

Détection de régions pathologiques du massif facial à partir d'images 3D CT par une approche *a contrario* étendue

F. MKADMINI¹, S. FAISAN², N. KHLIFA¹, F. ROUSSEAU², K. HAMROUNI¹, C. MEYER³

¹UR-Traitement du Signal, Traitement des Images, Reconnaissance des Formes - ENIT
BP 37, Le Belvedere, 1002 Tunis, Tunisie

²Laboratoire des Sciences, de l'Image, de l'Informatique et de la Télédétection, UMR CNRS/UDS 7005
Pôle API, Bd Sébastien Brant, BP10413, 67412 Illkirch, France

³ Centre Hospitalier Universitaire de Besançon,
Hôpital Jean Minjot, 3 bd Alexandre Fleming, 25000 Besançon, France*

mkadmini.firas@hotmail.com, faisan@lsiit.u-strasbg.fr, khelifa_nawres@yahoo.com,
rousseau@lsiit.u-strasbg.fr, kamel.hamrouni@enit.rnu.tn, et
christophe.meyer@chu-besancon.fr

Résumé – Si les approches *a contrario* sont fortement reliées au cadre classique des tests d'hypothèse, elles présentent également le bénéfice de permettre la prise de décision par région en fusionnant l'information provenant de plusieurs mesures et différents seuils de significativité. L'apport principal concerne ici l'extension de ces approches à la comparaison individu vs groupe. Dans ce cadre, l'approche développée offre de nombreux avantages sur les approches traditionnelles de comparaison : certains sont directement hérités des approches *a contrario* standards alors que d'autres sont obtenus grâce à l'existence du corpus d'apprentissage (le groupe) qui permet l'estimation fidèle d'un certain nombre de paramètres et l'utilisation de règles de fusion quelconques (*i.e.* adaptables suivant l'application). Les premiers résultats obtenus dans le cadre de la détection de régions pathologiques du massif facial montrent l'intérêt de la méthode et ouvrent de nouvelles perspectives méthodologiques.

Abstract – The *a contrario* framework is related to hypothesis testing and presents major advantages over traditional methods: a decision is taken for a region using simple fusion rules from several measures and thresholds. This framework is here extended to deal with single-subject vs group analysis. In this context, we show that the *a contrario* approach has a very nice property when a learning set is available, namely, the possibility to estimate different distributions without using poor approximations as in the classical *a contrario* approach. Preliminary results obtained in the context of pathological region detection of the facial maxillary demonstrate the interest of the method and open new mathematical insights into the *a contrario* approach.

1 Introduction

Le bilan des disharmonies crano-faciales dans le cadre de la chirurgie orthognatique reste encore largement basé sur des analyses céphalométriques 2D. Ces analyses fastidieuses nécessitent le placement de points anatomiques souvent difficiles à repérer et ne donnent aucune information volumique. Dans ce contexte, nous présentons ici un travail relatif à la création d'un modèle statistique de référence 3D du massif crano-facial permettant de faciliter l'analyse des disharmonies par comparaison automatique à ce modèle ou atlas.

Les atlas peuvent être classés suivant deux grandes familles. Tout d'abord, les méthodes univariées considèrent les voxels mutuellement indépendants : aussi bien l'estimation des paramètres du modèle que la comparaison sujet vs groupe s'effectuent voxel par voxel. Le modèle linéaire général est souvent utilisé dans de telles approches [1]. Enfin, les méthodes multivariées prennent en considération la corrélation qui peut exister entre les voxels. Il n'y a plus un modèle pour chaque voxel mais un unique modèle pour l'ensemble de l'image. La stratégie de ces approches consiste à réduire la dimensionnalité du

problème en plongeant les données d'origine dans un espace de plus petite dimension (analyse en composantes principales, isomap [2]...). La comparaison entre le sujet et le groupe peut ensuite être réalisée pour l'ensemble de l'image en effectuant l'analyse statistique dans l'espace de petite dimension ou être réalisée voxel par voxel en passant au préalable par une étape de reconstruction.

Dans le présent travail, la comparaison ne peut être effectuée ni voxel par voxel, ni pour l'image toute entière, en raison de l'application chirurgicale sous-jacente : les chirurgiens voudraient connaître la normalité du sujet dans des régions spécifiques où ils peuvent intervenir. La comparaison statistique doit donc être réalisée région par région. Compte tenu du nombre limité de sujets à notre disposition et de la complexité du massif facial, l'utilisation d'une approche globale pour chaque région s'avère difficile (l'estimation de la relation entre l'espace réduit et l'espace d'origine nécessite un nombre important d'exemples). Il est à noter qu'un atlas du massif facial a été obtenu dans [3] avec un groupe de 6 sujets. Cependant, les données ont été réduites avant même la phase d'apprentissage en extrayant des lignes de crête. Au contraire, nous proposons une méthode basée région qui ne nécessite pas de réduire la dimensionnalité du problème. La méthode proposée

*Ce travail a été soutenu par le projet CMCU franco-tunisien (08G115) de collaboration entre le LSIIT et l'UR-TSTIRF

repose sur l'approche *a contrario* [4] qui est étendue ici à la comparaison individu vs groupe.

Le papier est organisé de la manière suivante : la méthode proposée est décrite dans la section 2, des résultats préliminaires sont données section 3. Des perspectives et la conclusion sont finalement apportées section 4.

2 Méthode proposée

2.1 L'approche *a contrario*

Dans l'approche *a contrario*, le modèle d'observation que nous désirions rejeter est le modèle *a contrario*. Il correspond à l'hypothèse \mathcal{H}_0 dans le cadre classique des tests d'hypothèse et on utilisera indifféremment par la suite les termes modèle *a contrario* et hypothèse \mathcal{H}_0 .

L'approche *a contrario* est une approche *région* qui se base sur la détection d'événements rares. Suivant l'application, ils peuvent être associés à la détection de mouvement dans une scène [5], à la détection d'alignement dans une série de points [4], etc. Dans tous les cas, ces événements rares sont définis comme étant des événements de faible probabilité d'occurrence sous \mathcal{H}_0 . Rejeter \mathcal{H}_0 au voxel v si un événement rare y est observé peut conduire à un nombre important de fausses alarmes en raison du problème de comparaisons multiples (le nombre de voxels pouvant être très grand). Pour contourner cette limitation, \mathcal{H}_0 est rejetée dans l'approche *a contrario* pour une région si le nombre d'événements rares y est important.

Ceci se formalise de la manière suivante. Soient X une variable aléatoire et x sa réalisation à chaque voxel. Le modèle est ici défini de telle sorte que la mesure X doit être importante pour rejeter le modèle *a contrario*. Soient :

- $\mu_j, j = 1, \dots, N_\mu$ un ensemble de seuils (un seuil permet de définir la rareté d'un événement : l'observation au voxel v d'une mesure plus grande que μ_j sera considérée comme un événement rare pour le seuil μ_j),
- R , une région composée de n voxels indépendants,
- k_j le nombre d'événements rares observés dans une région R pour le seuil μ_j .

Il est à noter que la significativité d'un événement rare est directement reliée au seuil μ_j et à la distribution de X sous \mathcal{H}_0 . Le modèle *a contrario* est rejeté pour une région R si plusieurs événements rares y sont observés. Par conséquent, on considère l'évènement : $E_{\mu_j, R} = \{ \text{au moins } k_j \text{ événements rares sont observés dans } R \text{ pour le seuil } \mu_j \}$. La probabilité de l'évènement $E_{\mu_j, R}$ correspond à la queue de la distribution binomiale :

$$P(E_{\mu_j, R}) = \sum_{k=k_j}^n C_j^k P(X > \mu_j)^k (1 - P(X > \mu_j))^{n-k} \quad (1)$$

En considérant N_R régions, une région R est dite ε -significative si

$$N_R \cdot N_\mu \cdot \min_j P(E_{\mu_j, R}) < \varepsilon. \quad (2)$$

La quantité $N_R \cdot N_\mu \cdot \min_j P(E_{\mu_j, R})$ correspond au nombre attendu de fausses alarmes sous \mathcal{H}_0 . Dit autrement, ε étant fixé, si l'on rejette \mathcal{H}_0 pour chaque région dès lors que la quantité $N_R \cdot N_\mu \cdot \min_j P(E_{\mu_j, R})$ est strictement inférieure à ε , le nombre attendu de régions pour lesquelles le modèle *a contrario* va être rejeté à tort est inférieur à ε . Comme il a déjà été

mentionné, les approches *a contrario* sont directement reliés au cadre classique de tests d'hypothèses [6] et en particulier, le terme $N_R \cdot N_\mu \cdot \min_j P(E_{\mu_j, R})$ correspond à la correction de Bonferroni [6].

Il est à noter que l'ensemble des p-valeurs $P(E_{\mu_j, R})$ est combiné en prenant le minimum de cet ensemble, ce qui peut être vu comme un opérateur OU puisque l'on considère uniquement le terme qui permet de rejeter au mieux \mathcal{H}_0 .

2.2 Extension de l'approche *a contrario* à la comparaison individu vs groupe

A notre connaissance, le modèle *a contrario* est toujours estimé directement à partir de l'image à traiter en négligeant les données de l'image qui ne sont pas représentatives de \mathcal{H}_0 . Nous étendons ici l'approche *a contrario* à la comparaison individu vs groupe en apprenant le modèle *a contrario* à partir d'un corpus d'apprentissage, à savoir ici, les individus du groupe. L'approche peut être divisée en deux étapes : une étape d'apprentissage du modèle et une étape de test d'un nouvel individu. Ces deux étapes sont utilisées itérativement dans un cadre de permutation dont l'intérêt sera discuté.

2.2.1 Apprentissage du modèle *a contrario*

L'information morphologique d'un individu est extraite en recalant son image 3-D avec une image de référence. Dans un premier temps, un recalage rigide est utilisé pour supprimer les variations relatives au positionnement du sujet durant l'acquisition. Ensuite, un recalage non-rigide fournit un champ de déformation h qui représente les différences anatomiques entre les deux sujets. Le choix du sujet de référence n'a pas une grande influence sur les résultats puisqu'un algorithme de recalage par groupe est utilisé[7].

En pratique, soit $N + 1$ le nombre de sujets. $N_R=17$ régions ont été délimitées sur l'image de référence par un expert (Fig. 1(a) et (b)). Ces régions sont des surfaces qui ont été définies en se basant sur des critères chirurgicaux. N transformations h_n ont été estimées par recalage non-rigide entre chaque sujet et la référence. La transformation h_n permet d'associer à chaque point de l'image de référence v ses coordonnées dans l'image du $n^{\text{ième}}$ sujet comme étant $h_n(v) + v$. Ainsi, pour chaque voxel v de l'image de référence, on obtient N vecteurs 3-D $h_n(v)$ (Fig. 1(c)) à partir desquelles on peut estimer différentes mesures telles que des valeurs locales du jacobien, la norme du vecteur de déplacement...

Soit $X_{mi}, m = 1, \dots, M$, la variable aléatoire reliée à une mesure m au voxel i . Un événement rare est défini comme étant l'observation d'une trop petite ou d'une trop grande valeur de x_{mi} (x_{mi} est une réalisation de X_{mi}) sous le modèle *a contrario*. Pour estimer le seuil à partir duquel un événement pourra être considéré comme rare, on commence par estimer la distribution de X_{mi} sous \mathcal{H}_0 . Sous l'hypothèse que $X_{mi} \sim \mathcal{N}(\mu_{mi}, \sigma_{mi}^2)$, les paramètres sont estimés à partir des N réalisations disponibles. Ensuite, on définit un ensemble de N_μ p-valeurs $\{p'_j, j = 1, \dots, N_\mu\}$. Finalement, l'observation x_{mi} au voxel i est considérée comme un événement rare pour la p-valeur p'_j et la mesure m si $x_{mi} > \mu_{jim}^\uparrow$ ou si $x_{mi} < \mu_{jim}^\downarrow$

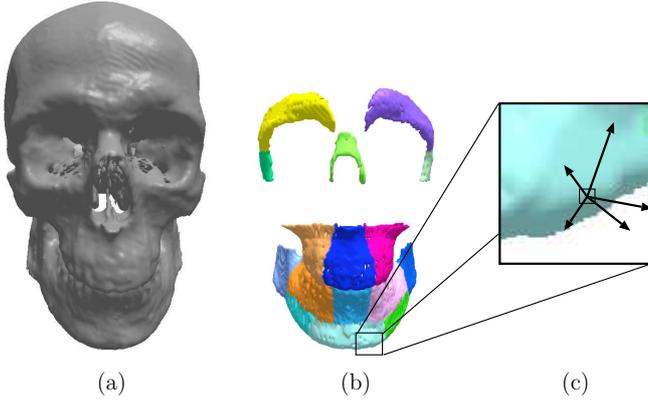


FIG. 1 – Extraction de la connaissance anatomique : (a) image de référence, (b) régions d'intérêt, (c) $N=5$ vecteurs 3-D $h_n(v)$ associés à v .

où μ_{jim}^\uparrow et μ_{jim}^\downarrow vérifient :

$$\begin{aligned} p'_j &= 2 \int_{x=\mu_{jim}^\uparrow}^{+\infty} \mathcal{N}(x; \mu_{mi}, \sigma_{mi}^2) dx \\ &= 2 \int_{x=-\infty}^{\mu_{jim}^\downarrow} \mathcal{N}(x; \mu_{mi}, \sigma_{mi}^2) dx \end{aligned} \quad (3)$$

L'ensemble $\mu_{jim}^\uparrow, \mu_{jim}^\downarrow$ ainsi que l'ensemble des p-valeurs associées $p'_j, j = 1, \dots, N_\mu$ sont les principaux paramètres qui définissent le modèle *a contrario*.

2.2.2 Test d'un nouvel individu

Tester la normalité d'un nouveau sujet consiste à effectuer les étapes suivantes :

- (i) recalculer l'image du sujet sur l'image de référence,
- (ii) déterminer les mesures x_{mi} à partir de la transformation estimée,
- (iii) comparer les mesures avec le modèle *a contrario*.

Les étapes (i) et (ii) se déroulent comme expliqué précédemment (Section 2.2.1). En ce qui concerne la troisième étape, on note k_{mjz} le nombre d'événements rares observés dans la région R_z pour la p-valeur p'_j et la mesure m .

On considère l'événement $E_{m,p'_j,z} = \{ \text{au moins } k_{mjz} \text{ événements rares ont été observés dans } R_z \text{ pour la p-valeur } p'_j \text{ et la mesure } m \}$. Sous \mathcal{H}_0 , on a :

$$P(E_{m,p'_j,z}) = \sum_{k=k_{mjz}}^{n_z} C_k^{n_z} p_j^k (1-p_j)^{n_z-k}, \quad (4)$$

où n_z est le nombre de voxels de R_z .

Le modèle *a contrario* est rejeté si la pathologie s'exprime pour au moins une p-valeur et une mesure si bien que l'on considère (une telle définition peut être perçue comme l'opérateur OU) :

$$r_z^\cup = \min_{j,m} P(E_{m,p'_j,z}). \quad (5)$$

A ce niveau, l'approche standard utilise la correction de Bonferroni en considérant qu'une région est ϵ -significative si l'on a $N_R N_\mu N_M r_z^\cup < \epsilon$. Cependant, une telle approche est conservatrice (d'autant plus que les tests sont ici fortement corrélés). De plus, les probabilités $P(E_{m,p'_j,z})$ sont calculées sous l'hypothèse que les voxels sont indépendants ce qui peut être considérée comme une approximation très grossière.

Pour résoudre ce problème, nous considérons r_z^\cup comme la réalisation d'une nouvelle variable aléatoire R_z^\cup . Estimer la distribution f_z de R_z^\cup sous \mathcal{H}_0 permet ensuite de déterminer la p-valeur $p_{r_z^\cup} = \int_0^{r_z^\cup} f_z(x) dx$ (on rejette \mathcal{H}_0 dès lors que r_z^\cup est faible). A noter que cette p-valeur est non-corrigée puisqu'elle ne prend pas en compte les N_R régions.

2.2.3 Estimation de f_z sous \mathcal{H}_0

Afin d'estimer f_z , nous utilisons une méthode de permutation dont le principe est d'itérer plusieurs fois les deux étapes présentées précédemment, à savoir l'apprentissage du modèle *a contrario* et le test d'un nouvel individu. A chaque itération, une nouvelle réalisation de R_z^\cup est obtenue. Comme cette réalisation doit être tirée sous \mathcal{H}_0 , le groupe utilisé pour l'apprentissage du modèle *a contrario* et le sujet à tester doivent être choisis avec soin.

Etant donné le groupe initial $S = \{S_1, S_2, \dots, S_N\}$ formé d'individus non pathologiques, le nouvel ensemble d'apprentissage à l'itération l est composé de $N-1$ sujets parmi les N disponibles $S - \{S_l\}$, et le sujet à tester est S_l . L'estimation de f_z est obtenue en lançant N fois la procédure, fournissant ainsi N réalisations de R_z^\cup (pour chaque région z). N étant faible, on utilise une représentation paramétrique : r_z^\cup est normalement distribué pour chaque région z et les paramètres de la distribution sont estimés avec les estimateurs traditionnels.

3 Résultats

La base d'apprentissage est actuellement constituée de $N=5$ sujets (l'acquisition de nouvelles images est en cours). Pour évaluer la méthode, une image d'un sujet dont la mandibule est pathologique (Fig. 2(a)) a été testée avec $N_M=3$, à savoir 3 mesures que sont la projection du vecteur $h_n(v)$ sur les axes x, y , et z , et avec $N_\mu=3, p'_1=0.01, p'_2=0.001$ et $p'_3=0.005$. Le résultat de détection est montré Fig. 2(b) : les régions détectées comme anormales sont indiquées sur l'image de référence.

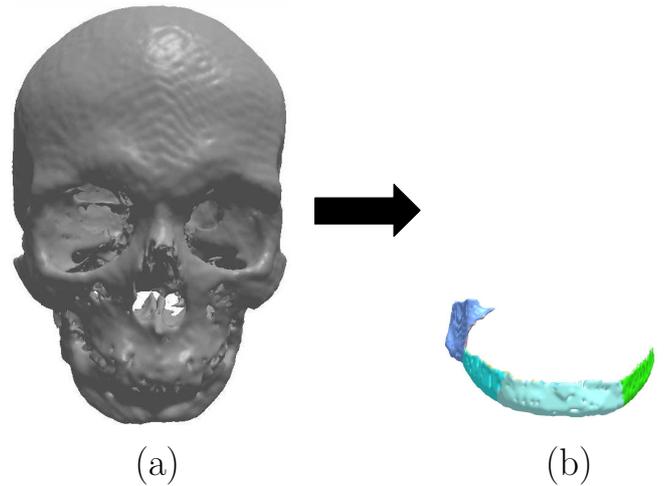


FIG. 2 – Résultats de détection : (a) sujet dont la mandibule est pathologique, (b) résultats de détection représentés sur l'image de référence.

Nous pouvons tout d'abord noter que les résultats sont cor-

rects et valident ainsi l'approche. Dans cet exemple, \mathcal{H}_0 a été rejeté pour une région z si $p_{r_z} < 0.01$ (voir Section 2.2.2). Le seuil de 0.01 est une valeur relativement importante. Cependant, on peut s'attendre à obtenir des résultats plus significatifs dès lors que la base d'apprentissage sera plus fournie. De plus, ces résultats sont d'autant plus encourageants que l'utilisation des modèles globaux décrits dans [8] ne permettent pas de détecter les régions pathologiques, et cela pour n'importe quel seuil. Ceci peut s'expliquer par la difficulté d'estimer correctement le lien entre l'espace d'origine et l'espace réduit, en raison du manque de données.

4 Discussion et conclusion

Cet article présente une extension de l'approche *a contrario* à la comparaison individu vs groupe. L'idée directrice est d'apprendre le modèle *a contrario* à partir d'un corpus d'apprentissage.

Ce travail met en lumière le potentiel de l'approche *a contrario* dès lors qu'un ensemble d'apprentissage est disponible et ouvre de nouvelles pistes de recherche. En effet, en utilisant la méthode de permutation décrite ci-avant, il est désormais possible d'estimer la distribution d'autres variables aléatoires. Par exemple, si l'ensemble des p-valeurs $P(E_{\mu_j, R})$ est habituellement résumé en prenant son minimum (opérateur OU), d'autres règles de fusion, beaucoup plus complexes, peuvent maintenant être envisagées pour combiner l'ensemble des p-valeurs telles que l'utilisation de la fonction de Tippett, de Fisher [9], ou de n'importe quelle fonction paramétrique ou non. Indépendamment de la fonction de fusion utilisée, sa distribution sous \mathcal{H}_0 pourra toujours être estimée avec l'approche de permutation. De la même manière, la probabilité d'observer au moins k événements rares a été exprimée comme la queue de la distribution binomiale (les voxels sont alors considérés mutuellement indépendants). Cette hypothèse n'est désormais plus nécessaire puisque la distribution du nombre des événements rares (pour une région, une mesure et une p-valeur donnée) peut être apprise sous \mathcal{H}_0 dans une première passe.

Finalement, l'approche *a contrario* dans le contexte de la comparaison individu vs groupe apparaît comme une approche puissante et versatile avec plusieurs avantages : approche région, prise en compte de plusieurs mesures, règle de fusion flexible.

Références

- [1] J. Ashburner and K.J. Friston, "Voxel-based morphometry – the methods," *NeuroImage*, vol. 11, pp. 805–821, 2000.
- [2] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, December 2000.
- [3] G. Subsol, J.-P. Thirion, and N. Ayache, "A scheme for automatically building 3D morphometric anatomical atlases : application to a skull atlas," *Medical Image Analysis*, vol. 2, no. 1, pp. 37–60, 1998.
- [4] A. Desolneux, L. Moisan, and J.-M. Morel, "A grouping principle and four applications," *IEEE Transactions on*

Pattern Analysis and Machine Intelligence, vol. 25, no. 4, pp. 508–513, April 2003.

- [5] T. Veit, F. Cao, and P. Bouthemy, "An a contrario decision framework for region-based motion detection," *International Journal on Computer Vision*, vol. 68, no. 2, pp. 163–178, 1998.
- [6] F. Rousseau, S. Faisan, F. Heitz, J.-P. Armspach, Y. Chevalier, F. Blanc, J. De Seze, and L. Rumbach, "An a contrario approach for change detection in 3D multimodal images : Application to multiple sclerosis in MRI," in *IEEE EMBS*, 2007.
- [7] V. Noblet, C. Heinrich, F. Heitz, and J.-P. Armspach, "Symmetric nonrigid image registration. application to average brain templates construction," in *Medical Image Computing and Computer Assisted Intervention*, september 2008, vol. 5242, pp. 897–904.
- [8] T. Vik, F. Heitz, and P. Charbonnier, "Robust pose estimation and recognition using non-gaussian modeling of appearance subspaces," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, no. 5, pp. 901–905, may 2007.
- [9] S. Hayasaka and T.E. Nichols, "Combining voxel intensity and cluster extent with permutation test framework," *NeuroImage*, vol. 23, pp. 54–63, 2004.