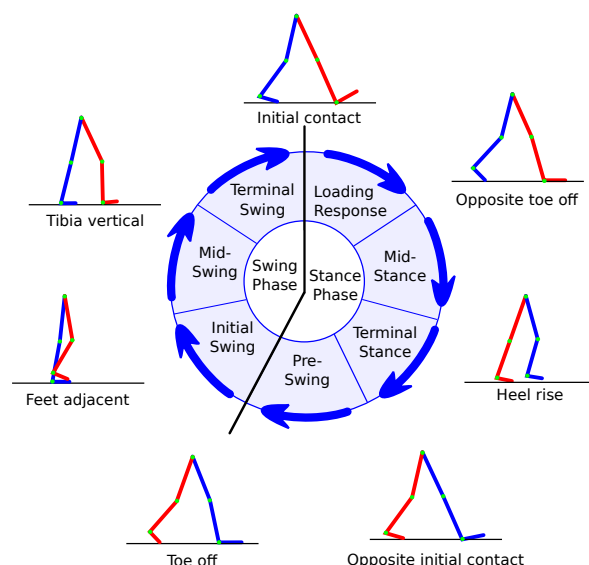


FIGURE 1.2 – Les différentes phases de la marche.

avec le sol) (Figure 1.3).

La description fonctionnelle propose d'associer à chacune de ces phases, une réalité biomécanique. Entre le contact initial et le décolllement du pied opposé, se trouve la phase de mise en charge où le pied devient un appui et supporte le poids du corps. La période allant du décolllement du pied opposé jusqu'au décolllement du talon, est la phase initiale de l'appui monopodal (Mid-stance). Elle est associée au balancement du pied opposé jusqu'au dépassement du tibia transformant l'énergie potentielle en énergie cinétique. La phase terminale de l'appui qui commence au décolllement du talon et va jusqu'au contact initial du talon opposé est une phase de propulsion (idem de l'autre côté). Durant les phases de simple appui, le corps est dans un état instable. Une coordination efficace des différents membres inférieurs et supérieurs doit être mise en oeuvre pour conserver la posture et éviter la chute. Ceci se traduit par une rotation croisée des ceintures scapulaires et pelviennes (ligne C de la figure 1.4).

Le plus souvent, les paramètres utilisés pour caractériser la marche d'un point de vue global sont la longueur et la fréquence du pas. De ces deux grandeurs, il est possible de déduire la vitesse de marche qui est le produit de la longueur et de la fréquence du pas. Pour avoir plus de précision, il est possible d'accéder à d'autres types d'informations : les angles aux articulations, les moments articulaires, les forces de réaction du sol, les activations musculaires via des systèmes électromyographiques... Toutes ces informations peuvent apporter des éléments pour mieux caractériser et quantifier un dysfonctionne-

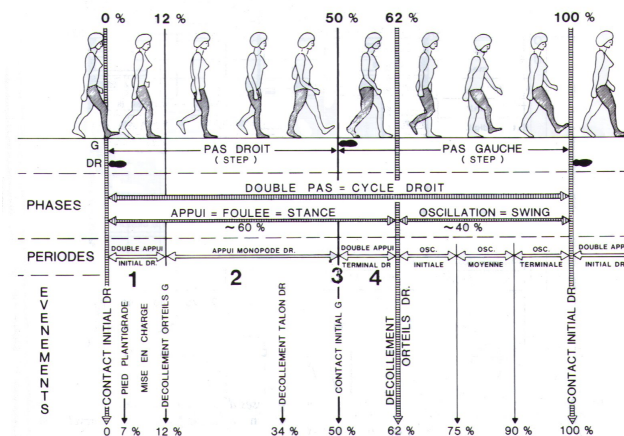


FIG. 2. — Découpage temporel du cycle de marche. La terminologie regroupe les principaux synonymes retrouvés dans la littérature francophone. Les pourcentages sont des moyennes pour une population normale ayant une marche de type adulte. (Adapté d'après Taillard et Blanc, 1988 [11]).

FIGURE 1.3 – Schéma du cycle de la marche (d'après Pélissier, 1994).

ment de tout ou partie du SNC ou du SNP. La difficulté est d'extraire les informations pertinentes parmi l'ensemble des variables disponibles, ce qui a donné lieu à de nombreux travaux et dont nombre de questions sont encore ouvertes du point de vue scientifique et clinique.

1.3 L'impact de la pathologie sur la marche

Si une pathologie venait à modifier le fonctionnement d'un système mis en jeu dans le contrôle de la motricité, le mouvement s'en retrouve directement affecté. Pour mieux comprendre l'impact de différentes pathologies sur le mouvement de marche, les principales approches consistent à définir une métrique pour calculer la distance entre les paramètres d'un patient (ou d'un groupe de patients) et une population témoin [41]. Le but recherché dans cette démarche est de quantifier la sévérité et de localiser de la pathologie du patient.

Prenons l'exemple d'un dysfonctionnement du système musculo-squelettique qui peut être dû à une asymétrie des segments osseux ou à une altération du fonctionnement des articulations, comme de l'arthrose. Dans le premier cas, Gurney relate dans [70], que la longueur du pas de la jambe la plus longue est augmentée par rapport à la jambe la plus courte. Il en est de même pour la durée de chaque pas, plus long pour la jambe la plus longue. En cas d'arthrose [163] le mouvement et les efforts au niveau de l'articulation atteinte sont modifiés par rapport à une articulation saine. De même, la cadence, la

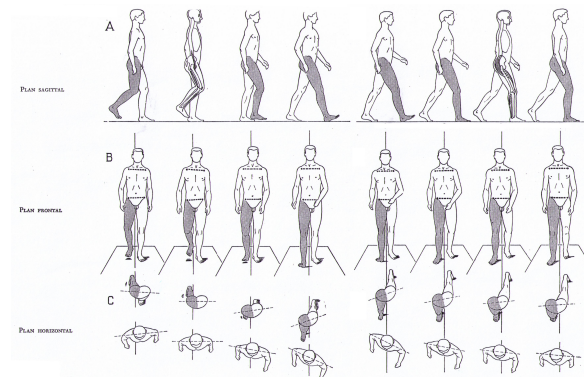


FIGURE 1.4 – Posture clé du mouvement de la marche pour un demi-pas (d’après Ducroquet)

vitesse, la longueur des pas sont diminuées de plus de 20% par rapport à un sujet sain. La durée du double support se retrouve augmentée de 55%. Dans tous ces cas, les paramètres globaux de la marche peuvent suffire à identifier le dysfonctionnement, mais une exploration plus avant des paramètres cinématiques, dynamiques et électromyographiques peut permettre de mieux mesurer l’impact et la gravité d’une pathologie.

Avec l’avance en âge, les pathologies du système nerveux se multiplient et peuvent avoir un effet considérable sur la fonction motrice [85]. Les maladies les plus connues dans ce domaine sont principalement Alzheimer et Parkinson qui provoquent des déformations quantifiables de la marche.

1.4 Relation entre la marche, l’âge et la chute

Grâce au progrès de la médecine, la durée de la vie s’est allongée. Malheureusement, la décroissance des capacités locomotrices apparaît avec l’avance en âge, du fait de l’affaiblissement des systèmes mis en jeu, et en particulier musculosquelettique, SNC et SNP. Pour le système musculo-squelettique, les exemples les plus courants sont la dégénérescence des tissus musculaires [47], l’arthrose pour les articulations [7] ou encore l’ostéoporose pour les os [65]. Le système nerveux central n’est pas épargné. Les maladies neurodégénératives attaquent les cellules du cerveau provoquant une perturbation de la motricité, et en particulier de l’équilibre [147], ou encore l’apparition des maladies d’Alzheimer ou de Parkinson [78].

Ainsi, ces troubles de la marche peuvent provoquer chez la personne âgée une perte d’équilibre entraînant la chute. Comme il est présenté plus en détail dans la première par-

tie de cette thèse, les conséquences peuvent être dramatiques lorsque la chute intervient au domicile d'une personne vivant seule. En cas de perte de connaissance, il survient alors une situation de détresse qui ne peut être évitée que par l'utilisation de systèmes automatiques de détection de chute.

Même si détecter la chute est important pour pouvoir porter secours à la personne, il est tout aussi pertinent d'identifier les populations à risque. Ceci peut s'effectuer en menant à bien une analyse quantifiée de la marche qui permet alors de diagnostiquer un dysfonctionnement pouvant entraîner la chute à plus ou moins long terme. Pour toucher un maximum de sujets, il est nécessaire que ces analyses puissent être menées à bien en dehors de grands centres de rééducation ou des hôpitaux. Cet usage en clinique s'accompagne de contraintes particulières : mesures simples d'utilisation et systèmes de mesure abordables.

1.5 Objectifs des travaux de la thèse

L'objectif de cette thèse est donc d'explorer une nouvelle approche de l'analyse clinique pour le diagnostic de pathologies associées à l'avance en âge. La prévention de la chute et l'analyse quantifiée de la marche sont particulièrement étudiées dans ce travail, car elles correspondent à deux mouvements largement étudiés dans les pathologies liées à l'avance en âge. Pour chacun de ces mouvements, nous dressons un état de l'art des méthodes d'analyse clinique en mettant en évidence la dichotomie qui existe entre systèmes fondés sur un point unique de mesure et ceux en utilisant plusieurs. Comme nous l'avons évoqué dans l'introduction, ces deux approches opposées, ou complémentaires, s'appuient sur un échantillonnage plus ou moins fin du corps du patient étudié. Le défi est de faire apparaître un ensemble très restreint d'informations pertinentes du point de vue clinique, alors que l'information disposée est de grande dimension. Les approches classiques consistent donc à systématiquement fournir les angles aux articulations pour ensuite effectuer des mesures de distance par rapport à une population de référence, ou à définir un ensemble d'indicateurs (dissymétrie, irrégularité, instabilités...).

Contrairement à ces approches qui se fondent sur une connaissance du squelette, et un échantillonnage plus ou moins fin du corps, pour évaluer des angles articulaires, nous proposons dans cette thèse de manipuler des données très finement échantillonnées (le volume ou la surface du sujet au cours du temps). L'idée principale est d'accéder directement à des paramètres cliniques, sans faire d'hypothèses fortes sur le squelette sous-

jacent. Ainsi, avec cette approche, plus besoin de rechercher les centres articulaires à partir d'un jeu de marqueurs, de modéliser les articulations de manière approximative, ou encore de définir des indices reposant sur une combinaison complexe d'angles. Les derniers travaux de la littérature en biomécanique tels que relatés par Fohanno dans [55], montrent d'ailleurs que ces méthodes d'estimation de trajectoires articulaires étaient fortement affectées par le jeu de marqueurs choisi, favorisant justement un grand nombre de marqueurs par segment pour fiabiliser le résultat. En rupture avec cette recherche de trajectoires articulaires de plus en plus précises, nous proposons de travailler directement sur les critères cliniques et sur leur observabilité dans le volume ou sur la surface tridimensionnelle du patient.

Une première partie s'intéresse donc à détecter la chute en milieu naturel, riche en occultations et en variétés de situations (cf. partie I). Cette première partie se fonde sur un système multi-caméras qui est alors exploité pour reconstruire un volume du patient et fournir un critère robuste de détection de chute. La deuxième partie propose d'utiliser une caméra de profondeur de type Kinect (produit de Microsoft) pour accéder à la surface du patient lors de la locomotion sur tapis roulant. Le critère clinique étudié concerne alors l'asymétrie de la marche d'un point de vue temporel et spatial, sans passer par une estimation d'angles aux articulations (cf. partie II). La thèse se termine par une discussion et une conclusion sur les contributions proposées.

Première partie

Étude de la chute

Chapitre 2

Détection de la chute

Sommaire

2.1	Relation entre l'âge et la chute	33
2.2	Quelques chiffres	34
2.3	État de l'art	34
2.3.1	Appareil d'alerte des secours en cas de chute	34
2.3.2	Dispositif de détection de chute automatique	35
2.4	Problématique	40
2.5	Méthode	40
2.5.1	Mode de représentation du volume	42
2.5.1.1	Représentation géométrique du volume	43
2.5.1.2	Représentation voxelique du volume	44
2.5.2	Discussion du choix de représentation du volume	45
2.5.3	Le calibrage	45
2.5.3.1	Calibrage des paramètres des caméras	46
2.5.3.2	Correction de la distorsion	46
2.5.4	Mise en œuvre	48
2.6	Conclusions	49

Avec l'allongement de la durée de la vie grâce au progrès de la médecine, différents troubles inhérents à l'usure du corps par les années apparaissent de plus en plus. L'un des effets associés au vieillissement est une régression du SNC et du système musculo-squelettique (arthrose, ostéoporose ...), qui peut se manifester par une perte de la qualité de la marche. L'une des conséquences directes est la chute de la personne âgée. En effet, face à l'émergence de cette problématique de santé publique, de nombreux pays industrialisés ont initié différents travaux de réflexion pour la prise en charge de ce fléau. Ceux-ci ont rapidement été rejoints par tous les autres pays du globe. Ces travaux de réflexion ont été menés dans tous les secteurs de recherche rattachés au système de santé, que cela soit de l'analyse au niveau épidémiologique, à la prise en charge directement du patient. La synergie de cette réflexion pluridisciplinaire a permis de définir plusieurs référentiels de bonnes pratiques permettant la formation des personnels de santé et l'information des populations.

2.1 Relation entre l'âge et la chute

La dégradation de la démarche du fait de l'âge a été décrite par Prince et al. dans sa revue de littérature [134]. Un premier indice de ce fait se trouve être l'accroissement de l'énergie consommée pour se propulser. En effet, même si la marche est un mouvement complexe, il est exécuté de manière à minimiser la consommation d'énergie [39]. Pour les personnes âgées, la consommation d'oxygène, donc d'énergie est plus importante même si la vitesse de marche est inférieure. Cette perte d'optimalité du mouvement laisse supposer une baisse de qualité des fonctions locomotrices.

Pour mieux comprendre l'effet de l'âge sur la marche, les différents paramètres ont été étudiés en fonction de l'âge. Les différents paramètres de la démarche déclinent avec l'accroissement de l'âge. Par exemple, la vitesse décroît par une réduction de la longueur et de la cadence des pas [134]. De nombreux travaux ont été réalisés dans le but de mieux comprendre les raisons pour lesquelles les aînées venaient à chuter. Tout d'abord, les liens entre la chute et différentes pathologies ont clairement été définis. Par exemple, les patients atteints de la maladie de Parkinson [25], du diabète [116] sont plus enclins à chuter.

2.2 Quelques chiffres

Chaque année, un tiers des personnes de plus de 65 ans chute [44], plaçant celle-ci comme le sixième facteur de mortalité par blessure de cette tranche d'âge aux États-Unis [120]. Environ 2.2 millions de chutes ont nécessité l'intervention des secours et le quart a entraîné une hospitalisation en 2009 [56], toujours dans ce pays. Ainsi, 40% des chutes sérieuses provoquent une hospitalisation et 30%-40% de ces hospitalisations entraînent un transfert en maison médicalisée [148] et d'après [137], chaque année 12033 hospitalisations de personnes de plus de 65 ans sont dues à une chute sur la période de 2000 à 2003 aux États-Unis. D'autre part, cet évènement provoque un traumatisme psychologique souvent à l'origine du syndrome de la peur de la chute, qui à lui seul entraîne une perte d'autonomie [121]. En 1995, le coût global des chutes en France a été évalué à plus de 1.07 M€[11]. D'après le centre pour le contrôle des maladies et de la prévention américain (CDC), la chute a coûté 19 milliards de dollars en 2000 et ce coût aurait atteint les 28 milliards en 2010 [156]. Il en résulte que les chutes et leurs conséquences sont identifiées comme l'un des problèmes majeurs de santé publique.

2.3 État de l'art

Pour répondre à ce besoin de santé publique, de nombreux travaux ont été menés dans différents domaines comme la prévention avec l'édition de guide de bonne pratique pour la France [27] et pour le Québec [61] ou encore la recherche de solutions techniques améliorant la sécurité des personnes à domicile. Le but d'une partie des solutions techniques consiste à permettre l'appel des services de secours dans le cas où la personne ayant chuté se trouve dans l'incapacité de se relever. Une brève présentation de ces types d'appareils d'alerte va maintenant être proposée.

2.3.1 Appareil d'alerte des secours en cas de chute

Pour permettre à la personne d'alerter les services de secours ou des proches lorsqu'elles ont chuté et qu'elles se trouvent dans l'incapacité de se relever, différentes solutions techniques ont été proposées. La première solution immédiate qui a été rapidement mise en œuvre, consiste à équiper la personne d'un dispositif d'appel portatif. Celle-ci se résume à porter un combiné téléphonique ou encore un bouton d'appel en permanence. Il nécessite que la personne soit toujours consciente après la chute, car malheureusement, en cas de perte de connaissance, il est impossible à la personne de déclencher l'alerte.

C'est pourquoi de nombreux travaux ont été mis en œuvre pour trouver le moyen de détecter automatiquement la chute sans l'intervention du sujet. Pour cela, différents types d'information et de technologie ont été mis à profit.

2.3.2 Dispositif de détection de chute automatique

Pour détecter la chute, différentes mesures physiques caractérisant la chute ont été mises à profit. Les caractéristiques de la chute recherchées sont parmi les informations d'évolution de paramètres spatiaux au cours du temps (position, vitesse, accélération), de l'interaction du sujet avec son environnement (par exemple l'impact avec le sol) ou encore de l'information audio (détection des cris).

Dans [166], Wu a défini des paramètres qui permettent de distinguer les chutes des activités quotidiennes. Principalement, la chute se distingue par des vitesses verticales et horizontales deux à trois fois supérieures à la vitesse normale et dans la plupart des cas l'augmentation de ces deux vitesses se fait de manière synchrone 300 à 400 millisecondes avant l'impact au sol.

On peut séparer en deux groupes les appareils de détection automatique de la chute. Le premier groupe contient les appareils ambulatoires qui nécessitent d'être portés comme les accéléromètres [28], les gyroscopes [101] ou une centrale inertielle qui combine les deux [128]. Ceux-ci enregistrent les accélérations au niveau d'une ou plusieurs parties du corps pour détecter le changement brusque de vitesse dû soit à la chute en elle-même par la perte d'altitude, soit par l'impact avec le sol ou soit par un changement brusque de l'angle du thorax. Seulement ces appareils nécessitent d'être portés, ce qui peut parfois être considéré comme intrusif dans les habitudes de vie de la personne.

L'autre groupe est constitué du reste des appareils qui sont installés de manière fixe dans l'habitat de la personne. Les événements recherchés sont, par exemple, l'impact avec le sol comme le font Alwan et al. dans [12] ou encore Zigel et al. qui y adjoignent l'information sonore dans [171]. Seulement dans le cas de chutes lentes, par exemple à partir d'une chaise, les auteurs précisent que la sensibilité de leur système est insuffisante. Enfin Li et al. dans [102] ne se basent que sur le son de l'impact pour détecter la chute.

Nyan et al. dans [127] ont utilisé un système de capture de mouvement classique avec des repères placés sur la cuisse, la ceinture et le thorax pour venir détecter la chute en fonction de l'angle de la cuisse avec le torse. Leur méthode arrive à détecter la chute avant l'impact, mais nécessite un système de mesure du mouvement opto-électronique comme un Vicon très coûteux.

À l'instar du développement de la vidéosurveillance, la caméra vidéo a été utilisée comme source d'information. La vidéo est principalement utilisée pour définir la position de la personne (debout, assise, allongée, ...) à partir des paramètres de la silhouette ou pour mesurer la vitesse de la personne.

La première approche consistant à analyser la position de la personne se base sur le fait que la plupart du temps la chute finit par une position allongée sur le sol comme le précise la définition de l'institut canadien d'information sur la santé [46]. Pour différencier cette position de toutes autres positions usuelles de la vie courante, des chercheurs [13, 160, 159] ont pour cela modélisé la silhouette de la personne à l'aide respectivement d'une boîte englobante et d'une ellipse. Ainsi lorsque le ratio hauteur/largeur ou respectivement l'orientation de l'ellipse sont anormaux tels que montrés sur la figure 2.1, une chute est alors détectée.

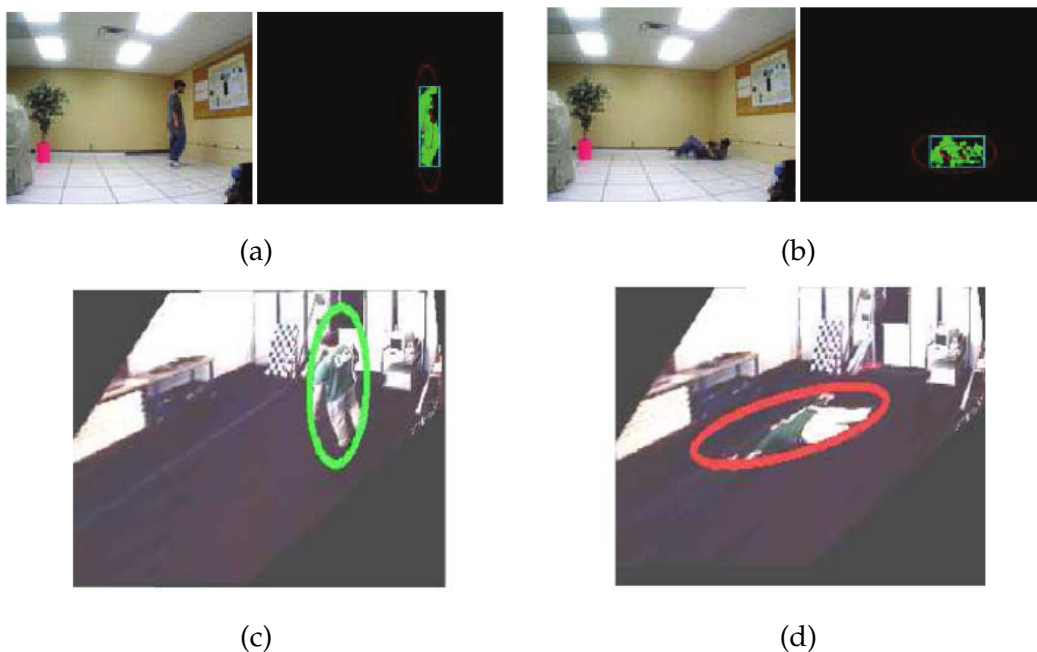


FIGURE 2.1 – Exemples de détection de chute avec les caractéristiques de la silhouette. Détection d'une personne debout (a) et allongée (b) par une boîte englobante et par une ellipse (c) et (d). Source [13, 160]

Seulement, dans certains cas de positionnement du corps dans l'axe de la caméra, la silhouette peut ne pas permettre la détection de la chute comme illustrée sur la figure 2.2.

Comme l'a décrit Wu dans [166], la vitesse est une information déterminante pour la reconnaissance de la chute. Pour mesurer cette vitesse, Sixsmith et Johnson [153] ont



FIGURE 2.2 – Exemples de mauvaise classification de posture à partir de la silhouette. La personne est considérée comme debout par la méthode proposée utilisant l'orientation de la silhouette. Source [160]

utilisé une caméra infrarouge sensible à la température du corps. Leur méthode consiste à mesurer le mouvement de la personne en suivant les objets de températures différentes. Il détecte alors une chute lorsque la vitesse verticale est trop importante. Seulement la projection du vecteur vitesse 3D sur le plan 2D de l'image de la caméra empêche de pouvoir la mesurer précisément. Pour pallier ce problème, Rougier et al. modélisent la tête avec une ellipse pour venir retrouver l'information de profondeur et ainsi estimer la vitesse réelle de la tête. Un exemple de cette méthode est montré sur la figure 2.3. Même si cette approximation de la profondeur permet de mieux estimer le vecteur vitesse 3D à partir de l'image 2D, il reste une erreur moyenne de 5%, soit 25 cm à 5 mètres de distance. Cette erreur peut rendre l'estimation de la vitesse déficiente dans certains cas et peut empêcher la détection d'une chute.

Pour remédier à cela, Lee et Mihailidis [99] et Nait-Charif et McKenna [123] ont placé leur caméra au plafond. Leur méthode détecte la chute en mesurant la vitesse de la silhouette ou de l'ellipse qui approxime celle-ci comme présenté figure 2.4. Mais le champ de vision de ce genre de caméra reste limité et il est très difficile de mesurer une vitesse verticale à l'aide de ce genre de caméra.

Il en résulte une impossibilité de mesurer précisément une vitesse 3D avec une seule caméra du fait de la nature 2D de l'information mesurée par la caméra et une sensibilité très importante aux occultations. Il est donc nécessaire d'utiliser plusieurs caméras pour pouvoir estimer plus finement les informations spatiales propres à la chute.

Pour la mise en œuvre de système multi-caméras, il existe plusieurs niveaux de col-

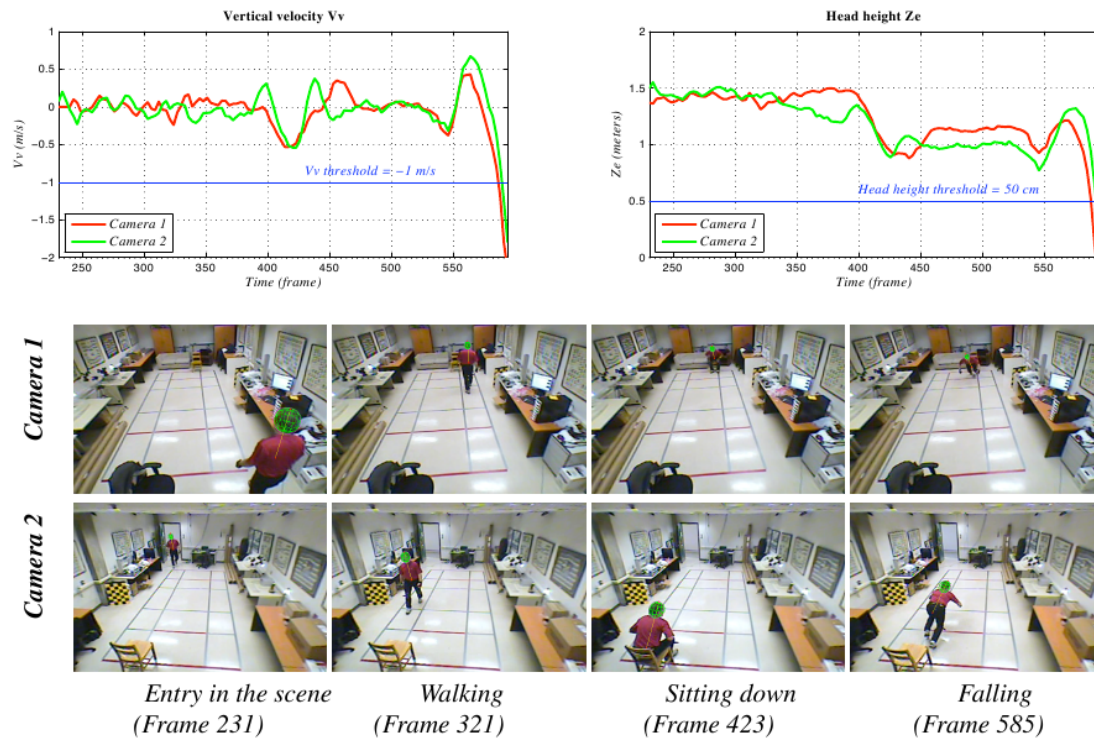


FIGURE 2.3 – Exemple du suivi de la tête et de la détection de la chute par estimation de la vitesse et de la hauteur. Source [138]

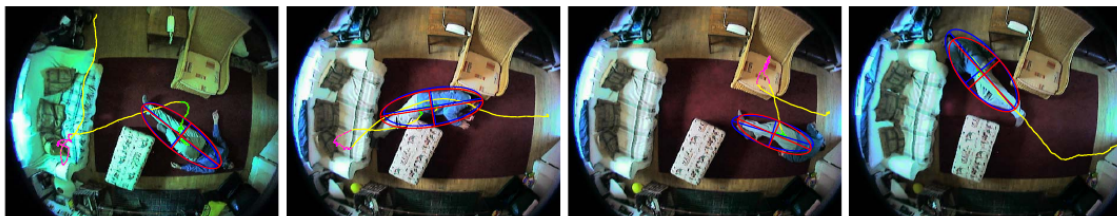


FIGURE 2.4 – Exemple de fonctionnement de la détection de la chute avec une caméra de plafond par estimation de la silhouette avec un ellipsoïde. Source [123]

laboration entre elles. Dans un premier temps, plusieurs caméras peuvent être mises à profit en fusionnant l'information de haut niveau calculée à partir de l'image de chacune des caméras pour inférer sur le résultat de détection de la chute provenant des différents points de vue. Dans ce cas le but est d'isoler l'information erronée par l'occultation. Thome et Miguet [160] intègrent les résultats d'une classification de la position de la silhouette provenant de plusieurs caméras à l'aide d'une fusion basée sur la logique floue. Rougier et al. présentent dans [143] une méthode qui utilise une classification par un modèle de mixture de gaussienne appliquée sur un indice calculé sur l'évolution du contour de la personne au cours du temps pour venir détecter une chute. La détection de la chute se fait par une méthode de vote sur le résultat de chacune des caméras.

Dans un deuxième temps, l'information de base est échangée entre les caméras pour combler les zones occultées dans chacune des caméras, mais la décision se fait toujours à partir d'information 2D. Cucchiara et al. dans [43] utilisent une classification à modèle markovien caché basée sur les histogrammes de projection de la silhouette de la personne selon l'axe horizontal et vertical de l'image pour définir la posture de la personne. Le partage d'informations entre les caméras se fait en projetant les pixels d'une des caméras vers l'image de l'autre à l'aide d'une projection homographique pour combler les portions occultées de la personne.

Dans un troisième temps, les informations de plusieurs caméras ont été fusionnées pour identifier des caractéristiques 3D de la position ou du mouvement de la personne. Hazelhoff et al. [75] fusionnent l'axe principal 2D de la personne dans chacune des deux caméras perpendiculaires et non calibrées pour estimer l'axe principal 3D de la personne dans la scène. Anderson et al. [14] utilisent une reconstruction 3D du volume de la personne effectuée par la méthode du visual hull et une analyse floue de l'axe principal de ce volume pour détecter la chute. Seulement la possibilité de résistance aux occultations n'a été notée que dans l'article de Rougier et al. [143] et Cucchiara et al. [43], mais sans la valider.

Jusqu'à maintenant, l'occultation a toujours été un problème soulevé dans les commentaires sans jamais avoir été intégrée comme une problématique majeure dont l'impact serait clairement pris en compte lors d'une étude.

2.4 Problématique

Comment parer à ce problème de l'occultation dans le cas de l'utilisation de caméras vidéos placées ailleurs qu'au plafond. Lorsqu'une personne se retrouve masquée dans une voir plusieurs images par un objet du mobilier (table, chaise, canapé), il devient difficile, voire impossible, de définir quelques paramètres que ce soit à propos de la personne. Pour résoudre ce problème, nous proposons d'utiliser plusieurs caméras comme sources d'information pour minimiser la probabilité d'occultation. Ce problème d'occultation est un problème clé qui se doit d'être étudié et résolu.

2.5 Méthode

Pour résoudre ce problème d'occultation, nous proposons d'utiliser plusieurs caméras. En effet, en multipliant le nombre de points de vue tout autour de la pièce, la probabilité qu'une partie de la personne soit totalement occultée diminue fortement. Chaque image provenant des caméras servira à reconstruire le volume de la personne. Ce volume sera l'information de base pour la détection de la chute au sol. Pour résister à l'occultation, nous proposons de nous baser sur une reconstruction du volume de la personne pouvant accepter l'absence de la personne dans une des caméras.

Pour reconstruire le volume d'une personne à partir de plusieurs caméras, différentes méthodes existent déjà.

La plus évidente consiste à mimer le système binoculaire de la vision humaine. C'est le principe de la stéréovision bi ou multi-caméras dont les principales méthodes sont présentées et évaluées par Seitz et al. dans [149]. Le principe de base est de mesurer une distance entre les projections d'un même point dans deux ou plusieurs caméras. Cette distance appelée disparité permet alors, connaissant la relation géométrique entre les caméras, d'obtenir la position de ce point dans l'espace commun par triangulation. La mise en correspondance des projections d'un point de l'espace dans chaque point de vue se base sur l'assomption que l'apparence de ces projections soit identique quel que soit le point de vue. De plus, il est nécessaire que l'aspect des points adjacents soit suffisamment différent pour pouvoir être correctement distingués. La figure 2.5 montre que ces conditions ne sont clairement pas respectées dans le cas de l'utilisation de nos caméras. En effet, dans le cas d'une surface homogène telle que les vêtements de la personne, les points correspondants sont difficilement distinguables les uns des autres. Ceci rend ce genre de méthode inutilisable.



FIGURE 2.5 – Exemples d’images provenant du dataset de chute. Ces images prises au même instant montrent une personne au début de la chute, avec pour chaque caméra l’image entière et une vue rapprochée.

Il existe une autre méthode basée sur la photocohérence : la méthode du sculpteur (ou “space carving”). Celle-ci consiste à découper l’espace étudié en cube élémentaire appelé voxel. Pour chaque voxel appartenant à l’enveloppe du volume total, la position de celui-ci dans chacune des images est calculée et si la cohérence des couleurs du pixel de chaque projection est correcte, le pixel est alors défini comme occupé. Sinon il est défini comme vide et est supprimé tel que montré figure 2.6(a). De manière itérative, tous les voxels appartenant à la surface externe sont alors testés et après plusieurs itérations, il ne reste que le volume minimal tel que montré figure 2.6(b). Mais cette méthode itérative est particulièrement coûteuse en temps de calcul et peut être mise en difficulté par le même problème d’uniformité de couleur dans la scène.

C’est pourquoi la méthode du visual hull déjà utilisée par Anderson dans [14] va être utilisée pour reconstruire la surface de la personne. Il est possible d’utiliser la méthode nommée “Shape from Silhouette” ou encore visual hull proposée dans [97]. Celle-ci correspond à segmenter la zone d’intérêt correspondant à l’objet dans chaque image. Ainsi en construisant une série de cônes infinis partant du centre optique et passant par le contour de chaque silhouette de chaque caméra, le volume d’intérêt se retrouve à l’intersection de ces cônes tel que présenté dans la figure 2.7. En effet, Il n’est pas nécessaire de reconstruire l’intégralité de la scène. Seul le volume de la personne sera reconstruit. De plus, la première partie de cette méthode correspondant à l’identification de la silhouette de la personne dans l’image est facilitée dans ce contexte. En effet, la personne étant en

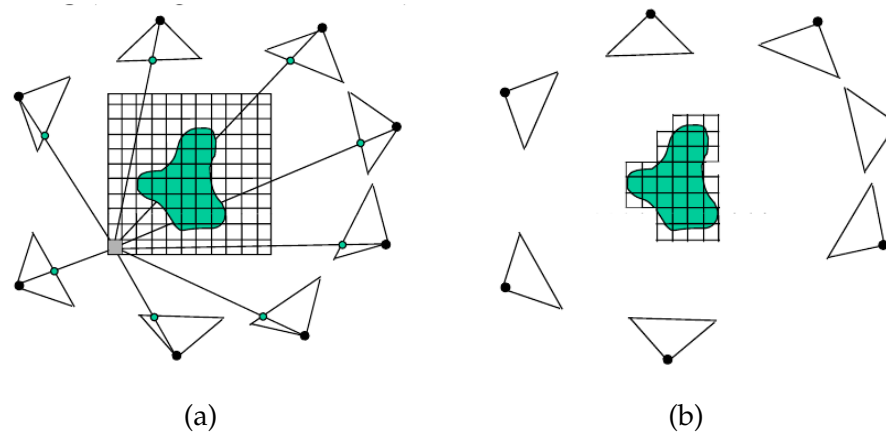


FIGURE 2.6 – Exemple du principe de l’algorithme du sculpteur. (a) Groupe de voxel initial où les projections dans chaque caméra du premier voxel ont été réalisées. (b) Groupe de voxels final après application de la méthode. Source [89]

mouvement en comparaison aux autres parties de l’image, la détection de la silhouette peut se faire avec une soustraction de fond, opération peu coûteuse. La reconstruction du volume est alors immédiate à l’aide de la projection de ces silhouettes.

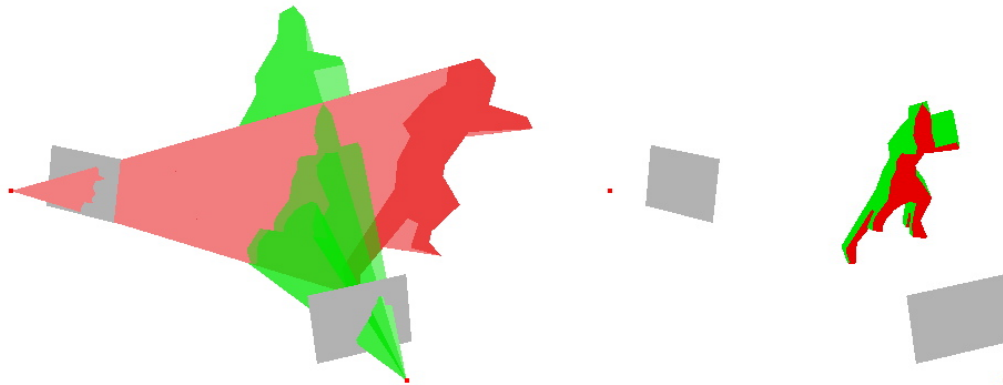


FIGURE 2.7 – Reconstruction 3D à partir de silhouettes en utilisant la méthode du visual hull. Le volume de l’objet est issu de l’intersection d’une série de cônes formés par le centre optique et le contour de la silhouette de chaque caméra. Source Wikipédia

2.5.1 Mode de représentation du volume

Il existe deux manières de mettre en œuvre cette projection. Elles se différencient au niveau du type d’espace de données utilisé. Le premier type est une espace géométrique,

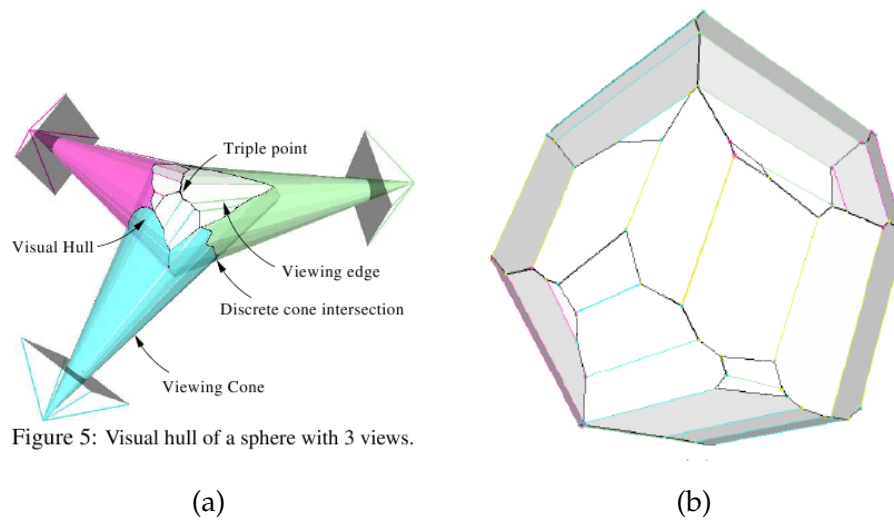


FIGURE 2.8 – Illustration de reconstruction d’une sphère par visual hull en géométrie computationnelle. Source Franco [57].

le second est un espace voxelique. Il est intéressant de voir comment la reconstruction du volume en cas d’occultation (c’est à dire que la contribution d’une caméra est manquante) et le calcul du volume en lui-même peut se faire avec ces deux modes de représentation.

2.5.1.1 Représentation géométrique du volume

Cette approche utilise une représentation du volume reconstruit sous forme d’un polyèdre géométrique comme décrit par Franco et Boyer dans [58]. Ce polyèdre est la forme spatiale issue de l’intersection des cônes géométriques formés par le centre optique de la caméra et le contour de la silhouette de chaque caméra. Le calcul de l’intersection se fait alors en utilisant des algorithmes issus de la géométrie computationnelle et donne un volume tel que montré figure 2.8. Cette méthode est simple dans le cas où toutes les caméras sont mises à contribution. Le polyèdre final Po_{Final} est alors l’intersection de tous les cônes C_i .

$$Po_{Final} = \bigcap_{i=0}^{i=n} C_i$$

Par contre dans le cas d’une reconstruction acceptant l’occultation, l’opération est alors plus compliquée. Le polyèdre final est l’union des polyèdres reconstruits en l’absence d’une caméra. En effet, dans le cas où il y a n caméras, soit Po_e le polyèdre issu de

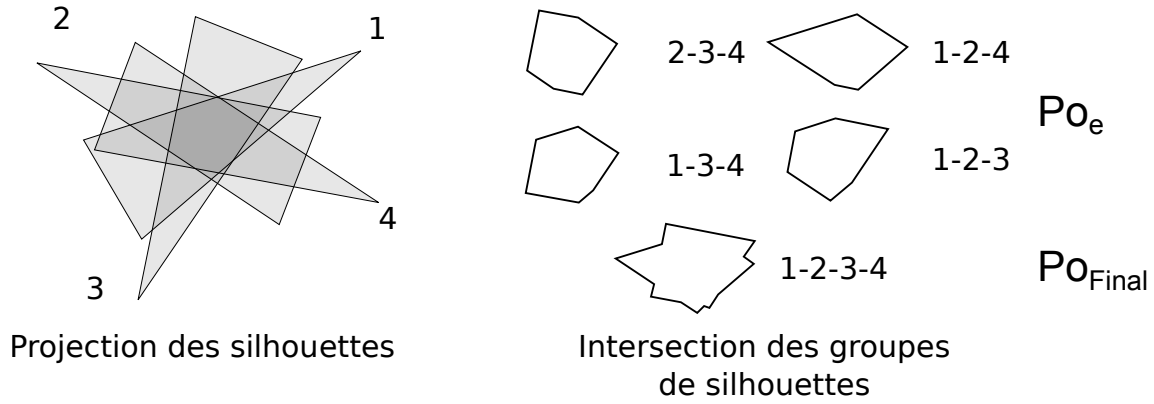


FIGURE 2.9 – Illustration d’intersections de polygones. Le polygone final Po_{Final} est égal à l’union des 4 polygones issus des intersections de 3 triangles.

l’intersection de toutes les caméras sauf la caméra e . Le polyèdre Po_{Final} se calcule alors :

$$Po_e = \bigcap_{i=0, i \neq e}^n C_i$$

$$Po_{Final} = \bigcup_{e=0}^n Po_e$$

La figure 2.9 permet de visualiser le problème dans le plan pour faciliter la représentation.

Comme le volume d’un polyèdre général n’a pas de formule algébrique, le calcul se fait alors en plusieurs étapes. Le volume d’un trièdre étant connu algébriquement, une décomposition du volume final en trièdre est réalisée en premier lieu. Puis la somme des volumes de tous ces trièdres est réalisée permettant ainsi d’obtenir le volume final.

2.5.1.2 Représentation voxelique du volume

Pour l’approche voxelique, l’espace de la scène est découpé en petites unités de dimension choisie nommées voxels. Pour réaliser la reconstruction du volume de la personne par la méthode de projection de silhouette, chaque voxel est projeté dans l’image de chaque caméra. Dès lors que toutes les projections se trouvent dans la silhouette de la personne dans chaque caméra, le voxel est défini comme occupé, sinon il est vide.

Dans le cas où il est nécessaire d’être robuste à l’occultation, il suffit d’utiliser la méthode “Extended Shape from Silhouette” proposée dans [113]. Celle-ci consiste à compter le nombre de caméras où le projeté du voxel appartient à la silhouette de la personne. Si cette somme est supérieure au nombre de caméras voulu alors le voxel est défini comme occupé.

Le calcul du volume est immédiat en sommant le nombre de voxels décrits comme occupés.

2.5.2 Discussion du choix de représentation du volume

Chacune de ces deux méthodes a ses avantages et ses inconvénients. La méthode utilisant une représentation géométrique utilise peu de points contrairement à la méthode voxélique, mais les algorithmes mis en œuvre sont complexes et sensibles aux erreurs d'arrondis. De l'autre côté, la représentation voxélique traite beaucoup de points, mais le traitement est plus évident. Nous avons donc fait le choix de la représentation voxélique pour plusieurs raisons. Elle semble plus adaptée à la problématique de reconstruction avec un nombre variable de caméras. Le calcul du volume est plus simple et efficace. Elle est plus rapidement maîtrisable et sa parallélisation sur carte graphique permet d'atteindre les contraintes de temps réel.

2.5.3 Le calibrage

Pour pouvoir calculer les coordonnées du projeté d'un voxel dans l'image d'une caméra, il est nécessaire de modéliser le comportement de la caméra. Le modèle le plus utilisé est celui du sténopé ("pin-hole" en anglais). Celui-ci se décompose en deux parties. D'abord la transformation d'un point 3D à partir de ses coordonnées dans le repère de la scène vers celui de la caméra à l'aide des paramètres externes qui relient le repère du monde à celui de la caméra. Puis la projection des points 3D dans le plan image de la caméra à l'aide des paramètres internes. Les paramètres externes sont composés des éléments classiques d'un changement de repère : une matrice de rotation $\mathbf{R}$ et un vecteur de translation $\mathbf{T}$. Les paramètres internes modélisent des éléments physiques de la caméra : la focale de la lentille f en millimètre, dx et dy les dimensions horizontales et verticales des pixels en millimètre, les coordonnées de projection de l'axe optique sur le capteur de la caméra $c(c_x, c_y)$ en pixels.

La mise en œuvre du calcul des coordonnées de la projection $\mathbf{x} = (x_p, y_p)$ dans l'image de la caméra d'un point $\mathbf{X}$ ayant pour coordonnées dans l'espace (X, Y, Z) se fait de la manière suivante :

Soit $\mathbf{X}_c = (X_c, Y_c, Z_c)$ la position dans le référentiel de la caméra

$$\mathbf{X}_c = \mathbf{R}\mathbf{X} + \mathbf{T} \quad (2.1)$$

Donc ceci intégré avec le modèle projectif du sténopé, nous avons l'équation suivante :

$$\begin{bmatrix} s.x_p \\ s.y_p \\ s \end{bmatrix} = \begin{bmatrix} f/dx & 0 & c_x & 0 \\ 0 & f/dy & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (2.2)$$

2.5.3.1 Calibrage des paramètres des caméras

Pour obtenir les paramètres d'une caméra, plusieurs méthodes existent. Les premières méthodes comme celle proposée par Tsai dans [161] utilisent un objet de référence tridimensionnelle. Un aperçu de ce type d'objet est montré par la figure 2.10. Seulement la construction précise de ce genre d'objet est difficile et son utilisation est quelque peu rendue difficile par son poids et son encombrement. C'est pourquoi plusieurs méthodes, telles que celle de Zhang relatée dans [169], se sont orientées vers une mire planaire. Pour cela, plusieurs images d'un échiquier orienté selon différents angles sont utilisées comme il est montré sur la figure 2.11. Une autre méthode se base sur un objet linéaire comme le propose Zhang dans [170]. Toutes ces méthodes nécessitent une information sur les positions des points de l'objet de référence. Certaines méthodes comme l'autocalibrage décrit par Hartley dans [72] et Faugeras dans [52] se servent de couples de points issus d'images d'une même scène, mais pris sous différents angles pour réaliser cette étape de calibrage.

Nous avons retenu la méthode planaire du fait qu'elle soit disponible sous différentes plateformes (Matlab et C) et qu'elle permette d'atteindre la précision nécessaire.

2.5.3.2 Correction de la distorsion

Le modèle de caméra projectif approxime au mieux le comportement d'une caméra. Certaines déformations ne sont pas considérées dans ce modèle comme celles de la distorsion. Celle-ci se décompose en deux formes de perturbations telles que présentées par Weng dans [164], la distorsion radiale et la distorsion tangentielle. La première peut prendre les formes caractéristiques de barillets ou de coussinets si le paramètre de la distorsion radiale est positif ou négatif. Ces deux déformations sont représentées sur la figure 2.12(2). La deuxième distorsion, dite tangentielle, est montrée sur la figure 2.12(3).

Tsai a montré dans [161] que la distorsion tangentielle est mineure devant la distorsion radiale. Ceci justifie que la distorsion tangentielle peut donc être négligée devant

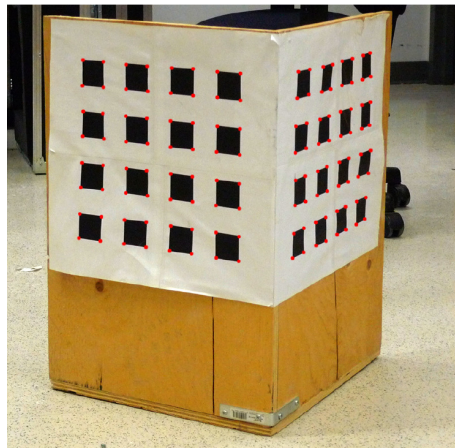


FIGURE 2.10 – Illustration d’une mire de calibrage 3D. Source [49]

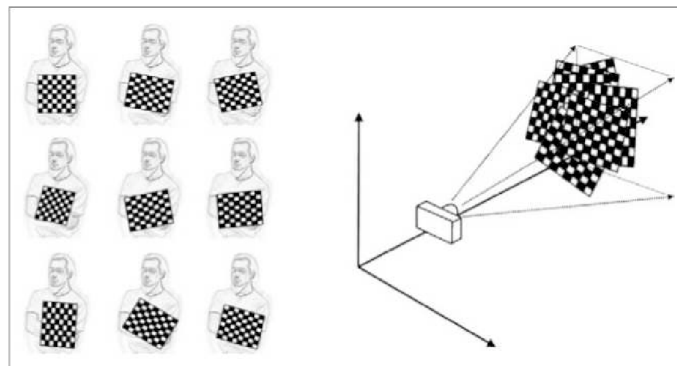


FIGURE 2.11 – Illustration de l’utilisation d’un damier pour le calibrage d’une caméra. Source [2]

la distorsion radiale. Les paramètres de distorsion peuvent y être introduits de la même manière que proposée par Heikkila dans [76]. La projection normalisée (x_n, y_n) du point $\mathbf{X}$ est :

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} X_c/Z_c \\ Y_c/Z_c \end{bmatrix} \quad (2.3)$$

Cette correction de la distorsion prend la forme suivante :

$$\begin{bmatrix} x_d \\ y_d \end{bmatrix} = \left(1 + k_1 r_n^2 + k_2 r_n^4 + k_5 r_n^6\right) \begin{bmatrix} x_n \\ y_n \end{bmatrix} + \begin{bmatrix} d_x \\ d_y \end{bmatrix} \quad (2.4)$$

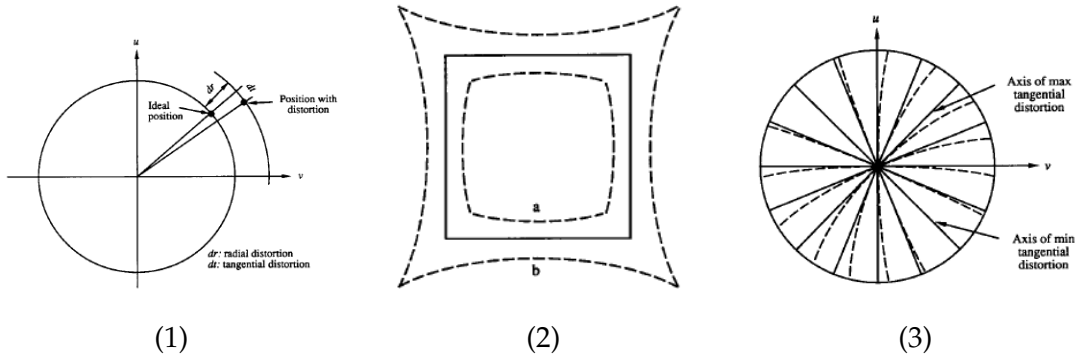


FIGURE 2.12 – Illustration de la paramétrisation de la distorsion radiale et tangentielle en (1) et de leurs effets respectifs sur l'image en (2) pour la distorsion radiale ((2)(a) pour la déformation négative, (2)(b) pour la distorsion radiale positive) et l'image (3) pour la distorsion tangentielle. Source Weng dans [164]

Où le vecteur de distorsion radiale vaut :

$$\begin{bmatrix} d_x \\ d_y \end{bmatrix} = \begin{bmatrix} 2k_3 x_n y_n + k_4 (3x_n^2 + y_n^2) \\ k_3 (x_n^2 + 3y_n^2) + 2k_4 x_n y_n \end{bmatrix} \quad (2.5)$$

Finalement, en multipliant ces nouvelles coordonnées normalisées avec la matrice de la caméra on obtient les coordonnées du pixel intégrant la distorsion.

$$\begin{bmatrix} x_p \\ y_p \end{bmatrix} = \begin{bmatrix} f/dx & 0 & c_x \\ 0 & f/dy & c_y \end{bmatrix} \begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix} \quad (2.6)$$

2.5.4 Mise en œuvre

Pour la réalisation des montages expérimentaux, la méthode planaire a été utilisée avec le toolbox de calibrage de caméra pour Matlab [26]. Cette démarche nous a permis d'obtenir les paramètres internes et externes de chaque caméra.

On obtient alors la scène entièrement calibrée telle que montrée figure 2.13. Pour permettre la reconstruction du volume à l'aide du visual hull maintenant que les paramètres de calibrage sont connus, il est nécessaire de pouvoir isoler la personne à l'aide de la segmentation avant/arrière-plan.

Cette opération consiste à définir si un pixel d'une image appartient à l'avant-plan (c'est à dire la personne) ou à l'arrière-plan. Dans notre étude, l'objectif étant de se focaliser sur l'occultation, les paramètres d'éclairage et de fond ont été ajustés pour être

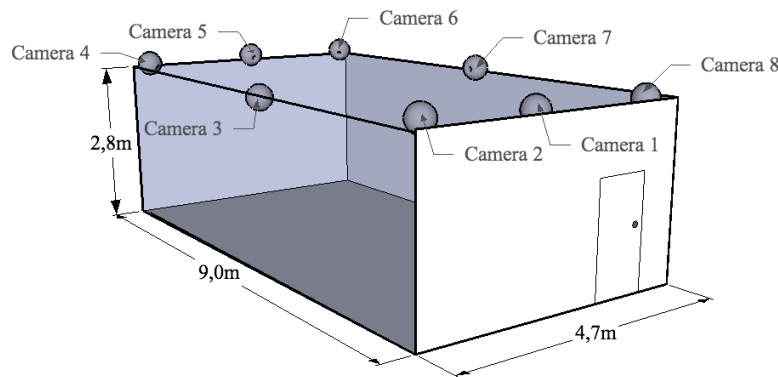


FIGURE 2.13 – Illustration du positionnement des caméras du montage expérimental multi-caméras pour la détection de la chute. Source [20].

stables. Donc une segmentation utilisant une méthode par soustraction de fond a été utilisée pour isoler la silhouette de la personne. Le modèle de fond utilisé est une image médiane calculée au début de chaque vidéo. Une opération de morphologie mathématique telle qu'une ouverture est alors appliquée pour affiner le résultat. Un exemple de segmentation est présenté sur la figure 3.3.

2.6 Conclusions

Ce chapitre a permis de montrer l'intérêt de détecter la chute à domicile. Dans ce contexte, l'utilisation de caméra vidéo permet à la personne d'être en sécurité sans avoir à porter un quelconque appareillage et d'apporter la possibilité au personnel médical de valider la chute en cas d'alarme pour éviter le départ d'une équipe en cas de fausse alarme. Dans ce cadre-ci, la problématique de la résistance à l'occultation est cruciale et nécessite d'être investiguée. Ceci est l'objectif que nous nous proposons de résoudre, avec les méthodes choisies pour les raisons explicitées tout au long de cette partie, au travers de l'article présenté dans le chapitre suivant.

Chapitre 3

Détection de la chute avec un système multi-caméras résistant aux occultations

Sommaire

3.1	abstract	54
3.2	Introduction	54
3.3	Method	57
3.3.1	Camera level	57
3.3.2	Data fusion level	61
3.3.3	Fall detection level	62
3.4	Materials and experiments	63
3.4.1	Experimental set-up	64
3.4.2	Fall scenarios	64
3.4.3	Data analysis	65
3.4.4	Statistical analysis	67
3.5	Results	68
3.5.1	VVDR behavior	68
3.5.2	Real-time implementation	69
3.6	Discussion	69
3.7	Conclusions	74

Article

Fall Detection With Multiple Cameras : An Occlusion-Resistant Method Based on 3-D Silhouette Vertical Distribution

Auteurs

Edouard Auvinet, Franck Multon, Alain Saint Arnaud, Jacqueline Rousseau, Jean Meunier

3.1 abstract

According to the demographic evolution in industrialized countries, more and more elderly people will experience falls at home and will require emergency services. The main problem comes from fall-prone elderly living alone at home. To resolve this lack of safety, we propose a new method to detect falls at home, based on a multiple cameras network for reconstructing the 3D shape of people. Fall events are detected by analyzing the volume distribution along the vertical axis, and an alarm is triggered when the major part of this distribution is abnormally near the floor during a predefined period of time, which implies that a person has fallen on the floor. This method was validated with videos of a healthy subject who performed 24 realistic scenarios showing 22 fall events and 24 cofounding events (11 crouching position, 9 sitting position, 4 lying on a sofa position) under several camera configurations and achieved 99.7% sensitivity and specificity or better with 4 cameras or more. A real-time implementation using GPU reached 10 frames per second (fps) with 8 cameras and 16 fps with 3 cameras.

3.2 Introduction

When approaching 65 years old, the risk of falling is rising. Indeed 30 percent of people over 65 and living in the community fall each year, and a fifth of fall incidents require medical attention [51]. Hence, falling is the most common cause of injury for elderly people [64]. It was the first cause of death by injury for elderly in 1997 and 1998 [80, 120]. Although most of the falls result in light injuries, 5% to 10% of falls in community-dwelling lead to serious injuries such as fractures, head injuries, or serious lacerations [60, 92]. An example of such injuries is hip fracture. Moreover 25% to 75% of "fallers" do not recover their pre-fracture level of movement and autonomy [105]. Besides, fear of falling appears and/or increases after falling which could increase the risk factor for future falls and reduce the quality of life [60]. This fall problem becomes more important for elderly people living alone because they cannot always call emergency services. Hence, many recent works have tried to develop easy-to-use and automatic techniques to detect falls in elderly people's houses [28, 91, 86, 141, 162, 14, 43, 112, 142]. The key question is : how to detect that a person has fallen in a house which contains many objects and where people can perform a wide range of activities ?

Indeed, one of the key problems is to recognize a fall among all the daily life activities. A description of the various phases of falling have been proposed in previous

studies [167],[126]. This classification provides us with physical features proper to fall movement which can be used to detect a fall in daily life. Falling is subdivided into four phases [126] : pre-fall (linked to daily life motions), critical (loss of balance), post-fall (final position after fall) and recovery (return to normal daily life) phases. The critical phase is extremely short (300-500 ms [166, 126]) composed of "free fall" and "impact with the floor" events. The former is associated with an increase of the body's velocity because of gravity. This velocity reaches abnormal maximal values for vertical and horizontal speeds compared to normal life activities [166] : typically 2-3 times higher values. At the "impact with the floor" event the speed decreases down to zero and a sudden inversion of acceleration polarity occurs. During the post-fall period the main features are an horizontal orientation of the body, a proximity to the floor, and commonly, lack of movements.

According to this description several approaches have been proposed to detect falls. These approaches mainly focus on the critical and post-fall phases. Wearable devices composed of accelerometers or gyroscopes directly placed on subjects' body parts (mainly chest [28], waist[91], or wrists[91]) enable to capture the high velocities which occur during the critical phase and the horizontal orientation during the post-fall phase. However these methods are based on the assumption that the subject wears the system at any time (with a warning by the system otherwise), and therefore if it is uncomfortable it could bother the user. Additionally such systems require recharging the battery frequently which could be a serious limitation for real application.

On the opposite, video systems enable an operator to rapidly check if an alarm is linked to an actual fall or not. Therefore cameras placed in the subject's environment were used to detect falls by measuring the movement or orientation of the body. A first approach consists in detecting abnormal horizontal and vertical speeds [141] or body silhouette changes [86, 142] associated with the critical phase. Another method consists in using body orientation features such as width and height of a silhouette by comparing a standing and a lying person [162, 103, 112]. In this case, the detection would be mainly based on post-fall shape or orientation features. As these approaches generally use only one camera, they could fail to detect falls in case of occlusions. These occlusions frequently occur in real situations at home because a room contains furnitures and objects that could be placed between the subject and the camera (as shown in Fig.3.1), contrary to easier and simplified experimental settings in laboratories. Dealing with occlusions is thus a key issue for using video systems in real situations in order to avoid misdetection and false alarms.

Using multiple cameras could overcome this limitation by offering several different points of view of the subject. It then becomes possible to extract a 3D silhouette of the subject. Some approaches use homography (a transformation between projective planes) to project silhouette (previously segmented with a foreground/background algorithm) on the ground and parallel planes for gathering information from different cameras and locate the person in the place (e.g [50][43][6][15]). Another method, the visual hull [97] consists in back-projecting silhouettes into space using camera models. Contrary to homography, camera models permit to represent more sophisticated situations such as lens distortion. The intersection of all those projections results in the final volume[97]. This method has been applied to fall detection in [14] to detect if the body is vertical or not during the post-fall phase. To this end, the method computes the centroid of the volume and its main axis using Principal Component Analysis. The authors did not report any information about the robustness of the system to occlusions. However, since the silhouettes coming from all the cameras were needed, when an occlusion occurs for one camera or more the reconstructed volume may become unreliable or unusable. To overcome this problem, it is possible to use an occlusion-resistant visual hull method [146]. This approach is able to reconstruct a volume even if one of the silhouette is not present for one of the cameras (such as when the body is occluded).

In this paper, we propose a method which is capable of dealing with several occlusions that could occur in personal houses. This method is based on two main ideas. Firstly, we use the occlusion-resistant algorithm introduced above in [19] in order to detect if a person is lying on the ground even if some occlusions occur. Secondly, we introduce the original and simple idea of "Vertical Volume Distribution Ratio" (denoted VVDR in the remaining of this paper). This ratio is obtained by dividing the volume that is below a given height by the total volume. For people lying on the ground, this ratio is high compared to when they are standing up. We assume that this feature is less sensitive to noise than methods based on principal axis of the reconstructed volume. VVDR has been successfully tested in a few occlusion-less situations [1]. In the present paper, we tested how this framework is able to manage occlusions in 24 realistic scenarios showing 22 fall events and 24 confounding events (11 crouching position, 9 sitting position, 4 lying on a sofa position) under several camera configurations. This unique dataset is documented [21] and made available to the scientific community through a website [20]. We also analyze theoretically the robustness of the method to occlusion by identifying the worst occlusion case and testing it with experimental videos. Finally a real-time im-

plementation using GPU is demonstrated.

The paper is organized as follows. Section 3.3 describes the theoretical background and the implementation of the method proposed in this paper in order to detect falls in these scenarios thanks to a multi-cameras system. Section 3.4 describes the experimental set-up that was used to generate the 24 scenarios with occlusions to validate the method. Section 3.5 provides some results and discussion about the performance of the method. Finally, section 3.7 gives conclusion and perspectives to this work.

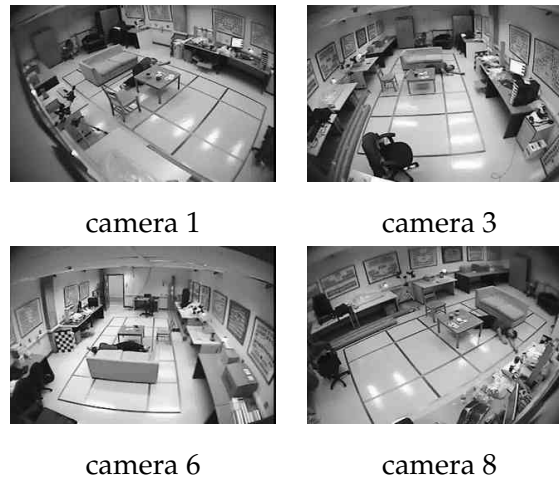


FIGURE 3.1 – Examples of real occlusions in our experimental setup.

3.3 Method

Briefly stated, the method involves essentially two main steps. First, with the set of cameras the 3D volume of the person is reconstructed with a shape-from-silhouette approach. Second an index (VVDR) is obtained by dividing the volume that is below 40 cm by the total volume of the person. Then a simple threshold is used to determine if this index reveals a fall or not. This section describes in more details these steps.

Our algorithm can be divided into 3 levels : camera and data fusion levels (step 1) and recognition level (step 2) as shown in Fig.3.2.

3.3.1 Camera level

In order to calculate the volume distribution of a subject in his environment the system must know the relationship between the camera coordinate system and the real 3D space. Thus, preliminary to the fall detection process, the cameras have to be calibrated.

TABLE 3.1 – Table of notation

$\mathbf{X} = (X, Y, Z)$	real world coordinates
$\mathbf{X}_c = (X_c, Y_c, Z_c)$	coordinates
$\mathbf{f} = (f_x, f_y)$	focal length (horizontal and vertical)
$\mathbf{c} = (c_x, c_y)$	optical center coordinates
$\mathbf{k} = (k_1, k_2, k_3, k_4, k_5)$	radial distortion parameters
$\mathbf{T}$	3D translation vector
$\mathbf{R}$	3D rotation matrix
(x_n, y_n)	normalized image projection
(d_x, d_y)	tangential distortion vector
r_n	radial distance
(x_d, y_d)	normalized image coordinates with radial distortion
(x_p, y_p)	pixel image coordinates
α	skew coefficient
i_j	image i of camera j
b_j	background model of camera j
s_j	binary image of the segmented foreground object for camera j
z_i	height of the horizontal plane i
$S_{i,j}$	projection of the image provided by camera j on the horizontal plane i
S_i	summation of the projection $S_{i,j}$ coming from n cameras
S_i^*	one slice of the 3D volume reconstructed
$VVD(i)$	Vertical Volume Distribution of the object at the i th slice
$VVDR$	Vertical Volume Distribution Ratio
Th	Segmentation threshold

Intrinsic parameters were computed using the chess board method[26] to define the focal distance $\mathbf{f} = (f_x, f_y)$, the optical center $\mathbf{c} = (c_x, c_y)$, the skew coefficient α and the radial distortion $\mathbf{k} = (k_1, k_2, k_3, k_4, k_5)$ as presented in [76]. The later parameters are

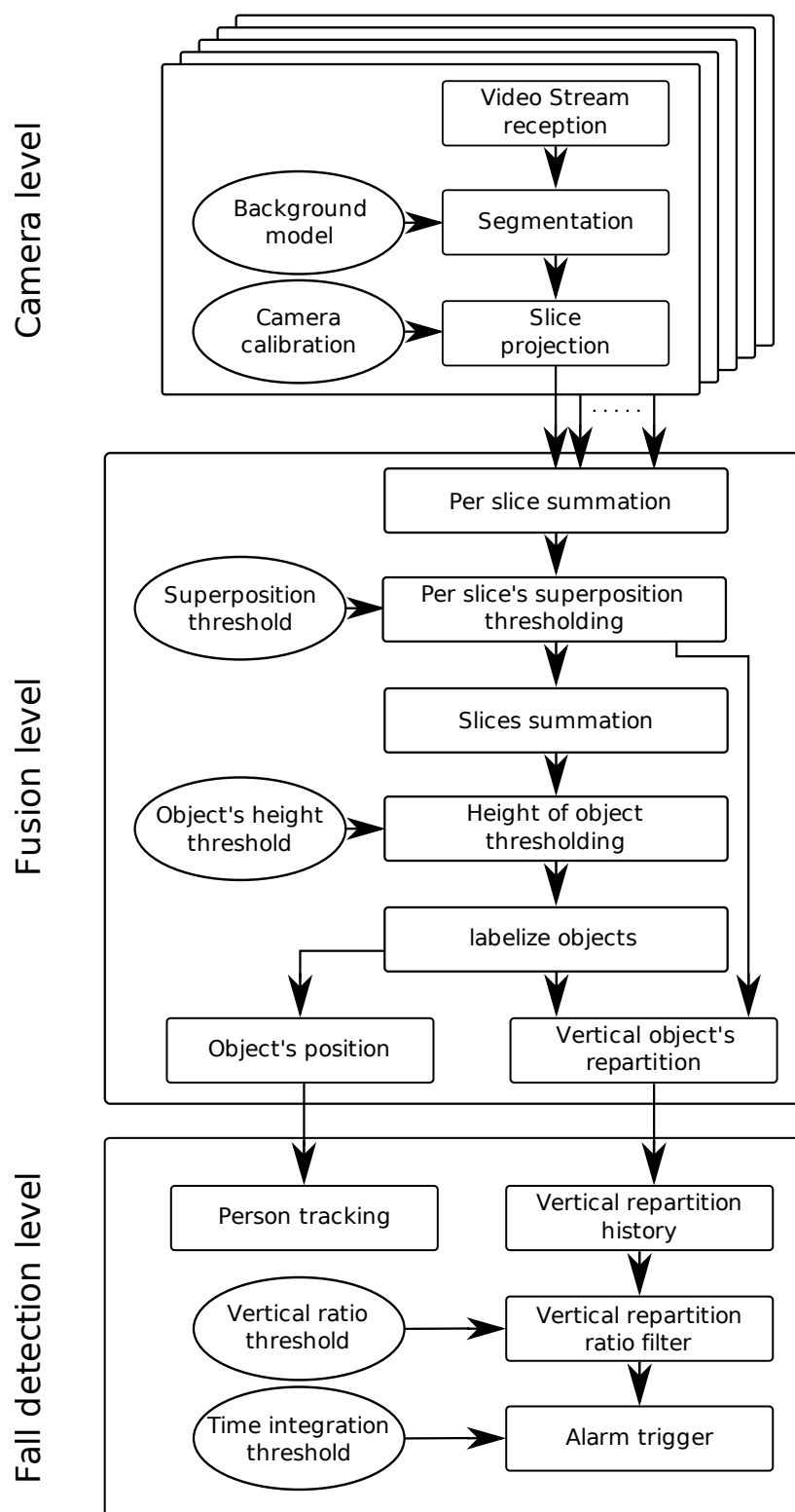


FIGURE 3.2 – Schematic representation of the entire process.

necessary because of non negligible radial distortion due to the large field of view of the camera lenses. External parameters, the rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{T}$ were calculated using feature points manually placed on the floor. Altogether, those parameters define the projective camera model described as follows. Let $\mathbf{X} = (X, Y, Z)$ be the real world vector of a 3D point and $\mathbf{X}_c = (X_c, Y_c, Z_c)$ his coordinates in the camera space then :

$$\mathbf{X}_c = \mathbf{R}\mathbf{X} + \mathbf{T}$$

The normalized image projection (x_n, y_n) is defined by :

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} X_c/Z_c \\ Y_c/Z_c \end{bmatrix}$$

The normalized point coordinates (x_d, y_d) with radial distortion become :

$$\begin{bmatrix} x_d \\ y_d \end{bmatrix} = (1 + k_1 r_n^2 + k_2 r_n^4 + k_5 r_n^6) \begin{bmatrix} x_n \\ y_n \end{bmatrix} + \begin{bmatrix} d_x \\ d_y \end{bmatrix}$$

Where the tangential distortion vector (d_x, d_y) is :

$$\begin{bmatrix} d_x \\ d_y \end{bmatrix} = \begin{bmatrix} 2k_3 x_n y_n + k_4 (3x_n^2 + y_n^2) \\ k_3 (x_n^2 + 3y_n^2) + 2k_4 x_n y_n \end{bmatrix}$$

and radial distance is : $r_n = \sqrt{x_n^2 + y_n^2}$

Finally, multiplying the normalized coordinates with the camera matrix gives pixel coordinates (x_p, y_p)

$$\begin{bmatrix} x_p \\ y_p \end{bmatrix} = \begin{bmatrix} f_x & \alpha \cdot f_x & c_x \\ 0 & f_y & c_y \end{bmatrix} \begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix}$$

where α is a skew coefficient. This function can be written as :

$$[x_p, y_p] = \phi(X, Y, Z, \mathbf{f}, \mathbf{c}, \mathbf{k}, \mathbf{R}, \mathbf{T}, \alpha)$$

In order to detect moving objects, each image of camera j , noted i_j , is subtracted from its own background model b_j obtained by computing a temporal median image of the sequence [131]. When the absolute difference of a pixel is higher than a previously defined threshold Th , it is registered as a foreground pixel, otherwise it is considered as a background pixel :

$$s_j(x_p, y_p) = \begin{cases} 1 & \text{if } |i_j(x_p, y_p) - b_j(x_p, y_p)| > Th \\ 0 & \text{otherwise} \end{cases}$$

Finally, in order to reduce noise detection and reinforce large surface detection, an opening morphological operation is done on s_j . An example of this segmentation is given in Fig.3.3.



FIGURE 3.3 – Result of the moving object segmentation process. From left to right, background median model b_j , current frame i_j , segmented picture s_j for camera j .

3.3.2 Data fusion level

This level aims at gathering projections of the 2D silhouette provided by each camera on horizontal slices in order to reconstruct the 3D volume of the subject. Let $S_{i,j}$ be the projection of the image provided by camera j on the horizontal plane i as follows :

$$S_{i,j}(X, Y) = s_j(\phi(X, Y, Z_i, \mathbf{f}_j, \mathbf{c}_j, \mathbf{k}_j, \mathbf{R}_j, \mathbf{T}_j, \alpha_j))$$

where Z_i is the height for the horizontal plane i , $\mathbf{f}_j, \mathbf{c}_j, \mathbf{k}_j, \mathbf{R}_j, \mathbf{T}_j, \alpha_j$ are the parameters for camera j .

For each horizontal slice i , S_i is the image corresponding to the summation of projection $S_{i,j}$ coming from n cameras :

$$S_i(X, Y) = \sum_{j=1}^n S_{i,j}(X, Y)$$

Where n is the total number of cameras. Therefore, $S_i(X, Y)$ takes values between 0 and n depending on the number of 2D silhouettes (from n cameras) contributing to the 3D reconstruction at position (X, Y) and at height Z_i . The distance between each slice was set arbitrarily to 10 cm in this study. Fig.3.4 illustrates an example of such kind of fusion.

Without occlusion, the person is visible from all cameras and consequently all positions (X, Y) where $S_i(X, Y) = n$ define the correct 3D reconstruction (slice by slice). To allow tolerance for one possible occlusion we simply add the positions where $S_i(X, Y) = n - 1$ at the expense of a slightly larger and coarser reconstruction. Therefore, by thresholding S_i at $n - 1$ we obtain the 3D reconstruction as a series of segmented slices S_i^* :

$$S_i^*(X, Y) = \begin{cases} 1 & \text{if } S_i(X, Y) \geq n - 1 \\ 0 & \text{otherwise} \end{cases}$$

As the threshold is applied individually to each position (X, Y) , we can also handle the case of multiples partial occlusions in different cameras if they are not affecting the same position (X, Y) . Notice that, reducing the threshold to accommodate more occlusions would result in unacceptable enlargement and innacuracy of the 3D reconstruction.

Let B be the set of pixels in each slice S belonging to the largest object. The Vertical Volume Distribution of this object at the i th slice denoted $VVD(i)$ is given by :

$$VVD(i) = \sum_{(X,Y) \in B} S_i^*(X, Y)$$

Examples of the resulting volume of a standing up (light gray) and lying down (dark gray) positions and their corresponding vertical volume distribution are presented in Fig.3.5 where the difference is clearly visible. Fig.3.6 represents the evolution of the vertical volume distribution (displayed with gray levels) of a subject obtained during a fall scenario.

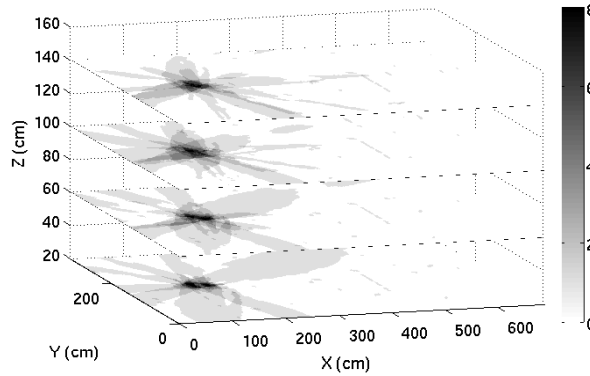


FIGURE 3.4 – Representation of 4 slices (S_i) where camera views were projected and summed (18 slices were used in practice with a 10 cm vertical interval)

3.3.3 Fall detection level

To detect a fall, an indicator based on the ratio between the sum of VVD values from the first 40 cm (5 slices starting from the floor) with respect to the whole volume ($m=18$ slices) is computed as follows :

$$VVD R = \frac{\sum_{i=1}^5 VVD(i)}{\sum_{i=1}^m VVD(i)} \quad (3.1)$$

This value 40 cm is justified by anthropometric data from [34]. In particular, for the

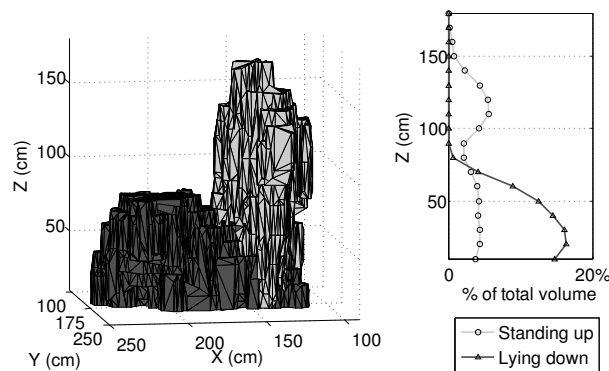


FIGURE 3.5 – 3D Reconstruction of a person after fusion of the different points of view and their corresponding vertical volume distribution on the right. Light gray color is attached to a standing up person, and dark gray for a lying on the ground person.

65 to 80 years old range, the shoulder width mean is 41.7 cm. This is the highest height to be expected for a lying down body.

A fall is detected if this ratio is above a preselected threshold during a predefined period of time (5 seconds in our case). This predefined period of time of 5 seconds is not a sensitive parameter and could be longer if needed. We chose 5 seconds because after that period, the subject stood up after a fall (we did not ask him to stay on the floor indefinitely) and the confounding events (e.g. crouching down) were lasting shorter periods of time. In practice this parameter should be chosen by the clinician considering the habits of the elderly person.

3.4 Materials and experiments

In order to evaluate the method proposed in this paper we have captured several videos containing a wide set of falls (see Fig.3.8). For each situation, we used several synchronized cameras. However it is impossible to capture real-life situations where people actually fall. This is why we have designed scenarios that were carried-out by an actor who performed the falls in our laboratory with appropriate protection (mattress). One has to notice that realism of the falling motion is not a key issue here as our approach focuses on the post-fall phase.

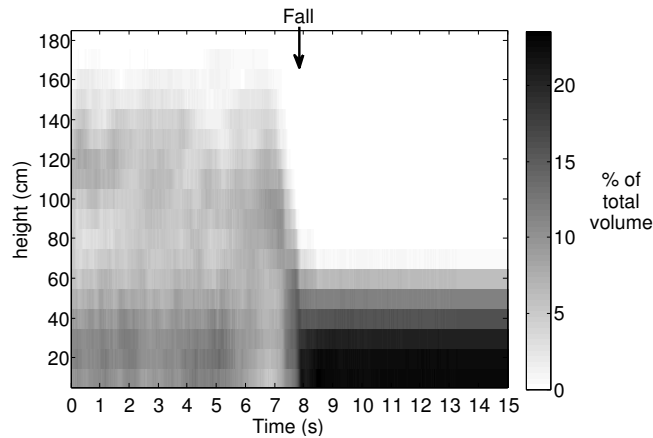


FIGURE 3.6 – Example of the vertical volume distribution during a fall scenario (displayed with gray levels).

3.4.1 Experimental set-up

The dimension of the area was 7m per 4m. A table, a chair and a sofa were introduced in the capture area in order to reproduce a normal room where people actually live. Adding such furnitures introduces occlusions in the videos for most of the scenarios. We assumed that a commercial system based on our technique would be made-up of IP video surveillance cameras with large field of view lenses. For all the scenarios we thus placed 8 such cameras (Gadspot 4600, 110 degree field of view) all around the area. They were attached to the ceiling at 2.5 meters height. Video streams (720 x 480 at 30 fps) were recorded and analyzed on a common desktop PC.

3.4.2 Fall scenarios

We decided to propose a wide range of realistic fall scenarios according to many previously published works (e.g. [54]). Each scenario is defined by a set of characteristics, such as the main falling direction (falling down, forward, backward, sideways) and the departure position (stand up, sit on a chair or a sofa). Each scenario is depicted in Fig. 3.8. Some situations which could lead to false alarms, such as occlusions due to furniture (see Fig. 1), crouching down on the floor and lying on a sofa (see Fig. 3.9) are also present to complexify the scenarios. Overall there were 24 realistic scenarios showing 22 fall events and 24 confounding events (11 crouching position, 9 sitting position, 4 lying on a sofa position) under several camera configurations. These scenarios captured with 8 cameras correspond to a total of 143472 frames (4782.4 seconds) to be analyzed by the system.

This unique dataset is documented in [21] and made available to the scientific community through a website [20].

Each scenario was performed once by one subject and approved by the local Institutional Review Board (IRB) authority. The subject in the videos is one of the author (A. S-A), a clinician whose research interests are elderly people affected by musculoskeletal and cognitive disorders living in the community. He is well aware of the different features of real falls in elderly people and took care of performing the simulated falls accordingly (e.g. slow motion, falls due to different disorders (loss of balance, blood pressure drop), abrupt sitting due to weakening of the ham-string muscles in elderly people etc.).

All the cameras were used to capture the fall. However, it was possible to test various camera configurations by using or not some of the video sequences during the analysis process. Hence, we tested configurations using 3 to 8 cameras. It enabled us to evaluate how our method was influenced by the number of the cameras used for the capture. For each scenario we tested 219 configurations : all the possible combinations when selecting 3 to 8 cameras among 8 cameras :

$$C_8^3 + C_8^4 + C_8^5 + C_8^6 + C_8^7 + C_8^8 = 219$$

Some of these scenarios involved occlusions due to furnitures placed in the environment.

In order to test further the ability of the system to tackle the problem of occlusions, we also introduced 2 artificial occlusions. The first one consists in completely deleting the contribution of one camera, corresponding to a full occlusion of this camera. The second one deletes only the contribution of one camera for the lowest 40 cm of the 3D volume. This correspond to the worst possible case because the volume of the lowest part becomes underestimated and consequently this reduces the VVDR value. This can be explained by the illustrative example in Fig.3.7. For a 4-camera setup without occlusion, the segmented slice S_i^* is larger (1-2-3-4) than with one occlusion (1-2-3,1-2-4,1-3-4,2-3-4). Therefore an occlusion may contribute to a higher rate of false negatives (failing to detect a real fall).

3.4.3 Data analysis

In this paper, we wish to evaluate the ability of the Vertical Volume Distribution Ratio to discriminate lying-on-the-floor position (corresponding to a fall) from others. To this end, we computed the VVDR for all the images coming from the sequences. As shown in Fig.3.10, VVDR for lying down positions is clearly different than others, such as standing

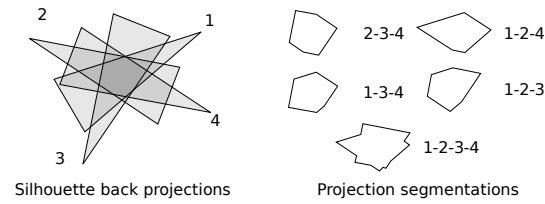


FIGURE 3.7 – Illustrative example with 4 cameras. For a given slice, the segmented surface S_i^* (and the reconstructed volume) is underestimated in case of a camera occlusion.

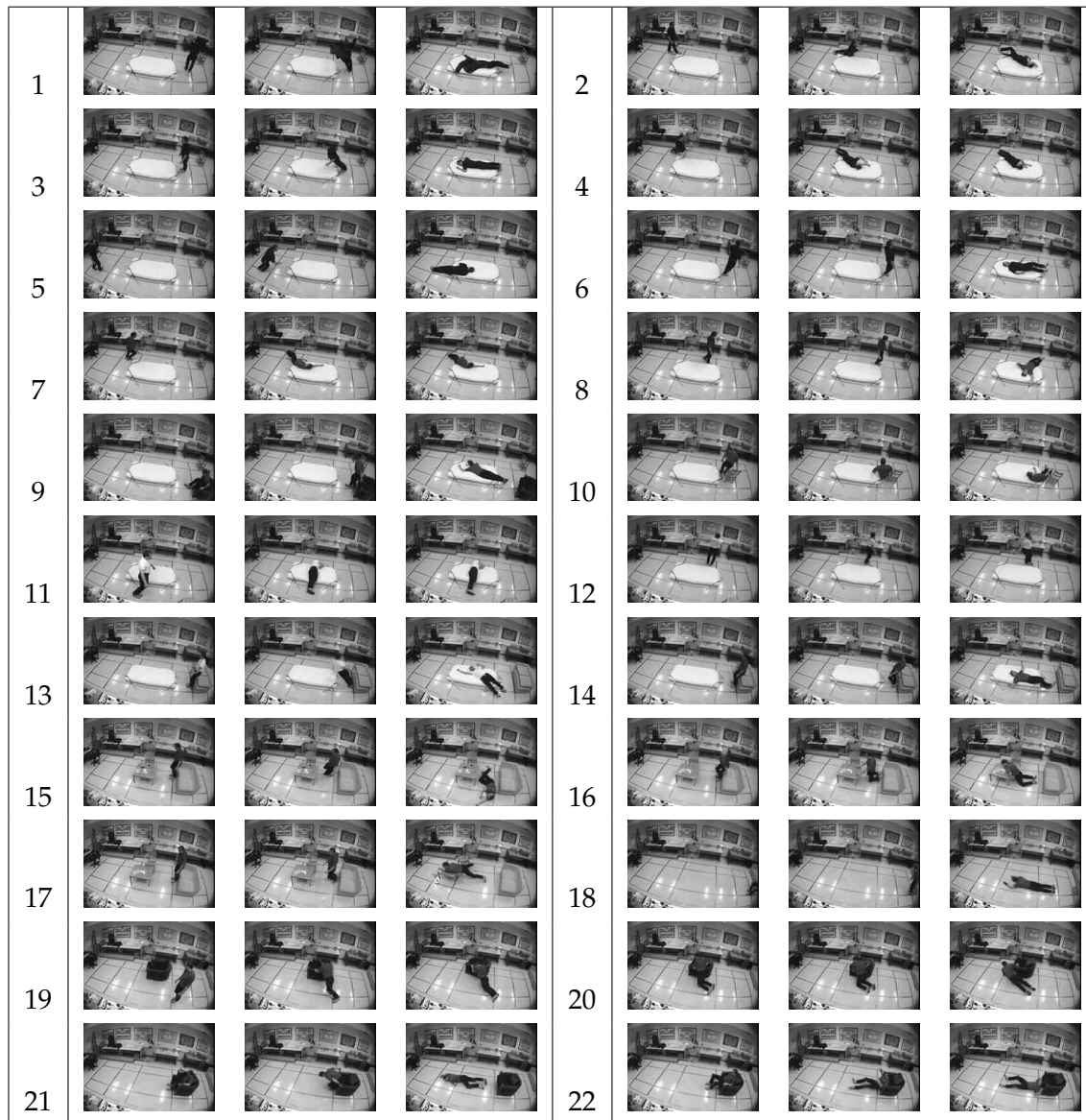


FIGURE 3.8 – Description of each scenarios

up, sitting down or crouched positions. This statement is true whatever the number of cameras and even with only 3 cameras where the separation remains acceptable.

The actual time where a fall occurs (denoted t_{fall}) is manually measured in the video sequences. This time is defined as the beginning of the post-fall period when the body hits the ground. If our method is able to detect a fall event after t_{fall} , the detection is supposed to be correct (True Positive). If the method doesn't detect any fall, it is supposed to have failed (False Negative). If it detects a fall event before t_{fall} this time interval it is supposed to have generated a false detection (False Positive). If no fall is detected before t_{fall} , it is then considered as True Negative.

3.4.4 Statistical analysis

The VVDR threshold to set a fall detection was simply taken as the 97.5% percentile of the no-fall region in Fig.3.10. This means that with VVDR alone, 2.5% of false positives (FP) will occur, but we will get most, if not all, the true positives (TP). This bias toward TP is reasonable since we prefer a few more FP to avoid some miss-detections of fall (risk minimization). Moreover, the predefined period of inactivity (5 seconds) after a potential fall will remove several others FPs.

We have tested this threshold with an unbiased leave-one-out strategy to compute the sensitivity and specificity of the complete system (including the period of inactivity of 5 seconds) in table 3.2. This means that for each scenario tested, we have computed the VVDR threshold corresponding to the 97.5% percentile of the no-fall region obtained from all the remaining (training) scenarios.

To analyze our recognition results, we compute the sensitivity and the specificity, as follows :

- Sensitivity : $Se = \frac{TP}{(TP+FN)}$
- Specificity : $Sp = \frac{TN}{(TN+FP)}$

Where

- True Positives (TP) : number of falls correctly detected (Among the 22 fall events multiplied by the total number of camera configurations)
- False Negatives (FN) : number of falls not detected.
- False Positives (FP) : number of normal activities detected as a fall (Among the 24 normal segments in each scenario multiplied by the total number of camera configurations)
- True Negatives (TN) : number of normal activities not detected as a fall.



FIGURE 3.9 – Examples of confounding events, from left to right, crouched down, lying on a sofa and sitted position.

3.5 Results

This section presents results obtained from experimentation with the dataset previously presented. In a first part, the ability of the VVDR to detect a fall is examined . Then the real-time constrain is tested with respect to the number of camera.

3.5.1 VVDR behavior

Results shown in Fig.3.10 prove that VVDR enables to discriminate lying-on-the-floor from others positions. Indeed, the 95% confidence intervals (gray areas) around the mean value of VVDR for lying-on-the-ground and others positions are very well separated with 4 cameras or more. The separation remains acceptable for the 3-camera setup, although some overlap appears between the confidence intervals.

Moreover, the distance between the two confidence intervals increases with the number of cameras which tends to show that the ability to detect lying positions increases with the number of cameras.

With four cameras or more the system achieved almost 100% sensitivity and specificity, as presented in the first part of Tab.3.2. The less favorable results were obtained with three cameras for which the sensitivity decreased down to 80.6%. Whatever the scenario was, simulating a partial occlusion of the lowest 40 cm above the ground (worst occlusions) in one camera led to an artificial decrease of VVDR (see Fig.3.11) resulting in a lower detection rate (55% sensitivity with 3 cameras), but also in the same way, a lower false detection rate (100% specificity) as shown in the second part of Tab.3.2. On the contrary, simulating total occlusion of one camera increased the VVDR (Fig.3.11) resulting in a higher detection rate (94.7% sensitivity with 3 cameras) at the expense of a higher false positive rate (95.6% specificity for 6 cameras and more) as shown in Tab.3.2.

The fact that the inflexion point is located at 40cm, when all contribution of one camera for the lower part of the body are occluded clearly, demonstrates that this is the worst occlusion case as explained by Fig.3.7. In this case, the numerator of Eq.3.1 is the

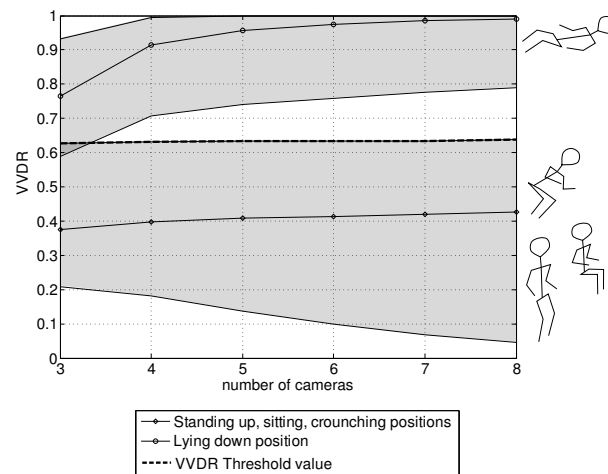


FIGURE 3.10 – Influence of the number of cameras on the capability of the VVDR to discriminate body postures obtained without artificial occlusions. The gray areas correspond to 95% confidence intervals and the solid lines are the medians.

most underestimated.

3.5.2 Real-time implementation

Three different implementations of this method have been tested. The first one used only the processor to deal with all the computations. The second one used the GPU for reconstructing the voxels while the remaining of the computations was performed by the processor. The last one used the GPU for reconstructing the voxels and segmenting the image. Computation times for these three methods are reported in Tab. 3.3. The main result is that the algorithm which used the GPU can go 18 times faster than the one with only the processor, for 3 cameras. This ratio increased up to 28 when using 8 cameras.

3.6 Discussion

Our results compare very favorably with those reported in the literature. For instance, Rougier et al. in [142] have developed a fall detection system with a single camera based on silhouette deformation of the subject with the same dataset used here. Their results gave a sensitivity and specificity of 95.5% and 96.4% respectively, that are lower than those obtained with our method, although this comparison is somewhat unfair because they used only one camera. Anderson et al. [14] used fuzzy logic with a multi-camera setup on a dataset containing 14 falls and 32 no-falls events. They obtained 100% true

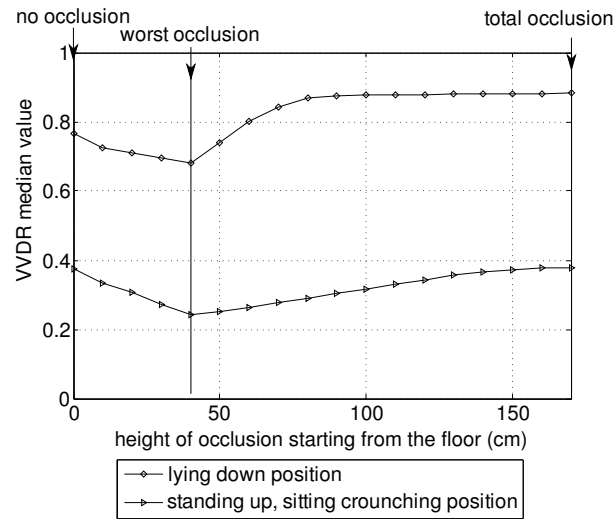


FIGURE 3.11 – VVDR for a standing up and lying down situations where one camera is gradually occluded (from the ground to the head of the subject) in a 3-cameras setup with 95% confidence interval in light gray.

positive detection and 6,25% false detection (sensitivity=100% and specificity=93.75%). However, they did not address the problem of occlusions and their approach requires the manual adjustment of several parameters. Cucchiara et al. [43] proposed a posture classification system that was able to achieve 97.23% accuracy with some occlusions and for four types of postures including lay down position. Notice that [14, 43] used datasets with different (unspecified number of) subjects while we used one (experimented) subject.

The method presented in this paper is able to deal with an occlusion of one camera without significantly decreasing the detection rate. In real life, situations where several cameras are occluded generally occur. However, the reconstruction algorithm is applied independently for each 3D position. Hence, the algorithm is able to deal with several occlusions, except if there is more than one occlusion for the same 3D position. In some real situations such as the one depicted in Fig.3.1 all the cameras may be partially occluded. A classical reconstruction method [14] may fail in recognizing a lying person in that case, as shown in Fig.3.12 whereas the method presented in this paper is able to reconstruct the volume of the actor. The quality of the resulting reconstructed volume is sufficient to compute VVDR and thus determine if the actor is lying down or not.

Another important feature of the method is the fact that it works without considering the speed or motion of the person. Indeed, by simply looking for abnormal volume dis-

TABLE 3.2 – Sensibility and specificity obtained with VVDR threshold set at 97.5 percentile of the no-fall region. Mean +- standard deviation of leave-one-out

Number of camera	original video	
	sensitivity	specificity
3	0.806 (+- 0.021)	1.000 (+- 0.000)
4	0.997 (+- 0.001)	0.998 (+- 0.000)
5	0.999 (+- 0.000)	1.000 (+- 0.000)
≥ 6	1.000 (+- 0.000)	1.000 (+- 0.000)
	worst occlusion	
	sensitivity	specificity
3	0.550 (+-0.022)	1.000 (+- 0.000)
4	0.895 (+- 0.019)	1.000 (+- 0.000)
5	0.933 (+- 0.016)	1.000 (+- 0.000)
≥ 6	0.954 (+- 0.011)	1.000 (+- 0.000)
	total occlusion	
	sensitivity	specificity
3	0.947 (+-0.007)	0.995 (+- 0.001)
4	0.999 (+- 0.000)	0.990 (+- 0.002)
5	1.000 (+- 0.000)	0.984 (+- 0.003)
≥ 6	1.000 (+- 0.000)	0.956 (+- 0.008)

tribution along the vertical axis, i.e. when the major part of the body is near the floor, fall detection is made possible. This point is important because motion is generally difficult to measure and usually needs more computer resources and a high and fixed frame rate to be accurate ; these requirements add complexity and could impair real-time processing.

Regarding the possibility that people just happen to have their bodies close to the floor for a long period of time (maybe to pick up something or to tie their shoelaces). This problem is usually tackled by the computation of the VVDR itself because a large part of the body remains above 40 cm, but in the unusual case where the elderly is very near the floor, the pre-defined limit of time (5 seconds in our case) is sufficient to avoid false positive. The limit could be increased to a higher value if necessary depending on the habits of the elderly person and is not a sensitive parameter. In this study, the 11

TABLE 3.3 – Influence on the implementation on computation time

Segmentation	CPU	CPU	GPU
Projection	CPU	GPU	GPU
Number of cameras	time (msec/frame)	time (msec/frame)	time (msec/frame)
3	1140	98	63
4	1516	100	72
5	1888	111	79
6	2258	122	88
7	2613	133	96
8	2980	145	105

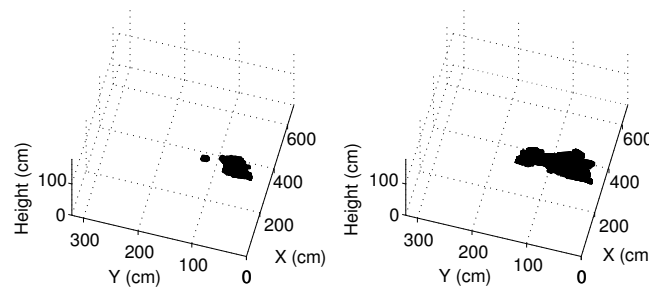


FIGURE 3.12 – Volume reconstructed in case of real occlusions with pictures shown in Fig3.1. Left : classical method. Right : occlusion-resistant method presented in this paper.

crouching down events were correctly identified as true negative (except for the 3 camera configuration). However, in the case where the subject finishes the fall onto an object (e.g. table, wall, or other furniture), a large part of the body could remain above 40 cm and thus, our system could fail. Adding some knowledge about the environment could help to detect these difficult cases. Similarly, a fall ending in a sitting posture (on the floor), could cause a miss detection.

Notice that another moving object entering/leaving the room, like a cat or dog would be ignored because the system analyses only the largest object (human) in the scene (see section Method). Very big dog are out of the question because of the additional risk of fall for an elderly person living alone.

To bring the system to a multi-room setup, a set of cameras needs to be installed in

every room. Fortunately, this does not require much more computer resources. Actually, the computer simply needs to know in what room the elderly person is and then process the data coming from this room only for fall detection. The presence of a person in a room can be easily and quickly monitored with simple background subtraction for all cameras (in all rooms) checked one after the other at a low frame rate per camera (e.g. 1 frame/sec). Only the cameras involved in the identified room would be processed at a higher frame rate for fall detection. More sophisticated alternatives are also possible and will be investigated in the future.

One shortcoming of a multi-room setup of our video system is the requirement for installation of adequate infrastructures that may cause a significant modification of the subject's home environment. Although this means certainly much work (e.g. compared to wearable devices), with the current miniaturization of cameras and reliable Wi-Fi technology we believe that such system is nevertheless realizable for real application. The cost of such system could be another problem for the user or provider ; however we think that the economic advantages will be noticeable when compared with traditional intervention i.e. placing the elderly in a specialized establishment (instead of the home setting). Another concern about video systems could be the intimacy and privacy of the user. For this reason, this system should use a closed circuit : the system will be activated to send an alarm signal toward an outside resource (e.g. via a phone or Internet) if and only if a fall is detected, then the images for that event could be accessed (with a password) by the designated persons (e.g. the main caregiver or an emergency call centre). Moreover, in some areas (e.g. bathroom) the images could be processed (blurring, pixelization or silhouette extraction) to ensure some privacy. Finally it is worth mentioning that in a recent study on the perception of intelligent videomonitoring system by elderly people [104], 96% of participants were favorable or partially favorable to such system for fall detection at home.

Finally, the quality of the images was rather poor here due to large field of view lenses and compression artifacts of low-cost cameras, resulting in noise on segmented pictures. Such noise may lead to errors in the silhouette of moving objects. However missing part of the silhouette could be considered as partial occlusions that our method is able to overcome. Imperfect segmentation are thus partially compensated by the method but improvement in the segmentation algorithm would certainly be desirable in the future to compensate for the limitations associated with low-cost cameras. Today's higher end cameras will certainly become more affordable in the future and could also contribute to

better performance of the system.

3.7 Conclusions

The results presented in this paper had shown that a multi-camera system is reliable in order to detect falls even if some occlusions occur. This result is valid even with only three cameras, but four or more cameras will offer better performances.

This research has led to five contributions : (1) VVDR, a simple and robust feature for fall detection, (2) application of occlusion-resistant volumetric reconstruction to fall detection, (3) creation of a unique dataset that is now documented [21] and made available to the scientific community through a website [20] (4) theoretical analysis of the robustness of the method to occlusion by identifying the worst occlusion case and testing it with experimental videos (5) real-time implementation with GPU.

One of the major contributions of this paper is the design of a simple index, the VVDR, which focuses on the change of shape vertical distribution of the subject (from standing up to lying down on the ground). VVDR is robust to some inaccuracies that could occur for a few images (because of too multiple occlusions or segmentation errors). It also means that using a lower frame rate could be considered without affecting the performance of the system since only the shape distribution at each frame is considered. Hence, we could imagine that a unique system could be used to monitor several rooms at a low frame rate for each camera. In this way, an entire home for autonomous people or multiple resident in a community dwelling could be monitored thanks to a unique system composed of a network of cameras and only one computer. In this paper, we have also shown that this type of detection process could be real-time if necessary by simply using a GPU.

The reconstruction method proposed in this paper could be applied to other types of applications, such as quantifying daily life activities which is a key issue of our modern society. As for detecting falls, real situations involve many occlusions and classical methods based on multiple cameras (e.g [14]) generally fail in solving this problem. More generally, this method should be useful for applications involving spatial location and activity classification depending on shape of subjects. Contrary to approaches mainly based on image analysis, dealing with 3D volumes in space brings richer information that should be useful to address complex monitoring processes.

Acknowledgment

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and MITACS.

Chapitre

4

Discussion sur la chute

Sommaire

4.1	Commentaires sur la résistance à l'occultation	79
4.2	Propositions d'amélioration	79
4.2.1	Amélioration de l'étape de calibrage	80
4.2.2	Amélioration de l'algorithme de segmentation	80
4.2.3	Utilisation de caméras de profondeur	82
4.3	Conclusion	82

Le travail précédent a permis de présenter la méthode de détection de la chute par un système multi-caméras. Différents points particuliers méritent quelques commentaires. Ceux-ci sont à propos de la résistance à l’occultation, et des différentes possibilités d’améliorations telles que la méthode de segmentation et de calibrage, et l’utilisation de carte de profondeur.

4.1 Commentaires sur la résistance à l’occultation

Il est intéressant de noter le fait que cette résistance à l’occultation se réalise au niveau du voxel. En effet, le voxel est défini comme occupé lorsque celui-ci est visible par $n - 1$ caméras.

Ainsi, l’absence de contribution d’un pixel peut effectivement provenir de l’occultation de ce pixel, car non visible par la caméra, mais aussi par un défaut de segmentation. Il en résulte donc que cette méthode peut aussi se montrer résistante au bruit de segmentation.

Pour la démonstration de la résistance à l’occultation dans cet article, seulement l’occultation d’une seule caméra été étudiée. Il est tout à fait possible d’imaginer une version où $n - 2$ contributions de caméras soient nécessaires pour définir le voxel comme occupé. Ceci pourrait rendre la méthode plus résistante au détriment d’une reconstruction plus grossière du volume et potentiellement du taux de fausses alarmes qui pourrait alors augmenter.

Il pourrait être intéressant d’utiliser le formalisme probabiliste développé par Franco dans [57], où chaque occupation de voxel est estimée à partir des probabilités que le projeté de ce voxel dans l’image de chacune des caméras appartienne au sujet. Cette méthode de reconstruction pourrait, par son formalisme probabiliste, être tout aussi, voir plus, résistante à l’occultation, mais sans avoir à définir le seuil du nombre de contributions de caméras nécessaires pour rendre un voxel occupé.

4.2 Propositions d’amélioration

Cette méthode utilise deux étapes fondamentales qui sont le calibrage de la scène et la segmentation avant/arrière-plan pour effectuer la reconstruction de la surface du sujet. Différentes pistes d’améliorations sont possibles pour la détection de la chute avec cette méthode. Premièrement, les deux étapes cruciales du calibrage et de la segmentation peuvent être améliorées avec d’autres méthodes plus complexes. D’autre part, l’utilisa-

tion de caméras de profondeur peut être considérée comme une solution de mesure très pertinente. Ces différents points vont maintenant être traités plus en détail.

4.2.1 Amélioration de l'étape de calibrage

Dans le travail présenté, le calibrage des caméras a été réalisé à l'aide de points de référence. Cette étape de préparation peut être remplacée par une méthode de calibrage dynamique qui essaye de faire correspondre en un volume complet et régulier les différentes silhouettes présentes sur chaque caméra. Cette méthode proposée par Boyer dans [29] utilise les silhouettes de la personne pour faire correspondre les sommets des silhouettes et minimiser les parties de chaque silhouette non recouverte par la reprojexion du volume reconstruit avec les paramètres de calibrage. Les paramètres de calibrage sont alors optimisés pour minimiser cette différence.

4.2.2 Amélioration de l'algorithme de segmentation

Dans notre expérimentation nous avons utilisé un modèle simple de segmentation de fond. Les paramètres d'éclairage étant contrôlés, et le fond totalement fixe, il n'apparaissait pas nécessaire d'utiliser une méthode plus compliquée. Il est évident qu'en situation réelle, ce type de segmentation d'arrière-plan serait inutilisable, et cela pour plusieurs raisons. La première raison provient de la gestion des ombres. Celles-ci projetées au sol vont augmenter artificiellement le volume de la personne aux niveaux proches du sol, ce qui pourrait provoquer des fausses alarmes plus facilement. La seconde, toujours associée à la lumière, provient de la variation de la lumière tout au long de la journée, modifiant ainsi la couleur de l'image de fond. Dans ce cas, il est nécessaire d'utiliser une mise à jour du modèle de fond qui va intégrer au fur et à mesure les variations de la lumière, tout comme les modifications de la scène du fait des déplacements d'objets.

Pour cela, il existe de nombreuses méthodes de soustraction de fond classique comme présenté par Piccardi dans [131] ou par Beneth et al. dans [23]. La plus généralement utilisée est la mixture de gaussienne, où pour chaque pixel, les couleurs prises par le fond sont modélisées par plusieurs gaussiennes. Il existe une méthode plus récente et similaire qui se base sur les "Codebook", publiée par Kim dans [93]. Elle modélise les mêmes couleurs le plus souvent prises par un pixel à l'aide d'une liste de code de couleur. Ces codes de couleurs rassemblent des propriétés de couleur comme une couleur moyenne, la luminosité maximale et minimale et des propriétés temporelles, par exemple la fréquence d'occurrence.

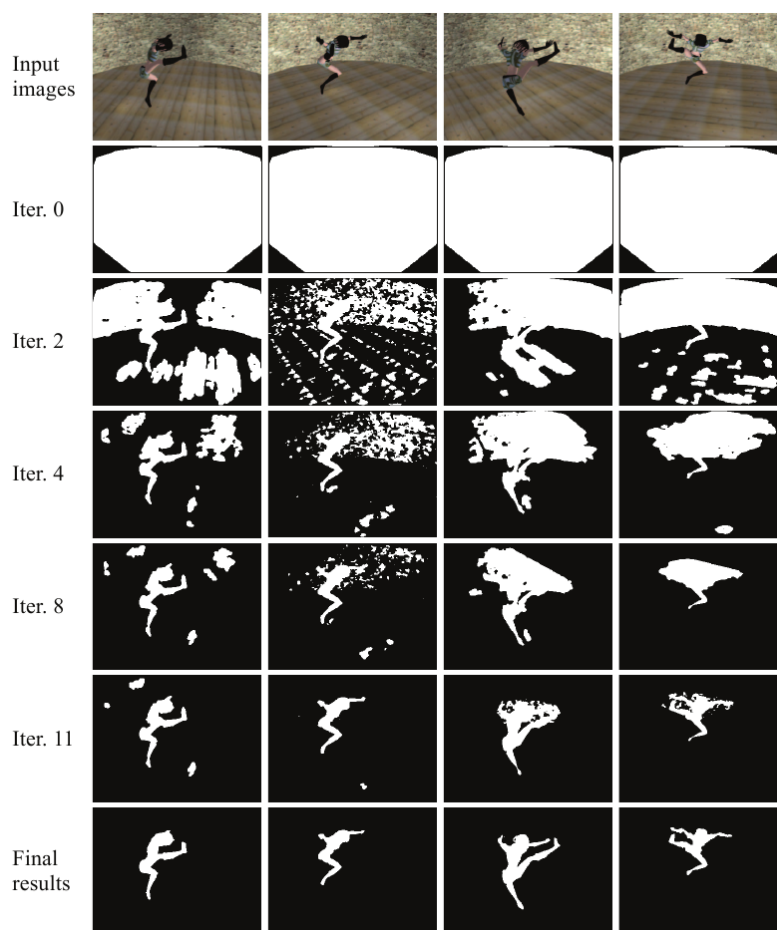


FIGURE 4.1 – Illustration du processus de segmentation multivues itératif. Chaque colonne représente une caméra. La première ligne montre l’image en couleur de la caméra puis chaque ligne une itération de la segmentation. Source [100].

Pour l’instant seules des méthodes de segmentation prenant en compte une seule caméra ont été explicitées. Comme l’ont proposé Wee et al. dans [100], l’étape de segmentation et de reconstruction du volume peuvent être réalisées de manières itératives pour fusionner l’information de chaque caméra dès l’étape de segmentation. Cette méthode reprojète le volume reconstruit dans l’image de chaque caméra pour en améliorer la segmentation avec cette information supplémentaire. En réalisant cela plusieurs fois, la reconstruction s’en retrouve améliorée comme le montre la figure 4.1

Le problème de la segmentation d’une image issue d’une caméra classique provient du fait que seule l’information de couleur au cours du temps permet de définir un pixel comme appartenant au premier plan ou à l’arrière-plan. Ce type d’information ressemble

à un problème mal posé. La classification à partir de la couleur dépend du fait que le fond et l'objet au premier plan vont avoir une couleur différente, ce qui est une condition qui n'a pas de rapport avec la nature de la classification. Il peut être intéressant de changer le type d'information pour se tourner vers une information qui a un rapport plus étroit avec la nature de la classification, par exemple la profondeur.

4.2.3 Utilisation de caméras de profondeur

Lorsqu'un objet apparaît dans une scène observée par une caméra de profondeur, il va forcément modifier la profondeur mesurée. En effet, si l'objet est derrière un autre, la profondeur apparente restera la même que la profondeur du fond. Dans le cas contraire, la profondeur sera forcément inférieure si l'objet est entre le fond et la caméra. La nature de l'information utilisée pour la classification est maintenant directement reliée au phénomène de classification.

Il existe maintenant des caméras de profondeur performantes et bon marché qui n'étaient pas disponibles lors de la réalisation de ce travail. Comme il le sera précisé plus en détail dans la section 5.5, ces caméras utilisent une information qu'elles projettent dans la scène qu'elles mesurent à l'aide d'une caméra permettant ainsi de calculer la distance entre la caméra et les objets visibles de la scène.

Différents travaux ont été réalisés dans ce sens, par exemple Rougier et al. dans [140] et Mastorakis et al. dans [109] qui après avoir effectué une segmentation de la personne basée sur ce principe, proposent de mesurer la hauteur du centroïde et la vitesse de la silhouette pour détecter la chute et ceci à l'aide d'une caméra de profondeur. On peut noter qu'un autre intérêt apparaît, le respect de la vie privée. Puisque l'information de profondeur permet de reconnaître la silhouette et la position de la personne tout en respectant son intimité.

4.3 Conclusion

Comme il a été présenté, cette méthode basée sur l'analyse de la surface issue d'une reconstruction multi-caméras est robuste aux occultations pour détecter la chute d'une personne au sol. Cette méthode peut être effectivement améliorée en plusieurs points pour être utilisée sur le terrain. Cette contribution permet de participer à la réflexion sur la prise en charge de la chute à domicile. Ce problème qui devient maintenant un problème universel. En effet, la revue de littérature de Kalula dans [90] et l'organisation

mondiale pour la santé à travers son rapport sur ce sujet dans [129] montrent que ce problème est aussi soulevé dans les pays émergents.

Mais la principale amélioration reste par rapport à l'évènement de la chute en lui-même. Serait-il possible de prévenir cette chute en mesurant la prédisposition de la personne à la chute avant que la chute n'intervienne? Effectivement, différents travaux montrent l'implication de la perturbation de la démarche dans la prédisposition à la chute, qu'elle soit due à l'âge ou à la maladie comme précisée par Richardson dans [136]. Ceci montre l'intérêt qui résiderait dans la capacité de mesure des paramètres de la marche en clinique.

Deuxième partie

Étude de la marche

Chapitre 5

Analyse de la marche à partir d'une carte de profondeur

Sommaire

5.1	Introduction	89
5.2	Paramètres cliniques de la marche	91
5.3	La mesure quantitative de la marche	93
5.3.1	Mesure du mouvement monopoint	94
5.3.2	Mesure du mouvement multipoints	96
5.3.2.1	Mesure dynamique	97
5.3.2.2	Mesure cinématique	98
5.3.3	Bilan sur les méthodes de mesure	102
5.4	Mesure de la marche à l'aide de caméras	103
5.4.1	Système mono-caméra	104
5.4.2	Système multi-caméras	105
5.5	Système utilisant une caméra de profondeur	107
5.5.1	Système stéréoscopique	107
5.5.2	Structure temporelle	108
5.5.3	Structure spatiale	109
5.6	Utilisation des caméras de profondeur pour l'étude de la marche	112
5.7	Conclusions	114

5.1 Introduction

La marche est l'un des mouvements les plus utilisés au quotidien, et qui est acquis généralement dès la première année de notre vie. Malgré cette acquisition précoce, la marche est un mouvement extrêmement complexe comme l'attestent les très nombreux travaux scientifiques qui continuent à l'explorer encore de nos jours. En particulier, c'est un mouvement qui utilise très judicieusement un enchaînement de déséquilibres vers l'avant pour minimiser les efforts à produire pour se propulser. Comme cela est décrit dans la partie 1.1, ce mouvement est rendu possible grâce à différents systèmes qui fonctionnent étroitement ensemble. Le système musculo-squelettique et le système nerveux central et le système nerveux périphérique sont tous trois mis en oeuvre pour réaliser ce mouvement. Il est donc normal qu'une pathologie touchant l'un de ces systèmes ait une répercussion sur la marche réalisée comme l'attestent de nombreux travaux.

Dans le contexte de vieillissement de la population, les troubles liés à l'avance en âge doivent être diagnostiqués au plus tôt pour améliorer leur prise en charge, à la fois pour améliorer la qualité de vie, mais aussi pour que la société puisse faire face au surcoût que cela engendre. En particulier des pathologies ou des dégénérescences de l'un des systèmes doivent être facilement détectées lors de simples visites cliniques régulières. Cela peut se réaliser via une analyse quantifiée de la marche, mais celle-ci doit pouvoir s'effectuer sans avoir recours à des systèmes de mesure complexes et coûteux. Dans le milieu clinique, les approches quantitatives sont souvent remplacées par des échelles et des appréciations qualitatives. Des auteurs [31] ont montré que la démarche de 30 patients observée sur vidéo par 10 médecins conduisait à des diagnostics parfois peu concordants.

Cette expérience montre la nécessité de pouvoir effectuer une mesure quantitative des paramètres de la marche sans perturbation subjective due au niveau d'expérience du praticien. Des protocoles de mesures adaptés à l'exercice clinique ont été proposés par le passé. Ceux-ci se basent sur un appareillage léger qui peut être rapidement mis en oeuvre, tel que des accéléromètres lombaires ou encore des semelles dynamométriques. Malheureusement l'information retournée est mesurée de manière ponctuelle sur le corps, et donne donc un reflet global de la marche avec un point unique de mesure, si bien qu'une partie de l'information significative peut être ignorée. D'autres protocoles, plus lourds à mettre en oeuvre, permettent les mesures multipoints de la marche, mais restent réservés aux laboratoires de recherches. Ils consistent principalement en un tapis d'analyse de pressions exercées au niveau des pieds (GaitRite [158]) ou d'un système opto-

électronique mesurant le mouvement de mires réfléchissantes positionnées sur le corps (système Vicon par exemple [5]). Cependant, ces systèmes sont relativement coûteux et nécessitent des compétences techniques particulières qui engendrent généralement le recrutement d'un ingénieur spécialisé. Il est donc difficile de les utiliser à grande échelle dans une clinique courante et ces systèmes restent donc réservés à de grands centres de rééducation. Même si l'information disponible avec ces systèmes est plus riche qu'avec un unique accéléromètre, l'idée reste toutefois d'échantillonner l'information disponible de manière plus ou moins détaillée. Encore une fois, le choix de se focaliser sur certains points de mesure peut conduire à des pertes d'informations. Les systèmes opto-électroniques reposent en effet sur un modèle biomécanique pour choisir le placement des mires réfléchissantes. Ce modèle simplifie le corps humain en solides rigides connectés par des articulations parfaites. De plus, ces modèles ne tiennent généralement pas compte des 200 os du corps humain, mais définissent des segments corporels simplifiés (en particulier pour le tronc, les mains, les épaules et l'avant-bras).

En dehors de ces problèmes d'échantillonnage de points de mesure, il est important de rappeler l'objectif final de ces analyses quantifiées de la marche. Cet objectif est d'effectuer un diagnostic ou un suivi de la pathologie d'un patient. Il est donc important d'avoir des critères quantifiés, mais aussi en nombre limité, afin de comparer les patients entre eux, ou à eux-mêmes entre deux traitements. On retrouve ici la contradiction entre la multiplication de l'information mesurée et la concision de l'information nécessaire pour classer les patients. Les systèmes couramment utilisés suivent un processus complexe qui consiste à calculer les angles articulaires liés au modèle biomécanique étudié, puis à effectuer des régressions statistiques afin de réduire l'information à un critère simple d'asymétrie, de régularité... En effet, mener une analyse objective pour classer deux patients sur la base de dizaines de critères quantifiés (les angles, les vitesses, les positions, les efforts...) est quasiment impossible. Revenir à des paramètres simples à partir de toutes ces mesures permet d'exploiter plus facilement les informations pour le diagnostic. Cependant, chaque étape du processus introduit son lot d'imprécisions et de modèles, car le critère recherché n'est pas directement mesuré, mais estimé à partir de mesures elles-mêmes indirectes (les angles articulaires obtenus à partir de mesures de positions de mires externes).

Dans cette partie de la thèse, nous proposons d'employer une méthode en rupture avec ces approches. L'idée fondamentale est d'identifier les critères pertinents utilisés en clinique et les obtenir directement à partir d'une mesure complète du mouvement du pa-

tient, sans hypothèse particulière d'échantillonnage. Dans une première section 5.2, une réflexion est menée sur les altérations de la marche introduites par l'avance en âge, ce qui nous amènera à cibler l'un des paramètres cliniques pertinents à mesurer : l'asymétrie. Puis nous abordons une brève revue des différents systèmes réalisant une mesure quantitative de la marche 5.3 avant de décrire plus précisément la problématique du travail présenté dans cette partie.

5.2 Paramètres cliniques de la marche

Pour décrire le mouvement de la marche, plusieurs avenues existent. La plus simple et la plus usitée est de définir le mouvement par la résultante de celui-ci. Ainsi, il sera fait état de vitesse, de longueur de pas, de cadence et de temps de doubles appuis. La marche se décrit généralement à l'aide de ces paramètres d'après Wittle dans [165]. D'un autre point de vue, la démarche peut être caractérisée par ce qui en est la source, les mouvements angulaires des articulations. Le mouvement est alors spécifié par la valeur des angles de chaque articulation au cours du temps. Ces informations de base peuvent être utilisées pour calculer des indices de plus haut niveau, par exemple la régularité et la symétrie de la démarche. Le paramètre de la régularité se base sur le fait que la marche est un mouvement rythmique. Il caractérise alors la différence qui peut exister d'un cycle à l'autre pour les variables spatio-temporelles ou pour des variables de cinématique et de dynamique. La symétrie reflète le principe de symétrie générale du corps humain dans le plan sagittal qui soit se retrouver dans la ressemblance du mouvement du pas gauche avec le mouvement du pas droit.

Tout au long du vieillissement, le corps change lentement. L'une des conséquences est une détérioration du système musculo-squelettique, du système nerveux central et du système nerveux périphérique. Il en résulte une modification de la démarche. Plus spécifiquement, des anomalies liées à la dégénérescence de tissus en relation avec l'avance dans l'âge sont connues. La dégénérescence de certains tissus provoque des modifications fonctionnelles y compris en l'absence de pathologies. Différentes variables de la démarche sont modifiées par ces changements. Sadeghi a publié dans [144] le résultat de la comparaison des démarches mesurées par un système opto-électronique de 18 sujets âgés sains comparées à celles de 18 sujets jeunes. Comme montrés dans les tableaux 5.1 et 5.2, les paramètres spatio-temporels de la marche sont tous affectés par l'âge.

De nombreuses maladies dégénératives modifient plus encore la démarche. Par

TABLE 5.1 – Tableau récapitulatif des déformations des principaux paramètres de la marche du fait du vieillissement. * $p < 0.05$. Source [144]

Paramètres	Seniors		Jeunes	
	Moyenne	Écart-type	Moyenne	Écart-type
Vitesse (m/s)	0.93 *	0.20	1.30 *	0.12
Phase d'appuis (%)	63.00 *	3.00	60.70 *	1.70
Longueur du pas (m)	0.61 *	0.06	0.73 *	0.04
Longueur de la foulée (m)	1.23 *	0.10	1.45 *	0.07
Cadence (foulée/min)	91.70 *	12.20	106.50 *	7.03

TABLE 5.2 – Tableau récapitulatif des coefficients de variation du moment développé par les muscles et la puissance musculaire aux articulations. Source [144]

Articulation	Plan	Seniors (CV) %		Jeunes (CV) %	
		Couple	Puissance	Couple	Puissance
Hanche	Sagittal	55.90	72.50	110.10	125.70
	Frontal	29.17	145.50	108.80	104.20
	Transverse	73.00	307.60	110.00	145.00
Genoux	Sagittal	117.20	100.50	167.70	109.00
	Frontal	110.10	320.20	131.60	399.40
	Transverse	162.90	281.30	105.90	127.60
Cheville	Sagittal	40.10	83.30	32.40	57.20
	Frontal	121.70	223.30	171.00	230.00
	Transverse	74.40	217.30	293.60	225.00

exemple l'arthrose, ou encore arthropathie chronique dégénérative, provoque l'apparition de fissures dans les cartilages des articulations. Par exemple Baam et al. dans [22] a noté une modification de la démarche lors de l'atteinte des pieds et des chevilles. Il remarque que généralement la déformation consiste en une vitesse plus faible, un temps de double support plus long et un évitement des positions extrêmes. Plus généralement, Weiss et al. ont étudié l'impact de l'arthrose rhumatoïde sur le mouvement des membres inférieurs dans [163]. Ses résultats montrent que tous les paramètres spatiaux temporels et de cinématiques articulaires sont modifiés significativement par la maladie. De même

lorsqu'une maladie dégénérative atteint le système ostéo-articulaire ou le système musculaire, elle entraîne une modification de la démarche. Par exemple, d'après Fielding et al. dans [53], l'un des symptômes reconnus de la sarcopenia est une vitesse de marche inférieure à 1 m/s associée à une faible masse musculaire constatée. Il est immédiat que lorsque la symétrie fonctionnelle n'est plus assurée du fait de l'atteinte d'un de ces systèmes par une pathologie seulement sur un des côtés, la symétrie du mouvement de la démarche sera altérée. Par exemple, Hodt-Billington a montré dans [79] que la mesure de l'asymétrie de la chute du pied et de l'accélération du tronc lors du mouvement de la marche permet de différencier les patients atteints d'arthrose de la hanche par rapport à des sujets sains. La pose d'une prothèse de hanche provoque aussi la modification de la marche en modifiant la symétrie du mouvement [111]. Il y a aussi la différence de longueur des membres inférieurs qui provoque une altération de la démarche comme le montre Gurney dans [70]. Il remarque une asymétrie de la longueur et de la durée des pas d'un côté par rapport à l'autre. En effet le pas est plus court en temps et en distance pour la jambe la plus courte.

Le système nerveux central qui subit aussi l'attaque de maladie neuro-dégénérative qui s'exprime par des modifications des mêmes paramètres de la marche. Par exemple, la maladie de parkinson modifie les paramètres de la longueur des foulées et de la vitesse, mais pas celui de la cadence, comme le décrivent Sofuwa et al. dans [154]. Par contre, cette maladie se caractérise par une variabilité plus importante par rapport aux sujets sains. Plotnik et al. ont montré dans [133] que la symétrie est aussi un élément à prendre en compte pour cette maladie, mais seulement dans le cas de "freezing of gait". De même dans le cas d'un accident vasculaire cérébral, le mouvement de la marche est modifié au niveau de l'asymétrie comme le décrivent Hsu et al. dans [81].

L'asymétrie est un indice d'évolution de la pathologie lorsque celle-ci est latéralisée. Cette information quantifiée peut aider le médecin lors du diagnostic et le suivi de la pathologie. Il en résulte une nécessité de mesurer cette information d'asymétrie du mouvement de la marche. Pour que cela soit vraiment efficace, il est impératif de considérer les contraintes de temps et de coût de l'utilisation d'un tel outil en clinique quotidienne.

5.3 La mesure quantitative de la marche

Depuis les travaux de JE Marey à la fin du 19ème siècle [106], les méthodes de mesure ont évolué, mais les fondements restent les mêmes : la nécessité d'étudier la marche à la

fois du point de vue cinématique et dynamique. Pour cela différents types de capteurs résumés dans le tableau 5.3 sont maintenant techniquement utilisables pour réaliser la mesure des paramètres de la marche. Surer et Kose [157] ont proposé une classification des méthodes de l'analyse de la marche généralement utilisées.

Les différents appareils de mesures qui ont été utilisés pour mesurer la démarche vont maintenant être présentés. Tout d'abord seront présentés les protocoles utilisant un seul point de mesure. Puis suivront ceux réalisant une mesure multipoint. Un bref résumé des différents types de mesure est fait dans le tableau 5.3.

TABLE 5.3 – Tableau récapitulatif des méthodes de mesure quantitatives de la marche.

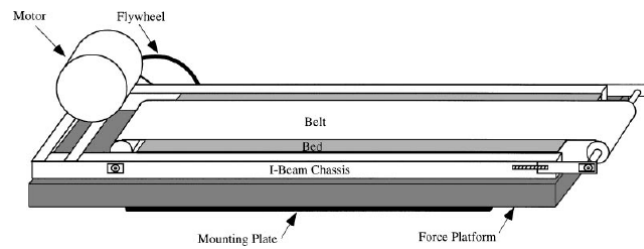
Nb points	Moyen physique	Exemples
Unique	Force	Plateforme de force, Jauge in vivo, ...
	Accélération	Accéléromètre
Multiple	Mécanique	Exo-squelette, locomètre de Bessou
	Cinématique	goniomètres
	Accélération	Accéléromètres
	Magnétique	Magnétomètre
	Optique	Caméra

5.3.1 Mesure du mouvement monopoint

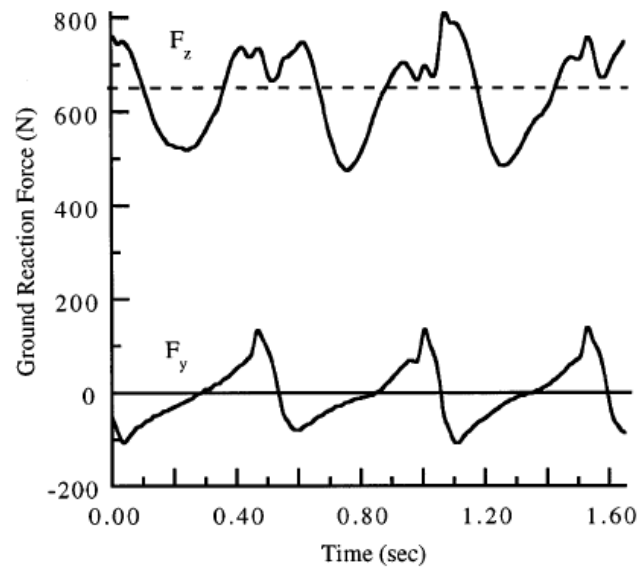
La représentation du mouvement de la marche peut se qualifier dans sa globalité selon un point de mesure. Comme il sera présenté plus tard, l'information mesurée est alors une résultante des forces de contact avec le sol ou l'accélération globale du corps. Puis le traitement de cette information pour le calcul de la symétrie sera présenté pour en cerner les limites.

Dans ce cas de figure, la démarche est alors mesurée d'un point de vue dynamique à l'aide d'une plateforme de force ou encore avec un accéléromètre placé sur un endroit du corps. Lorsqu'un seul point de mesure est disponible pour quantifier la symétrie, il est alors nécessaire d'enregistrer une mesure représentant la globalité du mouvement du corps. Une première approche fut proposée par Kram et al. dans [96]. Celle-ci consiste à installer une plateforme de force sur un tapis roulant comme illustré sur la figure 5.1. Comme montré sur la figure 5.1(a), en utilisant une plateforme de force avec 4 jauges de

contraintes, une à chaque coin, il est possible de mesurer les trois composantes de la force (F_x , F_y , F_z) exercée par le pied sur la plate-forme, les composantes antéro-postérieures (F_y) et verticales (F_z) sont illustrées sur la figure 5.1(b). Les forces mesurées sont alors inversées pour déduire la force de réaction du sol subie par le sujet lors de la marche.



(a)



(b)

FIGURE 5.1 – Illustration d'un tapis roulant équipé d'une plateforme de force (a) et du signal résultant de quelques foulées (b) pour la composante verticale (F_z) et la composante antéro-postérieure (F_y). (source [96]).

Dans le cas où un capteur est placé sur le corps, il est nécessaire de le placer sur le plan de symétrie du corps pour pouvoir réaliser une mesure de symétrie. Ainsi, grâce à la miniaturisation des capteurs sur MEMS (Micro Electro Mechanical Systems), il est possible de disposer d'accéléromètres de petites dimensions utilisables de manière ambulatoire. Ainsi B. Auvinet et al. proposent dans [16] de placer ce type de capteur au niveau des lombaires pour mesurer l'accélération du tronc au plus proche du centre de gravité du

corps. Ainsi l'accélération globale du corps peut alors être estimée comme le montre R. Le Bris dans [30].



FIGURE 5.2 – Illustration du placement de l'accéléromètre en position lombaire.

Pour calculer un index de symétrie à partir d'une mesure globale du mouvement, il est nécessaire de travailler sur le signal au niveau temporel. Par exemple, B. Auvinet et al. proposent dans [16] une méthode de calcul de l'asymétrie se faisant sur le signal de l'accélération verticale. L'index de symétrie est le résultat de l'autocorrélation de ce signal décalé d'une demi-période de la foulée. Seulement, dans le cas d'une forte variabilité du mouvement ou tout simplement si les pas droit et gauche n'ont pas la même durée, les cycles ne sont pas en phase. Ceci entraîne une dégradation du sens de la symétrie qui est mesurée.

Ceci pourrait être amélioré en utilisant la méthode "Piecewise alignment", telle que présentée par Helwig et al. dans [77], permettrait d'isoler chaque cycle et d'en estimer la similarité. Seulement cette information représente l'asymétrie globale du mouvement de la marche sans permettre d'isoler spatialement la cause de l'asymétrie. Pour distinguer cette perturbation, il est nécessaire de mesurer chaque partie du corps indépendamment.

5.3.2 Mesure du mouvement multipoints

Pour mieux expliquer le mouvement global, il est intéressant de connaître les mouvements internes relatifs entre les membres. Pour cela, la mesure de la dynamique et de la cinématique du mouvement de plusieurs membres est importante. Dans un premier temps seront présentés les moyens de mesurer la dynamique du mouvement puis ceux mesurant sa cinématique.

5.3.2.1 Mesure dynamique

L'une des informations dynamiques du mouvement est le contact avec le sol. Pour les plateformes de force, l'installation d'une plateforme pour chaque pied permet de mesurer la dynamique de l'appui de manière indépendante. Comme montré sur la figure 5.1, il est alors possible de différencier l'anormalité en comparant les signaux de chaque pas. Il existe des modèles de tapis roulant où plusieurs plateformes de forces peuvent être installées de manière parallèle telle qu'illustrée sur la figure 5.4. Cette installation rend possible l'enregistrement des appuis de manière continue.

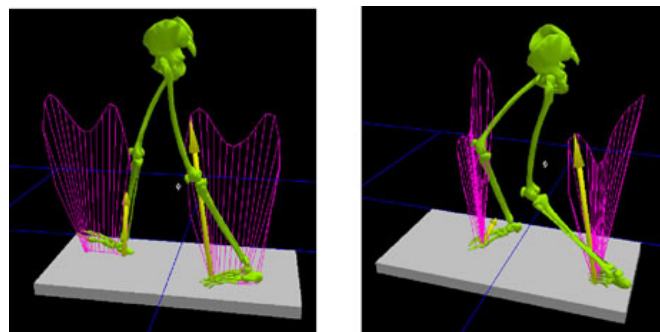


FIGURE 5.3 – Image de synthèse d'une plateforme de force avec le tracé du vecteur de la réaction du sol sur le pied . Un sujet sain à gauche et pathologique à droite. le signal représentant la force en jaune est identique pour chaque pas dans le cas du sujet sain et différent dans le cas du sujet pathologique. (source [83]).

Pour permettre de mesurer la réaction du sol sur plusieurs foulées et en dehors d'un tapis roulant, des matrices de capteurs de force ont été développées et peuvent ainsi être intégrées dans les chaussures du sujet ou dans un tapis de marche. Les forces mesurées sont généralement limitées à la composante orthogonale au capteur, donnant ainsi plutôt une mesure de pression que de force à proprement parler, comme le montre la figure 5.5.

En biomécanique, le corps humain est vu comme un système de corps rigides poly-articulés. Les efforts de liaison aux articulations sont très riches en informations sur les contraintes mécaniques subies par le squelette et indiquent aussi indirectement la résultante des actions des muscles sur ce squelette. Pour mesurer les efforts au niveau d'une articulation, il est possible d'implanter in vivo une jauge de contrainte dans l'articulation comme mentionné par Koni dans son état de l'art [95]. Mais cela serait difficilement envisageable du fait qu'une opération chirurgicale soit nécessaire. Dès lors, la seule solution est d'utiliser la seconde loi de newton [124] qui permet d'évaluer la force avec l'accélé-

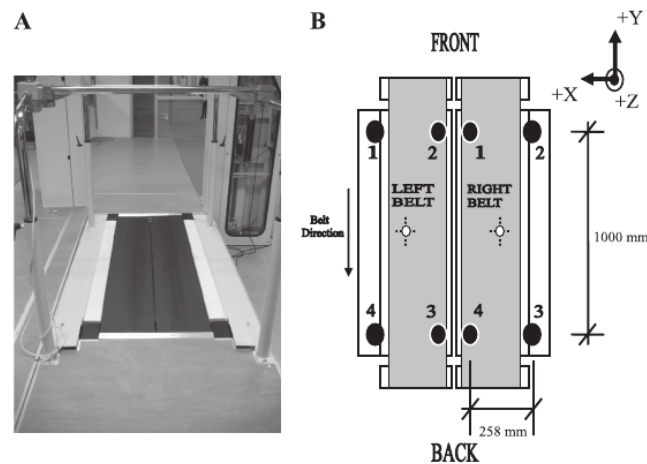


FIGURE 5.4 – Illustration d'un tapis roulant équipé de deux plateformes de force parallèles. (source [98]).

ration à un facteur de masse près. Par exemple, 4 accéléromètres et un gyroscope ont été positionnés sur chaque segment des jambes par Mayagoitia et al. dans [110] pour mesurer le mouvement des jambes. Mais dans ce cas une information dynamique pure n'est pas utilisable directement sans l'aide de l'information de vitesse de rotation donnée par les gyroscopes. Il est tout de même possible d'accéder à l'accélération en dérivant successivement la position de chaque point pour obtenir l'accélération. Pour réaliser cette mesure de position, les méthodes cinématiques doivent être utilisées.

5.3.2.2 Mesure cinématique

L'analyse cinématique de la marche s'intéresse à mesurer les trajectoires articulaires (angulaires ou cartésiennes) sans se soucier des causes mécaniques qui en sont responsables. Les premières études ont utilisé la chronophotographie, reprenant une méthode développée par Muybridge dans [122]. Étienne-Jules Marey l'appliqua à l'étude de nombreux mouvements tel que celui de la locomotion humaine. Cette méthode qu'il explique dans [106] consiste à prendre une série de photographies à intervalles réguliers d'une personne effectuant un mouvement. Grâce à des points de repère sur le corps, il est alors possible de mesurer la position des articulations au cours du temps tel que montré sur la figure 5.6.

Ce principe a été largement amélioré et accéléré grâce à l'utilisation de l'électronique. On peut distinguer maintenant les systèmes utilisant une mesure mécanique (transfor-

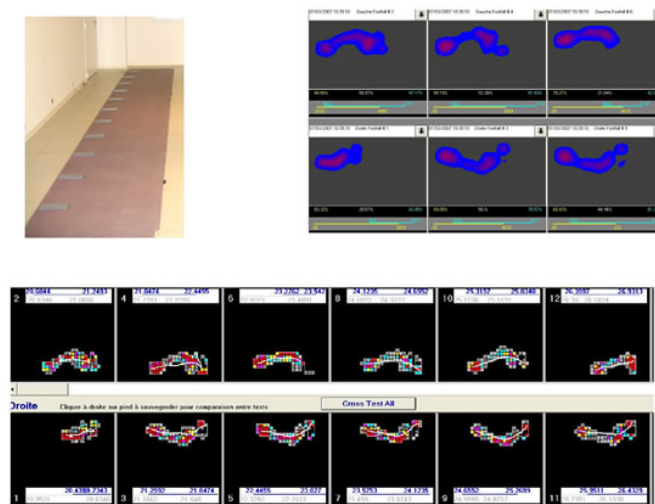


FIGURE 5.5 – Illustration du tapis de mesure GaitRite et des résultats calculés pour chacun des pas. (source [83]).

mée en signal électrique), inertielle ou visuelle. Nous développons maintenant ces trois systèmes de mesure.

5.3.2.2.a Mesure mécanique

L'utilisation d'un exo-squelette tel que celui proposé par AnimaZoo [82] permet d'enregistrer le mouvement grâce à des capteurs sensibles aux angles rapportés par la structure solidaire de chaque membre. Le sujet porte donc l'exo-squelette sur lui, solidairement attaché, et ce dernier enregistre en continu les transformations géométriques qu'il subit à chaque articulation.

Une autre approche développée par Bessou [24], nommé le locomètre, permet de faire des mesures cinématiques de points particuliers en les reliant physiquement à une base de mesures. Le mouvement est enregistré à l'aide de câbles attachés au niveau des points de mesure et reliés à des dérouleurs ayant la capacité d'enregistrer leurs propres rotations. Le locomètre est principalement utilisé pour enregistrer le déplacement des pieds au cours du temps et, donc, de déduire la longueur et la fréquence de chaque pas. Déduire un indice d'asymétrie ou un déficit de longueur des pas est ainsi quasiment immédiat. Cependant, la mesure ne porte que sur les pieds alors que l'asymétrie peut intervenir à plusieurs autres niveaux du corps. D'un point de vue pratique, ce système est limité aux déplacements en ligne droite, dans l'axe de l'appareil, et sur quelques mètres seulement.

Ces systèmes mécaniques sont généralement invasifs et impliquent d'équiper le pa-

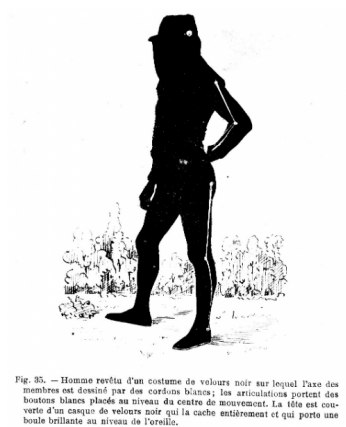


Fig. 35. — Homme revêtu d'un costume de velours noir sur lequel l'axe des membres est dessiné par des cordons blancs; les articulations portent des boutons blancs placés au niveau du centre de mouvement. La tête est couverte d'un casque de velours noir qui la cache entièrement et qui porte une boule brillante au niveau de l'oreille.

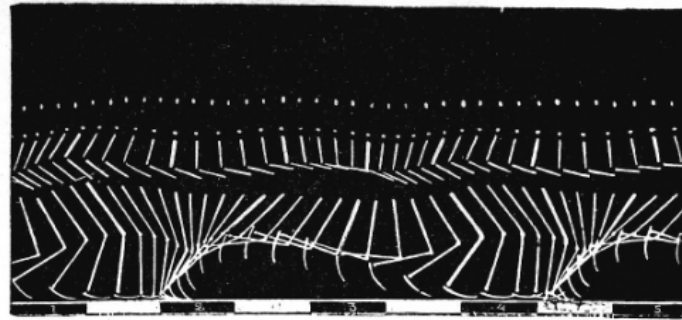


Fig. 36. — Images chrono-photographiques d'un coureur. En bas de la figure, une échelle dont les divisions ont 0m,50 de longueur sert à apprécier l'étendue des mouvements.

FIGURE 5.6 – Illustration provenant des travaux de Jules-Etienne Marey montrant l'utilisation de la chronophotographie pour la mesure du mouvement de la course. À gauche est représenté l'habit où sont attachés les marqueurs. À droite, le résultat de l'exposition de la plaque photographique à plusieurs moments du mouvement.(source : [107])



FIGURE 5.7 – Illustration du locomètre de Bessou. (source [4]).

tient. De plus, ils sont très limités en terme de mesure et de points d'intérêt, comme nous l'avons montré pour le locomètre.

5.3.2.2.b Mesure magnétique, centrales inertielles, fusion de capteurs

Si on cherche à augmenter le nombre de points de mesures pour suivre le mouvement de plusieurs segments corporels, les capteurs positionnés sur le corps offrent des fonctionnalités intéressantes. Qu'ils soient fondés sur des accéléromètres tri-axiaux, des magnétomètres, des gyroscopes, ou une combinaison de ces capteurs, ces systèmes retournent généralement leur orientation dans un repère cartésien. Par exemple, 4 accéléromètres et un gyroscope ont été positionnés sur chaque segment par Mayagoitia et al. dans [110] pour mesurer le mouvement des jambes. Les dernières technologies, comme celle proposée par XSens (www.xsens.com) ou Movea (www.movea.com), intègrent une

batterie et un système de transmission sans fil qui offre une grande liberté de mouvement. Toutefois, il est nécessaire de calibrer le système, et plus particulièrement le modèle biomécanique du sujet. En effet, ces systèmes reposent sur un modèle de squelette du sujet qui est généralement figé. Les orientations absolues de chaque capteur sont alors traduites en angles articulaires qui peuvent ensuite être utilisés pour l'analyse clinique.

Comme tout système de mesure cinématique s'appuyant sur une multitude de points de mesure, la principale difficulté consiste à extraire l'information pertinente d'un point de vue clinique, à partir de la somme de mesures disponibles. Comment constituer un indice d'asymétrie unique à partir d'un vecteur d'angles articulaires pouvant aller à plus de 50 degrés de liberté. Gouwanda et al. [68] ont proposé d'utiliser ces systèmes pour calculer un nouvel indice d'irrégularité de la marche. Cependant, cet indice se fonde sur un nombre très limité de points de mesures et peut passer à côté d'informations importantes.

5.3.2.2.c Mesure optique

Afin d'éviter que le sujet ne doive porter un équipement complexe sur lui, comme une combinaison faisant circuler des câbles chez Xsens, il est possible de remplacer les capteurs ci-dessus par des marqueurs passifs. Les capteurs sont alors externalisés et la méthode consiste à suivre la trajectoire de ces marqueurs. C'est le cas des systèmes vidéo et opto-électroniques. Ces derniers utilisent des caméras infrarouges et des stroboscopes afin d'isoler dans l'image les marqueurs réfléchissants. Les marqueurs, positionnés sur des repères anatomiques, s'ils sont vus par au moins deux caméras, peuvent être reconstruits en 3D, comme le montre Cappozzo et al. dans [32]. Le système opto-électronique, tel que montré sur la figure 5.8, est le plus précis disponible actuellement, mais aussi le plus coûteux. La précision de localisation des marqueurs dans l'espace de mesure peut être submillimétrique. Seulement, comme le système mécanique d'exo-squelette, il nécessite du temps pour être mis en œuvre, ce qui est inadéquat pour une large diffusion en utilisation clinique.

Ces systèmes fournissent directement la position 3D de repères anatomiques. Cette donnée de base est très éloignée des paramètres pertinents d'un point de vue clinique. Il existe donc un grand nombre de manières d'analyser ces données comme le précise Chau dans [35, 36] et nombreux travaux définissant des critères cliniques à partir de ces mesures, en particulier le Gait Index [41] ou les nombreux calculs d'asymétrie [130, 68]. La méthode consiste alors à estimer les centres articulaires, à reconstruire le squelette sous-jacent selon un modèle biomécanique, calculer les angles articulaires, puis les ex-

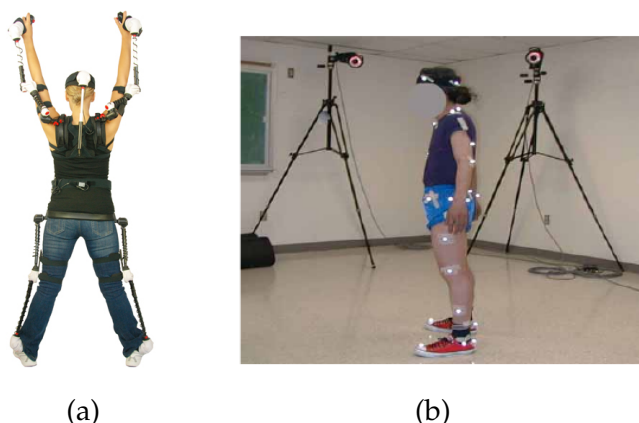


FIGURE 5.8 – (a) - Illustration du système de capture du mouvement par un système exo-squelette. (source [82]). (b) - Illustration du système de capture du mouvement opto-électronique. (D'après Wittle 2007).

exploiter pour obtenir le paramètre clinique. Chacune de ces étapes repose sur un modèle, des hypothèses et des simplifications qui apportent leur lot d'imprécisions et de difficultés. Il n'est pas rare d'utiliser plus de 50 marqueurs pour obtenir ces informations, ce qui revient à échantillonner le corps de manière relativement précise, si on compare aux autres approches. Cependant, les méthodes utilisées dans ces systèmes reposent très fortement sur les hypothèses du modèle biomécanique sous-jacent, ce qui peut être un véritable problème dans certains cas pathologiques. De plus, les informations en sortie sont extrêmement riches et il est très difficile pour le clinicien de proposer un diagnostic avec un tel volume d'information.

5.3.3 Bilan sur les méthodes de mesure

Pour chacune des méthodes précédemment citées, leur mise en place ne correspond pas aux contraintes d'une utilisation clinique quotidienne pour l'analyse de la marche. En effet, soit l'information est réduite à un seul type de point de mesure (GaitRite, plateforme de force, accéléromètre lombaire) ce qui rend l'information retournée incomplète, soit celle-ci est échantillonnée uniquement au niveau des centres articulaires. De plus, pour ces dernières méthodes multi-segmentaires, il est nécessaire d'équiper le sujet avec des marqueurs/ capteurs (Vicon, accéléromètres/ goniomètres, exo-squelette) ce qui rend le protocole de mesure inadéquate pour la clinique de tous les jours.

Pour pallier cela, des systèmes sans marqueur, utilisant des caméras vidéos ont été développés. Ils permettent une mesure du mouvement sans contact et sans que le su-

jet doit porter de dispositif sur le corps. La caméra vidéo permet d'enregistrer toute l'information de couleurs visibles d'une personne. Il n'y a donc pas d'effet de sous-échantillonnage avec ce capteur. La donnée de base ne tient pas compte non plus du modèle biomécanique sous-jacent et représente donc la totalité de l'information de ce qu'a effectué le sujet. Ce domaine de recherche est particulièrement prolifique depuis ces 20 dernières années, grâce au progrès des techniques d'analyse et des capacités de calcul. Les conditions échantillonnage et de rapidité que nous nous sommes posés semblent être en adéquation avec ces méthodes et nous proposons donc d'examiner de plus près cette approche dans la section suivante.

5.4 Mesure de la marche à l'aide de caméras

Depuis longtemps, la marche a été étudiée à l'aide de séquence d'images. D'abord sous forme de chrono-photographie, puis avec des caméras électroniques lorsque celles-ci apparurent. Dès lors que le signal vidéo et son traitement numérique furent disponibles, l'analyse automatique par ordinateur s'est rapidement imposée. En effet l'intérêt majeur de ce type de capteur vient du fait que l'information mesurée est obtenue sur toute la partie du corps visible et non plus sur un sous-échantillonnage de points particuliers. Mais le revers de la médaille est que toute l'information de la scène est enregistrée sans distinction entre le sujet et le fond de celle-ci. Seule l'information de couleur est utilisable pour réaliser cette distinction. De nombreux travaux ont été réalisés pour estimer la pose d'un sujet dans une ou des images à partir de ce type d'information de couleur. Plusieurs revues de littérature de ces différentes méthodes ont été réalisées par Moeslund [114, 115], Aggarwal[8], Gavrilă[62]. Ces méthodes présentées dans ces revues bibliographiques ont comme objectif principal, la reconstruction de poses pour la reconnaissance gestuelle. Une pose est un vecteur d'état à un moment donné, regroupant toutes les informations liées aux configurations articulaires du sujet. En règle générale, la plupart de ces méthodes traitent de conditions génériques de pose dans des environnements quelconques. Ces travaux ont souvent une finalité de reconnaissance de geste ou capture de mouvement pour une identification ou pour une utilisation ludique ou artistique.

Plus particulièrement, dans le cadre de la vidéosurveillance, de nombreux travaux ont été effectués sur la reconnaissance de la marche dans le but d'identifier une personne. En effet en supposant que chaque individu a sa propre démarche, de nombreux travaux portent sur l'analyse d'une signature de marche propre à chacun. Par exemple, Han et

Bhanu proposent dans [71] une méthode sommant une série de silhouettes réalignées d'une personne marchant pour identifier la personne à travers une analyse statistique. De même que Sivapalan et al. proposent dans [152] une adaptation de cette méthode à des informations tridimensionnelles pour la reconnaissance d'activité. Vu notre objectif d'analyse clinique de la marche, et vu l'ampleur des travaux dans le domaine de l'analyse vidéo du mouvement, nous focalisons la revue de littérature ci-dessous sur les méthodes ayant pour but l'analyse clinique de la marche ou s'en rapprochant. On peut toutefois distinguer deux grandes familles d'approche : mono ou multi-caméras.

5.4.1 Système mono-caméra

Un premier exemple d'application de cette méthode pour l'analyse de la marche a été réalisé par Courtney et Paor dans [38]. Les différentes parties du corps y sont localisées en modélisant chacune par une ellipse. Le corps filmé de profil est alors modélisé en 4 ellipses pour la tête, le tronc, la cuisse et le tibia. Cette méthode permet de mesurer la cadence, la vitesse, la longueur des pas, la rotation de la hanche et du genou durant quelques pas. Seulement, ces rotations sont uniquement mesurées dans le plan sagittal, ne fournissant ainsi qu'une information partielle. De plus, du fait de l'autoocclusion du corps dans le plan sagittal, les deux jambes ne peuvent être mesurées en même temps. Nghiem et al. proposent d'utiliser dans [125] les bords de chaque jambe comme source d'information. Il est alors possible de comparer les angles des genoux durant le cycle de marche sauf lorsque l'occlusion est trop importante.

Ces différents travaux nécessitent un axe de marche perpendiculaire à l'axe optique de la caméra. Pour palier à cela, Goffredo et al. réalisent une rectification selon l'axe de marche dans [66]. Seulement la méthode nécessite que la personne marche de manière rectiligne comme le montre la figure 5.9. Jean et al. dans [87] procèdent à une rectification qui s'adapte à l'évolution de ce vecteur de marche tout au long du parcours permettant ainsi de s'abstraire de la contrainte de parcours rectiligne telle que montrée dans la figure 5.10.

Ces systèmes mono-caméras sont intéressants pour l'obtention de mesures cinématiques de base, comme les paramètres spatiotemporels. Mais pour une mesure d'une information cinématique précise et complète, le problème est mal posé. En effet, les rotations dans le plan frontal ou transversal sont difficiles à mesurer, surtout lorsque celles-ci sont de faibles amplitudes. De plus l'auto-occlusion réduit l'utilisation possible pour la comparaison droite gauche, ce qui est un problème majeur dans l'analyse clinique de



FIGURE 5.9 – Exemple de résultats de la méthode de Goffredo et al. Source Goffredo et Al [66].

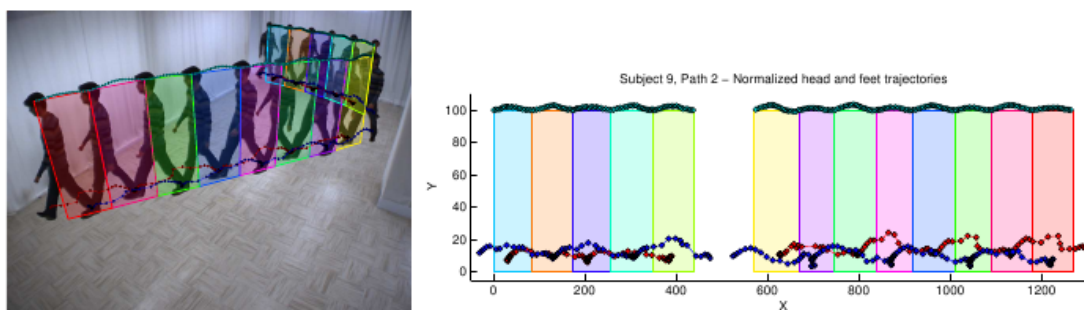


FIGURE 5.10 – Exemple du résultat de mesure de l'emplacement des pieds et de la tête dans le plan sagittal rectifié tout au long du parcours. Source Jean et al. [87].

l'asymétrie, comme c'est le cas de cette thèse.

Dans ce cas, il est nécessaire d'utiliser une information 3D pour dépasser ces problèmes.

5.4.2 Système multi-caméras

Avec l'augmentation des capacités de calcul, d'enregistrement et de communication, les systèmes multi-caméras sont devenus possibles et abordables. Plusieurs manières de fusionner l'information à partir des différents points de vue ont été explorées.

Marzani et al. proposèrent une méthode dans [108] qui fusionne l'information des bords des membres inférieurs pour suivre la cuisse et le tibia modélisés avec des ellipsoïdes. Rapidement, la fusion d'information a été réalisée grâce à la technique du visual hull déjà présentée dans la section 2.5. Ces techniques permettent de reconstruire l'enveloppe d'un sujet. Cette reconstruction à partir de la méthode initialement proposée par

Laurentini [97] a été largement améliorée et spécialisée à la reconstruction du modèle humain par Boyer, pour les étapes de calibrage [29], de segmentation [100] ou encore de l'analyse de la pose.

Seulement ces travaux ont été réalisés dans un but de capture de mouvements génériques, sans finalité spécifique à la marche, ou encore dans le cadre clinique. Pour évaluer la possibilité de mesurer les paramètres de la marche dans un but clinique, à partir d'une surface reconstruite à par visual hull, Mündermann et al. ont étudié dans [119] l'impact du nombre de caméras sur l'erreur de reconstruction de la surface du corps humain. Ils ont évalué qu'un minimum de 8 caméras était nécessaires pour obtenir une erreur moyenne absolue de 11 mm sur tout le corps lors d'une pose statique telle que montrée figure 5.11. Dans [118], ils ont montré qu'un minimum de 8 caméras est nécessaire pour obtenir une reconstruction acceptable pour un mouvement de marche sur 2 pas dans un volume de 3m par 1.5m par 2m, conseillant plutôt l'utilisation de 16 caméras installées régulièrement de manière circulaire ou hémisphérique autour du sujet.

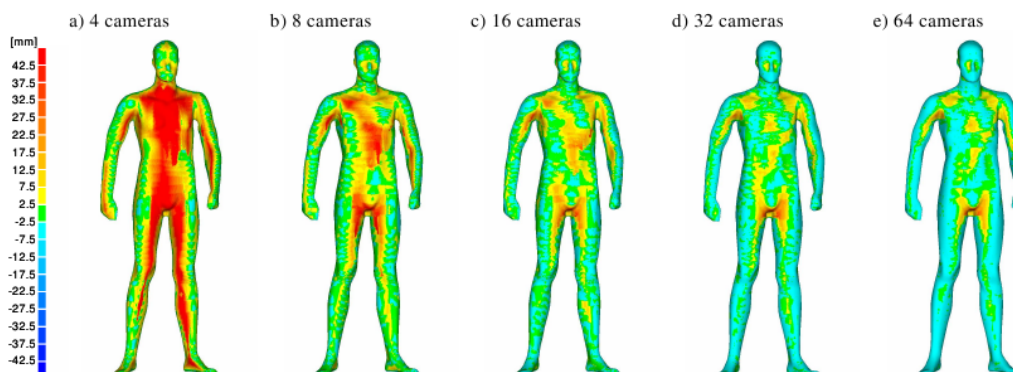


FIGURE 5.11 – Exemple de résultat de mesure de l'erreur de reconstruction obtenue par Mündermann dans [118]. Source Mündermann et al. [118].

Finalement, Corazza et al. ont montré dans [37] qu'il était possible de localiser le centre articulaire de la hanche avec une succession de surface décrivant le mouvement de la jambe autour de cette articulation. Leur méthode utilise un recalage de toutes les surfaces sur celle du premier instant en minimisant leur différence par un algorithme ICP "Iterative Closest Point". Les résultats obtenus par simulation ont montré que l'erreur de localisation était comparable à celle obtenue en utilisant des marqueurs placés sur la peau. Mündermann et al. ont montré dans [117] que leur méthode permettait d'estimer les angles de flexion et d'abduction du genou avec une précision respective de 2.3 degrés

et 1.6 degrés.

Pour une application clinique, il en résulte une impossibilité d'utilisation, du fait de la complexité de l'installation (salle suffisamment grande) et le prix de toute l'installation.

Le point commun de toutes ces méthodes utilisant une ou plusieurs caméras vidéos est leur utilisation du contour de la silhouette comme source principale d'information. Par rapport à notre problématique de mesure du mouvement à l'aide du maximum d'information, on peut noter le progrès réalisé grâce à la vidéo sur le nombre de points de mesure décrivant le mouvement. En effet dans le cas où la personne a une hauteur de 500 pixels dans une image, son contour est alors décrit au minimum par 2 points par ligne, ce qui représente un minimum de 1000 points de mesure par caméra. Mais pour définir ce contour, seule l'information de couleur est disponible. Comme nous l'avons évoqué dans la section 4.2.2, le problème de la segmentation de la personne dans l'image est mal posé par rapport au type d'information qu'est la couleur. Même si dans ce cas, il est possible d'adapter la scène et l'habillage du sujet en conséquence pour faciliter cette opération.

Retrouvant la même problématique de segmentation du sujet dans l'image que pour la première partie concernant la chute, nous nous sommes orientés vers les caméras de profondeur comme outils de mesure.

5.5 Système utilisant une caméra de profondeur

Les caméras de profondeur sont des appareils autonomes qui permettent de mesurer directement la profondeur des objets dans une scène. L'information de profondeur retournée décrit la surface visible par la caméra. Trois différents systèmes sont actuellement disponibles sur le marché pour mesurer de manière relativement précise une information de profondeur. Ceux-ci sont la stéréo-vision, la mesure du temps de parcours de la lumière et la reconstruction à l'aide de lumière spatialement structurée. Les principes de fonctionnement de ces caméras de profondeur sont étudiés par Kolb et al. dans [94] et très bien présentés par Castena et Navab dans [33].

5.5.1 Système stéréoscopique

La plus ancienne méthode de construction de carte de profondeur consiste à reproduire le système binoculaire humain. Le principe de la stéréovision repose sur la connaissance de la position d'un point de l'espace dans l'image de chaque caméra et la relation géométrique qui relie celles-ci. Par triangulation il est alors possible de retrouver la po-

sition 3D de ce point. Cette méthode est très performante, mais dépend beaucoup de la qualité de localisation dans l'image de chaque caméra de la position d'un même point de l'espace. C'est souvent la faiblesse de cette méthode puisque la qualité de reconstruction de la scène dépend de la complexité de la scène elle-même. En effet, si la scène présente des points suffisamment caractéristiques les uns des autres, il est alors possible de les associer. Si une scène présente des objets de textures uniformes (par exemple : la peau), leur reconstruction sera alors de piètre qualité, car il sera difficile de distinguer correctement les couples de points dans chaque caméra.

Pour résoudre ce problème de dépendance de la scène, les nouveaux systèmes qui viennent d'apparaître sur le marché se fondent sur le principe de la lumière structurée. Cette méthode utilise une information ajoutée dans la scène pour permettre la reconstruction de la carte de profondeur. Pour reconstruire la profondeur, la structure de l'information ajoutée dans la scène est soit de nature temporelle soit de nature spatiale. Ces deux méthodes sont développées plus en détail dans les deux sous-sections suivantes.

5.5.2 Structure temporelle

La première méthode se fonde sur le temps mis par le rayon lumineux pour réaliser le chemin émetteur-scène-récepteur. C'est pour cela qu'elle est appelée "Temps de vol" ou "Time Of Flight" (TOF). Un signal lumineux mesuré par chaque pixel est analysé pour retrouver la structure temporelle introduite à l'émission. Il est ainsi possible de mesurer le temps de parcours, et donc, d'en déduire la profondeur. Pour cela, deux structures temporelles de lumière sont possibles. La première est un codage temporel en amplitude de la lumière, illustré en figure 5.12. La lumière est envoyée de manière pulsée. Le récepteur mesure alors le temps entre l'envoi de l'impulsion lumineuse et la détection de celle-ci sur le capteur de la caméra. Seulement les temps à mesurer sont très courts (de l'ordre de la dizaine de pico seconde) et nécessite une précision de mesure très coûteuse.

La deuxième méthode utilise un codage en fréquence de l'information. En appliquant une modulation de la fréquence sur la lumière projetée dans la scène, le temps de parcours est alors accessible en mesurant le déphasage entre la lumière envoyée et la lumière reçue.

Pour effectuer une mesure, il est nécessaire d'intégrer les informations dans le temps. Ceci peut devenir problématique en cas de mouvement. En effet, un flou peut apparaître dans l'image lorsque les objets dans la scène bougent trop vite. Ce genre de capteur est toujours en cours de développement et en constante amélioration. Seulement leur prix

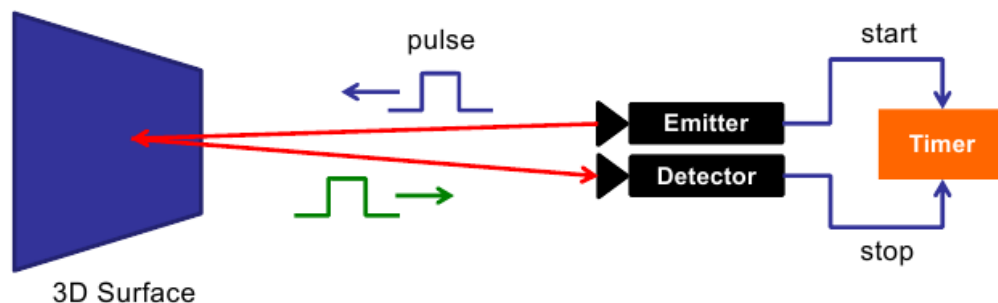


FIGURE 5.12 – Schéma du principe de fonctionnement du Time Of Flight avec une structure temporelle pulsé. Source Castena et Navab dans [33].

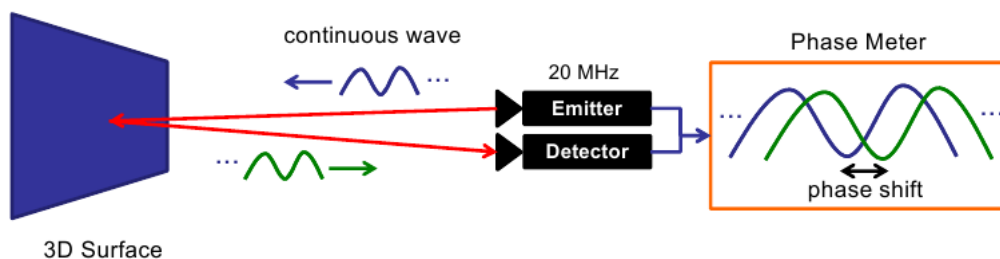


FIGURE 5.13 – Schéma du principe de fonctionnement du Time Of Flight avec une structure temporelle pulsé. Source Castena et Navab dans [33].

(plusieurs milliers de dollars) et leur faible résolution (160 par 120 pixels) rendent leur accessibilité et leur utilisation difficile. Pour résoudre ce problème, une information lumineuse à structure spatiale peut alors être projetée dans la scène permettant ainsi d'utiliser des caméras conventionnelles de haute résolution et de moindre coût.

5.5.3 Structure spatiale

Le principe de cette structure est de projeter une information, typiquement une image, dont les caractéristiques de voisinage permettent de reconnaître chaque pixel de cette image. Ainsi, à l'instar de la stéréovision, le projecteur devient alors une caméra inversée qui projette cette image connue dans la scène. Cette information projetée permet d'identifier les couples de pixels correspondants entre la caméra et le projecteur. À partir de ces couples de pixels et de la relation géométrique connue entre la caméra et le projecteur, la profondeur est calculée, par triangulation. Il est donc nécessaire d'être capable

d'identifier l'information projetée dans l'image de la caméra.

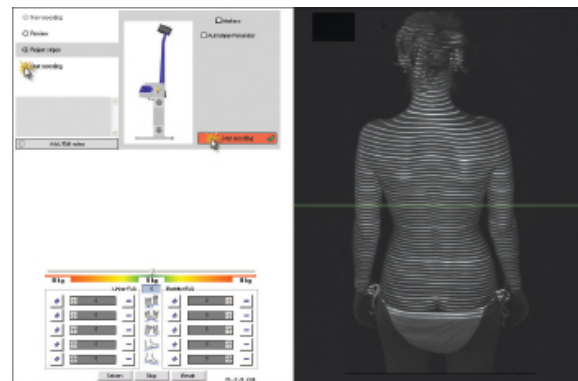


FIGURE 5.14 – Exemple de reconstruction de la surface du dos par lumière structurée. Source [63].

Pour reconstruire une information dense de la scène, le lien entre chaque pixel de la caméra et son correspondant dans l'image projetée doit pouvoir être obtenu. Comme le nombre de niveaux de gris mesurables ou projetables est limité, la structure spatiale doit alors être périodique. Différents motifs sont utilisés pour cela, par exemple, des lignes horizontales sont utilisées par l'appareil Formetric 4D [63] pour reconstruire la surface du dos, comme montrée sur la figure 5.14. Seulement dans ce cas, la plage de profondeur mesurable est définie par la périodicité du motif. En effet en cas de discontinuité de profondeur, il peut apparaître une ambiguïté au niveau de la reconnaissance du motif. Dans le cas de la reconstruction du dos, la plage de profondeur étant de quelques centimètres, cette méthode est utilisable. Pour le cas de la marche, où toute la surface du corps doit être numérisée, le problème reste ouvert.

Pour lever cette ambiguïté, de l'information supplémentaire peut être ajoutée par l'envoi de plusieurs motifs successifs permettant de construire une signature temporelle pour chaque pixel. La recherche en fonction de cette évolution permet ainsi de retrouver les pixels de la caméra et du projecteur correspondants. C'est la méthode utilisée par l'outil MegaCapturor 3D digitizer de Creaform [40] pour la reconstruction de la surface du corps humain en position quasi statique.

Dans le cas de mouvements importants, ceci ne peut être fait sans risquer des perturbations dues au flou causé par le mouvement, réduisant ainsi la précision de mesure.

Un compromis doit donc être fait entre densité et précision de mesure. Une solution consiste à réaliser la mesure de profondeur avec un motif fixe dans le temps contenant seulement un certain nombre de points (mais suffisamment grand pour pouvoir en inter-

poler une carte dense de précision suffisante, mais de résolution spatiale moindre) mais de faire en sorte que ces points puissent être identifiables quelque soit leur position dans l'image capturée.

Ce type de solution a été retenu pour le développement de la Kinect. La caméra projette un motif de points. Celui-ci a la particularité de donner pour chaque point une signature unique calculée à partir des positions relatives des points du voisinage immédiat. Cela rend ainsi chaque point identifiable par rapport aux autres. Le motif projeté est visible sur la figure 5.15. La profondeur est alors calculée en mesurant la distance entre la position dans l'image de la caméra et celle du projecteur. Cela revient à un problème de stéréovision et de mesure de profondeur par triangulation, mais cette fois sans dépendre des textures présentes dans la scène.



FIGURE 5.15 – Image du motif de lumière structurée par une Kinect. Source [132]

Cette méthode permet d'être résistant au bruit de mouvement du fait qu'une seule image soit nécessaire pour réaliser le calcul de la profondeur. Ceci se fait au prix d'une information parcellaire de la surface tridimensionnelle de la scène.

Lorsque les informations de profondeur sont mesurées, la partie suivante va présenter les méthodes proposées pour analyser la marche à partir de ce genre d'information.

5.6 Utilisation des caméras de profondeur pour l'étude de la marche

L'information de profondeur rendue disponible grâce à ce genre de caméra a été exploitée de différente manière pour mesurer le mouvement d'une personne.

Pour reconstruire celui-ci, différents travaux ont été réalisés pour venir reconstruire la pose à partir de cette information de profondeur. Par exemple Shotton et al. proposent dans [151], une méthode de classification d'appartenance d'un pixel à une partie du corps basée sur la comparaison de profondeur entre ce pixel et de son voisinage. Lorsque chaque pixel a été classifié comme appartenant à une partie du corps, il est alors possible de déduire les positions des centres articulaires comme illustrés sur la Figure 5.16. De même que Rasmus et al. dans [88] utilisent une optimisation d'un squelette dans la carte de profondeur telle que montrée par la figure 5.17.

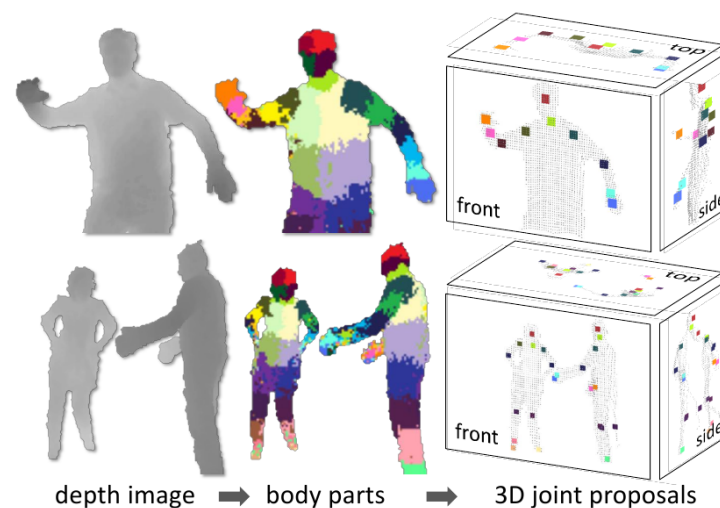


FIGURE 5.16 – Schéma de fonctionnement de la localisation du squelette selon la méthode de Shotton et al. dans [151]. De gauche à droite sont représentées les cartes de profondeur, puis au milieu, les résultats de classification des pixels en fonction de leur appartenance à une partie du corps (par exemples : main, poignet, avant-bras, coude, bras, ...) et enfin à droite la localisation des centres articulaires. Source : Shotton et al. dans [151]

Dans ces deux cas, l'information sur le mouvement est calculée en fonction d'un modèle articulé. Ce genre de méthode est sensible à l'état d'initialisation de la longueur des segments. Mais il pourrait être imaginé que ces mesures puissent être faites rapidement avant le test et rentrées manuellement. Seulement le problème de la quantification de la

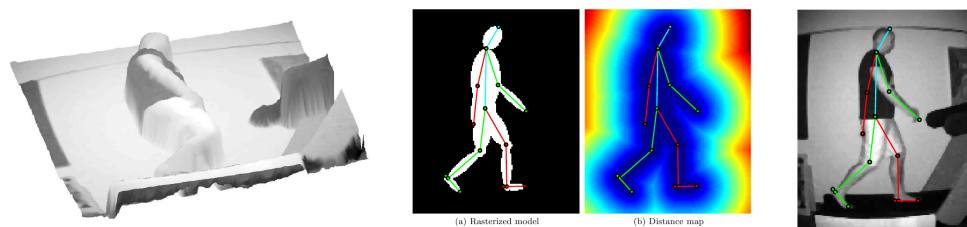


FIGURE 5.17 – Images de la méthode de Rasmus et al. dans [88]. Source : Rasmus et al. dans [88]

symétrie à partir des résultats d'un modèle squelettique reste ouvert. Dans le cadre de notre travail sur la symétrie, nous souhaitons proposer une méthode simple qui pourrait s'abstraire d'un quelconque modèle articulé, ceci en traitant directement l'information de base : la surface.

Pour valider cette hypothèse de calcul d'index à partir des informations de la surface, nous avons réalisé une première expérimentation telle que présenté dans [139] et dans [18]. Dans la première méthode, le sujet est observé de dos. Après avoir été segmentées en fonction de la différence des profondeurs entre chaque image et un modèle de fond, les surfaces sont réalignées pour que tout leur centroïde coïncide. Ceci donne le résultat proposé figure 5.18. Puis l'index de symétrie est calculé en sommant les différences entre cette surface moyenne et son symétrique par rapport à l'axe vertical passant par le centroïde commun. Lors du calcul de la carte de profondeur moyenne, la propriété projective des images n'était pas rectifiée. En effet, la carte de profondeur de la Kinect donne une information de profondeur en coordonnées métriques alors que les informations de coordonnées verticales et de coordonnées horizontales sont en pixel. Or comme il a été vu dans la partie 2.5.3, le modèle projectif rend ces deux coordonnées dépendantes de la profondeur pour la transposer en métrique. De ce fait un même pixel dans la carte de profondeur ne pointe pas un point à la même altitude en fonction de la profondeur.

Il en retourne que faire la moyenne des valeurs de profondeur prises par un même pixel au cours du temps peut introduire des erreurs pour cette raison.

L'autre méthode que nous avons proposée dans [18] utilise plusieurs caméras de profondeur. Dans ce cas comme chaque point était reprojecté dans l'espace, le modèle projectif était bien respecté. Chaque volume du sujet reconstruit au cours du temps était recalé de sorte à faire coïncider les centroïdes de chacun. Or Gurney précise dans [70] que la différence de longueur de segments des membres inférieurs impacte sur le mouvement du centre de gravité. Il en découle qu'une autre méthode de recalage doit être envisagée.

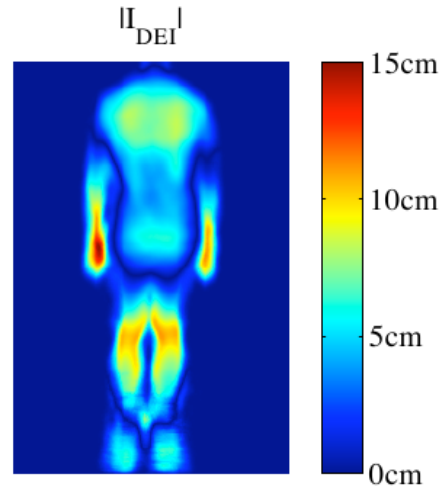


FIGURE 5.18 – Exemple de sommation de carte de profondeur d'une personne vue de dos pendant le mouvement de la marche. Source [139].

5.7 Conclusions

Cette introduction a permis de préciser les différents points de réflexions qui ont orienté le contexte de travail. Tout d'abord le contexte médical montre clairement la nécessité de mesurer l'asymétrie en clinique routinière. Le but est de pouvoir diagnostiquer ou suivre l'évolution de pathologies impactant sur la symétrie de la marche. Par exemple, des maladies latéralisées sur un côté du corps ou encore une maladie du système nerveux central comme la maladie de Parkinson. L'objectif d'utilisation en clinique impose des contraintes de simplicité du protocole (sans marqueur et sans contact pour une rapidité de réalisation) et de faibles coûts (pour une large diffusion). La revue des différents appareils disponibles a rapidement orienté le choix vers les caméras de profondeur à lumière spatialement structurée de type Kinect. Nous allons proposer dans la partie suivante comment utiliser cette information de profondeur sans avoir recours à un quelconque modèle biomécanique pour réaliser la mesure d'un index de symétrie du mouvement de membres inférieurs dans le cas où une différence de longueur de jambe existe.

Chapitre 6

Mesure de l'asymétrie des membres inférieurs avec une caméra de profondeur

Sommaire

6.1	abstract	118
6.2	Introduction	118
6.3	Previous works	120
6.4	Method	121
6.4.1	Depth map pretreatment	122
6.4.1.1	Transformation from depth map to 3D points	122
6.4.1.2	Clipping and orthographic projection	123
6.4.2	Key depth maps	123
6.4.2.1	Step cycle detection	123
6.4.2.2	Assignment of frames to key depth map	124
6.4.2.3	Key depth map construction	124
6.4.3	Key depth map analysis for asymmetry measurement	124
6.4.3.1	Leg segmentation	124
6.4.3.2	Asymmetry measurement	125
6.5	Experimentation and validation	126
6.5.1	Experimentation	126
6.5.1.1	Description of the population	126
6.5.1.2	Description of the protocol	127

6.5.1.3	Setup equipment	128
6.5.2	Validation	128
6.6	Results	128
6.7	Discussion	130
6.8	Conclusion	133

Article

Lower limb movement asymmetry during the gait cycle with a depth camera

Auteurs

Edouard Auvinet, Franck Multon, Jean Meunier

6.1 abstract

Human walking is a complex neural and biomechanical process that is still studied in many domains. Many processes are involved in controlling gait so that many diseases or injuries can be observed in gait pattern alterations. For example, gait asymmetry can be related to neurological diseases, segment length differences or joint deficiencies. Hence gait asymmetry is widely studied in clinical analysis to diagnose pathologies and to evaluate the performance of treatments. However designing a system to measure asymmetry in clinics is still a challenge as it requires designing fast and easy-to-use systems that could be used without the help of specialized technicians. To tackle this problem we propose a novel method to analyze asymmetry of the lower limb movements thanks to a depth camera placed in front of a treadmill. The method can isolate lateral and antero-posterior symmetry information for the lower limb movements within a gait cycle. 20 healthy subjects participated in an experiment where they had to walk under conditions that could affect gait symmetry, by placing a 5 cm thick sole below one of their feet. The results show that the method was able to statistically distinguish asymmetrical gaits. These promising results tend to show that this low-cost system could be used in any clinical setting as it doesn't need complex devices and no specific preparation time.

6.2 Introduction

Despite its inherent complexity, gait is one of the most used motion in our daily life. Bipedal walking is composed with sequences of double supports and stance periods leading to series of forward falling motions. Hence bipedal walking is a challenge from the balance point of view which leads humans to be the only primates who use this locomotion in an exclusive manner. Balance maintenance and gait patterns involve to deal with functional perceptual information (vision, proprioception, plantar pressure and inertial information provided by the vestibular system), a correct integration and interpretation (through the Central Nervous System), and a correct execution (through the musculoskeletal system). If some problems occur at least for one of these processes, because of diseases or injuries, gait pattern can be altered. Hence clinical gait analysis, through quantified measurements, is widely used to diagnose pathologies of these processes.

Among all the parameters used to quantify gait disorders, gait asymmetry is very famous, especially when the pathology have direct consequences on one side of the body (named hemisyndroms). Typically, stroke injures a part of the brain which could lead to

hemiplegia which clearly leads to changes in gait pattern at least for one side of the body. The link between stroke and gait asymmetry has been widely studied (see [9], among many other works). Another example which will be more precisely presented in the remaining of this paper is Leg Length Discrepancy. In that case gait asymmetry is mainly due to the difference that exists between the two leg lengths as shown in [70]. In the same way, the efficiency of joint replacement that occurs after injuries or with elderly people is generally evaluated through gait asymmetry analysis.

However studying gait asymmetry is practically complex. Indeed, gait asymmetry could be viewed in many different body parts and parameters, such as the lateral and vertical oscillations of the pelvis, the knee, hip and ankle angular trajectories, the head and shoulder motion. . . Among all these parameters, the clinician needs to get only relevant asymmetry information, such as a unique clinical parameter. Hence, two main approaches arise. The first approach consists in focusing on one point on the body, such as placing a unique accelerometer on the patient's pelvis to measure changes between two successive steps within a gait cycle [16]. The second approach consists in selecting a wide range of points, such as all the joint centers associated to a proposed biomechanical model. For the latter, the most common system consists in measuring the 3D displacement of external markers and to deduce the 3D angular trajectories at each joint. Statistical data reduction and comparison to reference values are used to obtain a unique value to quantify gait disorders, including asymmetry [41]. However, this approach relies on very expansive systems, is strongly associated with an imposed biomechanical model, and requires high technical skills to perform this type of evaluation. Thus it cannot be used by clinicians in small hospitals.

In this paper we propose an original approach to tackle these problems. The proposed system relies on a treadmill and a low-cost depth camera that don't require high technical skills and that doesn't require to place devices or markers on the patient's body. The system is also free of biomechanical model that could lead to uncertainties, especially for disable people. It is cheap and easy to use in order to be used even in very small doctor's offices.

In the next section 6.3, we provide the reader with a state of the art of systems previously used to measure gait asymmetry. Section 6.4 describes the method proposed in this paper to measure gait asymmetry with a treadmill and a unique depth camera placed in front of the patient. The method is then evaluated thanks to an experiment which methods and protocols are described in section 6.5. The results are given in section 6.6

and discussed in section 6.7.

6.3 Previous works

Different works have already been reported with sophisticated and specialized systems to investigate gait asymmetry. For example an adapted treadmill with force sensors was used to measure asymmetry gait parameters [48]. However ground reaction forces are limited to the foot and no information on other segments are provided by this system. To measure multiple segment information, Sadeghi [145] reviewed works done with electromyography measurements (EMG) and joints kinematic measurement with optoelectronic systems. But all those systems still need to put numerous (wearable) sensors on the subject which is quite difficult to implement in daily clinical practice.

On the contrary, video systems do not have these drawbacks. For example, Green and Guan [69] used 2D sagittal plane images from a video camera to recognize Parkinson patients based on asymmetry information during the gait cycle. Video analysis could also take advantages of 3D reconstruction using multiple cameras system for gait analysis. Mundermann et al. in [118] experimented a system where subject's volume reconstruction was done using multiple video cameras and Visual Hull computation. Unfortunately, this system needs at least 8 cameras to obtain satisfactory results which means high cost and high technical knowledge to maintain it operational (e.g. camera internal and external calibration). Volume information could be obtain from depth measurements too. This depth information is typically obtained with at least two cameras and one can find stereo-cameras for this purpose. However this technology is not always robust, particularly if the scene does not offer a sufficiently textured image. Recently new types of depth camera have appeared on the marketplace. They are based mainly on two different methods to measure depth : structured light [59] and time of flight [67]. The latter is still in development with low resolution and high price which makes it useless for large scale diffusion for the moment. But the first one is already in millions of home [1] and is very cheap, reliable and easy to setup. This is why we propose here a method that uses this kind of depth camera to record the depth information of a walking subject for gait analysis.

Hu et al. [84] used a depth camera placed on a trolley to track the lower limbs of an elderly subject. However the fact that the subject needed to push the trolley that supported the depth camera impairs the regular gait. This is why we propose to use a treadmill to

get a regular and stabilized gait during our tests (after a period of adaptation) and then to record the depth information to measure gait parameters of left/right asymmetry.

In our previous work [139] we used a depth camera to measure asymmetry induced by a heel cup under one foot. The depth camera was placed behind the subject facing his back. Accumulated depth images with centroid registration were computed to locate differences between the right and left sides of the body. However our method did not correct perspective distortions of the depth map that could lead to misalignment of the body parts. In [18], we used three depth cameras to reconstruct the entire body volume during the walk and summed these volumes after centroid registration to obtain the full occupied space for the whole gait cycle. Asymmetries were observed when a heel cup impaired the normal walk. These two studies integrated the data over several gait cycles and it was therefore not possible to distinguish precisely when in the gait cycle that asymmetry was maximal or minimal. In addition, Song et al. [155] showed that the difference of knee moments between the longer and the shorter limbs is not constant during the movement. This is an incentive to quantify the asymmetry all along the gait cycle.

Moreover according to Gurney [70], a difference of lower limb lengths increases the vertical movement of the center of mass because the subject has to vault over the longer leg and step down onto the shorter one. Considering our aim to focus on the legs, registering depth maps with the centroid could introduce erroneous displacements.

In the following, we explain how we take into account all these observations to improve gait assessment with a single depth camera.

6.4 Method

Our method analyzes the depth maps returned by the depth camera placed in front of the subject in 3 distinct steps detailed in the scheme of Fig.6.1 and in this section. The first step described in subsection 6.4.1 aims to identify region of the depth map belonging to the subject and transform it from perspective projection which is camera-dependent to an orthographic view that is more convenient for registration. Then section 6.4.2 proposes to recover the gait double support instants from the knee motion and construct key depth maps describing the body key poses during the gait cycle. The next subsections present how right and left legs are segmented and compared for each instant of the gait cycle.

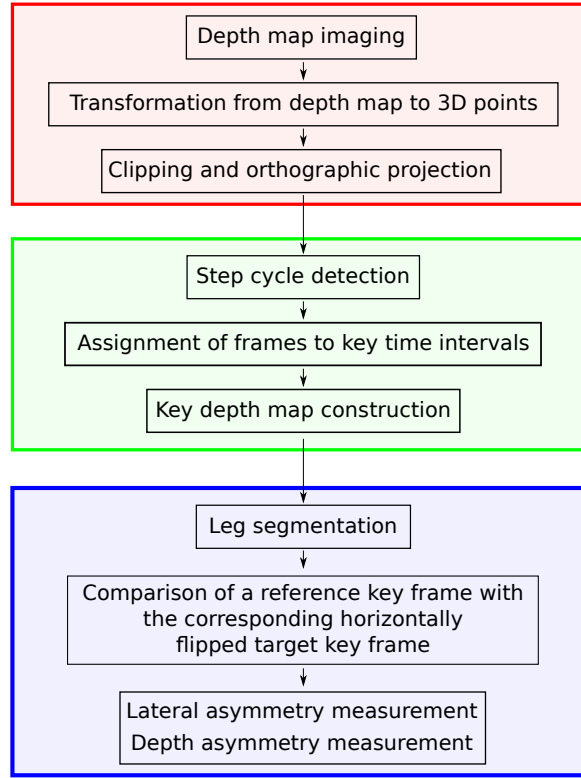


FIGURE 6.1 – Schematic representation of the method. In the red box are the depth map pretreatment operations corresponding to section 6.4.1. In the green box are key depth map construction steps which are detailed in 6.4.2. Finally in the blue box are the different operations needed to construct the asymmetry index explained in the section 6.4.3.

6.4.1 Depth map pretreatment

This first stage of the method recovers 3D surfaces from depth maps returned by the sensor and selects the region of interest corresponding to the subject.

6.4.1.1 Transformation from depth map to 3D points

The depth camera is considered as a pinhole camera model with its optical axis intersecting the center of the image plane and the focal $f = 575.82$ pixels given by the drivers from the manufacturer [3]. Then 3D points (X, Y, Z) are reconstructed with the following equations using the depth information Z returned for each pixel (x, y) :

$$X = \left(x - \frac{w}{2}\right) \frac{Z}{f} \quad Y = \left(y - \frac{h}{2}\right) \frac{Z}{f} \quad Z = Z$$

where w and h are respectively the width and the height of the depth map (in pixels).

6.4.1.2 Clipping and orthographic projection

Once the scene is represented as a cloud of 3D points (X, Y, Z) , the subject is localized in space within a 2.5 meters (height) x 2 meters (depth) x 1 meters (width) bounding box centered on the treadmill belt surface. Only points inside this region are kept for further processing. The result can be seen in Fig. 6.3.

Once the 3D points of the subject are correctly clipped, the depth of this surface is projected orthographically on a plane perpendicular to the optical axis of the depth camera as illustrated in Fig. 6.2. The spatial resolution of this orthogonal image was set to 0.5 cm per pixel in this study. The depth value of each pixel is computed with an inverse distance interpolation as described by Shepard in [150].

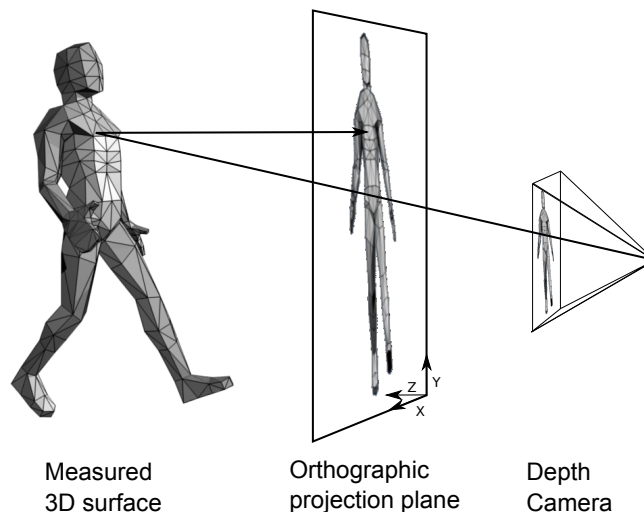


FIGURE 6.2 – Schematic representation of the orthographic back-projection.

6.4.2 Key depth maps

Once each depth map is corrected, the next step is to compute different key depth maps to describe the gait cycle.

6.4.2.1 Step cycle detection

The step (half cycle) starting and ending frames are identified by double support detection with the antero-posterior distance between knees. This distance is equal to the difference between the depth of each knee localized with the anthropometric ratio $0.26 * H$ where H is the mean height of the subjects (see Fig. 6.3). The depth pixels of each knee

are separated into two groups (left and right knees) with a K-means algorithm based on their x and z values. Then the minimum depth of each leg is measured to compute the distance shown in Fig. 6.4. This signal is low-pass filtered at 5Hz (typical cut-off frequency for gait analysis) and the maximum and minimum are detected to get the double support instants.

6.4.2.2 Assignment of frames to key depth map

Each step (half gait cycle) is divided into 15 equal intervals in time. Each step is processed separately since their duration might be different. Then each depth map frame is assigned (temporal registration) to the corresponding nearest interval.

6.4.2.3 Key depth map construction

All frames of all gait cycles belonging to the same time interval are averaged to create a key depth map. Before averaging, some operations are needed to compensate for the lateral and depth movements of the subject on the treadmill. First, a lateral registration (cross-correlation maximization) of the frames is performed in the leg region of interest corresponding to the area below $0.43 * H$ [34] and 10 cm above the treadmill belt (Fig. 6.3). Then global depth movements were cancelled with registration of the mean depth of the lower part of the trunk. This area corresponds to a 20 cm square below the elbow height [34] (Fig. 6.3). These operations generate 30 key depth maps for the gait cycle (15 for each step). The result of these operations is represented in Fig. 6.5 and Fig. 6.6.

6.4.3 Key depth map analysis for asymmetry measurement

The left and right legs are then identified in each key depth map to permit the comparison of a reference key frame with the corresponding target key frame one half stride later in the gait cycle. Horizontal flipping of the target key frame is performed to compare right and left sides based on their natural symmetry.

6.4.3.1 Leg segmentation

The legs are identified on each horizontal line separately. For each line, only pixels belonging to the body for at least 80% of the frames are considered. Again the left and right legs are separated thanks to a K-means algorithm in the same way as for double support detection.

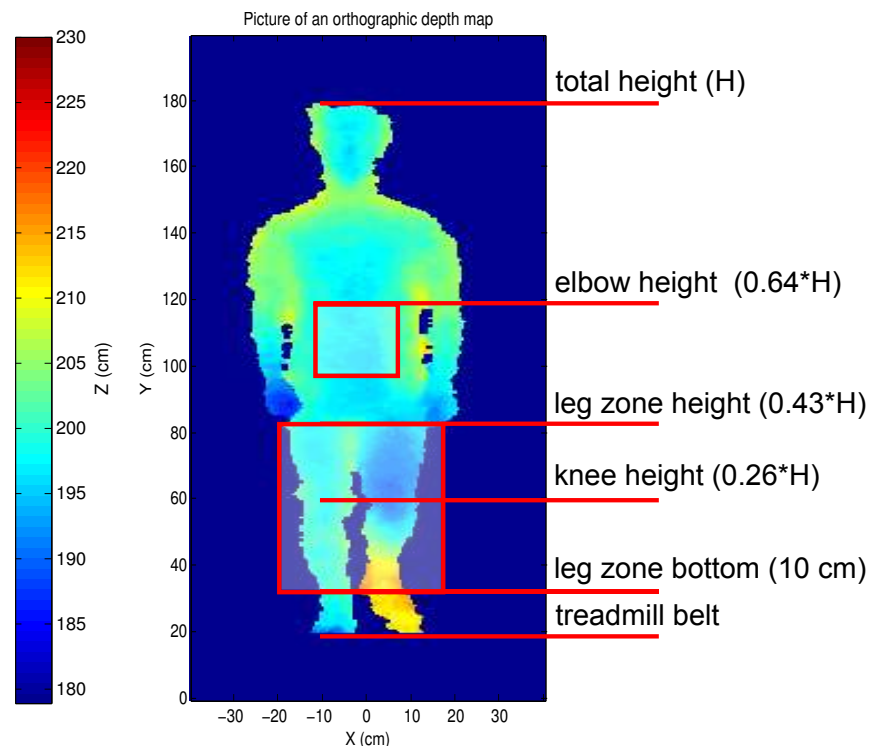


FIGURE 6.3 – Orthographic depth map with the anthropomorphic ratios and leg region of interest.

6.4.3.2 Asymmetry measurement

The leg motion asymmetry is evaluated thanks to the following procedure where each legs of the reference key frame will be compared with the corresponding ones in the target key frame as shown on Fig.6.7. The result of those comparisons is the lateral and depth asymmetry index.

First, for each pair (reference and target) of key frames, the lateral shift between legs is computed line by line using a cross-correlation. This permits to register laterally the left (δ_l) (right (δ_r)) leg of the reference key image with the horizontally flipped right (left) target key image.

Once legs are laterally registered, the depth difference can now be computed without perturbation from the lateral asymmetry. The mean depth difference between the reference and target key depth map is then computed line by line for the left (Δ_l) and right (Δ_r) legs respectively. The final lateral (δ) and depth (Δ) asymmetry indexes are obtained by averaging the left and the right contributions on all the leg zone. The meaning of these indexes are respectively the differences between the lower limb poses in space in

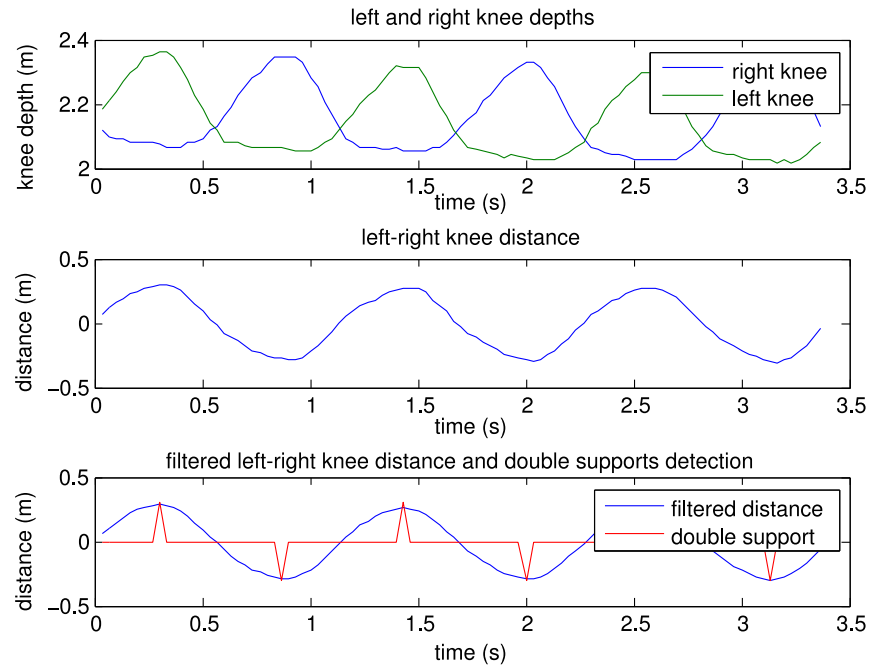


FIGURE 6.4 – Sequence of signals for the double support detection process. On the top the original signal of left/right knee distance. In the middle, the filtered version of this signal. At the bottom, the results of local minimum and maximum detection.

the antero-posterior (depth) direction (Δ) and lateral direction (δ) at different instants in the gait cycle.

6.5 Experimentation and validation

6.5.1 Experimentation

6.5.1.1 Description of the population

An experiment was conducted with the approval of the local ethical committee. 21 subjects accepted to walk on a treadmill (LifeFitness F3) at their comfort speed. The population was composed of young and healthy people with no apparent asymmetry for 20 of them. One of the subject was dismissed due to a visible gait asymmetry during normal gait trials. The subjects' mean height was 177 cm (± 8 cm) and mean weight was 75.11 kg (± 13.8 kg).

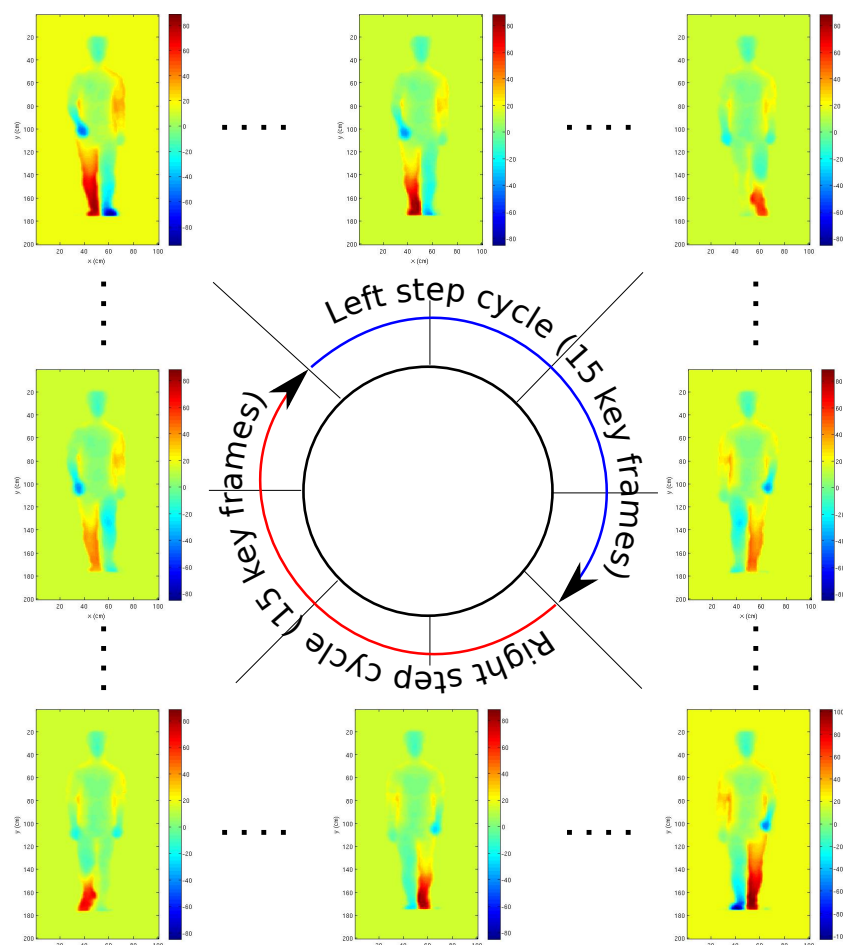


FIGURE 6.5 – Representation of the gait cycle with the key depth maps.

6.5.1.2 Description of the protocol

For each subject, 3 depth map recordings were done, the first for a normal gait, the second with a 5 cm thick sole under the left foot and then the third with the sole under the right foot. Each session began with an adaptation period before image acquisition. The recording period began once the subject felt comfortable on the treadmill. Each recording period was about 5 minutes. The analysis was done on the last 180 gait cycles. The choice of a 5 cm sole was dictated by the work of Gurney [70] who demonstrated that a minimum discrepancy of 49 mm was necessary to generate externally visible effects. This can be classified as a moderate leg length discrepancy (LLD) according to Reid and Smith [135].

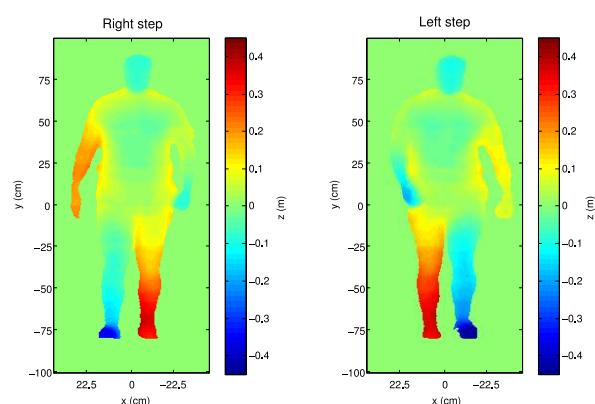


FIGURE 6.6 – Key depth map for the right and left step at the heel strike instant.

6.5.1.3 Setup equipment

The depth images were recorded with a depth camera Kinect at 30 frames per second with a resolution 640 per 480 pixels. It was placed in front of the subject and rotated 90 degrees around the optical axis in order to make the width of the image correspond to the height of the subject, as shown on Fig.6.8. The optical axis of the camera was parallel to the translation vector of the treadmill belt.

6.5.2 Validation

The validation of the asymmetry indexes consisted in two statistical studies. The first was for estimating the ability of this index to detect gait deformation within subjects. For that, a paired Student t test was applied on with-sole gait trials compared with the corresponding normal gait trials subject by subject. The second statistical test was done for assessing the classification ability of these indexes when unpaired populations were considered. Some precautions are necessary for this stage. The 20 subjects must be separated in independent subgroups because of the paired nature of normal and simulated pathological subjects. Therefore the 20 subjects were assigned to 2 independent groups of 10 subjects each. Then all group combinations were tested with an unpaired student t test.

6.6 Results

In this section, we present results regarding the symmetry of a normal and an impaired walk (with a 5 cm sole) during the gait cycle with lateral (Fig.6.10) and depth

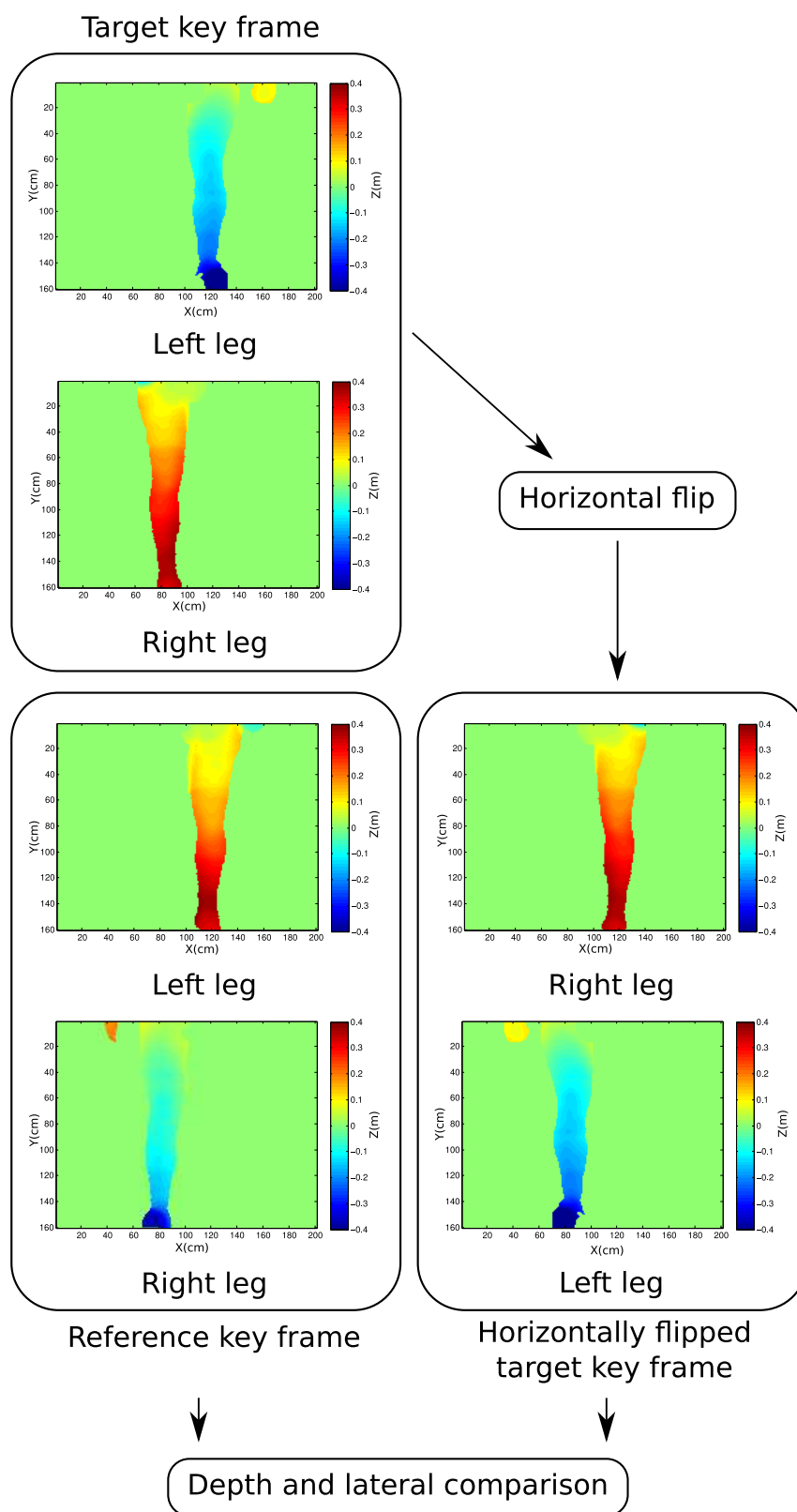


FIGURE 6.7 – Representation of the operations done on the reference and target key frames for lateral and depth asymmetry evaluation.



FIGURE 6.8 – Picture of the experimental setup where the depth camera (1), the recording computer (2) and the treadmill (3) are visible.

(Fig.6.9) asymmetry indexes. It was clearly visible that the depth asymmetry index was impacted by the deformation introduced by the sole. However there was no significant perturbation of the lateral symmetry.

The statistical difference for the paired t test were highly significant for all instants of the gait cycle ($p < 0.01$). This means that when compared subject by subject, the asymmetry index was very sensitive in this case. For unpaired inter-group statistical analysis, all groups had a p -value $p < 0.05$ for the first 20% and the last 10% of the step cycle for both right and left deformations. The worst case was in the middle of the gait cycle where 77% of group assignments were significantly different ($p < 0.05$).

6.7 Discussion

Our results demonstrate that the method is able to compute a relevant asymmetry index for the lower limbs.

The lateral asymmetry measurement was not especially relevant for leg discrepancy as shown in Fig.6.10. However some others pathological diseases like hemiplegia [81] could involve more lateral asymmetry.

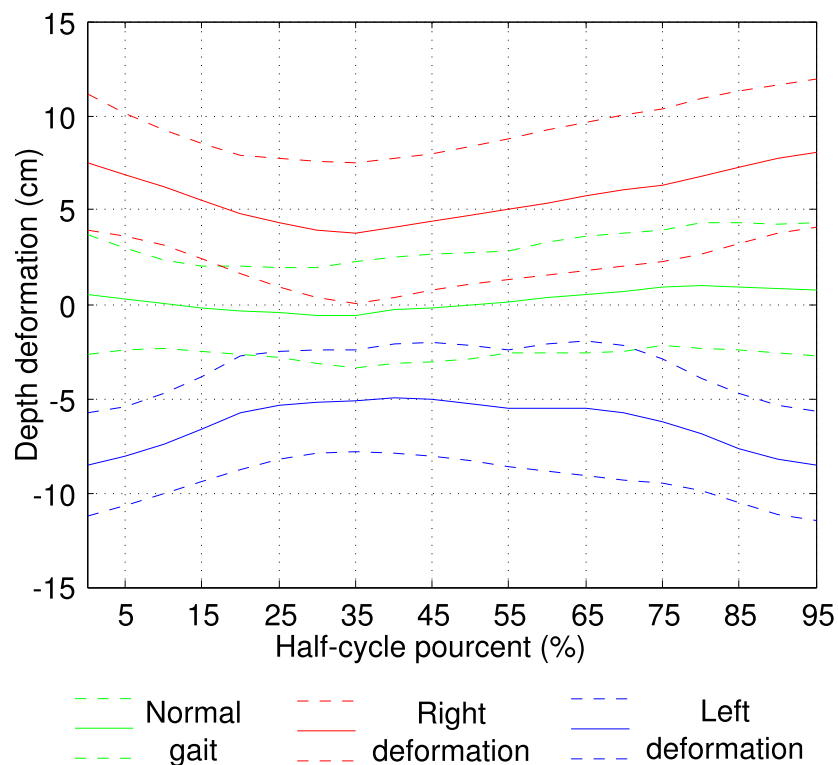


FIGURE 6.9 – Depth asymmetry index (Δ) for normal, right and left deformed legs during the step. The dashed lines define one standard deviation intervals.

Another major contribution is asymmetry quantification *within* the step cycle. The variation with time of the mean asymmetry for each population is presented in Fig.6.9 and shows in particular that the three populations were better separated at the beginning and the end of the cycle.

However, our method has three parameters to set : (1) the minimum number of frames (80% in this paper) for leg segmentation, (2) the number of key frames (15 per step), (3) the number of strides needed for averaging the asymmetry index for one step (180 in this study). We conducted a sensitivity study of these parameters with values ranging from 50% to 90%, 5 to 20 key frames per step, and 60 to 180 strides without noticing any significant changes. This demonstrates the robustness of our approach.

The computation efficiency could be an important factor for clinical practice. In this study, the mean time to extract the depth maps, compute the key frames and generate the asymmetry index was less than 2 minutes on an Intel Quad Core CPU 2.0 Ghz.

In the future we plan to test different leg discrepancies with different sole thicknesses

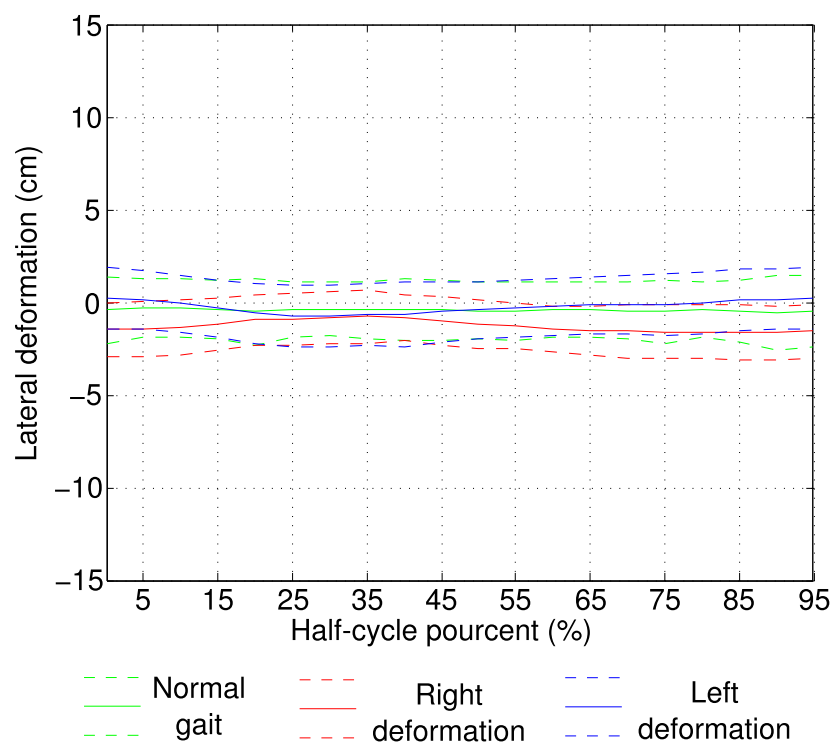


FIGURE 6.10 – Lateral asymmetry index (δ) for normal, right and left deformed legs during the step. The dashed lines define one standard deviation intervals.

and perform a regression analysis of the results to see if the method could infer the severity of the asymmetry.

There are still some open questions to solve with such methodology since there is no biomechanical model involved. For instance we do not know how a subject was adapting his walk (e.g. with hip motion) to the sole perturbation. Therefore, for the moment, our method could be used as a screening test prior to more sophisticated tests involving, for instance classical biomechanical model analysis with opto-electronics systems in specialized gait analysis lab in order to diagnose the cause of the detected gait asymmetry.

6.8 Conclusion

The results obtained with the proposed experimentation have shown that our method is reliable. A 5 cm leg length discrepancy introduced a clearly visible depth difference that can be measured with a depth camera facing a subject walking on a treadmill. This method might be usable for other pathology as well where asymmetry is a symptom. In addition to screening, this method could permit to make a follow-up of patients after a surgery (e.g. for joint replacement) and to measure the recovery after a stroke etc. Due to the low-cost of the components needed and its fast experimental protocol, this kind of measurement equipment could be promising for a wide distribution in clinical settings.

Acknowledgement

The authors wish to thank Caroline Rougier for her help with the experimentation, the FQRNT (Québec) and the ANRT for the CIFRE 146/2009 funding (France).

Chapitre 7

Discussion sur la mesure des paramètres de la marche

Sommaire

7.1	Mesure d'un index de la marche sans modèle biomécanique	137
7.2	Problème dû au capteur	139
7.2.1	Erreur due au codage de la profondeur	139
7.2.2	Erreur due au principe de mesure de la profondeur	141
7.3	Amélioration du protocole expérimental	142
7.4	En route vers la reconstruction 3D	142
7.5	Conclusions	144

L'article précédemment exposé montre l'intérêt de l'utilisation d'une caméra de profondeur pour la mesure de l'asymétrie de la marche sur un tapis roulant. Les résultats présentés prouvent que la méthode est capable de discerner à l'aide d'un index de symétrie les passages de sujet avec une asymétrie simulée. Quelques points méritent toutefois d'être discutés plus attentivement comme l'absence de modèle biomécanique dans la méthode, l'amélioration du protocole expérimental et les problèmes dus à la caméra de profondeur. Finalement, une solution intégrant plusieurs caméras de profondeur sera abordée.

7.1 Mesure d'un index de la marche sans modèle biomécanique

Cette méthode de mesure d'un index de symétrie du mouvement des membres inférieurs est plus directe. L'une des raisons de ce fait tient de l'absence d'analyse des données à l'aide d'un modèle biomécanique. En effet, comme introduit dans la section 5.3, des méthodes utilisent un modèle biomécanique pour analyser l'information de base issue de la mesure. Cette information peut prendre différentes formes comme un jeu de marqueurs ou de capteurs fixés à des endroits stratégiques du corps, ou bien des pixels issus d'une image représentant la couleur ou la profondeur. Dans le premier cas, le mouvement est sous-échantillonné, mais chaque point a une signification biomécanique très riche (localisation anatomique connue et identification constante dans le temps), dans l'autre cas, un grand nombre de points sont échantillonnés, mais au détriment de leur signification biomécanique (localisation imprécise sur le corps, aucun lien temporel entre un point particulier d'une image à l'autre).

L'utilisation d'un modèle biomécanique permet d'extraire l'information nécessaire à partir de capteurs dynamiques [110, 68], de positions de marqueurs [32] ou encore à partir d'images de couleur [114, 115, 8, 62] ou de profondeur [151, 88]. Ceci est fait à l'aide de contraintes mécaniques provenant de la modélisation de la chaîne cinématique du corps humain [168] et des informations de longueur des segments propres au sujet. Ce groupe de contraintes permet d'obtenir les meilleurs résultats de mesure du mouvement possible à ce jour et introduit une résistance aux bruits pouvant perturber les mesures (par exemple les bruits électroniques d'amplification pour les accéléromètres et gyroscopes, l'occultation de marqueur ou d'une partie de l'image du sujet). Seulement, des erreurs dans la mesure de ces paramètres de longueur de segment peuvent impacter, dans certaines circonstances, la qualité des résultats comme l'indique [42]. Dans le cadre d'un

laboratoire spécialisé, cette mesure peut se faire manuellement par un expert ou encore une opération dite de “Range of Motion”, permet de calibrer le modèle avec les données de l’appareil et une série de mouvements prédéterminée. Dans ce cadre d’environnement spécialisé, le temps et la compétence sont disponibles, ce qui n’est pas forcément le cas des consultations cliniques de médecine générale.

Dans un usage de clinique routinière, il est plus difficile de consacrer du temps et du personnel qualifié à ce genre de mesures préparatoires. C’est en cela que l’absence d’utilisation d’un modèle biomécanique dans la méthode proposée est intéressante. Elle simplifie le protocole expérimental en supprimant la période préparatoire de calibrage. Mais surtout, elle simplifie la méthode de calcul en supprimant l’étape d’analyse des données de base à travers ce modèle de contrainte. La sélection des informations pertinentes est alors réalisée en utilisant quelques informations a priori comme les rapports anthropométriques et les caractéristiques techniques du matériel utilisé (dimensions du tapis roulant). Une information sur la symétrie du mouvement des jambes est alors calculée directement depuis les cartes de profondeur.

Il est très important de noter que l’information retournée par cette méthode sans modèle biomécanique est moins riche en information que ce que pourrait donner une méthode classique basée sur l’utilisation d’un modèle biomécanique. En effet, dans le cas où une asymétrie est constatée, il n’est pour l’instant pas possible de pouvoir expliquer la cause de cette déformation. Avec un modèle biomécanique, il est possible d’isoler le mouvement de chaque articulation et ainsi de déduire l’endroit de la chaîne dont le mouvement est le plus affecté par l’asymétrie. Même si dans certains cas, la redondance et le couplage entre les différents angles introduisent une difficulté pour isoler dans ces grands nombres d’informations disponibles, celles qui sont pertinentes. Dans notre étude, il n’est pour l’instant pas possible d’estimer à l’aide de notre méthode quel genre de stratégie d’adaptation à la déformation a été adoptée par les sujets. En effet, dans la littérature, plusieurs stratégies sont mentionnées. Par exemple, Song dans [155] parle de marche sur les doigts de pied, le mouvement de circonduction ou encore la flexion constante de la jambe la plus longue. Pour l’instant, ce travail préliminaire montre qu’il est possible de mesurer un indice d’asymétrie du mouvement des jambes sans modèle biomécanique. Des travaux supplémentaires suivant cette démarche pourront évaluer la possibilité de mesurer d’autres paramètres telle que ces stratégies d’adaptation.

De plus, cette méthode dispose d’informations qui sont bruitées, et ceci pour les raisons qui sont présentées dans la section suivante.

7.2 Problème dû au capteur

La caméra de profondeur que nous avons utilisée est une caméra Kinect développée par Primesense et diffusée par Microsoft et Asus. Cette caméra peu coûteuse permet de rapidement disposer d'une carte de profondeur d'une scène directement calculée par un processeur dans la caméra. Seulement il existe différents problèmes dus, premièrement au codage de l'information, et deuxièmement à la nature de la mesure par lumière structurée ponctuelle.

7.2.1 Erreur due au codage de la profondeur

L'information retournée par la Kinect vers l'ordinateur est une carte de disparité. Dans un premier temps, la seule information disponible était une approximation que nous avons réalisée à l'aide d'un rail pouvant se déplacer sur commande avec une précision millimétrique. Nous avons obtenu l'équation suivante :

$$d = \frac{1}{-0.0032936 D + 3.5463}$$

où d est la profondeur en mètre et D la disparité retournée par la librairie Libfreenect.

Le fait que la disparité transmise sous forme entière soit au dénominateur laisse supposer que les plans ne seront pas espacés de la même distance les uns des autres. Dans un deuxième temps, nous sommes passés à la librairie OpenNI qui retourne une valeur métrique directement. Ne connaissant plus comment la profondeur en métrique est obtenue à partir de la disparité, une expérience consistant à mesurer une scène avec un objet en mouvement a été réalisée pour estimer le comportement de la mesure de profondeur. En traçant l'histogramme de la valeur prise par tous les pixels de toutes images de la vidéo, les profondeurs possibles apparaissent comme montrées par la figure 7.1. Le mouvement d'un plan de dimension de 1.2 par 0.80 m et évoluant de 1 à 6 mètres du capteur assure que toutes les valeurs de profondeur soient bien disponibles dans la scène filmée. La distance entre les plans donne une idée de l'erreur d'échantillonnage due au codage de transmission et à la méthode d'identification des points du motif projeté. En effet, un point étant à mi-chemin entre deux plans de mesure aura alors une erreur de mesure égale à la demi-distance entre les plans. La variation de cette distance entre les plans, tracée figure 7.2, allant du millimètre à 70 cm du capteur jusqu'à 25 cm à plus de 9 m. Un agrandissement de la zone allant de 1.7 m à 2.6 m du capteur, permet d'évaluer la distance entre les valeurs d'échantillonnage de la profondeur en général et dans la zone d'utilisation durant nos expériences en particulier. Comme la distance entre chaque plan

est inférieure à 2 cm dans cette zone de 1.7 à 2.6, on peut donc en déduire que l'erreur de mesure à cause de ce phénomène de communication est inférieure à 1 cm.

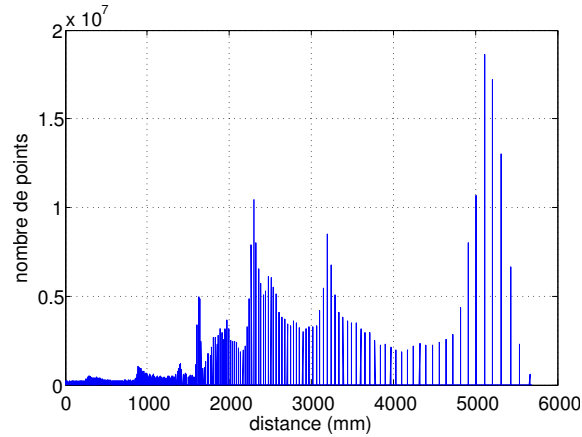


FIGURE 7.1 – Tracé de l'histogramme des profondeurs mesurables par la Kinect.

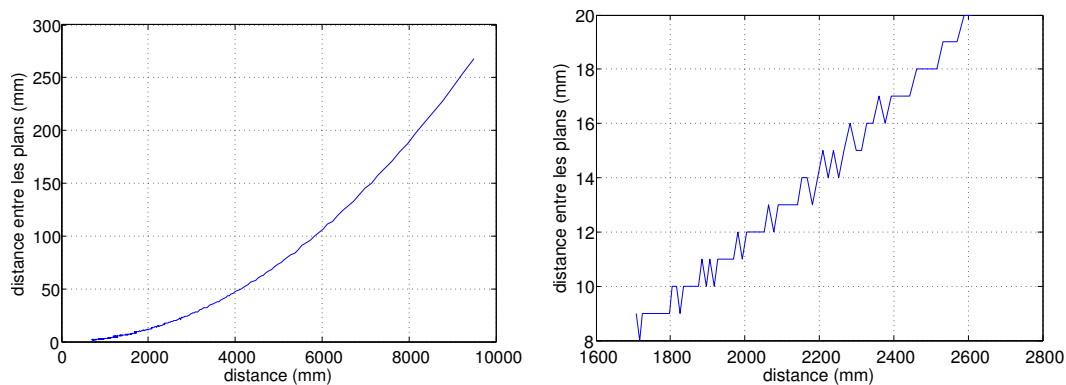


FIGURE 7.2 – Tracé de la distance entre les plans en fonction de la distance avec la Kinect de 0 à 9.5 m à gauche et sur la plage de profondeur utilisée lors de nos expériences à droite.

L'erreur du capteur que nous avons utilisé durant nos acquisitions est de l'ordre du centimètre dans la zone utilisée. De plus, l'erreur précédemment décrite peut être considérée comme une erreur de quantification avec comme erreur maximale de quantification q de cet ordre. Or lors du calcul d'une valeur moyenne avec n mesures, l'erreur résultante autour de la moyenne est égale à $\frac{q}{\sqrt{n}}$. Le cas le plus difficile ici est alors l'utilisation de seulement 60 cycles de marche. Dans ce cas, l'erreur est de 0.13 centimètre ($\frac{1}{\sqrt{60}}$). Cette erreur est inférieure à l'ordre de grandeur du phénomène étudié.

7.2.2 Erreur due au principe de mesure de la profondeur

La mesure de la profondeur se fait de manière parcellaire en estimant le décalage de chaque point par rapport à sa position de référence. Entre ces points de mesure, le programme embarqué de la Kinect interpole les profondeurs pour représenter une image complète. La reconnaissance de chaque point se fait par l'identification de son voisinage. Or lorsque la forme du voisinage est trop perturbée par un brusque changement de profondeur, il arrive que le point ne soit pas reconnu et empêche donc la mesure de la profondeur correcte pour la zone forçant l'appareil à interpoler. De plus, comme tout système apparenté à la stéréoscopie (ici le projecteur est considéré comme la deuxième caméra), il existe des zones où la profondeur n'est pas calculée du fait que la caméra observe une zone non accessible par le projecteur. Ces artefacts sont tout à fait visibles sur la figure 7.3.

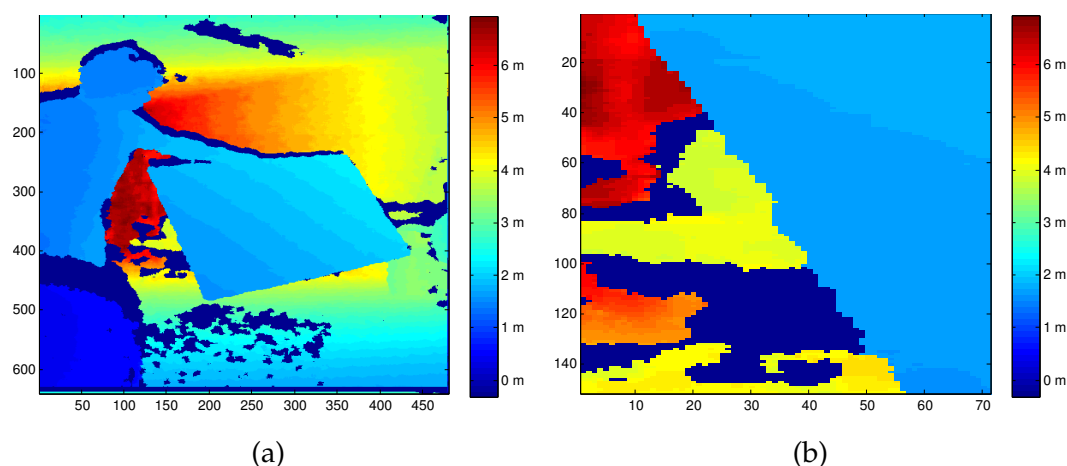


FIGURE 7.3 – Illustration du problème de mesure de profondeur sur le bord d'un plan. À gauche (a) l'image montre un plan incliné de 1.20 par 0.80 mètres où l'on peut voir la zone d'ombre au-dessus de chaque objet. Et à droite (b) une vue rapprochée d'un des bords montre qu'il n'est pas rectiligne.

La méthode que nous avons proposée a tout de même été capable de différencier la démarche de personne ayant été perturbée avec une semelle de 5 centimètres. Il serait juste de se poser la question sur le résultat qui pourrait être obtenue avec des semelles de hauteurs différentes.

7.3 Amélioration du protocole expérimental

Les résultats de l'index de symétrie proposé ont été validés avec l'utilisation d'une semelle de 5 cm de hauteur. Il serait toutefois intéressant de compléter cette étude par une nouvelle expérimentation faite cette fois avec plusieurs semelles de hauteurs différentes à chaque passage. Cela permettrait d'évaluer la courbe de réponse de l'index en fonction de la hauteur de déformation.

La dernière étape serait de valider la méthode en condition clinique en recrutant une cohorte de patients. Cette expérience pourrait être réalisée conjointement avec une mesure opto-électronique pour permettre d'évaluer la possibilité de déduire la stratégie d'adaptation employée par le patient par rapport à la déformation.

Il serait très intéressant de proposer le même genre d'étude pour quantifier l'asymétrie du mouvement de la partie supérieure du corps. Ceci donnerait alors une information sur l'asymétrie générale.

Dans cette méthode, une seule caméra de profondeur a été utilisée pour enregistrer la surface avant de la personne. Il se pourrait que la fusion des surfaces issues de plusieurs caméras de profondeur placées toutes autour du sujet puissent apporter plus d'informations.

7.4 En route vers la reconstruction 3D

L'utilisation d'une Kinect permet de reconstruire la surface de la personne visible par la caméra de profondeur. Pour permettre de rejoindre l'hypothèse d'échantillonnage du maximum de points, il serait intéressant de disposer de la totalité de la surface du sujet. Pour cela, nous avons exploré cette possibilité en plaçant plusieurs caméras de profondeur Kinect autour du sujet. Le problème de précision de mesure des bords comme présenté dans la partie précédente empêchant l'utilisation des méthodes classiques de calibrage. Un premier travail fut réalisé pour développer une nouvelle méthode de calibrage externe de caméra de profondeur à l'aide de plans. Cette méthode calcule un nuage de points virtuels à partir d'intersections de plans réels mesurés à différents instants par chaque caméra. Comme les caméras sont synchronisées, les nuages de points de chaque caméra correspondent et permettent alors retrouver la relation géométrique entre chaque caméra. Dans la publication Auvinet et al. [17], le mode opératoire du calibrage multi-caméras de profondeur est défini plus en détail. Cet article présente aussi une adaptation de la méthode de reconstruction du volume par visual hull aux données provenant de

caméras de profondeur. Cette méthode utilise, en plus de l'information des silhouettes comme dans une situation avec des caméras classiques, l'information de profondeur en définissant la zone entre la caméra et la surface mesurée comme inoccupée. Une illustration de la méthode est proposée sur la figure 7.4. De plus, une étude réalisée dans cet article a permis de comparer les résultats de reconstruction obtenus en simulant le mouvement de la marche avec un mannequin numérique. Les résultats ont montré que seulement 3 caméras de profondeur permettent d'atteindre la précision d'au moins 20 caméras conventionnelles.

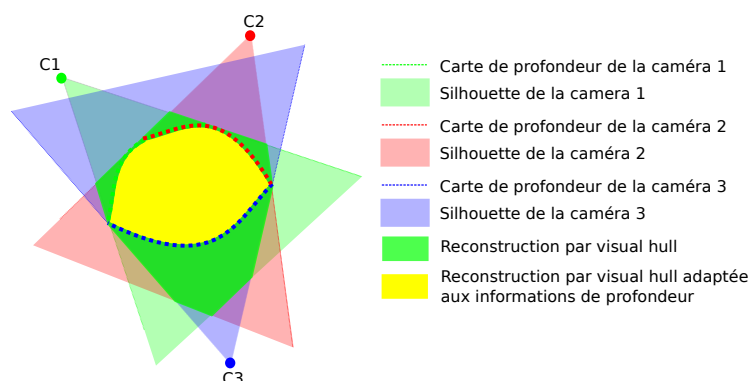


FIGURE 7.4 – Illustration de la méthode de reconstruction du volume par visual hull adaptée aux caméras de profondeur.

Une fois le calibrage externe possible, nous avons réalisé une expérience de mesure d'asymétrie par reconstruction de l'enveloppe externe avec 3 caméras de profondeur. En plaçant une talonnette d'une hauteur de 2.5 cm sous le pied droit puis sous le pied gauche, le mouvement de la marche ainsi déformé a pu être comparé au mouvement normal. La méthode calcule le volume total occupé dans le temps en sommant les volumes reconstruits à chaque instant. Ce volume total illustré figure 7.5 a été analysé pour mesurer la symétrie de la marche. L'asymétrie introduite artificiellement a pu être détectée grâce à l'allongement du pas du côté où la talonnette était placée tel que cela a été publié dans Auvinet et al. [18].

Seulement, les reconstructions tridimensionnelles effectuées avec cette méthode souffrent des interférences entre les motifs projetés au niveau des bords du corps du sujet. Cela provoque une difficulté supplémentaire au problème déjà soulevé sur la méthode de reconstruction de la profondeur à partir de l'identification de point tel que montré précédemment par la figure 7.3. L'impact de ces deux problèmes couplés à l'imprécision de mesure donne les erreurs de reconstruction que l'on peut remarquer sur la figure 7.6.

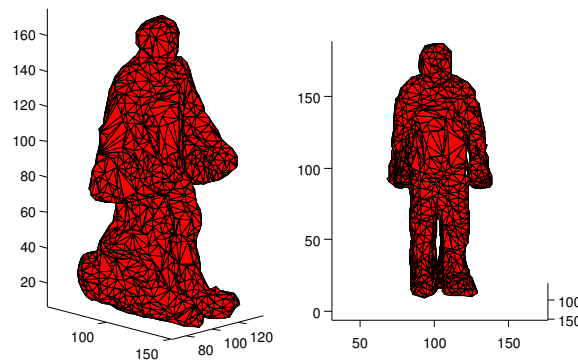


FIGURE 7.5 – Illustration du volume total occupé pendant plusieurs cycles de marche reconstruit par un système à trois caméras de profondeur. (les axes sont en centimètre)

Cette coupe horizontale du volume où sont reprojctées les silhouettes à partir des surfaces montre que les zones occupées par les jambes (en rouge) sont très irrégulières.

Différentes voies d'amélioration pourraient exploiter la conservation de paramètres constants dans le temps comme le volume ou la forme générale de chaque membre. Il peut aussi être envisagé d'utiliser une autre méthode de reconstruction de carte de profondeur comme par exemple la stéréovision active.

Cette voie pourrait être explorée conjointement à un système opto-électronique, car cela pourrait apporter une information supplémentaire telle que le volume de chaque segment. La connaissance de celui-ci permettrait de mieux estimer le moment d'inertie de chaque membre et rendre ainsi le calcul des forces internes de manière plus précise.

7.5 Conclusions

Cette partie a présenté une nouvelle méthode de mesure de la symétrie de la marche basée sur l'utilisation d'une caméra de profondeur et un tapis roulant. Le calcul de cet index de symétrie du mouvement des jambes se fait en comparant les surfaces de chaque jambe à la jambe respective du pas opposé. La simplicité de cette méthode passe par l'absence de nécessité d'utiliser un modèle articulé ou de marqueur, ce qui permet de se dédouaner d'une étape de calibrage de longueur de segment et de préparation du sujet. C'est l'une des raisons pour laquelle le protocole de mesure répond tout à fait aux exigences de l'exercice clinique. De plus, la totalité du matériel utilisé coûte à ce jour moins de 5000\$, ce qui peut faciliter l'équipement de petite structure médicale. De plus amples expérimentations seront nécessaires pour finaliser la validation de cet index de

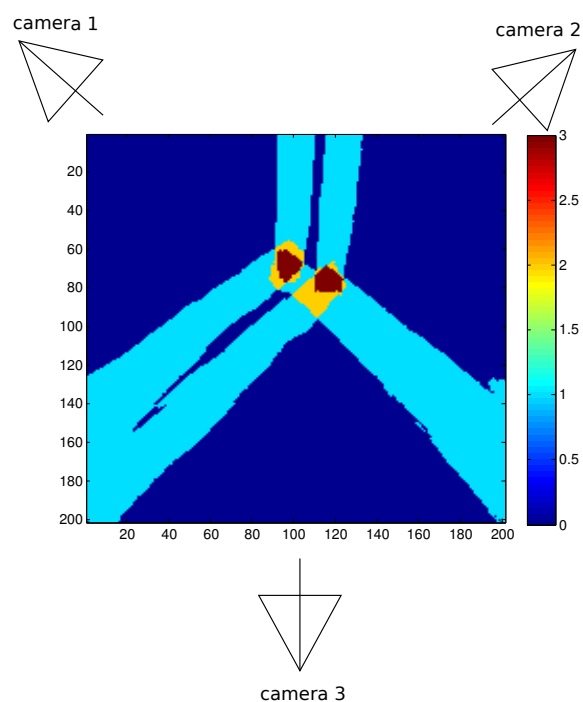


FIGURE 7.6 – Coupe horizontale au niveau des jambes de la reprojection des surfaces issues des cartes de profondeur pour reconstruire le volume du sujet. Les jambes sont en rouge sur cette image (la couleur spécifie le nombre de projections superposées).

symétrie pour permettre de passer à la validation clinique avec des patients. De futurs travaux vont être envisagés pour définir la faisabilité de prendre en compte la totalité du corps et de quantifier la variabilité du mouvement de la marche.

Chapitre 8

Discussion et Conclusion

Sommaire

8.1	Contribution pour la communauté scientifique	149
8.2	Contribution pour la communauté médicale	150
8.3	Contribution pour la société	151

Lors de cette thèse, le domaine de l'analyse de la marche a été abordé selon deux axes de travail différents : la détection de la chute et l'analyse de la marche à proprement parler.

Dans la première partie traitant de la chute, l'objectif réalisé est la mise au point d'une méthode de détection de la chute capable de résister aux occultations. Le problème d'occultation étant rarement au centre de la problématique, il nous semblait important d'y remédier. Pour cela, une méthode analysant le volume reconstruit grâce à un système multi-caméras, nous a permis de valider la possibilité d'obtenir des résultats intéressants même en cas d'occultations.

Dans la seconde partie, une nouvelle méthode de mesure de l'asymétrie de la marche a été proposée. Celle-ci s'organise autour de l'information de profondeur fournie par une caméra de profondeur placée au-devant d'une personne marchant sur un tapis roulant. L'analyse des différences spatiales entre les jambes au même moment dans le cycle du demi-pas permet de calculer un index d'asymétrie. Les résultats de cet index appliqué sur la marche normale et les marches déformées ont permis de montrer qu'il était possible de détecter la perturbation d'une semelle de 5 centimètres placée sous un pied.

Ces deux preuves de concept validées expérimentalement sont les contributions majeures réalisées durant ce travail de thèse.

Il peut être intéressant de regarder sous différentes perspectives ces différents apports pour mieux cerner l'intérêt qu'ils peuvent représenter pour les différentes communautés concernées selon les sujets abordés. En effet, les communautés scientifiques, médicales et toute la société sont concernées par les différentes problématiques présentées.

8.1 Contribution pour la communauté scientifique

Les deux projets de recherche réalisés ont comme point commun l'analyse d'une information tridimensionnelle pour venir détecter des anomalies dans le mouvement de la marche. Ces anomalies sur lesquelles nous avons porté notre attention sont la chute et l'asymétrie de la marche. Elles ont été détectées à l'aide d'un indicateur unique, le ratio de la répartition verticale du volume $VVDR$ pour la chute et l'index d'asymétrie pour la marche. Ces deux indicateurs ont été calculés à partir d'une information tridimensionnelle, le volume du corps pour la chute et la surface avant pour l'asymétrie. Ceci a été réalisé en n'utilisant que des rapports anatomiques généraux comme la largeur des épaules pour la chute et des ratios de hauteurs anatomiques pour la marche. Ceci a per-

mis de répondre à l’une des problématiques de la thèse qui était d’étudier la possibilité de réaliser ces deux tâches sans utiliser de modèle biomécanique.

La question sur la pertinence de mesurer une information sur le mouvement de la marche sans utiliser un modèle biomécanique peut se poser. Nous suggérons d’aborder la question en fonction de l’objectif escompté. Dans le but d’indiquer la présence d’un problème, ce genre de démarche peut se montrer intéressant. Dans le travail réalisé sur la détection de la chute, seul le fait que la personne soit au sol est identifié, et non la raison pour laquelle la personne est tombée (perte de connaissance, accrochage du pied dans un objet, ...). Dans le cadre de l’étude de la marche, l’information calculée permet de détecter la présence de l’asymétrie, mais ne permet pas pour l’instant d’en identifier la cause. Pour identifier la cause de ces problèmes, une analyse de la marche classique avec un modèle biomécanique reste aujourd’hui nécessaire.

D’un autre côté, plusieurs contributions plus techniques ont été apportées. En effet, dans le souci de proposer une méthodologie utilisable en condition réelle, différents choix techniques et d’organisation des calculs ont été réalisés pour permettre de respecter les contraintes liées à chaque problématique. Ceci est particulièrement critique dans le cadre de la chute, où la contrainte majeure est de pouvoir exécuter les calculs en temps réel. Ceci a été réalisé en organisant les calculs de manière à les exécuter sur les unités les plus appropriées. La condition “temps réel” est presque atteinte grâce à une parallélisation massive du traitement des données d’image sur GPU, laissant au CPU les tâches de gestion de communication et de synthèse d’information sur les volumes.

Une dernière contribution pour le domaine de la chute est le partage de nos données expérimentales. En effet, la mise à disposition de nos vidéos de chutes et des paramètres de calibrage à la communauté permet aux autres équipes de disposer de données qui pourraient leur être impossible de réaliser, mais aussi d’évaluer leur méthode par rapport aux autres sur une base de données identique.

8.2 Contribution pour la communauté médicale

Pour la communauté médicale, les apports des travaux réalisés sont en fonction des problématiques étudiées.

Pour la chute, le fait d’augmenter la robustesse de détection des chutes à domicile permet d’améliorer la sécurité des personnes âgées susceptibles de chuter chez elles. De plus, une information précise sur le nombre de chutes peut alors être évaluée. Celle-ci

peut être importante dans l'évaluation du diagnostic, et parfois peut être mal évaluée par le patient du fait du déni d'importance de la chute, du déni de la pathologie, ou encore par simple oubli à cause de la maladie.

Pour l'étude de la marche, l'intérêt majeur du travail réalisé est d'apporter un outil de mesure d'un indicateur de la symétrie du mouvement des membres inférieurs, facile à mettre en œuvre lors d'une consultation clinique et ne nécessitant que peu d'investissement et de formation. L'intérêt est donc multiple. Pour le praticien, il pourrait disposer d'une information sur la symétrie du mouvement de la marche et intégrer facilement cela dans l'exercice de sa clinique journalière. Il pourrait utiliser cette information pour du dépistage ou pour le suivi de l'évolution d'une pathologie après un traitement. Il est intéressant de souligner l'intérêt pour la recherche clinique. En effet, un protocole plus simple et un équipement moins coûteux permettent d'investiguer un phénomène plus facilement sur une plus grande cohorte. Pour l'organisation du système de santé, l'utilisation en tant qu'outil de dépistage permet de ne pas focaliser l'activité des services avec des équipements de pointe pour l'étude de la marche pour la détermination de la pathologie.

Avec le vieillissement, des troubles de l'équilibre et de la démarche peuvent apparaître et augmenter le risque de chute. Par exemple, Hausdorff [74] a montré que la variabilité de la durée du demi-pas pouvait être un facteur de risque de la chute. Il est possible d'imaginer qu'un appareil utilisant la méthode présentée puisse être utilisé en routine clinique comme source d'information sur les paramètres de la marche. De plus, d'autres paramètres comme la variabilité du mouvement de la marche pourraient alors être quantifiables avec ce genre de système. Ceci pourra favoriser le dépistage précoce de troubles susceptibles de provoquer un risque de chute et améliorer la prise en charge du patient. De plus, d'après Alexander et Hausdorff [10], il existerait des liens de plus en plus étroits entre la marche, la chute et les fonctions cognitives. Ceci pourrait potentiellement aider le praticien à dépister les troubles de la fonction cognitive.

8.3 Contribution pour la société

Le bénéfice pour la société pour ce genre de travaux est multiple. Dans le domaine de la détection de la chute, la possibilité de pouvoir être secouru en cas de chute permet aux personnes âgées susceptibles de chuter de conserver une vie autonome à la maison plus longtemps. Ceci est un élément important pour améliorer leur qualité de vie et la

sécurité.

Pour la quantification de l'asymétrie de la marche, l'apparition d'un outil de dépistage de ce genre pourrait participer à l'amélioration de la qualité de soin en permettant une meilleure prise en charge de certaines pathologies par une détection plus précoce et un meilleur suivi de l'évolution d'une pathologie ou d'une convalescence.

Finalement, le fait de retarder le départ des personnes âgées vers les maisons de retraite et l'amélioration de la prise en charge médicale tout en réduisant les frais d'équipement, participe à l'effort général de réduction des frais du système de santé.

Ces améliorations participent à l'augmentation de la durée ainsi qu'à l'amélioration de la qualité de vie des aînés.

Bibliographie

- [1] Microsoft kinect - <http://www.xbox.com/kinect>.
- [2] The opencv library - <http://sourceforge.net/projects/opencvlibrary/>.
- [3] Primesense openni - <http://www.openni.org>.
- [4] Technoconcept - walkmeter - <http://www.technoconcept.fr>.
- [5] Vicon motion capture system - <http://www.vicon.com>.
- [6] D. Ganesan A. Hanso A. Williams. Aging in place : fall detection and localization in a distributed smart camera network. Dans *ACM MM*, 2007.
- [7] A. Abyad et J.T. Boyer. Arthritis and aging. *Current Opinion in Rheumatology*, 4(2) :153–162, 1992.
- [8] J.K. Aggarwal et Q. Cai. Human motion analysis : A review. *Computer Vision and Image Understanding*, 73(3) :428–440, 1999.
- [9] L.D. Alexander, S.E. Black, K.K. Patterson, F. Gao, C.J. Danells, et W.E. McIlroy. Association between gait asymmetry and brain lesion location in stroke patients. *Stroke*, 40(2) :537–544, 2009.
- [10] N.B. Alexander et J.M. Hausdorff. Guest editorial : linking thinking, walking, and falling. *The journals of gerontology Series A Biological sciences and medical sciences*, 63(12) :1325–1328, 2008.
- [11] M. Andrieux J.M. Westerloppe J. Allard. Le coût économique de la chute peut-il être estimé ? *L'année de la gérontologie*, 379, 1995.
- [12] M. Alwan, P.J. Rajendran, S. Kell, D. Mack, S. Dalal, M. Wolfe, et R. Felder. A smart and passive floor-vibration based fall detector for elderly. Dans *Information and Communication Technologies*, 2006. *ICTTA '06. 2nd*, volume 1, pages 1003–1007, 0-0 2006.

- [13] D. Anderson, J.M. Keller, M. Skubic, Xi Chen, et Zhihai He. Recognizing falls from silhouettes. Dans *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, pages 6388–6391, 30 2006-sept. 3 2006.
- [14] D. Anderson, R.H. Luke, J.M. Keller, M. Skubic, M. Rantz, et M. Aud. Linguistic summarization of video for fall detection using voxel person and fuzzy logic. *Computer Vision and Image Understanding*, 113(1) :80–89, 2009.
- [15] D. Arsic, M. Hofmann, B. Schuller, et G. Rigoll. Multi-camera person tracking and left luggage detection applying homographic transformation. Dans *Proceedings 10th IEEE International Workshop on PETS, Rio de Janeiro*, 2007.
- [16] B. Auvinet, G. Berrut, C. Touzard, L. Moutel, N. Collet, D. Chaleil, et E. Barrey. Reference data for normal subjects obtained with an accelerometric device. *Gait & Posture*, 16(2) :124–134, 2002.
- [17] E. Auvinet, J. Meunier, et F. Multon. Multiple depth cameras calibration and body volume reconstruction for gait analysis. Dans *The 11th International Conference on Information Sciences, Signal Processing and their Applications : Main Tracks (ISSPA2012 - Tracks)*, Montreal, Canada, Canada, Juillet 2012.
- [18] E. Auvinet, F. Multon, et J. Meunier. Gait analysis with multiple depth cameras. Dans *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 6265–6268, 30 2011-sept. 3 2011.
- [19] E. Auvinet, L. Reveret, A. St-Arnaud, J. Rousseau, et J. Meunier. Fall detection using multiple cameras. Dans *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 2554–2557, aug. 2008.
- [20] E. Auvinet, C. Rougier, J. Meunier, A. St-Arnaud, et J. Rousseau. Multiple cameras fall data set. Dans *Technical report no. 1350, University of Montreal, Canada*, 2010.
- [21] E. Auvinet, C. Rougier, A. St-Arnaud, J. Rousseau, et J. Meunier. Multi camera fall dataset, 2010.
- [22] H. Baan, R. Dubbeldam, A.V. Nene, et M.A.F.J. van de Laar. Gait analysis of the lower limb in patients with rheumatoid arthritis : A systematic review. *Seminars in Arthritis and Rheumatism*, (0) :–, 2012.
- [23] Y. Benezeth, P.M. Jodoin, B. Emile, H. Laurent, et C. Rosenberger. Review and evaluation of commonly-implemented background subtraction algorithms. Dans

-
- Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4, dec. 2008.
- [24] P. Bessou, P. Dupui, et R. Montoya. Simultaneous recording of longitudinal displacements of both feet during human walking. *Journal de physiologie*, 2(83) :102–10, 1988-1989.
 - [25] B. R. Bloem, J.M. Hausdorff, J.E. Visser, et N. Giladi. Falls and freezing of gait in parkinson's disease : A review of two interconnected, episodic phenomena. *Movement Disorders*, 19(8) :871–884, 2004.
 - [26] J.Y. Bouguet. Matlab calibration toolbox, 2007.
 - [27] H. Bourdessol et S. Pin. *Prévention des chutes chez les personnes âgées à domicile*. Réseau francophone de prévention des traumatismes et de promotion de la sécurité, 2005.
 - [28] A.K. Bourke, J.V. O'Brien, et G.M. Lyons. Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. *Gait & Posture*, 26(2) :194–199, 2007.
 - [29] E. Boyer. On using silhouettes for camera calibration. Dans *In Proceedings of the seventh Asian Conference on Computer Vision*, LNCS. Springer, January 2006.
 - [30] R. Le Bris. *Etude biomécanique de la course à pied par une méthode accélérométrique. Comparaisons méthodologiques et applications*. PhD thesis, Université de Rennes 2, 2006.
 - [31] J.J. Brunnekreef, C.J. Van Uden, S. Van Moorsel, et J.G. Kooloos. Reliability of videotaped observational gait analysis in patients with orthopedic impairments. *BMC Musculoskeletal Disorders*, 6(17) :17, 2005.
 - [32] A. Cappozzo, U. Della Croce, A. Leardini, et L. Chiari. Human movement analysis using stereophotogrammetry : Part 1 : theoretical background. *Gait & Posture*, 21(2) :186–196, 2005.
 - [33] V. Castaneda et N. Navab. Time-of-flight and kinect imaging, 2011.
 - [34] E. Ceriez et R. Motmans. Anthropometry table ergonomie rc. Rapport technique, DinBelg, 2005.
 - [35] Tom Chau. A review of analytical techniques for gait data. part 1 : fuzzy, statistical and fractal methods. *Gait & Posture*, 13(1) :49–66, 2001.
 - [36] Tom Chau. A review of analytical techniques for gait data. part 2 : neural network and wavelet methods. *Gait & Posture*, 13(2) :102–120, 2001.
-

- [37] S. Corazza, L. Mündermann, et T. Andriacchi. A framework for the functional identification of joint centers using markerless motion capture, validation for the hip joint. *Journal of Biomechanics*, 40(15) :3510–3515, 2007.
- [38] J. Courtney et A.M. de Paor. A monocular marker-free gait measurement system. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 18(4) :453–460, aug. 2010.
- [39] R. Craik et C.A. Oatis. *Gait analysis : theory and application*. Mosby, 1995.
- [40] Creaform.
- [41] A. Cretual, K. Bervet, et L. Ballaz. Gillette gait index in adults. *Gait & Posture*, 32(3) :307–310, 2010.
- [42] U. Della Croce, A. Leardini, L. Chiari, et A. Cappozzo. Human movement analysis using stereophotogrammetry : Part 4 : assessment of anatomical landmark misplacement and its effects on joint kinematics. *Gait & Posture*, 21(2) :226–237, 2005.
- [43] R. Cucchiara, A. Prati, et R. Vezzani. A multi-camera vision system for fall detection and alarm generation. *Expert Systems*, 24(5) :334–345, 2007.
- [44] R.G. Cumming. Epidemiology of medication-related falls and fractures in the elderly. *Drugs & Aging*, 12(1) :43–53, 1998.
- [45] S.L. Delp et J.P. Loan. A graphics-based software system to develop and analyze models of musculoskeletal structures. *Computers in Biology and Medicine*, 25(1) :21–34, 1995.
- [46] Institut Canadien d’Information sur la Santé. Rai-home care (rai-hc) manual, 2002.
- [47] T.J. Doherty. Invited review : Aging and sarcopenia. *Journal of Applied Physiology*, 95(4) :1717–1727, 2003.
- [48] E.R.C. Draper. A treadmill-based system for measuring symmetry of gait. *Medical engineering & physics*, 22(3) :215–222, April 2000.
- [49] J. Draréni. *Exploitation de contraintes photométriques et géométriques en vision : application au suivi, au calibrage et à la reconstruction*. PhD thesis, Université de Montréal - Université de Grenoble, 2010.
- [50] E. Grossmann C. Rougier M. Dahmane E. Auvinet et J. Meunier. Left-luggage detection using homographies and simple heuristics. Dans *Proceedings of the 9th IEEE International Workshop on Performance Evaluation in Tracking and Surveillance (PETS ’06)*, Eurographics 2006 Short Papers Preceedings, pages 51–58, 2006.

-
- [51] W. Gillespie M. Robertson et al. L. Gillespie. Interventions for preventing falls in elderly people. *Cochrane Database of Systematic Reviews*, 2003.
 - [52] O. Faugeras, Q. Luong, et S. Maybank. Camera self-calibration : Theory and experiments. Dans G. Sandini, editeur, *Computer Vision — ECCV'92*, volume 588 de *Lecture Notes in Computer Science*, pages 321–334. Springer Berlin / Heidelberg, 1992.
 - [53] R.A. Fielding, B. Vellas, W.J. Evans, S. Bhasin, J.E. Morley, A.B. Newman, G. Abellan van Kan, S. Andrieu, J. Bauer, D. Breuille, T. Cederholm, J. Chandler, C. De Meynard, L. Donini, T. Harris, A. Kannt, F. K. Guibert, G. Onder, D. Papanicolaou, Y. Rolland, D. Rooks, C. Sieber, E. Souhami, S. Verlaan, et M. Zamboni. Sarcopenia : An undiagnosed condition in older adults. current consensus definition : Prevalence, etiology, and consequences. international working group on sarcopenia. *Journal of the American Medical Directors Association*, 12(4) :249–256, 2011.
 - [54] A. Fleury, P. Rumeau, A. Bourke, G. Laighin, V. Rialle, N. Noury, et J. Lundy. Fall detection - principles and methods. Dans *Conference of the IEEE EMBS*, 2007.
 - [55] V. Fohanno. *Contribution à l'estimation de la cinématique articulaire par des méthodes numériques d'optimisation. Application à la gestuelle du kayakiste*. PhD thesis, Université de Poitiers, 2011.
 - [56] Centers for Disease Control, National Center for Injury Prevention Prevention, et Control. Web-based injury statistics query and reporting system (wisqars), november 2010.
 - [57] J.S. Franco. *Modélisation tridimensionnelle à partir de silhouettes*. These, Institut National Polytechnique de Grenoble - INPG, Dec 2005.
 - [58] J.S. Franco et E. Boyer. Efficient polyhedral modeling from silhouettes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(3) :414–427, march 2009.
 - [59] B. Freedman, A. Shpunt, M. Machlinem, et Y. Arieli. Depth mapping using projected patterns. *American Patent US 2010/0118123*, 2010.
 - [60] S.M. Friedman, B. Munoz, S.K. West, G.S. Rubin, et L.P. Fried. Falls and fear of falling : which comes first ? a longitudinal prediction model suggests strategies for primary and secondary prevention. *Journal of American Geriatrics Society*, 50(8), 2002.
 - [61] C. Gagnon et M. Lafrance. Prévention des chutes auprès des personnes âgées vivant à domicile : analyse des données scientifiques et recommandations préliminaires à l'élaboration d'un guide de pratique clinique, 2011.
-

- [62] D.M Gavrilu. The visual analysis of human movement : A survey. *Computer Vision and Image Understanding*, 73(1) :82–98, 1999.
- [63] J. Gibeault. Rasterstereography : Radiation-free technology for spine and pelvis analysis. *Canadian Chiropractor*.
- [64] L.D. Gillespie, W.J. Gillespie, C.M. Robertson, S.E. Lamb, R.G. Cumming, et B.H. Rowe. Interventions for preventing falls in elderly people. *Cochrane Database of Systematic Reviews*, 2, 2009.
- [65] L. Ginaldi, M. Di Benedetto, et M. De Martinis. Osteoporosis, inflammation and ageing. *Immunity & Ageing*, 2(1) :14, 2005.
- [66] M. Goffredo, R.D. Seely, J.N. Carter, et M.S. Nixon. Markerless view independent gait analysis with self-camera calibration. Dans *Automatic Face Gesture Recognition, 2008. FG '08. 8th IEEE International Conference on*, pages 1–6, sept. 2008.
- [67] S.B. Gokturk, H. Yalcin, et C. Bamji. A time-of-flight depth sensor - system description, issues and solutions. Dans *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW '04. Conference on*, page 35, june 2004.
- [68] D. Gouwanda et S.M.N.A. Senanayake. Identifying gait asymmetry using gyroscopes—a cross-correlation and normalized symmetry index approach. *Journal of Biomechanics*, 44(5) :972–978, 2011.
- [69] R.D. Green et Ling Guan. Quantifying and recognizing human movement patterns from monocular video images-part ii : applications to biometrics. *Circuits and Systems for Video Technology, IEEE Transactions on*, 14(2) :191–198, feb. 2004.
- [70] B. Gurney. Leg length discrepancy. *Gait & posture*, 15(2) :195–206, April 2002.
- [71] J. Han et B. Bhanu. Individual recognition using gait energy image. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(2) :316–322, Février 2006.
- [72] R.I. Hartley. An algorithm for self calibration from several views. Dans *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on*, pages 908–912, 1994.
- [73] J.M. Hausdorff et N.B. Alexander. *Gait Disorders : Evaluation and Management*. Neurological Disease and Therapy. Taylor & Francis, 2005.
- [74] J.M. Hausdorff, D.A. Rios, et H.K. Edelberg. Gait variability and fall risk in community-living older adults : A 1-year prospective study. *Archives of Physical Medicine and Rehabilitation*, 82(8) :1050–1056, 2001.

-
- [75] L. Hazelhoff, J. Han, et P.H. With. Video-based fall detection in the home using principal component analysis. Dans *Proceedings of the 10th International Conference on Advanced Concepts for Intelligent Vision Systems, ACIVS '08*, pages 298–309, Berlin, Heidelberg, 2008. Springer-Verlag.
- [76] J. Heikkila et O. Silven. A four-step camera calibration procedure with implicit image correction. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 0 :1106, 1997.
- [77] N.E. Helwig, S.Hong, E.T. Hsiao-Wecksler, et J.D. Polk. Methods to temporally align gait cycle data. *Journal of Biomechanics*, 44(3) :561–566, 2011.
- [78] J.V. Hindle. Ageing, neurodegeneration and parkinson’s disease. *Age and Ageing*, 39(2) :156–161, 2010.
- [79] C. Hodt-Billington, J.L. Helbostad, W. Vervaat, T. Rognsvag, et R. Moe-Nilssen. Criteria of gait asymmetry in patients with hip osteoarthritis. *Physiotherapy Theory and Practice*, 28(2) :134–141, 2012.
- [80] D.L. Hoyert, K.D. Kochanek, et S.L. Murphy. Deaths : final data for 1997. *National Vital Statistic Reports*, 47(19), 1999.
- [81] A. Hsu, P. Tang, et M. Jan. Analysis of impairments influencing gait velocity and asymmetry of hemiplegic patients after mild to moderate stroke. *Archives of Physical Medicine and Rehabilitation*, 84(8) :1185–1193, 2003.
- [82] <http://www.animazoo.fr/> <http://www.animazoo.fr/> <http://www.animazoo.fr/> <http://www.animazoo.fr/>. Animazoo mechanical motion capture system.
- [83] <http://www.clinique-du-dos.com> <http://www.clinique-du-dos.com> <http://www.clinique-du-dos.com> <http://www.clinique-du-dos.com>. Institut rossetti - la clinique du dos.
- [84] R.Z.-L. Hu, A. Hartfiel, J. Tung, A. Fakih, J. Hoey, et P. Poupart. 3d pose tracking of walker users’ lower limb with a structured-light camera on a moving platform. Dans *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on*, pages 29–36, june 2011.
- [85] K. Jahn, A. Zwergal, et R. Schniepp. Gait Disturbances in Old Age : Classification, Diagnosis, and Treatment From a Neurological Perspective. *Dtsch Arztebl International*, 107(17) :306–316, 2010.
-

- [86] B. Jansen, F. Temmermans, et R. Deklerck. 3d human pose recognition for home monitoring of elderly. Dans *Conference of the IEEE EMBS*, Août 2007.
- [87] F. Jean, A. Branzan-Albu, et R. Bergevin. Towards view-invariant gait modeling : Computing view-normalized body part trajectories. *Pattern Recognition*, 42(11) :2936–2949, November 2009.
- [88] R. Jensen, R. Paulsen, et R. Larsen. Analysis of gait using a treadmill and a time-of-flight camera. Dans Andreas Kolb et Reinhard Koch, éditeurs, *Dynamic 3D Imaging*, volume 5742 de *Lecture Notes in Computer Science*, pages 154–166. Springer Berlin / Heidelberg, 2009. 10.1007/978-3-642-03778-8_12.
- [89] P. Jodoin. Vision par ordinateur imn 559, 2011.
- [90] S. Zimba Kalula, Vicky S., A. Dowd, et K. Brodrick. Falls and fall prevention programmes in developing countries : Environmental scan for the adaptation of the canadian falls prevention curriculum for developing countries. *Journal of Safety Research*, 42(6) :461–472, 2011.
- [91] M. Kangas, A. Konttila, P. Lindgren, I. Winblad, et T. Jämsä. Comparison of low-complexity fall detection algorithms for body attached accelerometers. *Gait & Posture*, 28(2) :285–291, 2008.
- [92] P. Kannus, S. Niemi, M. Palvanen, et J. Parkkari. Fall-induced injuries among elderly people. *The Lancet*, 350(9085) :1174, 1997.
- [93] K. Kim, T.H. Chalidabhongse, D. Harwood, et L. Davis. Background modeling and subtraction by codebook construction. Dans *In International Conference on Image Processing*, pages 3061–3064, 2004.
- [94] A. Kolb, E. Barth, R. Koch, et R. Larsen. Time-of-flight cameras in computer graphics. *Computer Graphics Forum*, 29(1) :141–159, 2010.
- [95] P.V. Komi. Relevance of in vivo force measurements to human biomechanics. *Journal of Biomechanics*, 23, Supplement 1(0) :23–34, 1990.
- [96] R. Kram, T.M. Griffin, J. Maxwell Donelan, et Y. Hui Chang. Force treadmill for measuring vertical and horizontal ground reaction forces. *Journal of Applied Physiology*, 85(2) :764–769, 1998.
- [97] A. Laurentini. The visual hull concept for silhouette-based image understanding. *IEEE Trans. Pattern Anal. Mach. Intell.*, 16 :150–162, February 1994.
- [98] S.J. Lee et J. Hidler. Biomechanics of overground vs. treadmill walking in healthy individuals. *Journal of Applied Physiology*, 104(3) :747–755, 2008.

-
- [99] T. Lee et A. Mihailidis. An intelligent emergency response system : preliminary development and testing of automated fall detection. *Journal of Telemedicine and Telecare*, 11(4) :194–198, 2005.
 - [100] W. Lee, W. Wontack, et E. Boyer. Silhouette segmentation in multiple views. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010.
 - [101] Q. Li, J.A. Stankovic, M.A. Hanson, A.T. Barth, J. Lach, et G. Zhou. Accurate, fast fall detection using gyroscopes and accelerometer-derived posture information. Dans *Wearable and Implantable Body Sensor Networks, 2009. BSN 2009. Sixth International Workshop on*, pages 138–143, june 2009.
 - [102] Y. Li, K. C. Ho, et M. Popescu. A microphone array system for automatic fall detection. *Biomedical Engineering, IEEE Transactions on*, 59(5) :1291–1301, may 2012.
 - [103] C.W. Lin, Z.H. Ling, Y.C. Chang, et C.J. Kuo. Compressed-domain fall incident detection for intelligent homecare. *Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology*, 49(3), 2007.
 - [104] R.H. Luke, J.M. Keller, M. bic bic bic bic, M.J. Rantz, et M.A. Aud D. Anderson. Modeling human activity from voxel person using fuzzy logic. *IEEE Transactions on Fuzzy Systems*, 17, 2009.
 - [105] J. Magaziner, E.M. Simonsick, T.M. Kashner, J.R. LHebel, et J.E. Kenzora. Predictors of functional recovery one year following hospital discharge for hip fracture : A prospective study. *Journal of Gerontology*, 45(3), 1990.
 - [106] E. Marey. La photographie du mouvement. *La nature*, 477, 1882.
 - [107] É.J. Marey. Étude de la locomotion animale par la chrono-photographie. *la Revue Scientifique*, pages 124–134, november 1886.
 - [108] F. Marzani, E. Calais, et L. Legrand. A 3-d marker-free system for the analysis of movement disabilities - an application to the legs. *Information Technology in Biomedicine, IEEE Transactions on*, 5(1) :18–26, march 2001.
 - [109] G. Mastorakis et D. Makris. Fall detection system using kinect’s infrared sensor. *Journal of Real-Time Image Processing*, pages 1–12, 2010. 10.1007/s11554-012-0246-9.
 - [110] R.E. Mayagoitia, A.V. Nene, et P.H. Veltink. Accelerometer and rate gyroscope measurement of kinematics : an inexpensive alternative to optical motion analysis systems. *Journal of Biomechanics*, 35(4) :537–542, 2002.
 - [111] J.L. McCrory, S.C. White, et R.M. Lifeso. Vertical ground reaction forces : objective measures of gait following hip arthroplasty. *Gait & Posture*, 14(2) :104–109, 2001.
-

- [112] S.G. Miaou, P.H. Shung, et C.Y. Huang. A customized human fall detection system using omni-camera images and personal information. Dans *D2H2 2006 (1st Transdisciplinary Conf. on Distributed Diagnosis and Home Care)*, 2006.
- [113] B. Michoud, E. Guillou, et S. Bouakaz. Shape From Silhouette : Towards a Solution for Partial Visibility Problem. Dans D.Fellner C.Hansen, editeur, *Eurographics 2006*, Eurographics 2006 Short Papers Preceedings, pages 13–16, Septembre 2006.
- [114] G. Moeslund et E. Granum. A survey of computer vision-based human motion capture. *Computer Vision and Image Understanding*, 81(3) :231–268, 2001.
- [115] T.B. Moeslund, A. Hilton, et V. Krüger. A survey of advances in vision-based human motion capture and analysis. *Comput. Vis. Image Underst.*, 104(2) :90–126, Novembre 2006.
- [116] S. Morrison, S.R. Colberg, H.K. Parson, et A.I. Vinik. Relation between risk of falling and postural sway complexity in diabetes. *Gait & Posture*, 35(4) :662–668, 2012.
- [117] L. Mundermann, S. Corazza, et T. Andriacchi. The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications. *Journal of NeuroEngineering and Rehabilitation*, 3(1) :6, 2006.
- [118] L. Mundermann, S. Corazza, A.M. Chaudhari, E.J. Alexander, et T.P. Andriacchi. Most favorable camera configuration for a shape-from-silhouette markerless motion capture system for biomechanical analysis. *SPIE-IS&T Electronic Imaging*, 5665 :278–287, 2005.
- [119] L. Mundermann, A. Mundermann, A.M. Chaudhari, et T.P. Andriacchi. Conditions that influence the accuracy of anthropometric parameter estimation for human body segments using shape-from-silhouette. *SPIE-IS&T Electronic Imaging*, 5665 :268–277, 2005.
- [120] S.L. Murphy. Deaths : final data for 1998. *National Vital Statistic Reports*, 48(11), 2000.
- [121] S.L. Murphy, J.A. Dubin, et T.M. Gill. The development of fear of falling among community-living older women : Predisposing factors and subsequent fall events. *The Journals of Gerontology Series A : Biological Sciences and Medical Sciences*, 58(10) :M943–M947, 2003.
- [122] E. Muybridge. Animal locomotion. 1887.

-
- [123] H. Nait-Charif et S.J. McKenna. Activity summarisation and fall detection in a supportive home environment. Dans *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 4, pages 323–326 Vol.4, aug. 2004.
 - [124] S.I. Newton, T. Leseur, et F. Jacquier. *Philosophiae naturalis principia mathematica*, volume 1. Excudit Georgius Brookman ; impensis TT et J. Tegg, 1833.
 - [125] A. Nghiem, E. Auvinet, F. Multon, et J. Meunier. Contactless abnormal gait detection. Dans *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 5076–5079, 30 2011-sept. 3 2011.
 - [126] N. Noury, P. Rumeau, A.K. Bourke, G. ÓLaighin, et J.E. Lundy. A proposal for the classification and evaluation of fall detectors. *IRBM*, 29, 2008.
 - [127] M.N. Nyan, Francis E.H. Tay, et Matthew Z.E. Mah. Application of motion analysis system in pre-impact fall detection. *Journal of Biomechanics*, 41(10) :2297–2304, 2008.
 - [128] M.N. Nyan, Francis E.H. Tay, et E. Murugasu. A wearable system for pre-impact fall detection. *Journal of Biomechanics*, 41(16) :3475–3481, 2008.
 - [129] World Health Organization. Global report on falls prevention in older age, 2008.
 - [130] K.K. Patterson, W.H. Gage, D. Brooks, S.E. Black, et W.E. McIlroy. Evaluation of gait symmetry after stroke : A comparison of current methods and recommendations for standardization. *Gait & Posture*, 31(2) :241–246, 2010.
 - [131] M. Piccardi. Background subtraction techniques : a review. Dans *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, volume 4, pages 3099–3104 vol.4, oct. 2004.
 - [132] Living place Hamburg.
 - [133] M. Plotnik, N. Giladi, Y. Balash, C. Peretz, et J.M. Hausdorff. Is freezing of gait in parkinson’s disease related to asymmetric motor function ? *Ann Neurol*, 57(5) :656–63, 2005.
 - [134] F. Prince, H. Corriveau, R. Hébert, et D. A. Winter. Gait in the elderly. *Gait & Posture*, 5(2) :128–135, 1997.
 - [135] D.C. Reid et B. Smith. Leglength inequality : a review of etiology and management. *Physiotherapy Canada*, 36(4) :177–182, 1984.
 - [136] J.K. Richardson, SB. Thies, T.K. DeMott, et JA. Ashton-Miller. Gait analysis in a challenging environment differentiates between fallers and nonfallers among older patients with peripheral neuropathy. *Archives of Physical Medicine and Rehabilitation*, 86(8) :1539–1544, 2005.
-

- [137] Y. Robitaille et J. Gratton. Les chutes chez les adultes âgés : vers une surveillance plus fine des données d'hospitalisation. Rapport technique, Institut National de Santé du Québec, 2005.
- [138] C. Rougier. *Vidéosurveillance intelligente pour la détection de chutes chez les personnes âgées*. PhD thesis, Université de Montréal, 2010.
- [139] C. Rougier, E. Auvinet, J. Meunier, M. Mignotte, et J.A. de Guise. Depth energy image for gait symmetry quantification. Dans *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 5136–5139, 30 2011-sept. 3 2011.
- [140] C. Rougier, E. Auvinet, J. Rousseau, M. Mignotte, et J. Meunier. Fall detection from depth map video sequences. Dans Bessam Abdulrazak, Sylvain Giroux, Bruno Bouchard, Hélène Pigot, et Mounir Mokhtari, éditeurs, *ICOST*, volume 6719 de *Lecture Notes in Computer Science*, pages 121–128. Springer, 2011.
- [141] C. Rougier, J. Meunier, A. St-Arnaud, et J. Rousseau. Monocular 3d head tracking to detect falls of elderly people. Dans *Conference of the IEEE EMBS*, Août 2006.
- [142] C. Rougier, J. Meunier, A. St-Arnaud, et J. Rousseau. Procrustes shape analysis for fall detection. Dans *The Eighth International Workshop on Visual Surveillance*, 2008.
- [143] C. Rougier, J. Meunier, A. St-Arnaud, et J. Rousseau. Robust video surveillance for fall detection based on human shape deformation. *Circuits and Systems for Video Technology, IEEE Transactions on*, 21(5) :611–622, may 2011.
- [144] H. Sadeghi. Dynamic stability and propulsion during gait of the healthy elderly and young subjects using three dimensional kinetic data. *World Journal of Sport Sciences*, 5(3) :205–215, 2011.
- [145] H. Sadeghi, P. Allard, F. Prince, et H. Labelle. Symmetry and limb dominance in able-bodied gait : a review. *Gait & Posture*, 12(1) :34–45, 2000.
- [146] R. Sakamoto, I. Kitahara, N. Orman, T. Toriyama, K. Kogure, et H. Kim. Compensated visual hull for defective segmentation and occlusion. Dans *International Conference on Artificial Reality and Telexistence*, 2007.
- [147] B. Salzman. Gait and balance disorders in older adults. *American Family Physician*, 82(1) :61–69, 2010.
- [148] R.W. Sattin, D.A. Lambert Huber, C.A. Devito, J.G. Rodriguez, A. Ros, et S. Bacchelli. The incidence of fall injury events among the elderly in a defined population. *American Journal of Epidemiology*, 131(5) :1028–1037, 1990.

-
- [149] S.M. Seitz, B. Curless, J. Diebel, D. Scharstein, et R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. Dans *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 519–528, june 2006.
 - [150] D. Shepard. A two-dimensional interpolation function for irregularly-spaced data. Dans *Proceedings of the 1968 23rd ACM national conference, ACM '68*, pages 517–524, New York, NY, USA, 1968. ACM.
 - [151] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, et A. Blake. Real-time human pose recognition in parts from single depth images. Dans *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 1297–1304, june 2011.
 - [152] S. Sivapalan, D. Chen, S. Denman, S. Sridharan, et C. Fookes. Gait energy volumes and frontal gait recognition using depth images. Dans *Biometrics (IJCB), 2011 International Joint Conference on*, pages 1–6, oct. 2011.
 - [153] A. Sixsmith et N. Johnson. A smart sensor to detect the falls of the elderly. *Pervasive Computing, IEEE*, 3(2) :42–47, april-june 2004.
 - [154] O. Sofuwa, A. Nieuwboer, K. Desloovere, A.M. Willems, F. Chavret, et I. Jonkers. Quantitative gait analysis in parkinson's disease : Comparison with a healthy control group. *Archives of Physical Medicine and Rehabilitation*, 86(5) :1007–1013, 2005.
 - [155] K.M. Song, S.E. Halliday, et D.G. Little. The effect of limb-length discrepancy on gait. *The Journal of Bone & Joint Surgery*, 79 :1690–1698, 1997.
 - [156] J.A. Stevens, P.S. Corso, E.A. Finkelstein, et T.R. Miller. The costs of fatal and non-fatal falls among older adults. *Inj Prev*, 12(5) :290–5, 2006.
 - [157] E. Surer et A. Kose. Methods and technologies for gait analysis. Dans Albert Ali Salah et Theo Gevers, éditeurs, *Computer Analysis of Human Behavior*, pages 105–123. Springer London, 2011. 10.1007/978-0-85729-994-9_5.
 - [158] GaitRite System.
 - [159] N. Thome, S. Miguët, et S. Ambellouis. A real-time, multiview fall detection system : A lhmm-based approach. *Circuits and Systems for Video Technology, IEEE Transactions on*, 18(11) :1522–1532, nov. 2008.
 - [160] B.U. Toreyin, Y. Dedeoglu, et A.E. Cetin. Hmm based falling person detection using both audio and video. Dans *Signal Processing and Communications Applications, 2006 IEEE 14th*, pages 1–4, april 2006.
-

- [161] R.Y. Tsai. Radiometry. chapitre A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses, pages 221–244. Jones and Bartlett Publishers, Inc., USA, 1992.
- [162] M. Turjo, M.F. Wong, M. Wang, J. Tao, et Y.-P. Tan. Fall incidents detection for intelligent video surveillance. Dans *Conference on Information, Communications and Signal Processing*, Decembre 2005.
- [163] R.J. Weiss, P. Wretenberg, A. Stark, K. Palmblad, P. Larsson, L. Gröndal, et E. Brström. Gait pattern in rheumatoid arthritis. *Gait & Posture*, 28(2) :229–234, 2008.
- [164] J. Weng, P. Cohen, et M. Herniou. Camera calibration with distortion models and accuracy evaluation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 14(10) :965–980, oct 1992.
- [165] M. Whittle. *Gait analysis : an introduction*. Elsevier, 3 édition, 2002.
- [166] Ge Wu. Distinguishing fall activities from normal activities by velocity characteristics. *Journal of Biomechanics*, 33(11) :1497–1500, 2000.
- [167] Y. Xinguo. Approches and principles of fall detection for elderly and patient. Dans *10th IEEE Intl. Conf on e-Health Networking, Application and Service*, 2008.
- [168] V.M. Zatsiorsky. *Kinematics of Human Motion*, volume 1. Human Kinetics, 1998.
- [169] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. Dans *in ICCV*, pages 666–673, 1999.
- [170] Z. Zhang. Camera calibration with one-dimensional objects. Dans *Proceedings of the 7th European Conference on Computer Vision-Part IV, ECCV '02*, pages 161–174, London, UK, UK, 2002. Springer-Verlag.
- [171] Y. Zigel, D. Litvak, et I. Gannot. A method for automatic fall detection of elderly people using floor vibrations and sound : Proof of concept on human mimicking doll falls. *Biomedical Engineering, IEEE Transactions on*, 56(12) :2858–2867, dec. 2009.