

Reconstruction d'humanoïde à partir de plusieurs vues

B. Michoud, E. Guillou et S. Bouakaz

Laboratoire LIRIS-CNRS UMR 5205 Université Claude Bernard Lyon 1

brice.michoud, erwan.guillou, saida.bouakaz@liris.cnrs.fr

Résumé : *Nous présentons une méthode qui permet l'acquisition automatique d'un personnage ainsi que sa posture à partir de plusieurs vues, fournies par des caméras calibrées et synchronisées. Pour représenter la forme humaine nous adoptons un modèle à base de squelette articulé (squelette d'animation). Un squelette est un ensemble de segments liés par des articulations. Les articulations sont généralement des transformations de type rotations auxquelles sont associées des contraintes (anatomiques) de variations d'angles. Nous supposons que les silhouettes du personnage sont extraites de chaque vue. Ces silhouettes sont alors utilisées comme entrée pour l'extraction des différentes parties du corps, pour la localisation des articulations du squelette et la modélisation de l'individu filmé.*

Notre approche s'appuie sur la reconstruction d'un avatar, qui soit une représentation du personnage filmé. L'originalité de notre approche est l'extraction automatique du squelette (d'animation) à partir des images de silhouettes et ceci sans imposer de contraintes de pose "type". Nous avons défini l'habillage géométrique de l'avatar comme étant un ensemble de primitives géométriques centrées sur le squelette.

Contrairement à la plupart des méthodes existantes, notre méthode n'impose aucune contrainte sauf le fait que la personne filmée doit être en équilibre stable. Les résultats obtenus dans des cas réels ont permis de valider notre approche.

Mots-clés : Squelette d'animation, posture, Reconstruction 3D, Incrustation.

1 Introduction

Ce travail se situe dans le cadre d'un projet "ACI masses de données : Cyber 2" Notre but final est de réaliser une insertion visuellement réaliste d'un personnage en mouvement (un présentateur d'émission de télévision par exemple) dans un environnement virtuel ou réel et d'assurer une bonne interaction entre le personnage et l'environnement "hôte". Pour cela nous devons être capable d'estimer la posture du personnage, ainsi qu'un modèle volumique le représentant. Ce qui permettra d'une part de gérer des phénomènes d'ombres portées indispensables à un rendu réaliste, et d'autre part de reconnaître une gestuelle et/ou de la reproduire sur un avatar. Compte tenu de la diversité des travaux réalisés dans ce domaine il est difficile de faire un état de l'art exhaustif, pour cela nous ne citons ici que ceux qui se rapprochent de nos travaux. Plusieurs articles de synthèse ont été proposés [8], [1], pour un travail plus récent on peut se reporter à [14]. Les travaux dans ce domaine concernent aussi bien la reconstruction de modèle et de posture à partir d'une seule séquence vidéo [5], que le recours à plusieurs séquences vidéo [4, 18]. Différents modèles ont été proposés [16]. Suivant le but poursuivi et les applications visées, on peut classer les différentes approches en deux catégories.

La première catégorie regroupe les approches qui proposent une reconstruction détaillée du modèle [16], [12], [15] Ces méthodes sont motivées par un souci d'esthétique pour un rendu réaliste [10] et requièrent des algorithmes coûteux et complexes. La plupart de ces méthodes s'appuient sur une comparaison dense entre le modèle et la posture prédits, et les données réelles. La recherche de la solution est généralement exprimée en termes d'optimisation d'une fonction dans un espace de grande dimension [11, 19, 2]. Différentes fonctions de comparaison ont été proposées et différentes méthodes d'optimisation ont été utilisées [17, 5, 21], [6]. Bien que l'acquisition et certaines étapes de pré-traitement soient réalisées en temps réel, les étapes d'ajustement de modèle sont très coûteuses et, à moins d'avoir recours au calcul parallèle, elles ne peuvent être utilisées dans des applications qui nécessitent le temps réel.

Dans la seconde catégorie, on retrouve les méthodes qui visent le temps réel, mais en retour fournissent des modèles peu détaillés. L'une des reconstructions les plus populaires est la reconstruction qui se base sur le modèle voxelique [9], [13]. Cheung et al. ont été parmi les premiers à proposer cette reconstruction en temps réel [3]. Après extraction de la silhouette, l'idée est de sculpter un volume dans l'espace. À partir d'un volume initial composé de cellules élémentaires, les voxels, on élimine les cellules qui se projettent à l'extérieur de la silhouette dans au moins une des différentes images considérées. L'utilisation d'une technique de type marching cube permet ensuite de lisser la surface d'un tel ensemble de voxels, pour estimer la surface du personnage. Bien que de telles méthodes fonctionnent en temps réel, la qualité du rendu, ajoutée au manque de souplesse de manipulation¹ de ces modèles, limitent leurs applications.

D'autres travaux proposent des algorithmes de suivi de mouvement ou d'estimation de posture qui s'exécutent en temps réel [14]. Ces approches s'appuient -pour la plupart- sur une estimation de la posture initiale au début du mouvement. Dans le cas où cette estimation n'est pas réalisée manuellement, on impose à l'individu filmé, de prendre des poses types prédéfinies.

¹Ces modèles ne sont pas déformables et ne se prêtent pas au mouvement

Nous présentons ici une méthode automatique d'initialisation de posture. Notre méthode s'appuie sur l'utilisation de plusieurs caméras et d'un modèle générique articulé pour représenter le personnage.

Il s'agit d'un modèle articulé paramétré qui inclut un squelette et un habillage volumique. Un squelette est un ensemble de segments joints par des articulations. L'ensemble des dépendances entre les articulations définit une structure hiérarchique. La base de notre approche est l'extraction d'un véritable squelette d'animation à partir des différentes vues. L'habillage géométrique du squelette est réalisé par un ensemble de primitives géométriques. Il permet de spécifier la forme de l'humanoïde reconstruit.

Cet article commence par quelques notions préliminaires concernant le squelette que nous voulons détecter, après avoir introduit la notion de Représentation Condensée (RC) [2]. La section 3 et la section 4 décrivent les étapes de recherche des différentes articulations à partir des silhouettes extraites de chaque vue. Cette étape est suivie par la reconstruction du squelette, et l'attribution d'une certaine sémantique aux différents membres. Un habillage du squelette par des primitives géométriques permet ensuite de retrouver une forme volumique du personnage. Cette étape est décrite dans la section 5. Enfin les résultats et la conclusion suivront dans les sections 6 et 7

2 Estimation de la posture d'un humanoïde

Généralement, dans le domaine de l'infographie, le corps humain est représenté par un squelette d'animation ; des primitives géométriques centrées sur le squelette lui donne sa forme volumique. Cette représentation donne beaucoup de souplesse pour la réalisation du mouvement. De ce point de vue, déterminer la posture d'une personne revient à déterminer la position des différentes articulations (voir fig. 2(a)) du squelette.

Comme nous ne disposons pas de la forme du personnage pour extraire le squelette ; nous utilisons les silhouettes pour simuler cette information. Pour cela nous travaillons sur un boîtier englobant B de l'espace occupé par la personne, cette boîte pouvant être déterminée à partir des pyramides de vision des différentes caméras. Les silhouettes sont extraites par la méthode décrite dans [9].

2.1 Coupes

Soit π un plan défini par un point P du plan de vecteur normal \vec{n} . Nous appellerons **coupe** l'image 2d représentant l'intersection entre π et la forme 3d du personnage.

Comme nous ne disposons pas de cette forme, les coupes sont calculées en utilisant des méthodes de *projective texture mapping* [20] (PTM), les masques de silhouette issus des caméras étant utilisés comme textures projetées. Cette projection s'effectue sur un quadrangle de support π et les textures projetées sont combinées en utilisant un opérateur logique *ET* (voir fig. 1(a)).

Notons que l'ensemble des coupes estimées sur la boîte englobante peut être assimilé à une représentation voxellique (Vox) du sujet.

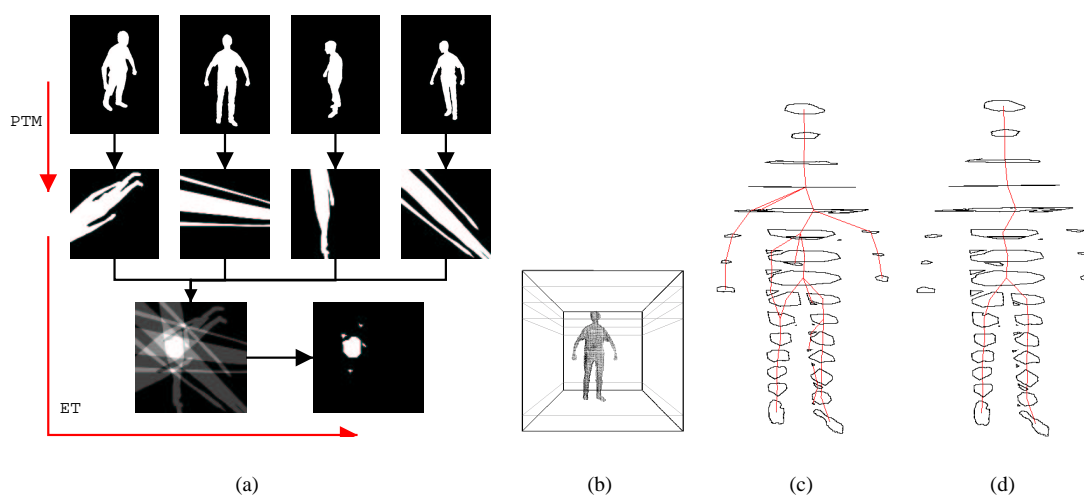


FIG. 1 – (a) Etapes de la construction des coupes en utilisant les masques de silhouette. Coupes et représentation condensée : (b) boîte englobante et coupes successives ; (c) représentation condensée ; (d) α -parcours de la RC décrite en (c) ($\alpha=0.3$).

Soit C_i la coupe de support $\pi_i(P_i, \vec{n})$, C_{i+1} décrit la coupe suivante de π_i dans la direction de \vec{n} .

Les articulations du squelette seront extraites à partir de **points représentatifs** sélectionnés sur les coupes estimées dans la boîte B (voir fig. 1(b)). De plus, on définit un chaînage entre les points représentatifs issus de coupes successives (voir sec.2.2).

2.2 Représentation condensée

Soit Ψ l'ensemble des J composantes connexes de la coupe C et $\zeta_j \in \Psi$, l'une de ces composantes connexes. Nous appellerons point représentatif d'une composante connexe ζ_j , le quadruplet $Pr_j = \{G_j, A_j, F_j, E_j\}$ avec :

- G_j la position 3d du centre de gravité ζ_j .
- A_j l'aire de ζ_j .
- F_j l'ensemble des points frontières de ζ_j .
- E_j les paramètres de l'ellipse d'inertie associée à ζ_j .

Nous définissons la notion de **Représentation Condensée** (notée RC par la suite) d'un ensemble de coupes parallèles S par l'ensemble de tous les points représentatifs Pr_i de toutes les coupes de S .

Sur cet ensemble de points représentatifs, on définit un chaînage orienté dans le sens de construction des coupes (voir fig. 1(c)) défini par : deux points représentatifs Pr_i et Pr_j seront chaînés si

1. ils sont issus de deux coupes successives ;
2. les composantes connexes associées à Pr_i et Pr_j sont connexes dans l'espace voxelique défini par Vox .

Il est à noter qu'un point représentatif peut être origine (ou extrémité) de plusieurs liens.

Afin de réduire l'ensemble des points représentatifs étudiés lors de l'estimation des positions des articulations, il est nécessaire de pouvoir ne traiter que les branches significatives de la RC. Pour cela, nous introduisons le concept d' **α -représentativité** : un point représentatif Pr_i est α -représentatif si l'aire A_i de la composante connexe qui lui est associée vérifie le prédicat 2.1.

$$\frac{A_i}{\sum_{j=1}^{j=J} A_j | (\zeta_i, \zeta_j) \in \Psi^2} > \alpha \text{ où } \alpha \in [0, 1] \quad (2.1)$$

Par extension, nous définissons un **α -parcours** comme étant le parcours suivant la chaîne orientée, des points α -représentatifs de la RC.

2.3 Ordre de détermination des articulations

D'après l'hypothèse définie précédemment, selon laquelle le sujet est debout en équilibre stable, les parties les plus contraintes (en terme de degrés de liberté) de son corps correspondront aux jambes. Ainsi, la première étape de la méthode consiste à déterminer les positions des articulations basses du corps (bassin, hanches, genoux, chevilles et pieds). Connaissant ces positions, il sera alors possible d'estimer celles des articulations supérieures (tête, cou, épaules, coudes, poignets et mains). La figure 2(b) représente cet ordre de détermination, pour chaque groupes d'articulations.

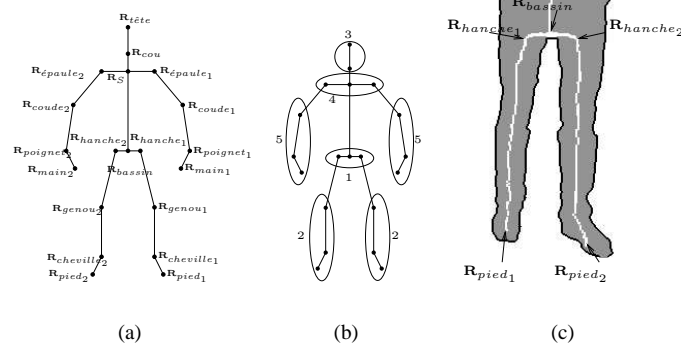


FIG. 2 – Squelette d'animation : (a) étiquetage des articulations ; (b) ordre d'extraction des articulations. Relation entre le bassin et les hanches (d).

3 Détermination des articulations basses

Notre méthode est basée sur une approche prédiction / correction. Dans une première étape, les positions du bassin et des hanches sont estimées. Ces estimations sont ensuite utilisées pour calculer les positions des articulations des deux jambes. Enfin, les mesures anthropométriques des jambes permettent de corriger les estimations faites lors de la première étape.

3.1 Estimation du bassin et des hanches

La recherche du bassin dans la RC revient à retrouver une configuration telle que présentée dans la figure 2(c) : un point représentatif où le chaînage local des jambes vient se connecter au chaînage local du bassin. Dans cette configuration le point de fusion de deux sous-chaînes de la RC correspond à une estimation de la position du bassin (notée \mathbf{R}_{bassin}). Pour trouver ce point, nous réalisons un α -parcours de la chaîne principale² de la RC depuis ses points les plus bas (notés \mathbf{Pr}_{pied_1} et \mathbf{Pr}_{pied_2} dans la fig. 2(c)). Ce parcours se terminera sur le premier confluent détecté dans la chaîne. D'après les résultats expérimentaux, une valeur $\alpha = 0.3$ (voir fig. 1(d)) donne de bons résultats en permettant de supprimer la plupart des artefacts dus à la construction des coupes.

Le point de confluence des deux sous-chaînes sera utilisé comme estimation de la position du bassin. Les points précédents (dans chacune des sous-chaînes) seront, quant à eux, utilisés pour définir les articulations des hanches (notées \mathbf{R}_{hanche_1} and \mathbf{R}_{hanche_2} dans la fig. 2(c)).

Nous verrons, lors de la détermination des jambes, une méthode permettant d'identifier le coté gauche du coté droit dans le squelette extrait.

3.2 Détermination des articulations dans un membre générique

Nous proposons une méthode de détermination des articulations dans un membre quelconque. Cette méthode se base sur l'extraction de certains points représentatifs d'une RC locale au membre à traiter, points qui seront ensuite utilisés comme approximations des positions des articulations du membre. Connaissant deux points extrêmes A et B dans le membre, nous construisons une RC locale \mathbf{CH}_{membre} sur la boîte englobante du membre. Cette RC est orientée dans la direction du vecteur \overrightarrow{AB} , A est un point de connexion entre le corps et le membre et B est un point à l'autre extrémité du membre (sur la main ou le pied selon le cas à traiter).

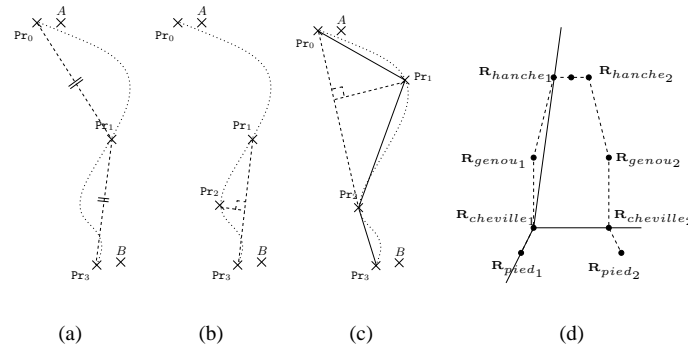


FIG. 3 – Etapes de la méthode générique d'extraction des articulations d'un membre (a); Trièdre défini par les articulations basses. Dans ce cas, les valeurs d'indice 1 définissent des articulations du coté droit (b)

Nous commençons par choisir le point \mathbf{Pr}_3 , une estimation du point extrême du membre (utilisé pour le pied ou la main), comme étant le point représentatif (de la RC locale) le plus proche de B. Puis nous déterminons une première estimation du point \mathbf{Pr}_1 (utilisé pour le genou ou le coude). Ce point est défini comme étant le point de \mathbf{CH}_{membre} situé à égale distance des points \mathbf{Pr}_0 (point représentatif de \mathbf{CH}_{membre} le plus proche de A) et \mathbf{Pr}_3 (voir fig. 3(a)), défini par l'équation 3.1.

$$\overline{\mathbf{Pr}_0\mathbf{Pr}_1} \simeq \overline{\mathbf{Pr}_1\mathbf{Pr}_3} \quad (3.1)$$

Nous déterminons ensuite le point représentatif \mathbf{Pr}_2 (utilisé pour la cheville ou le poignet) comme étant le point de \mathbf{CH}_{membre} le plus éloigné³ de la droite ($\mathbf{Pr}_1\mathbf{Pr}_3$) (voir fig. 3(b)) en utilisant l'équation 3.2 où \mathbf{Pr} est un point

²Chaîne reliant le maximum de points représentatifs.

³au sens de la distance euclidienne entre un point et une droite

représentatif du sous-chaînage de CH_{membre} défini entre les points Pr_1 et Pr_3 .

$$d_e(Pr_2, (Pr_1Pr_3)) = \max d_e(Pr, (Pr_1Pr_3)) \quad (3.2)$$

Enfin, le point Pr_1 est redéfini comme étant le point de CH_{membre} le plus éloigné de la droite (Pr_0Pr_2) (voir fig. 3(c)), en utilisant l'équation 3.2 appliquée au sous-chaînage de CH_{membre} entre les points Pr_0 et Pr_2 .

3.3 Estimation des articulations des jambes

Nous déterminons les articulations d'une jambe par la méthode proposée au 3.2 en utilisant R_{hanche} et Pr_{pied} comme initialisations des points A et B. Ceci permet de déterminer les articulations R_{genou} , $R_{cheville}$ et R_{pied} , dont les positions sont initialisées à l'aide des points représentatifs Pr_1 , Pr_2 et Pr_3 . Notons que l'articulation R_{hanche} est redéfinie par la position du point représentatif Pr_0 de la méthode générique.

Pour déterminer l'autre jambe, on utilise la même méthode, en choisissant l'autre extrémité basse de la chaîne principale comme estimation du second pied.

Connaissant les positions des différentes articulations dans la jambe, il est possible de corriger les estimations les moins fiables en utilisant des estimations statistiques faites sur un échantillon représentatif d'êtres humains. Par exemple, la longueur de la jambe est sensiblement proche de la moitié de la hauteur totale de la personne [7].

De plus, les estimations sont plus précises lorsque la jambe est pliée. Ainsi, pour avoir la même précision sur la position des articulations des deux jambes, nous utilisons les longueurs des parties rigides de la jambe la plus pliée pour corriger la position de l'articulation R_{genou} de l'autre jambe.

3.4 Correction des positions des hanches et du bassin

Les estimations faites (voir sec. 3.2) pour les articulations du bassin et des hanches se situent au niveau de la séparation visible entre les deux jambes. Ainsi, ces estimations sont très sensibles aux vêtements que peut porter la personne.

Pourtant l'orientation générale des cuisses, fournie par les vecteurs $\overrightarrow{R_{genou_1}R_{bassin}}$ et $\overrightarrow{R_{genou_2}R_{bassin}}$, est une information exploitable. Soit ℓ_t la longueur du tibia et ℓ_p la hauteur du pied. Alors l'articulation R_{hanche} sera déplacée en utilisant l'équation 3.3.

$$R_{hanche} = R_{genou} + \frac{\overrightarrow{R_{genou}R_{hanche}}}{\overrightarrow{R_{genou}R_{hanche}}}(\ell_t + \ell_p) \quad (3.3)$$

Ceci est fait pour chacune des deux jambes, la position de R_{bassin} est ensuite déplacée au centre du segment $[R_{hanche_1}R_{hanche_2}]$.

3.5 Détermination gauche-droite

Nous venons de voir comment estimer les positions des articulations basses du corps humain, nous allons maintenant voir comment lever l'ambiguïté de choix entre les articulations du coté gauche et celle du coté droit.

Soit T le trièdre ayant $R_{cheville_1}$ pour sommet et une base définie par les articulations R_{hanche_1} , R_{pied_1} et $R_{cheville_2}$ (voir fig. 3(d)).

Si T est direct, alors les articulations indicées 1 (resp. 2) définiront des articulations sur le coté droit (resp. gauche).

Il est à noter que cette méthode nous permet de calculer $\overrightarrow{V_{face}}$, le vecteur définissant l'avant de la personne.

4 Extraction des articulations supérieures

Maintenant que nous connaissons une position (estimée) des articulations basses du corps, nous allons pouvoir traiter les articulations hautes, depuis le bassin jusqu'à la tête en incluant les bras.

Dans un premier temps (section 4.1), nous travaillerons sur les articulations du buste et de la tête. Pour cela, les points représentatifs associés aux bras et aux jambes ne nous sont d'aucune utilité (voir fig. 4(a)). C'est pourquoi il est nécessaire de réaliser un nouveau filtrage sur les points de la RC globale, pour ne considérer que les points ayant une signification particulière. Ainsi, nous définissons CH_{buste} comme étant la sous-chaîne issue de R_{bassin} et ne contenant que les points représentatifs ayant une aire maximale sur chacune des coupes (voir fig. 4(b)). Notons alors que CH_{buste} relie aux maximum un point représentatif par coupe.

Pour déterminer les articulations hautes du sujet, nous commençons par calculer la position de la tête. En utilisant cette information, nous estimons la position des deux épaules qui permettront de déterminer les articulations des bras du sujet.

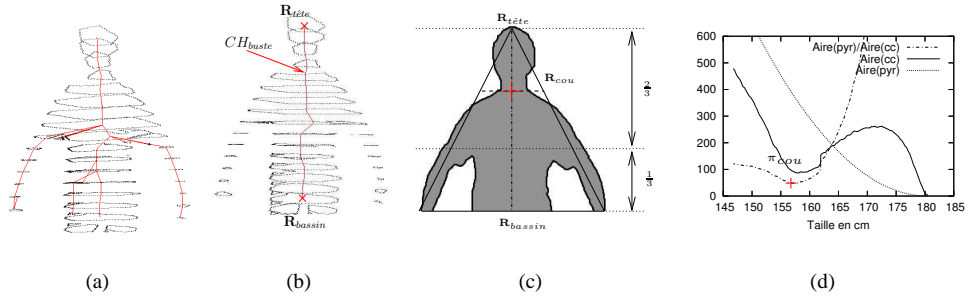


FIG. 4 – (a) chaînage principal de la partie supérieure du corps (b) sous-chaîne CH_{buste} de détermination du cou ; (c) représentation 2d de \mathbf{R}_{cou} and $\mathbf{R}_{tête}$; (d) Courbes de variation des aires décrivant la recherche de \mathbf{R}_{cou}

4.1 Détermination du cou et de la tête

Soit $\text{Pr}_{tête}$ le point représentatif terminal de la chaîne CH_{buste} , qui correspond au point observable le plus haut de la personne. Ce point sera utilisé comme estimation de l'articulation supérieure du corps $\mathbf{R}_{tête}$. Nous devons déterminer la position de l'articulation \mathbf{R}_{cou} , jonction entre la tête et le buste (ie l'un des points du cou, voir fig. 4(b)). L'ensemble I_{cou} des points représentatifs candidats pour la position du cou est défini par l'équation 4.1 où Pr est un point représentatif de la chaîne CH_{buste} et ℓ_h est la distance euclidienne entre $\mathbf{R}_{tête}$ et \mathbf{R}_{bassin} (voir fig. 4(c)).

$$I_{cou} = \left\{ \text{Pr}, \frac{2\ell_h}{3} < \overline{\mathbf{R}_{bassin}\text{Pr}} < \ell_h \right\} \quad (4.1)$$

Soit Pyr la pyramide régulière orientée suivant $\overrightarrow{\mathbf{R}_{tête}\mathbf{R}_{bassin}}$, de sommet $\mathbf{R}_{tête}$ et de base carrée (centrée sur \mathbf{R}_{bassin} et de côté ℓ_h). La position de l'articulation \mathbf{R}_{cou} sera définie par le point représentatif Pr_{cou} vérifiant le prédicat 4.2, où π_i est le plan de support sur lequel a été estimée la coupe C_i (voir fig. 4(d)).

$$\frac{A_{cou}}{\text{Aire}(\pi_{cou} \cap Pyr)} = \min_{\text{Pr}_i \in I_{cou}} \frac{A_i}{\text{Aire}(\pi_i \cap Pyr)} \quad (4.2)$$

4.2 Estimation de la position des épaules

En utilisant les mêmes estimations statistiques que lors de la correction des positions des articulations des jambes (voir sec. 3.3) ou les rapports proposés par Leonard de Vinci, nous supposons que : (1) les épaules se situent aux deux tiers de la hauteur du tronc ; (2) la largeur du corps au niveau des épaules est égal au quart de la taille de l'individu.

A partir de ces observations, nous choisissons la composante connexe ζ principale⁴ sur une coupe ayant pour plan de support $\pi_{épaule}$ parallèle au plan de référence du sol et dont l'un des points E est défini par l'équation 4.3.

$$\mathbf{E} \in [\mathbf{R}_{bassin}\mathbf{R}_{tête}] \text{ et } \overline{\mathbf{R}_{bassin}\mathbf{E}} = \frac{2}{3} \overline{\mathbf{R}_{bassin}\mathbf{R}_{tête}} \quad (4.3)$$

Le point représentatif Pr associé à la composante connexe ζ sera utilisé comme estimation de la position de l'articulation \mathbf{R}_E , point de connexion entre les épaules et le buste. Les articulations $\mathbf{R}_{épaule_g}$ ($\mathbf{R}_{é_g}$) et $\mathbf{R}_{épaule_d}$ ($\mathbf{R}_{é_d}$) seront alors calculées comme les points de l'axe principal d'inertie de ζ selon l'équation 4.4 où ℓ_H représente la hauteur totale estimée de la personne.

$$\overline{\mathbf{R}_{é_g}\mathbf{R}_{é_d}} = \frac{\ell_H}{5} \text{ et } \mathbf{R}_E = \frac{\mathbf{R}_{é_g} + \mathbf{R}_{é_d}}{2} \quad (4.4)$$

4.3 Détermination des articulations des bras

Nous estimons les articulations des bras par la méthode générique (voir sec. 3.2). Le point A est associé à l'articulation $\mathbf{R}_{épaule}$, nous devons déterminer le point P_{main} sur l'autre extrémité du bras.

La première étape consiste à trouver un ensemble de points candidats. Soit π le plan défini par \mathbf{R}_{bassin} and $\overrightarrow{V_{face} \wedge \mathbf{R}_{bassin}\mathbf{R}_{tête}}$. De la représentation condensée principale privée des points représentatifs associés aux

⁴celle ayant une aire maximale

jambes et à la tête (notée \widetilde{RC}), nous construisons deux ensembles S_g et S_d de points candidats. Soit P un point frontière de \widetilde{RC} , P appartiendra à l'ensemble S_d si il est au-dessus (dans la direction de la normale) du plan π et à S_g dans le cas contraire (voir fig. 5(a)).

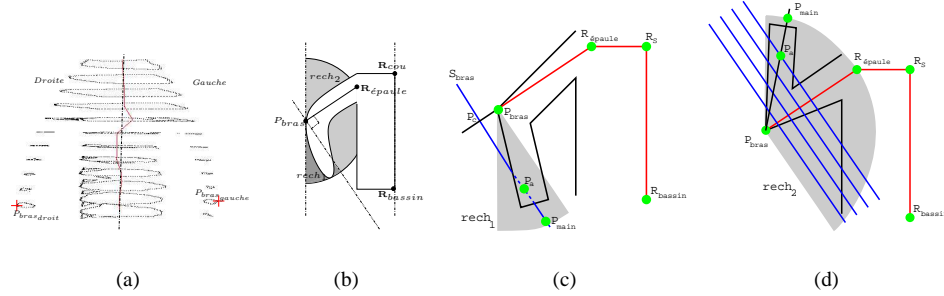


FIG. 5 – (a) ensembles des points candidats ; (b) espaces de recherche de P_a ; (c) première étape de la recherche de P_a ; (d) seconde étape

Dans la suite, nous nous attacherons à expliquer le principe de la recherche pour l'un des bras ainsi, et par souci de clarté, les indices g et d ne seront plus utilisés. Soit P_{bras} un point de S défini selon l'équation 4.5.

$$d_e(P_{bras}, (\mathbf{R}_{cou} \mathbf{R}_{bassin})) = \max_{P \in S} d_e(P, (\mathbf{R}_{cou} \mathbf{R}_{bassin})) \quad (4.5)$$

Soit $I([ab])$ la proportion du segment $[ab]$ se trouvant à l'intérieur de la forme 3d de la personne. Cette proportion peut être calculée à partir des points représentatifs de la RC.

– P_{bras} est une estimation de \mathbf{R}_{main} si il vérifie l'un des prédicats 4.6 ou 4.7.

$$\overline{P_{bras} \mathbf{R}_{épaule}} \simeq \overline{\mathbf{R}_{hanche} \mathbf{R}_{genou}} + \overline{\mathbf{R}_{genou} \mathbf{R}_{ped}} \quad (4.6)$$

$$I([\mathbf{R}_{épaule} P_{bras}]) < 0.9 \quad (4.7)$$

– Dans le cas contraire, P_{bras} est une estimation de la position du coude. L'articulation \mathbf{R}_{main} se trouvera alors dans l'un des deux espaces de recherche définis dans la figure 5(b) selon l'angle formé par le bras et l'avant-bras. Pour cela, nous cherchons un point P_a situé sur l'avant-bras.

Dans un premier temps, cette recherche s'effectue dans $rech_1$ (comme décrit dans la fig. 5(c)). Soit P_c un point de la droite $(P_{bras} \mathbf{R}_{épaule})$ défini par l'équation 4.8 où ϵ correspond à la distance minimale entre deux points de la RC.

$$P_c = \mathbf{R}_S + (1 + \epsilon) \overrightarrow{\mathbf{R}_S P_{bras}} \quad (4.8)$$

Soit C_{bras} la coupe calculée sur le plan défini par P_c et $\overrightarrow{P_{bras} \mathbf{R}_{épaule}}$. Deux cas sont possibles :

1. soit il existe sur C_{bras} une composante connexe ζ_i dont le centre de gravité G_i vérifi l'équation 4.9, alors Pr_i sera utilisé comme estimation de la position de P_a .

$$I([G_i P_{bras}]) > 0.9 \text{ et } I([G_i \mathbf{R}_{épaule}]) < 0.7 \quad (4.9)$$

2. dans le cas contraire, nous devons chercher P_a dans $rech_2$ (voir fig. 5(d)). Nous estimons des coupes successives C_i depuis le point P_{bras} , dans la direction de $\overrightarrow{P_{bras} \mathbf{R}_{épaule}}$. La première coupe nous fournissant au moins deux composantes connexes : la première $\zeta_{j,0}$ correspond au bras ; la deuxième $\zeta_{j,1}$ a son centre de gravité $G_{j,1}$ qui vérifi l'équation 4.10. Dans ce cas, $G_{j,1}$ sera utilisé comme une estimation de la position de P_a .

$$I([G_{j,1} P_{bras}]) > 0.9 \text{ et } I([G_{j,1} \mathbf{R}_{épaule}]) < 0.7 \quad (4.10)$$

Nous disposons de deux points P_{bras} and P_a définissant l'orientation moyenne de l'avant-bras. Le point P_{main} est alors calculé, en utilisant le fait que les longueurs du bras et de l'avant-bras sont sensiblement identiques, par l'équation 4.11.

$$P_{main} = P_{bras} + \frac{\overline{P_{bras} \mathbf{R}_{épaule}}}{\overline{P_{bras} P_a}} \overrightarrow{P_{bras} P_a} \quad (4.11)$$

Les articulations internes du bras (\mathbf{R}_{coude} et $\mathbf{R}_{poignet}$), peuvent maintenant être estimées par la méthode générique (voir sec. 3.2). Le point A (resp. B) est associé à $\mathbf{R}_{épaule}$ (resp. P_{main}). L'utilisation de cette méthode nous fournit \mathbf{R}_{coude} , $\mathbf{R}_{poignet}$ et \mathbf{R}_{main} qui seront estimés à partir des points représentatifs Pr_1 , Pr_2 et Pr_3 .

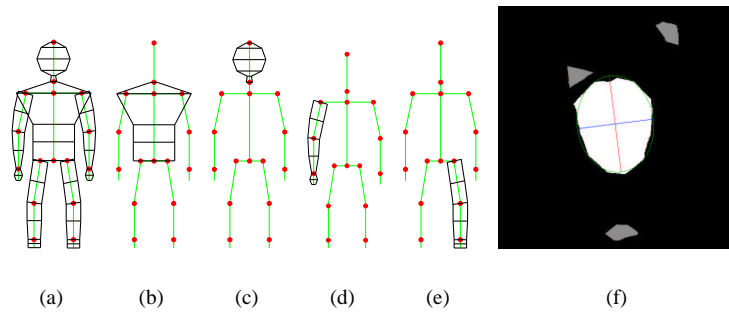


FIG. 6 – Représentation 2d de l’enveloppe volumique d’un personnage (a)(b)(c)(d)(e) sur-imprimée sur le squelette estimé à l’étape précédente ; Estimation d’une ellipse sur l’image d’une coupe (f).

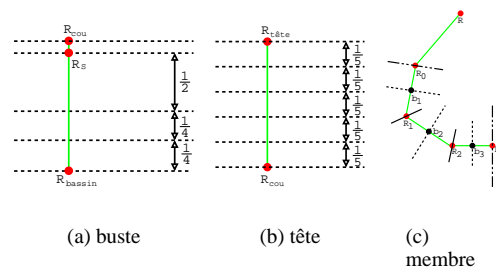


FIG. 7 – Plans de coupe utilisés pour la modélisation 3d du personnage.

5 Estimation de l’enveloppe d’un humanoïde

Nous venons de voir comment il est possible d’estimer, à partir d’un ensemble d’images, les positions des articulations d’un squelette d’animation afin de décrire la posture d’un personnage réel. Nous allons maintenant voir comment déterminer les paramètres d’un ensemble de primitives 3d qui vont définir l’enveloppe volumique du sujet. Cette enveloppe sera utilisée pour habiller le squelette d’animation lors de la phase de rendu du personnage.

5.1 Représentation de l’enveloppe

Dans les méthodes de reconstruction habituellement utilisées (voxélisation, enveloppe visuelle, ...), le modèle obtenu est très souvent décorrélé de toute information sémantique : il n’est pas possible de savoir à quelle partie du corps correspond telle ou telle facette (ou voxel). De plus, lors d’une reconstruction de mouvement à partir de séquences vidéos, les facettes (ou voxels) définissant le volume du sujet évoluent : il est impossible de suivre l’évolution d’une même facette au fil de la séquence.

Le squelette d’animation estimé à l’étape précédente est décrit par une hiérarchie d’articulations. Ces articulations sont reliées entre elles par des parties rigides. Nous avons choisi de décrire l’enveloppe volumique du sujet en représentant les volumes qui décrivent ces parties rigides. Ces volumes seront définis par un ensemble de cônes tronqués. La figure 6(a) représente l’enveloppe volumique globale du personnage, les figures 6(b) à 6(e) présentent les reconstructions des différentes parties du corps humain.

pour déterminer chaque cône tronqué nous réalisons une coupe au point p souhaité de normale \vec{n} . Sur cette coupe nous déterminons la composante connexe principale⁵ (il s’agit de la composante connexe blanche sur la figure 6(f)). Cette composante connexe décrit la partie du modèle que nous souhaitons reconstruire et peut être approchée par une ellipse centrée sur le centre de gravité de la composante connexe et dont les dimensions sont définies par celles de son ellipse d’inertie.

5.2 Reconstruction de l’enveloppe volumique

L’estimation du modèle géométrique représentant notre personnage se fait alors en 3 temps :

⁵la composante connexe dont le centre de gravité est le plus proche du centre de l’image associée à cette coupe.

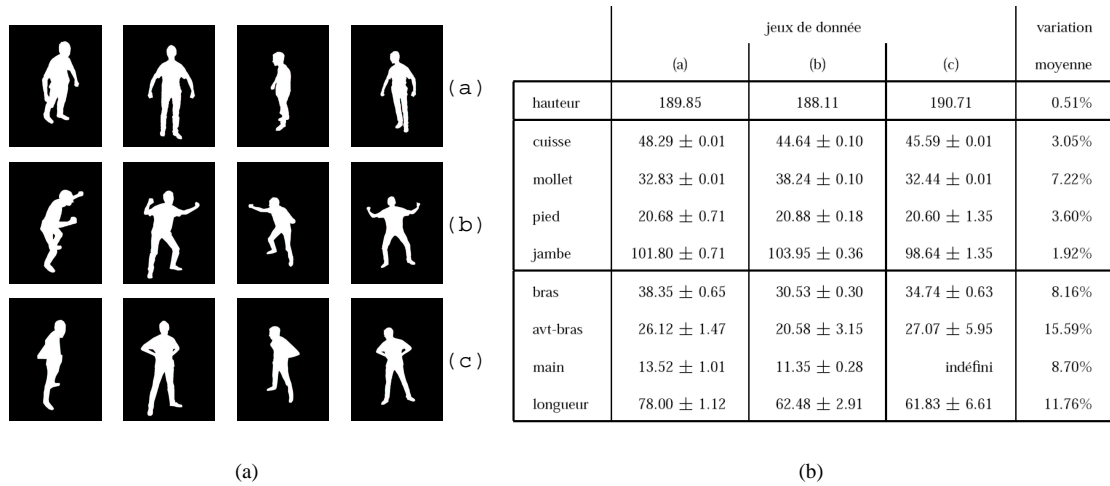


FIG. 8 – Silhouettes de trois différents jeux de données (a); Mesures anthropométriques estimées (données en *cm*)(b)

- Le buste est modélisé par 4 cônes généralisés (voir figure 6(b)) présenté plus en détail figure 7(a).
- La tête est modélisée par 5 cônes (voir figure 6(c)) dont les bases sont estimées sur 6 coupes régulièrement espacées entre l’articulation R_{cou} et l’articulation $R_{tête}$ et parallèles au plan au sol (voir figure 7(b)).
- La modélisation géométrique des différents membres (jambes et bras) se fait de manière générique. Chaque membre est décrit par quatre articulations R_i reliées entre elles par les parties rigides décrivant les sections du membre (cuisse, mollet et pied dans le cas d’une jambe) et liées au buste par une cinquième articulation R (voir figure 7(c)).

6 Résultats

La méthode proposée a été appliquée à différents jeux de données réels (voir fig 8(a)). Tous les calculs ont été réalisés sur un PIV 2.8 Ghz avec 512Mo de RAM et une carte vidéo NVIDIA GeForce FX 5900.

6.1 Estimation de la posture

Les résultats de l’estimation de la posture sont décrits dans la fig. 9. La deuxième colonne montre la reprojction du squelette d’animation sur les masques binaires de silhouette. Comme on peut le voir, les estimations donnent de bons résultats quant à l’orientation des parties rigides du corps, et ce même si les mains sont en contact avec le corps (dans le troisième jeu de donnée).

Tous ces jeux de donnée sont acquis à partir d’une même personne. Pour étudier la validité de notre méthode, nous comparons les différentes mesures anthropométriques obtenues sur ces jeux de donnée. Les résultats sont donnés dans le tableau 8(b).

Comme nous pouvons le voir, les mesures sur les jambes sont très stables (moins de 1% de variation). La différence maximale est obtenue sur les bras. Ceci s’explique par le fait que les bras sont fort peu contraints en terme de position par rapport au corps et la disposition des caméras dans la pièce (laissant certains ”angles morts”).

6.2 Estimation de l’enveloppe volumique

Dans la deuxième partie de notre méthode, nous estimons l’enveloppe volumique de la personne à partir de cônes généralisés. Des résultats sont donnés dans la figure 10(a), où nous comparons notre méthode à une reconstruction voxelique. Les deux modèles sont visualisés suivant les positions et orientations des caméras d’origine. Nous constatons qu’il existe peu de différence entre la reconstruction voxelique et les images originales. Notre méthode présente quelques variations, essentiellement situées au niveau des jonctions entre les membres et le buste. Ceci s’explique par le faible nombre de cônes utilisés ainsi que pas le fait que nous approchons une forme quelconque (les composantes connexes) à l’aide d’une ellipse.

L’un des avantages de notre approche, par rapport à la reconstruction voxelique, est décrit dans la figure 10(b), où les deux modèles sont visualisés depuis un nouveau point de vue. Nous constatons que les artefacts de reconstruction qui apparaissent dans le cas de la reconstruction voxelique ne sont pas présents dans notre méthode. De plus,

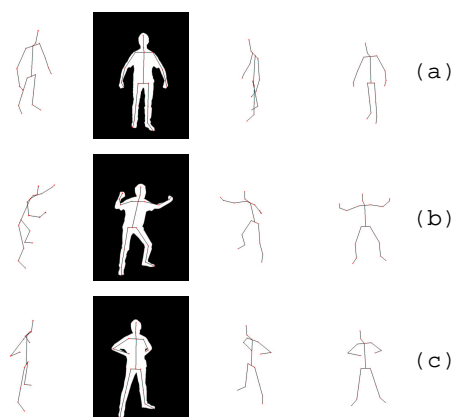


FIG. 9 – Visualisation des squelettes d’animation estimés sur les jeux de données de la fig. 8(a).

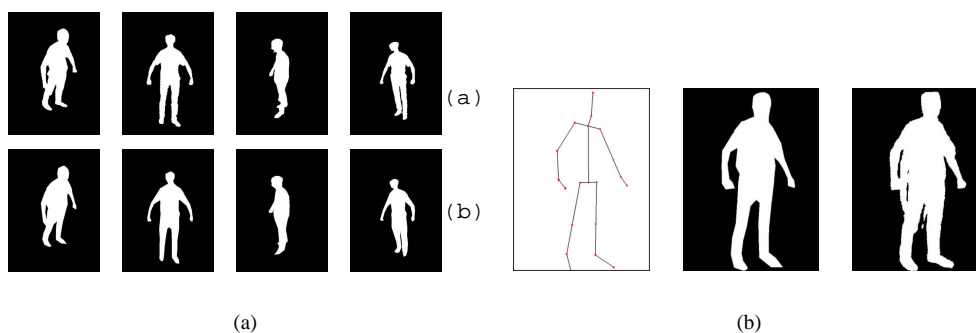


FIG. 10 – (a) Enveloppe volumique (pour le jeux (a) de la fig. 8(a)) : (a) obtenue par reconstruction voxelique de résolution 256^3 , temps de calcul : 580 ms ; (b) en utilisant notre méthode, temps de calcul : 44 ms. Rendu depuis un nouveau point de vue(b).

de part sa nature, notre approche permet de donner une signification sémantique aux différentes parties du modèle 3d obtenu.

7 Conclusion et perspectives

Nous avons présenté une méthode qui permet d’estimer de façon automatique la posture d’un personnage à partir de plusieurs vues. A l’issue de la première étape, nous obtenons un véritable squelette d’animation (conforme au standard H-anim). Un habillage par des primitives géométriques permet d’obtenir une forme volumique. La méthode proposée n’impose aucune contrainte au personnage filmé, en dehors du fait qu’il doit être en équilibre stable. La construction se fait en temps interactif (120ms) à partir de quatre caméras calibrées sur une architecture ”basique” (un seul PC). Les résultats obtenus sur des cas réels ont permis de valider la méthode. Telle quelle est développée, notre méthode n’est capable de traiter que des personnages séparés, nous travaillons actuellement sur une amélioration de la méthode pour pouvoir traiter le cas de plusieurs personnes présents dans le même espace et éventuellement en contact.

Références

- [1] J.K. Aggarwal and Q. Cai. Human motion analysis : a review. *Proc. IEEE Nonrigid and Articulated Motion Workshop*, pages 90–102, June 1997.
- [2] T.J Cham and J. Rehg. A multiple hypothesis approach to figure tracking. *Proc. Computer Vision and Pattern Recognition*, 2 :239–245, 1999.

- [3] K. M. Cheung, T. Kanade, J. Bouguet, and M. Holler. A real time system for robust 3d voxel reconstruction of human motions. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 714–720, 2000.
- [4] Kong Man Cheung, Simon Baker, and Takeo Kanade. Shape-from-silhouette across time part i : Theory and algorithms. *International Journal of Computer Vision*, 62(3) :221 – 247, May 2005.
- [5] B.Triggs C.Sminchisescu. Estimating articulated human motion with covariance scaled sampling. *International Journal of Robotics Research*, 22(6) :371 – 393, 2003.
- [6] Q. Delamarre and O.D. Faugeras. 3d articulated models and multiview tracking with physical forces. *Computer Vision and Image Understanding*, 81(3) :328–357, March 2001.
- [7] H. Dreyfuss and A. R. Tilley. *The Measure of Man and Woman : Human Factors in Design*. John Wiley & Sons, 2001.
- [8] D.M. Gavrilu. The visual analysis of human movement : a survey. *Computer Vision and Image Understanding*, 73(1) :82–98, January 1999.
- [9] Jean-Marc Hasenfratz, Marc Lapierre, Jean-Dominique Gascuel, and Edmond Boyer. Real-time capture, reconstruction and insertion into virtual world of human actors. In *Vision, Video and Graphics*, pages 49–56. Eurographics, Elsevier, 2003.
- [10] A. Hilton, D. Beresfoed, Thomas Gentils, R. Smith, W. Sun, and John Illingworth. Whole-body modelling of people from multiview images to populate virtual worlds. *Visual Computer*, 16(1) :411–436, 2000.
- [11] D. Hogg. Model-based vision : a program to see walking person. *Images Vision Computing*, 1 :5–20, 1983.
- [12] I. Kakadiaris and D. Metaxas. Model-based estimation of 3d human motion. *Trans. on PAMI*, 22(12) :1453–1459, 2000.
- [13] I. Mikic, Mohan Trivedi and E. Hunter, and P. Cosman. Muman body model acquisition and tracking using voxel data. *International Journal of Computer vision*, 53(1) :199–223, 2003.
- [14] T. Moeslund and E. Granum. A survey of computer vision-based human motion capture. *Computer vision and Image Understanding*, 81 :231–268, 2004.
- [15] D. Ormoneit, H. Sidenbladh, M.J. Black, T. Hastie, and D.J. Fleet. Learning and tracking human motion using functional analysis. *Proc. IEEE workshop Human Modeling Analysis and Synthesis*, pages 2–9, 2000.
- [16] A. Pentland and B. Horowitz. Recovery of non-rigid motion and structure. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(7) :730–742, July 1991.
- [17] Y. Perret, S. Bouakaz, and T. Excoffier. Real object parameters tracking : a geometric model based approach. In *12th Scandinavian Conference on Image Analysis, Bergen, Norway*, pages 501–507, June 2001.
- [18] R. Plänkers and P. Fua. Articulated soft objects for multi-view shape and motion capture. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10) :63–83, 2003.
- [19] J.M. Rehg and T. Kanade. Model-based tracking of self -occluding articulated objects. *Proc. 5th ICCV*, pages 612–617, 1995.
- [20] Mark Segal and al. Fast shadows and lighting effects using texture mapping. In *Proc. Siggraph'92*, pages 249–252, 1992.
- [21] Y. Yacoob and M.J. Black. Parametrized modeling and recognition of activities. *Computer Vision and Image Understanding*, 73(2) :232–247, 1999.