



**HAL**  
open science

## Recalage visuel de modèles 3D maxillo-mandibulaires

Raphaël Destrez, Benjamin Albouy-Kissi, Sylvie Treuillet, Yves Lucas,  
Arnaud Marchadier

► **To cite this version:**

Raphaël Destrez, Benjamin Albouy-Kissi, Sylvie Treuillet, Yves Lucas, Arnaud Marchadier. Recalage visuel de modèles 3D maxillo-mandibulaires. Grets2011, Sep 2011, Bordeaux, France. hal-00648737

**HAL Id: hal-00648737**

**<https://hal.science/hal-00648737v1>**

Submitted on 6 Dec 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Recalage visuel de modèles 3D maxillo-mandibulaires

RAPHAEL DESTREZ<sup>1,3</sup>, BENJAMIN ALBOUY-KISSI<sup>2</sup>, SYLVIE TREUILLET<sup>1</sup>,  
YVES LUCAS<sup>1</sup>, ARNAUD MARCHADIER<sup>3</sup>

<sup>1</sup> Laboratoire PRISME, Polytech'Orléans, 12 rue de blois, 45067 Orléans Cedex 2, France

<sup>2</sup> Laboratoire ISIT, Université d'Auvergne, IUT, BP 219, 43006 Le Puy en Velay cedex, France

<sup>3</sup> Useful Progress, 23 rue d'Anjou, 75008, Paris, France

<sup>1</sup>[Raphael.Destrez@etu.univ-orleans.fr](mailto:Raphael.Destrez@etu.univ-orleans.fr), [Sylvie.Treuillet@univ-orleans.fr](mailto:Sylvie.Treuillet@univ-orleans.fr), [Yves.Lucas@bourges.univ-orleans.fr](mailto:Yves.Lucas@bourges.univ-orleans.fr)

<sup>2</sup>[Benjamin.Albouy@iut.u-clermont1.fr](mailto:Benjamin.Albouy@iut.u-clermont1.fr),

<sup>3</sup>[Arnaud.Marchadier@gmail.fr](mailto:Arnaud.Marchadier@gmail.fr)

**Résumé** - Dans cet article, nous comparons des méthodes de recalage de modèles 3D par vision pour développer un outil essentiel en orthodontie : la mise en occlusion virtuelle des deux moulages des mâchoires. Pour cela, on utilise des photos de la bouche du patient et des points mis en correspondance entre les images et les modèles dentaires surfaciques. A partir d'un ensemble de correspondances 2D/3D disponibles établies sur les deux arcades, on calcule la matrice de projection, puis on effectue le recalage de la mandibule sous le maxillaire en estimant une transformation rigide. Deux approches sont comparées selon que l'on dispose ou non de la connaissance des paramètres intrinsèques de la caméra. Le recalage de la mandibule est ensuite estimé en minimisant l'erreur de reprojection. Les tests sont effectués sur des données virtuelles puis des images réelles. A partir d'un cas virtuel « parfait », nous évaluons la dégradation des performances avec l'ajout de bruit et l'intérêt de combiner plusieurs vues. Les matrices de projections estimées sont évaluées par les erreurs de reprojection et les recalages par les écarts constatés sur les 6 paramètres de la transformation rigide par rapport à un positionnement de référence.

**Abstract** - In this paper, we compare two vision registration methods of 3D models to develop a main tool for orthodontics based on the virtual occlusion of two dental casts. For that purpose, we process photos of the patient mouth and match points between pictures and dental geometrical models. From a set of 2D/3D matches of the two arcades, we calculate the projection matrix, and after we do the registration of the mandible under the maxillary through a rigid transformation. Two methods with or without the knowledge of camera intrinsic parameters are compared. Minimizing the reprojection errors, we perform the registration of the mandible. Tests are carried out on virtual data and real images. Using a virtual case, assumed as perfect, we evaluate the robustness against noise and the increase of performance using several views. Projection matrices and registration are evaluated, respectively, by reprojection errors and the recorded differences on the 6 parameters of rigid transformation compared with the reference pose.

## 1 Problématique

L'orthodontie est une spécialité dentaire dont l'objectif est de corriger les mauvaises positions des dents et d'aider au bon développement des bases osseuses dans un but esthétique ou fonctionnel (optimiser l'occlusion et offrir un contact efficace pour la mastication). Pour faire un diagnostic et planifier les soins [1], les orthodontistes utilisent encore massivement des modèles en plâtre obtenus par prise d'empreinte des deux arcades, parfois complétés par des photos prises en bouche grâce à des écarteurs (figure 1 **Erreur ! Source du renvoi introuvable.**). S'ils permettent au spécialiste une bonne observation et manipulation, ces modèles en plâtres présentent de sérieux inconvénients : stockage, fragilité lors des transports entre les différents intervenants, et surtout impossibilité de les inclure au dossier patient informatisé. Une nouvelle approche propose d'exploiter des modèles virtuels après numérisation des moulages par scanner laser ou tomographie X [2]. Ces techniques permettent d'obtenir des modèles

surfaciques d'excellente précision (de l'ordre de 100 mm) mais, chaque arcade étant scannée séparément, l'ajustement des modèles virtuels de la mandibule et du maxillaire nécessite de longues et délicates opérations manuelles par un technicien spécialisé. D'un autre côté, les techniques d'imagerie médicale qui permettraient d'accéder directement à la numérisation 3D des deux arcades en occlusion sur le patient (scanner X ou CT-scan) sont trop lourdes et trop coûteuses pour envisager une utilisation de routine. Notre objectif est d'apporter une solution automatique de mise en occlusion des modèles virtuels des arcades numérisées pour rendre cette procédure plus rapide, plus précise et plus indépendante de l'opérateur. L'approche proposée s'appuie sur les photos couleurs prises en bouche du patient, qui représentent la véritable occlusion « in vivo », selon plusieurs points de vue pour faire un recalage des modèles maxillo-mandibulaires. Le recalage visuel doit s'affranchir de plusieurs difficultés liées aux prises de vue (poses des caméras inconnues, calibrage partiel,...) ou à la

recherche de correspondances entre modèles 3D surfaciques et images couleurs.

## 2 Recalage par vues externes

Comme les modèles surfaciques 3D des arcades sont obtenus dans des repères différents, le recalage nécessite de déterminer la transformation rigide mettant correctement en occlusion la mandibule et le maxillaire. Cette transformation peut être estimée à partir d'un ensemble de correspondances entre les arcades.

Le recalage 2D et 3D est un problème qui a grandement alimenté la littérature scientifique (détection de primitive : sift, surf, Harris ; mise en correspondance ; algorithme ICP)[3][4]. Le cas particulier du recalage 3D à partir d'images 2D est principalement présent dans l'imagerie médicale [5] mais aussi dans d'autres domaines comme le suivi de mouvement[6][7].

Une première approche est donc de s'appuyer sur les images couleurs disponibles de la bouche du patient et d'un ensemble de correspondances 2D/3D disponibles sur les deux arcades. Plusieurs vues peuvent être utilisées simultanément pour le recalage. En s'appuyant sur un modèle de formation des images de type sténopé, le problème peut se résoudre en deux étapes: d'abord estimer la matrice de projection liée à chaque vue en s'appuyant sur le maxillaire, puis effectuer le recalage de la mandibule par des techniques analytiques ou de minimisation.

### 2.1 Estimation des matrices de projection

Pour estimer la matrice de projection, deux cas de figure se présentent selon que l'appareil photo est calibré, ou non. S'il est calibré, il s'agit de déterminer sa pose (3 angles d'Euler définissant l'orientation  $R_i$  et 3 composantes de translation  $t_i$ ), soit 6 ddl. Dans le cas contraire, c'est la matrice de projection dans sa globalité qui est à estimer, soit 11 ddl. En utilisant les coordonnées homogènes, chaque point  $M_j$  de l'espace est relié à ces projections  $m_{ij}$  dans chaque image  $i$  par la relation:

$$m_{ij} \propto P_i M_j \quad (1)$$

où  $\propto$  signifie « égal à un facteur d'échelle près ». La matrice de projection  $P_i$  de taille  $3 \times 4$  peut se décomposer en un produit de la matrice de calibrage  $K$  contenant les paramètres intrinsèques de l'appareil avec la matrice de pose (combinaison de rotation  $R_i$  et translation  $t_i$ ):

$$P_i = K[R_i | t_i] \quad (2)$$

Pour les appareils photographiques actuels, on peut considérer les pixels carrés et le point principal au centre des images, sans grande incidence sur la précision. Cette simplification permet de construire la matrice  $K$  à partir de la taille de l'image et de la distance focale exprimée en pixels. Pour l'estimation de la matrice de projection, on distingue donc deux cas de figure. Lorsque la focale est connue, l'algorithme

POSIT de Dementhon [8] permet une première estimation des 6 paramètres de pose, qui sont ensuite optimisés par la méthode de Levenberg-Marquardt [9], en minimisant l'erreur quadratique de reprojection pour  $n_i$  points localisés sur le maxillaire:

$$\sum_{j=1}^{n_i} [m_{ij} - \tilde{m}_{ij}]^2 \quad (3)$$

$\tilde{m}_{ij}$  désigne la reprojection par la matrice de projection  $P_i$  d'un point du modèle 3D et  $m_{ij}$  le point correspondant détecté dans l'image  $i$ . Lorsque la focale est inconnue, la matrice de projection  $P_i$  est estimée par minimisation au sens des moindres carrés du système d'équations issues de l'ensemble des correspondances 2D/3D du maxillaire, en fixant le dernier élément à 1. Ensuite, les 11 paramètres de la matrice  $P_i$  sont optimisés par l'algorithme de Levenberg-Marquardt comme précédemment.

### 2.2 Recalage maxillo-mandibulaire

Le repère de référence lié à chaque arcade a pour origine le centre de gravité du nuage de points, l'axe Y orienté vers le haut et l'axe Z parallèle aux deux rangées de molaires. Le recalage maxillo-mandibulaire consiste à calculer une transformation rigide  $T$ , soit 6 ddl supplémentaires (3 angles d'Euler  $\alpha$ ,  $\beta$  et  $\gamma$  et 3 translations selon X, Y et Z):

$$m_{i,mand} \propto P_i T M_{mand} \quad (4)$$

où  $M_{mand}$  est un point 3D de la mandibule et  $m_{i,mand}$  sa projection sur la photographie  $i$ . La transformation  $T$  est estimée en recherchant la meilleure correspondance entre des points localisés sur le modèle 3D de la mandibule et leurs projections dans les photos couleurs. Comme précédemment, l'équation 3 est minimisée par l'algorithme de Levenberg-Marquardt à partir de l'identité (figure 2) après une initialisation par une légère translation.



Figure 1 : Images réelles et virtuelles des vues de face et de profil

## 3 Expérimentations et résultats

L'objectif des expérimentations est d'évaluer la précision des deux approches. Les tests comparatifs sont faits sur des données virtuelles et des images réelles. Sur les modèles 3D issus de la numérisation d'arcades de

deux patients, dix points ont été sélectionnés manuellement sur chaque mâchoire.

Des captures d'écran du rendu 3D réalisé sous VTK [10] permettent d'obtenir des images virtuelles qui sont sans distorsions mais avec des caractéristiques identiques aux photos réelles (angle de vue, focale, etc...). Des images virtuelles peuvent ainsi être créées selon plusieurs points de vue en maîtrisant les caractéristiques de la prise de vue (figure 1). A partir de ces vues, on peut alors obtenir les projections parfaites sur les images des points 3D saisis et ainsi avoir des mises en correspondances 2D/3D parfaites.

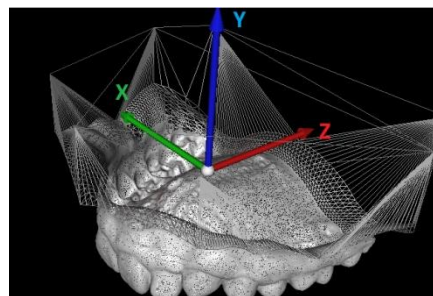


Figure 3 : Repère 3D de référence

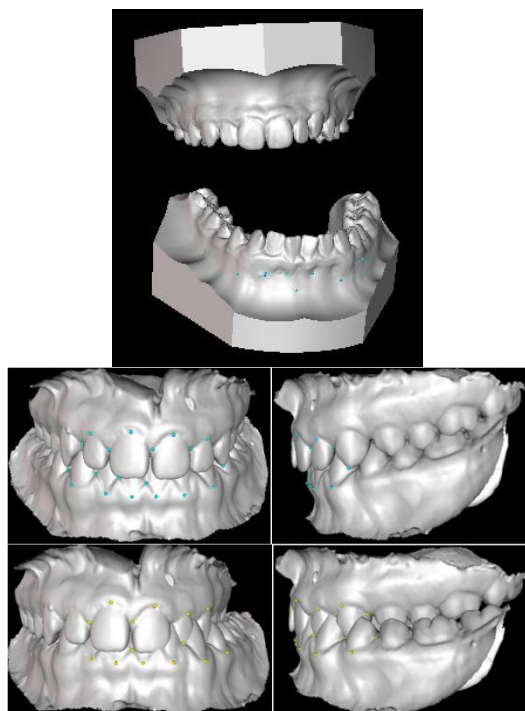


Figure 2 : Positionnement initial des deux arcades et mises en occlusion obtenues à partir des images réelles avec la focale calibrée (en haut) et inconnue (en bas).

De ce « cas parfait », on peut contrôler le niveau de bruit sur les coordonnées des points image. Les images réelles ont été prises à partir d'un appareil photo numérique Canon EOS 350D avec un objectif macro de 60 mm et un flash annulaire. Le recalage est réalisé à partir d'une même configuration initiale des arcades (figure 2), en s'appuyant sur une des deux vues (face ou profil) ou les deux vues simultanément.

Il est évalué par l'erreur quadratique moyenne en pixels mesurée dans les images et par les écarts constatés sur les six paramètres de la transformation rigide (trois angles d'Euler et trois translations). Le positionnement de référence a été fourni par l'expert orthodontiste.

Le repère 3D de référence est orienté et placé comme sur la figure 3. Le repère propre à la mandibule est orienté de manière identique et placé à son centre de gravité. Il faut noter que les écarts constatés sur les paramètres du recalage sont dépendants du repère propre à la mandibule.

Une partie des résultats sont présentés dans les tableaux 1 et 2. Pour les données virtuelles, les erreurs sont moyennées sur 100 réalisations aléatoires de bruit uniforme. La sélection manuelle pratiquée a également été testée en moyennant cinq essais répétitifs.

Tab 1 : Erreur quadratique moyenne de reprojection (en pixels) sur les points de la mandibule après recalage. Pour les données virtuelles, les coordonnées exactes des points dans l'image sont perturbées par un bruit uniforme de 2 pixels ou sélectionnés manuellement comme dans les images réelles.

<b>Virtuel bruité (2 pixels)</b>	<b>Focale calibrée</b>		<b>Focale inconnue</b>	
Vue de face	1,35		1,82	
Vues de profil	1,35		1,59	
2 vues simultanées	1,64		3,06	
<b>Sélection manuelle</b>	<b>Virtuel</b>	<b>Réel</b>	<b>Virtuel</b>	<b>Réel</b>
Vue de face	3,24	5,78	4,50	9,66
Vue de profil	3,80	4,50	4,65	4,83
2 vues simultanées	4,07	6,16	8,71	10,77

Tab 2 : Ecart après recalage sur les six paramètres de la transformation rigide (en degrés et mm) en s'appuyant sur les deux vues simultanées (face et profil). Pour les données virtuelles, les coordonnées exactes des points dans l'image sont perturbées par un bruit uniforme de 2 pixels ou sélectionnés manuellement comme dans les images réelles.

<b>Virtuel bruité (2 pixels)</b>	<b>Focale calibrée</b>		<b>Focale inconnue</b>	
$\alpha$	0,19		0,43	
$\beta$	0,22		0,42	
$\gamma$	0,08		0,26	
X	0,05		0,10	
Y	0,02		0,08	
Z	0,10		0,16	
<b>Sélection manuelle</b>	<b>Virtuel</b>	<b>Réel</b>	<b>Virtuel</b>	<b>Réel</b>
$\alpha$	0,60	1,26	1,36	1,20
$\beta$	0,33	0,95	0,75	3,74
$\gamma$	0,48	0,27	0,32	1,77
X	0,09	0,27	0,37	0,94
Y	0,17	0,56	0,23	1,27
Z	0,11	0,14	0,38	0,58

Les erreurs augmentent de façon linéaire avec le bruit. Les tableaux présentent le cas d'un bruit de 2 pixels.

L'utilisation de plusieurs vues dans la minimisation s'accompagne d'une augmentation de l'erreur finale dans les images due à un compromis entre les vues. Cependant elle permet la convergence et donne de meilleurs résultats observés sur les six paramètres du recalage.

La comparaison des deux méthodes (focale calibrée ou inconnue) est en faveur d'un calibrage de l'appareil. Cependant, les tests répétitifs montrent que les résultats sont assez sensibles à la sélection manuelle (équivalente à un bruit de 5 pixels environ). Les correspondances sont plus incertaines dans le cas des images réelles qui présentent des caractéristiques visuelles différentes des modèles surfaciques et pourraient être aidées par un plaquage de texture.

L'influence de la position des points est en cours d'étude, ainsi que l'impact des distorsions.

L'étape suivante des travaux est la détection et la mise en correspondance automatique de points et/ou courbes singulières s'appuyant sur des descripteurs robustes combinant les modèles surfaciques 3D et les textures couleur 2D. D'autres contraintes pourront également être prises en compte dans la minimisation en complément de l'erreur de reprojection, comme des distances entre points clé sur les arcades.

## Bibliographie

- [1] Ettorre, G. & al, Standards for digital photography in cranio-maxillo-facial surgery - Part I: Basic views and guidelines. *Journal of Cranio-Maxillofacial Surgery*, 34, pp.65-73 (2006).
- [2] <http://www.ortho-concept3d.com/>
- [3] Sugaya, Y. & Kanatani, K. , Extracting moving objects from a moving camera video sequence , *Mem. Fac. Eng. Okayama Univ*, 39, pp. 279–284 (2004)
- [4] Fitzgibbon, A. W., Robust registration of 2D and 3D point sets , *Image and Vision Computing*, 21(13-14), pp. 1145-1153( 2003)
- [5] Zhu, Y.-M. & Cochoff, S. M., Influence of Implementation Parameters on Registration of MR and SPECT Brain Images by Maximization of Mutual Information , *Journal of Nuclear Medicine*, 43(2), pp. 160 -166( 2002)
- [6] Henia, O. B., Hariti, M. & Bouakaz, S. , A two-step minimization algorithm for model-based hand tracking , 18th International Conference on Computer Graphics, Visualization and Computer Vision (WSCG)( 2010)
- [7] Ouhaddi, H. & Horain, P., 3D Hand Gesture Tracking by Model Registration , *Proc. IWSNHC3DI'99* (70–73) (1999)
- [8] Dementhon, D.F. & al, Model-based object pose in 25 lines of code, *Int. J. of Computer Vision*, 15(1), pp.123-141 (1995)
- [9] Marquardt, D., An algorithm for least-squares estimation of non linear parameters, *J. of Applied Math*, 11(2), pp.431-441 (1963)
- [10] <http://www.vtk.org/>