

Recalage variationnel non rigide d'images par f -information exclusive

Nicolas ROUGON, Caroline PETITJEAN, Françoise PRÉTEUX

GET / INT - Unité de Projets ARTEMIS
9, Rue Charles Fourier - 91011 Evry Cedex, France

Nicolas.Rougon@int-evry.fr, Caroline.Petitjean@int-evry.fr,
Francoise.Preteux@int-evry.fr

Résumé – Nous nous intéressons à la mise en correspondance non rigide dense d'images dans des contextes monomodaux avec fortes variations photométriques/texturales ou multimodaux, et étudions à cette fin des critères de similarité statistiques fondés sur des mesures d'information généralisées au sein de la classe d'Ali-Silvey. Nous introduisons une nouvelle classe de fonctionnelles, dénommées f -informations exclusives, et développons un cadre variationnel générique bien-posé pour leur optimisation sur des espaces de déformations non paramétriques et paramétriques, généralisant les méthodes par information mutuelle. Cette approche est appliquée à l'alignement aveugle robuste de visages sous éclairage arbitraire et pour des déformations faciales complexes.

Abstract – We address the problem of robust non rigid image registration and focus our interest on similarity criteria based on generalized information measures within the Ali-Silvey class. We introduce exclusive f -information as a novel class of similarity functionals, and develop a generic variational framework for their optimization over non-parametric and parametric transform spaces. This approach is applied to blind non-rigid alignment of face images with arbitrary lighting conditions and complex deformations.

1 Introduction

Le recalage non rigide d'images a suscité de nombreux travaux en imagerie médicale et en vision par ordinateur. Parmi ceux-ci, les approches iconiques, qui visent à maximiser la similarité inter-image au niveau pixel, ont reçu une attention particulière de par leur capacité à générer des cartes d'appariement denses sans pré-segmentation [3, 2, 6]. S'il existe aujourd'hui des techniques performantes pour des données monomodales présentant des variations stationnaires de photométrie et de géométrie, l'élaboration de méthodes de recalage robustes, opérant sur des données non stationnaires, hétérogènes ou incommensurables, demeure un problème largement ouvert.

Dans de tels cas, l'hypothèse faible d'une dépendance statistique entre signaux conduit à des mesures de similarité non locales issues de la théorie de l'information, telles l'entropie conjointe ou l'information mutuelle (IM) [12, 5]. Ces mesures sont le plus souvent maximisées sur des espaces de transformations paramétriques de faible dimension, soit par des méthodes directes utilisant des estimées empiriques [5, 8] ou paramétriques [4] des densités de probabilité, soit par des techniques variationnelles reposant sur des estimées analytiques [9] ou non paramétriques de type Parzen [12]. Cette dernière approche s'étend à des espaces de transformations non paramétriques, conduisant à des modèles bien-posés [1].

En dépit de bonnes performances, l'IM est sensible au recouvrement partiel des supports images, et peut présenter des minima locaux parasites. Plusieurs solutions ont été proposées pour tenter de résoudre ces limitations : modèle probabiliste *a priori* issu d'un apprentissage supervisé [4]; introduction d'information spatiale par pondération de l'IM par un terme dépendant du gradient [8]; recours à d'autres critères entropiques, tels

que des mesures d'information normalisées dérivées de l'entropie de Shannon [5, 10], ou des f -divergences [8].

Dans cet article, nous introduisons une nouvelle classe de mesures de similarité entropiques, appelées f -informations exclusives, et étudions leur pertinence pour le recalage variationnel non rigide robuste d'images. Nous présentons d'abord la classe d'informations généralisées d'Ali-Silvey, de laquelle les f -informations exclusives sont dérivées, puis calculons les flots qu'elles induisent sur des espaces de transformations non paramétriques. Nous illustrons enfin leurs performances pour l'alignement d'images de visages sous illumination arbitraire et pour des déformations complexes.

2 Mesures d'information généralisées

Soient X et Y des variables aléatoires (VA) continues sur un espace d'état Ω , de densités respectives p^X et p^Y . Soit $p^{X,Y}$ la densité de la VA jointe (X, Y) . Enfin, soit f (resp. ψ) une fonction continue convexe (resp. monotone) sur \mathbb{R}^+ . La notion d'entropie est fondamentalement définie pour les mesures de probabilité (MP) dont elle quantifie l'incertitude. Dans ce cadre, l'entropie d'une MP est définie relativement à une MP de référence, d'où le nom d'*entropie relative*. On distingue les *entropies intégrales* et les *entropies non intégrales* [11].

Définition 1 (f -entropie) L'entropie intégrale (resp. non intégrale) $H_{f,\nu}(\mu)$ (resp. $H_{\psi,\nu}(\mu)$) d'une MP μ relativement à la MP de référence ν est :

$$H_{f,\nu}(\mu) = - \int f \left(\frac{d\mu}{d\nu} \right) d\nu$$
$$H_{\psi,\nu}(\mu) = - \log \psi^{-1} \left(\int \frac{d\mu}{d\nu} \psi \left(\frac{d\mu}{d\nu} \right) d\nu \right) .$$

Ces définitions se spécialisent au cas de VA continues en posant $\mu \equiv X$ et ν borélienne, et en utilisant les dérivées de Radon-Nikodym $\frac{d\mu}{dx} \equiv p^X(x)$ et $\frac{d\nu}{dx} \equiv 1$. La mesure de Kullback $f_{\text{KL}}(x) := x \log x$ ou $\psi_{\text{KL}}(x) := \log x$ conduit à l'entropie de Shannon, notée H_S . De nombreuses entropies ont été proposées dans la littérature, offrant un large éventail de possibilités pour quantifier l'incertitude [11].

La similarité entre MP est formalisée par la notion de *divergence* qui dérive de l'entropie relative. La comparaison directe de MP (par opposition à celle de mélanges) conduit à la classe de divergences d'Ali-Silvey ou *f-divergences*.

Definition 2 (f-divergence) Une *f-divergence* D_f est formellement définie par: $D_f(\mu \parallel \nu) = -H_{f,\nu}(\mu)$.

Cette classe se subdivise en *f-divergences* intégrales et non intégrales. Le cas des VA continues se déduit en posant $\mu \equiv X$ et $\nu \equiv Y$ avec $\frac{d\mu}{dx} \equiv p^X(x)$ et $\frac{d\nu}{dx} \equiv p^Y(x)$. Le cas $f = f_{\text{KL}}$ ou $\psi = \psi_{\text{KL}}$ correspond à la divergence de Kullback-Leibler D_{KL} . Le choix $f(x) = \frac{x^\alpha - \alpha x + \alpha - 1}{\alpha(\alpha - 1)}$ ($\alpha \neq \{0, 1\}$) définit la I_α -divergence I_α qui vérifie: $\lim_{\alpha \rightarrow 1} I_\alpha = D_{\text{KL}}$.

La notion de gain d'information entre MP se formalise en mesurant l'information de la MP jointe à l'aide d'une *f-entropie* relativement à la MP produit associée au cas d'indépendance statistique. On aboutit ainsi à la *f-information*:

Definition 3 (f-information) Une *f-information* I_f est formellement définie par: $I_f(\mu, \nu) = D_f((\mu, \nu) \parallel \mu \times \nu)$.

La *f-information* quantifie donc la réduction d'incertitude de μ apportée par ν au sens de la métrique f . La mesure de Kullback conduit à l'IM, notée simplement I .

Une mesure de gain d'information alternative, définie par $Z(X, Y) := H_S(X, Y) - I(X, Y)$, a été proposée dans le cadre de Shannon [5]. Puisque $I(X, Y)$ quantifie le gain d'information résultant de l'observation conjointe de X et Y par rapport à leur observation indépendante, et que $H_S(X, Y)$ est l'information de (X, Y) , Z mesure l'information exclusivement contenue dans X et Y lorsque ces VA sont observées conjointement. Pour cette raison, nous dénommons Z *information exclusive*. Partant de sa définition, nous généralisons Z en utilisant les identités $H_S(X, Y) = H_S(X) + H_S(Y) - I(X, Y)$ et $H_S(X) = I(X, X)$, et en passant à f arbitraire. On aboutit ainsi à la *f-information exclusive*:

Definition 4 (f-information exclusive) Une *f-information exclusive* est une mesure Z_f telle que:

$$Z_f(\mu, \nu) := I_f(\mu, \mu) + I_f(\nu, \nu) - 2I_f(\mu, \nu) \quad (1)$$

Les définitions des *f-informations* (standards ou exclusives) intégrales et non intégrales pour les VA continues se déduisent en posant $\frac{d\mu}{dx dy} \equiv p^{X,Y}(x, y)$ et $\frac{d\nu}{dx dy} \equiv p^X(x)p^Y(y)$.

3 Recalage variationnel non rigide par *f-information exclusive*

Soient I_1 et I_2 , des images définies sur un domaine $D \subset \mathbb{R}^n$ et à valeurs dans un intervalle Ω . Recaler I_2 sur I_1 consiste à identifier une transformation $\phi : D \rightarrow D$ au sein d'un espace \mathcal{T} de fonctions régulières, telle que $I_2 \circ \phi$ soit similaire à I_1 au sens d'un critère prédéfini. Ceci équivaut à trouver un champ de déplacement $\mathbf{u} = \phi - \text{Id} \in \mathcal{T}$ tel que $I_2^{\mathbf{u}}(\mathbf{x}) := I_2(\mathbf{x} + \mathbf{u}(\mathbf{x}))$

soit similaire à $I_1(\mathbf{x})$ en tout point $\mathbf{x} \in D$. Ce problème est mal-posé et doit être régularisé.

Une approche classique définit une solution comme un minimiseur \mathbf{u}^* d'une fonctionnelle $\mathcal{J}(\mathbf{u}) := \mathcal{S}(\mathbf{u}) + \alpha \mathcal{R}(\mathbf{u})$ sur \mathcal{T} , combinant un terme de similarité \mathcal{S} quantifiant l'écart entre I_1 et $I_2^{\mathbf{u}}$, et un stabilisateur \mathcal{R} garantissant des solutions p.p. régulières, pondérés par un hyperparamètre $\alpha > 0$. Nous considérons ici le cas où $\mathcal{S}(\mathbf{u}) = Z_f(I_1, I_2^{\mathbf{u}})$, I_1 et $I_2^{\mathbf{u}}$ étant modélisées comme des VA de densités respectives p^{I_1} et $p^{I_2^{\mathbf{u}}}$ et de densité jointe $p^{I_1, I_2^{\mathbf{u}}}$.

La minimisation variationnelle de \mathcal{J} s'effectue en intégrant par descente de gradient les équations d'Euler obtenues en annulant sa dérivée de Gâteaux. La forme analytique du flot résultant dépend du contexte choisi pour modéliser le champ \mathbf{u} :

- Pour des déplacements non paramétriques, on obtient un flot sur D : $\frac{\partial \mathbf{u}}{\partial t}(\mathbf{x}) = -\partial_{\mathbf{u}} \mathcal{J}(\mathbf{u}(\mathbf{x}))$ où $\partial_{\mathbf{u}} \mathcal{J}$ est la dérivée variationnelle de \mathcal{J} .
- Pour des déplacements paramétriques $\mathbf{u} := \mathbf{u}^\Theta$ dépendant d'un paramètre $\Theta \in \mathbb{R}^p$, on obtient un flot sur \mathbb{R}^p : $\partial_t \Theta = -\partial_\Theta \mathcal{J}(\mathbf{u}^\Theta)$ où $\partial_\Theta \mathcal{J}$ est la dérivée variationnelle de \mathcal{J} . Nous nous restreignons dans la suite aux modèles linéaires pour lesquels $\mathbf{u}^\Theta(\mathbf{x}) = \mathbf{B}(\mathbf{x})\Theta$, et $\mathbf{B}(\mathbf{x})$ est une matrice associée à une base locale de \mathcal{T} .

Dans chaque cadre, le point central réside dans le calcul de la dérivée de Gâteaux d'une *f-information*. Dans ce but, nous utilisons des estimateurs de Parzen pour exprimer analytiquement les densités $p^{I_1, I_2^{\mathbf{u}}}$ et $p^{I_2^{\mathbf{u}}}$. La densité p^X d'une VA X sur D est ainsi approchée par: $p^X(X_0) = \frac{1}{\mu(D)} \int_D \mathcal{K}_d(X(\mathbf{x}) - X_0) d\mathbf{x}$, où \mathcal{K}_d est un noyau de même dimension d que X et $\mu(D)$ est le volume de D [12, 1].

3.1 Modèles de mouvement non paramétriques

On montre alors que [7]:

$$\begin{aligned} \partial_{\mathbf{u}} I_f(I_1, I_2^{\mathbf{u}}) &= V_1^{I_1, I_2^{\mathbf{u}}}(\mathbf{x}) \nabla I_2^{\mathbf{u}}(\mathbf{x}) \quad (2) \\ V_1^{I_1, I_2^{\mathbf{u}}}(\mathbf{x}) &:= K \left\{ \left[\mathcal{K}_2 \star \frac{\partial L_1^{\mathbf{u}}}{\partial i_2} \right] (I_1(\mathbf{x}), I_2^{\mathbf{u}}(\mathbf{x})) + \right. \\ &\quad \left. E_{I_1} \left(\left[\mathcal{K}_1 \star \frac{\partial L_2^{\mathbf{u} | i_1}}{\partial i_2} \right] (I_2^{\mathbf{u}}(\mathbf{x})) \right) \right\} \end{aligned}$$

où $E_{I_1}(\cdot)$ est l'espérance par rapport à I_1 et \star est la convolution 2D (resp. 1D) sur l'espace d'intensité Ω^2 (resp. Ω). Les symboles $K(I_1, I_2^{\mathbf{u}})$, $L_1^{\mathbf{u}}$ et $L_2^{\mathbf{u}}$ sont définis dans les Tableaux 1 et 2 avec $\rho^{\mathbf{u}}(i_1, i_2) := \frac{p^{I_1, I_2^{\mathbf{u}}}(i_1, i_2)}{p^{I_1}(i_1)p^{I_2^{\mathbf{u}}}(i_2)}$ et $L_2^{\mathbf{u} | i_1}(\cdot) := L_2^{\mathbf{u}}(i_1, \cdot)$ pour $i_1 \in \Omega$ donné. Le flot (2) est contrôlé localement par le contraste de $I_2^{\mathbf{u}}$, et la variation du couplage statistique entre images induite par \mathbf{u} , mesurée via les estimateurs $L_1^{\mathbf{u}}$ et $L_2^{\mathbf{u}}$ [7]. Cette équation généralise le résultat obtenu dans [1] pour l'IM, que l'on retrouve en posant $f = f_{\text{KL}}$ ou $\psi = \psi_{\text{KL}}$. Dans ce cas, l'espérance est nulle, rendant l'estimateur $L_2^{\mathbf{u}}$ non informatif. Par contraste, le choix $f \neq f_{\text{KL}}$ permet une exploitation plus exhaustive de la statistique de $I_2^{\mathbf{u}}$.

On montre de même que la dérivée variationnelle de l'auto *f-information* $I_f(I_2^{\mathbf{u}}, I_2^{\mathbf{u}})$ s'écrit:

$$\begin{aligned} \partial_{\mathbf{u}} I_f(I_2^{\mathbf{u}}, I_2^{\mathbf{u}}) &= V_2^{I_2^{\mathbf{u}}}(\mathbf{x}) \nabla I_2^{\mathbf{u}}(\mathbf{x}) \quad (3) \\ V_2^{I_2^{\mathbf{u}}}(\mathbf{x}) &:= K \left\{ \left[\mathcal{K}_1 \star \frac{\partial L_3^{\mathbf{u}}}{\partial i_2} \right] (I_2^{\mathbf{u}}(\mathbf{x})) \right\} \end{aligned}$$

TAB. 1: Structure analytique de la dérivée variationnelle des f -informations intégrales.

$I_f(I_1, I_2^{\mathbf{u}}) = \int_{\Omega^2} p^{I_1}(i_1) p^{I_2^{\mathbf{u}}}(i_2) f(\rho^{\mathbf{u}}(i_1, i_2)) di_1 di_2$	
$L_1^{\mathbf{u}}$	$f'(\rho^{\mathbf{u}})$
$L_2^{\mathbf{u}}$	$f(\rho^{\mathbf{u}}) - \rho^{\mathbf{u}} f'(\rho^{\mathbf{u}})$
$L_3^{\mathbf{u}}$	$2f\left(\frac{1}{p^{I_2^{\mathbf{u}}}}\right) p^{I_2^{\mathbf{u}}} - f'\left(\frac{1}{p^{I_2^{\mathbf{u}}}}\right)$
$K(I_1, I_2^{\mathbf{u}})$	$\frac{1}{\mu(D)}$

TAB. 2: Structure analytique de la dérivée variationnelle des f -informations non intégrales.

$I_{\psi}(I_1, I_2^{\mathbf{u}}) = \log \psi^{-1}(B_{\psi}(I_1, I_2^{\mathbf{u}}))$ $B_{\psi}(I_1, I_2^{\mathbf{u}}) = \int_{\Omega^2} p^{I_1, I_2^{\mathbf{u}}}(i_1, i_2) \psi(\rho^{\mathbf{u}}(i_1, i_2)) di_1 di_2$	
$L_1^{\mathbf{u}}$	$\psi(\rho^{\mathbf{u}}) + \rho^{\mathbf{u}} \psi'(\rho^{\mathbf{u}})$
$L_2^{\mathbf{u}}$	$-(\rho^{\mathbf{u}})^2 \psi'(\rho^{\mathbf{u}})$
$L_3^{\mathbf{u}}$	$\psi\left(\frac{1}{p^{I_2^{\mathbf{u}}}}\right) - \frac{1}{p^{I_2^{\mathbf{u}}}} \psi'\left(\frac{1}{p^{I_2^{\mathbf{u}}}}\right)$
$K(I_1, I_2^{\mathbf{u}})$	$\frac{(\psi^{-1})'(B_{\psi}(I_1, I_2^{\mathbf{u}}))}{\psi^{-1}(B_{\psi}(I_1, I_2^{\mathbf{u}}))}$

l'estimateur $L_3^{\mathbf{u}}$ étant défini dans les Tableaux 1 et 2. La structure de ce flot est similaire à (2), le contrôle statistique étant fondé sur la variation de l'information de $I_2^{\mathbf{u}}$, analysée par l'estimateur $L_3^{\mathbf{u}}$. En posant $V^{I_1, I_2^{\mathbf{u}}} := 2V_1^{I_1, I_2^{\mathbf{u}}} - V_2^{I_2^{\mathbf{u}}}$, on obtient finalement l'expression de la dérivée variationnelle d'une f -information exclusive $\mathcal{S}(\mathbf{u}) = Z_f(I_1, I_2^{\mathbf{u}})$:

$$\partial_{\mathbf{u}} \mathcal{S}(\mathbf{u}) = -V^{I_1, I_2^{\mathbf{u}}}(\mathbf{x}) \nabla I_2^{\mathbf{u}}(\mathbf{x}). \quad (4)$$

3.2 Modèles de mouvement paramétriques

Pour les mouvements paramétriques linéaires, des calculs analogues conduisent au résultat suivant :

$$\partial_{\Theta} \mathcal{S}(\mathbf{u}^{\Theta}) = - \int_D V^{I_1, I_2^{\mathbf{u}}}(\mathbf{x}) \left[(\nabla I_2^{\mathbf{u}})^t \mathbf{B} \right](\mathbf{x}) dx \quad (5)$$

Le flot dans l'espace des paramètres s'obtient ainsi par accumulation de contributions locales sur la projection $\left[(\nabla I_2^{\mathbf{u}})^t \mathbf{B} \right]$ de $\nabla I_2^{\mathbf{u}}$ sur la base \mathbf{B} de \mathcal{T} . Notons que pour $f = f_{\text{KL}}$, ce résultat diffère de l'approche de Viola [12]. Cette différence provient du fait que nous optimisons *exactement* la fonctionnelle d'IM et non une approximation ergodique comme dans [12].

3.3 Régularisation d'ordre 1

Les modèles de mouvement non paramétriques et paramétriques d'ordre élevé requièrent une régularisation explicite. Dans le contexte applicatif ciblé (cf. Section 4), nous avons eu recours au stabilisateur de Nagel-Enkelmann [1, 7], qui est un stabilisateur géométrique classique d'ordre 1, fonction du jacobien $\nabla \mathbf{u}$ du champ \mathbf{u} . Sa dérivée variationnelle est donnée par $\partial_{\mathbf{u}} R(\mathbf{u}) = -\nabla \cdot (T_{I_1} \nabla \mathbf{u})$, où :

$$T_{I_1} = \frac{1}{\|\nabla I_1\|^2 + 2\beta} \left[(\beta + \|\nabla I_1\|^2) \mathbb{I}d - \nabla I_1 \nabla I_1^t \right]$$

et $\beta > 0$ désigne un paramètre de contraste. Le tenseur T_{I_1} définit un projecteur orthogonalement au gradient ∇I_1 , permettant de privilégier une régularisation intra-régionale et ainsi de limiter l'influence de régions voisines sur le champ de déplacement estimé au sein d'un objet de l'image. Il en résulte des déformations nettement plus localisées que dans le cas de stabilisateurs standards de type élasticité linéaire.

4 Résultats expérimentaux

Les flots précédents se discrétisent simplement par différences finies explicites, noyaux de Parzen gaussiens et interpolation bilinéaire.

Cette approche a été appliquée à l'alignement d'images de visage présentant de grandes variations d'éclairage, de pose et d'expression. Dans ce cadre, nous avons eu recours à une stratégie multirésolution par pyramides gaussiennes pour les images et les déplacements [7]. Le modèle a tout d'abord été validé en recalant des images natives avec des versions de synthèse générées par transformation non linéaire de contraste puis déformation par un champ \mathbf{u}_0 prédéfini (Figure 1). Des tests approfondis ont démontré que le recalage par f -information exclusive fournit des résultats précis et de qualité au moins égale à ceux issus de schémas variationnels par IM, tout en autorisant une accélération de convergence significative en comparaison des f -informations pour f fixé.

La Figure 2a-c présente des résultats d'alignement intra-individu par I_{α} - et f_{KL} -informations exclusives, illustrant la robustesse du critère relativement à des variations de contraste fortement non linéaires et à des mouvements de grande amplitude avec occlusions partielles. La Figure 3 compare les performances des I_{α} -informations standards et exclusives: toutes choses étant égales, ces dernières conduisent à des résultats plus précis.

L'alignement inter-individu est illustré sur la Figure 2d-e. Afin d'évaluer la précision du recalage, 20 points caractéristiques parmi les paramètres de définition faciaux MPEG-4 ont été sélectionnés sur l'image source et recalés sur l'image cible via le champ \mathbf{u}^* . La comparaison avec leurs homologues sur la cible révèle un alignement morphologique précis (MSE = 2 ± 1.5 pixels).

5 Conclusion et perspectives

Dans cet article, nous avons étudié l'utilisation de mesures d'information généralisées pour le recalage non rigide robuste d'images. Nous avons introduit, au sein de la classe d'Ali-Silvey, une nouvelle famille de mesures de similarité statistiques, dénommées f -informations exclusives; puis, via l'estimation de Parzen, identifié l'expression générale de leur gradient variationnel sur des espaces de transformations régulières non paramétriques et paramétriques. Enfin, nous avons illustré la pertinence de ces critères dans le cadre de l'alignement non rigide de visages sous éclairage arbitraire, et pour des poses et déformations faciales complexes. Les développements en cours concernent l'extension de ces techniques au recalage 2D/3D non rigide orienté modèle.

Références

- [1] G. Hermosillo, *Variational methods for multimodal image matching*, Thèse de Doctorat, Université de Nice, Mai 2002.
- [2] D.L.G. Hill, P.G. Batchelor, M. Holden, D.J. Hawkes, "Medical image registration", *Physics in Medical Biology*, 46(3), 2001, R1:R45.
- [3] H. Lester, S.R. Arridge, "A survey of hierarchical non-linear medical image registration", *Pattern Recognition*, 32(1), Janvier 1999, 129:149.
- [4] M.E. Leventon, W.E.L. Crimson, "Multi-modal volume registration using joint intensity distributions", *Proceedings MICCAI'98, Cambridge, MA*, Octobre 1998, 1057:1066.
- [5] F. Maes, A. Collignon, D. Vandermuelen, G. Marchal, P. Suetens, "Multimodality image registration by maximization of mutual information", *IEEE Transactions on Medical Imaging*, 16(2), Avril 1997, 187:198.
- [6] T. Mäkelä, P. Clarysse, O. Sipilä, N. Pauna, Q.C. Pham, T. Katila, I.E. Magnin, "A review of cardiac image registration methods", *IEEE Transactions on Medical Imaging*, 21(9), 2002, 1011:1021.
- [7] C. Petitjean, N. Rougon, F. Prêteux, "Non rigid image registration using generalized information measures", *Research Report 02001 INT, France*, Mars 2002.
- [8] J.P.W. Pluim, *Mutual information based registration of medical images*, Ph.D. Dissertation Utrecht University, 2000.
- [9] D. Rueckert, L.I. Sonoda, C. Hayes, D.L. Hill, M.O. Leach, D.J. Hawkes, "Nonrigid registration using free-form deformations: application to breast MR images", *IEEE Transactions on Medical Imaging*, 18(8), August 1999, 712:721.
- [10] C. Studholme, D.L. Hill, D. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment", *Pattern Recognition*, 32(1), Janvier 1999, 71:86.
- [11] I. Vajda, *Theory of Statistical Evidence and Information*, Kluwer Academic Publishers, Dordrecht, 1989.
- [12] P. Viola, W.M. Wells III, "Alignment by maximization of mutual information", *IJCV*, 24(2), 1997, 137:154.

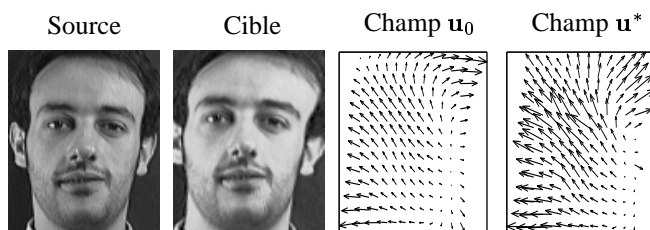


FIG. 1: Recalage non rigide par I_α -information exclusive pour une déformation synthétique ($\alpha = 1.5$).

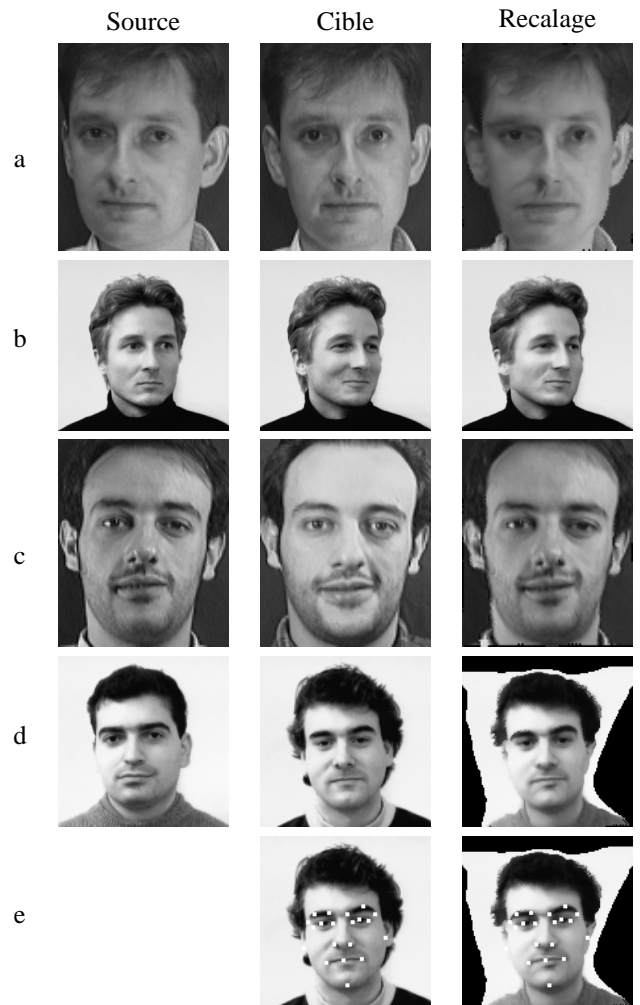


FIG. 2: Recalage non rigide de visages présentant de grandes variations de (a,b) pose, (c) éclairage et (d,e) morphologie par (a,b) $I_{1,1}$ - et (c,d,e) f_{KL} -information exclusive.

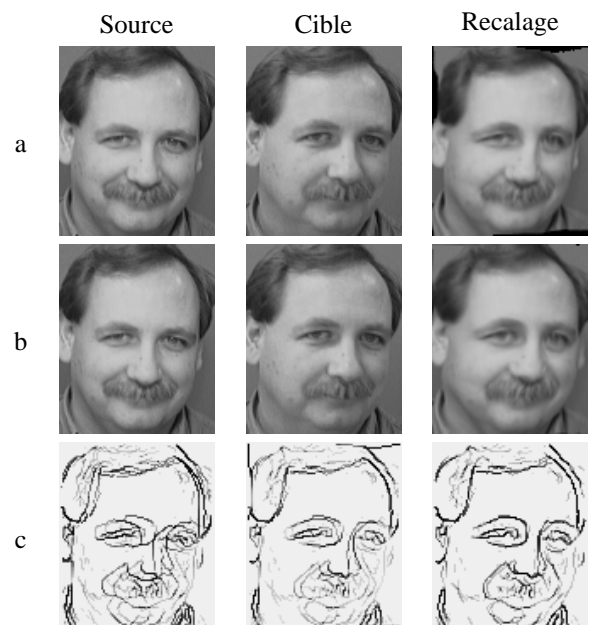


FIG. 3: Recalage non rigide par I_α -information (a) standard et (b) exclusive ($\alpha = 1.2$) en présence de grandes variations de pose. (c) Superposition des contours de la cible et de la source (gauche) / source recalée (centre et droite).