

Utilisation de l'information photométrique pour la sélection des hyperparamètres en recalage géométrique d'images

F. Brunet^{1,2}

A. Bartoli¹

N. Navab²

R. Malgouyres³

¹ ISIT - Université d'Auvergne, Clermont-Ferrand

² CAMPAR - Technische Universität München, Munich

³ LIMOS - Université d'Auvergne, Clermont-Ferrand

florent.brunet@u-clermont1.fr

Résumé

Cet article traite du recalage paramétrique d'images à partir de correspondances de points en environnement déformable. Dans ce problème, il est essentiel de déterminer des valeurs correctes pour les hyperparamètres tels que le nombre de points de contrôle du modèle de déformation, un paramètre de régularisation ou l'échelle d'un M-estimateur. Cela est souvent réalisé à la main par tâtonnement ou en optimisant un critère générique comme la validation croisée. Dans cet article, nous proposons un nouveau critère pour sélectionner différents hyperparamètres en combinant les avantages de l'approche géométrique et de l'approche photométrique au recalage d'images. Plus précisément, nous proposons de considérer les correspondances de points comme un jeu d'entraînement et la photométrie comme un jeu d'essai. L'approche proposée est robuste dans la mesure où elle résiste à la fois aux correspondances de points erronées et aux défauts des images comme les occultations ou les spécularités.

Mots clefs

Hyperparamètre, recalage, image, déformation.

1 Introduction

Le recalage d'images consiste à déterminer les paramètres (naturels) d'une déformation de manière à aligner une image source et une image cible. Outre les paramètres naturels, des valeurs correctes doivent être affectées aux hyperparamètres du problème afin d'obtenir un recalage de bonne qualité. Les hyperparamètres sont soit des paramètres additionnels du modèle de déformation (*hyperparamètres de modèle*), soit des paramètres de la fonction de coût à optimiser (*hyperparamètres de coût*). La figure 1 montre à quel point le choix des hyperparamètres est crucial en recalage d'images. Il existe deux approches principales au recalage d'images [1] : l'approche géométrique et l'approche photométrique (ou approche directe). Chacune de ces méthodes a ses propres avantages mais aucune d'entre elles ne permet d'estimer directement des hyperparamètres appropriés. Nous proposons donc de combi-

ner les forces respectives des approches géométriques et photométriques afin de construire une nouvelle méthode de sélection automatique des hyperparamètres en recalage d'images.

Soit $\mathcal{W} : \mathbb{R}^2 \times \mathbb{R}^l \rightarrow \mathbb{R}$ un modèle de déformation paramétré par un ensemble de l paramètres regroupés dans une matrice P ayant l coefficients. L'homographie [1, 4] est un exemple de déformation paramétré par les 8 coefficients indépendants de la matrice d'homographie. Les déformations libres [5] sont un autre exemple de modèle paramétré par $\frac{l}{2}$ points de contrôle 2D. Le nombre de points de contrôle d'une déformation libre ou la largeur de bande des fonctions à bases radiales [6] sont deux exemples parmi d'autres d'hyperparamètres de modèle.

Dans l'approche géométrique [1, 7], les images source et cible sont abstraites par un ensemble fini d'éléments saillants caractéristiques. Dans cet article, nous utilisons des points d'intérêt qui sont appariés pour former un ensemble de n correspondances de points $\{\mathbf{p}_i \leftrightarrow \mathbf{q}_i\}_{i=1}^n$. Le principe fondamental de l'approche géométrique consiste à minimiser l'erreur entre les points transformés de l'image source et les points de l'image cible. Une mesure robuste comme un M-estimateur peut être utilisée à cet effet. Des termes supplémentaires, comme un terme de régularisation [7], sont parfois ajoutés afin de faire face, par exemple, aux imprécisions des points ou à un manque de données dans certaines parties des images. Tous ces éléments sont regroupés dans la formulation suivante [7, 8] :

$$\min_P \mathcal{E}(P; \theta), \quad (1)$$

où θ est un vecteur contenant les hyperparamètres et \mathcal{E} est la fonction de coût ainsi définie :

$$\mathcal{E}(P; \theta) = \sum_{i=1}^n \rho(\mathcal{W}(\mathbf{p}_i; P) - \mathbf{q}_i; \gamma) + \lambda \mathcal{R}(P), \quad (2)$$

avec ρ un M-estimateur, γ son facteur d'échelle, \mathcal{R} un terme de régularisation¹ et λ un paramètre contrôlant le compromis entre attache aux données et régularisation. Les

1. Comme, par exemple, l'énergie de torsion (détaillée dans §4).

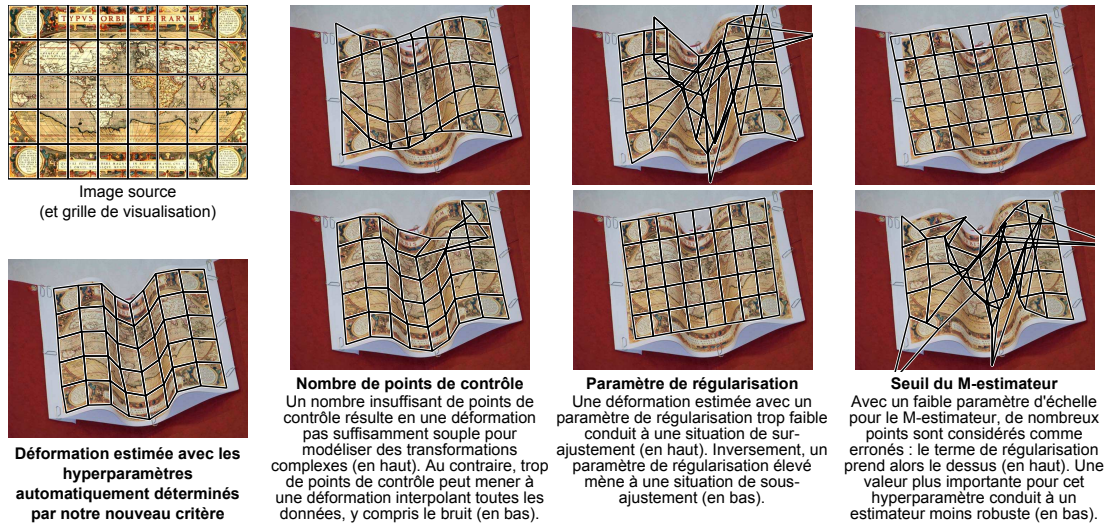


Figure 1 – Illustration de l'influence de quelques hyperparamètres sur le recalage d'images. Dans cet exemple, les paramètres naturels de la déformation sont estimés à partir de correspondances de points extraites avec SIFT [2, 3]. L'approche proposée dans cet article permet de déterminer automatiquement des valeurs adéquates pour les hyperparamètres en combinant de manière judicieuse les approches géométriques et photométriques.

valeurs γ et λ sont deux exemples d'hyperparamètres de coût. D'autres hyperparamètres apparaissent si l'on décide, par exemple, d'ajouter d'autres termes. Les atouts de l'approche géométrique sont qu'elle fonctionne lorsque la magnitude des déformations est importante et qu'elle est efficace en terme de temps de calcul². Cependant, l'approche géométrique ne permet pas de déterminer les hyperparamètres. En particulier, comme expliqué en §2, il est *impossible* d'inclure les hyperparamètres directement dans le problème (1), c.-à-d. $\min_{\mathbf{P}, \theta} \mathcal{E}(\mathbf{P}, \theta)$.

L'autre approche au recalage d'images est l'*approche photométrique* (ou *approche directe*) [10, 11] qui détermine les paramètres de la transformation en minimisant la différence de couleur entre les pixels de l'image cible transformée et ceux de l'image source. L'avantage principal de cette méthode réside dans l'importante densité des données (c.-à-d. les pixels) utilisées pour estimer les paramètres de la déformation. De la même manière qu'avec l'approche géométrique, l'approche photométrique à elle seule ne permet pas de déterminer les hyperparamètres.

Dans la mesure où les hyperparamètres ne peuvent pas être trivialement estimés, ils sont souvent fixés empiriquement et de manière définitive pour une application donnée. Il est aussi possible de les déterminer par tâtonnement pour chaque paire d'images à recalcer. Bien entendu, cette technique n'est pas satisfaisante à cause de son manque d'automatisme et de son manque d'« objectivité ». Il existe des méthodes génériques permettant de sélectionner automatiquement certains hyperparamètres. Elles consistent généralement à minimiser un critère qui évalue la qualité de la transformation estimée en fonction des hyperparamètres.

2. Cela est particulièrement vrai lorsque l'on utilise un détecteur de points efficace (comme SIFT [2] ou SURF [9]) et un bon appareteur (comme celui implémenté dans [3]).

Par exemple, ces critères peuvent mesurer la capacité d'une déformation à prédire de nouvelles données. Le critère informatif d'Akaike (*Akaike Information Criterion*) [12]), le critère C_P de Mallow (*Mallow's C_P*) [13], le critère de longueur minimale de description (*Minimum Description Length*) ou les techniques de validation croisée [8, 14, 15] (détaillées en §2) sont des exemples de telles approches. Notons néanmoins qu'aucune de ces méthodes n'est spécifique au problème du recalage d'images et que, par conséquent, aucune n'exploite pleinement les particularités des données de ce problème.

Les approches existantes pour la sélection des hyperparamètres ont la caractéristique commune de n'employer uniquement que les correspondances de points. Or, en recalage d'images, une autre information est disponible : la photométrie. Nous proposons donc un nouveau critère qui utilise *toute* l'information disponible : les correspondances de points sont utilisées comme un jeu d'entraînement et l'information photométrique est utilisée comme un jeu de test. Autrement dit, nous proposons de combiner les avantages des deux approches classiques : l'approche géométrique est employée pour déterminer les paramètres naturels de la déformation tandis que les principes de l'approche photométrique servent à la sélection des hyperparamètres. Notre critère est plus flexible que les approches statistiques classiques dans la mesure où des hyperparamètres de types différents (entiers ou réels) peuvent être simultanément recherchés. Des expérimentations sur données synthétiques et réelles sont menées en §4 pour différents hyperparamètres avec comme modèle de déformation des B-splines bidimensionnelles.

Notations. Les scalaires sont notés en italique (x), les vecteurs en gras (\mathbf{p}) et les matrices en caractères sans-

serif (M). La norme euclidienne d'un vecteur \mathbf{v} est notée $\|\mathbf{v}\|$. Les images, notées en fonte calligraphique (\mathcal{I}), sont considérées comme des fonctions de \mathbb{R}^2 dans \mathbb{R}^c avec c le nombre de canaux. Leur évaluation pour des coordonnées non entières est réalisée par interpolation bilinéaire.

2 Travaux antérieurs sur la sélection d'hyperparamètres

2.1 Sélection automatique

Différents hyperparamètres ont été présentés dans l'introduction. Il est important de comprendre que des résultats aberrants sont obtenus si l'on introduit les hyperparamètres directement dans le problème (1). Par exemple, si le terme de régularisation est un terme toujours positif, mettre λ à 0 est la manière la plus simple de diminuer sa contribution. De même, le coût est artificiellement diminué lorsque l'échelle γ du M-estimateur tend vers 0 puisque cela conduit à considérer comme erronés la quasi totalité des points (et que le coût assigné à de tels points tend vers 0 quand $\gamma \rightarrow 0$).

L'approche habituelle pour construire une méthode de sélection automatique des hyperparamètres consiste à créer un critère \mathcal{C} qui évalue la qualité d'un jeu donné d'hyperparamètres [8, 16]. L'estimation conjointe des paramètres et des hyperparamètres est alors réalisée en optimisant le problème de minimisation imbriqué suivant :

$$\min_{\mathbf{P}} \mathcal{E}(\mathbf{P}; \arg \min_{\boldsymbol{\theta}} \mathcal{C}(\boldsymbol{\theta})). \quad (3)$$

Remarquons bien que l'introduction du critère \mathcal{C} rend le problème (3) complètement différent du problème dégénéré $\min_{\mathbf{P}, \boldsymbol{\theta}} \mathcal{E}(\mathbf{P}; \boldsymbol{\theta})$.

2.2 La validation croisée

La validation croisée (VC) est un principe général à fondements statistiques communément utilisée pour déterminer de manière automatique des hyperparamètres [16]. Dans le contexte du recalage géométrique d'images, une procédure de VC consiste à minimiser un critère (fonction des hyperparamètres) mesurant la capacité qu'a une déformation (estimée pour un jeu donné d'hyperparamètres) à prédire de nouveaux points. Pour cela, l'ensemble des correspondances est partitionné. Chacune des parties est ensuite utilisée alternativement comme jeu d'entraînement et comme jeu de test afin de construire le critère de la VC. Ce critère a été utilisé par [8] dans le cadre du recalage d'images. Nous présentons maintenant deux variantes de la VC : la VC Ordinaire et la *V-fold Cross-Validation*.

Validation Croisée Ordinaire (VCO). Étant donné un jeu d'hyperparamètres $\boldsymbol{\theta}$, soit $\mathbf{P}_{\boldsymbol{\theta}}^{(k)}$ la matrice des paramètres de la déformation estimée en laissant la k -ème correspondance de côté. Le critère de la VCO est défini ainsi :

$$\mathcal{C}_{VCO}(\boldsymbol{\theta}) = \frac{1}{n} \sum_{k=1}^n \left\| \mathbf{q}_k - \mathcal{W}(\mathbf{p}_k; \mathbf{P}_{\boldsymbol{\theta}}^{(k)}) \right\|^2. \quad (4)$$

Sélectionner les hyperparamètres avec la VCO consiste à minimiser \mathcal{C}_{VCO} par rapport à $\boldsymbol{\theta}$. La VCO présente différents inconvénients. Premièrement, son calcul est extrêmement coûteux : l'évaluation avec (4) de \mathcal{C}_{VCO} pour une seule valeur de $\boldsymbol{\theta}$ nécessite l'estimation de chacune des n matrices $\{\mathbf{P}_{\boldsymbol{\theta}}^{(k)}\}_{k=1}^n$. Il existe des approximations de la formule (4) pour réduire les temps de calcul mais celles-ci ne sont valables que dans le cadre d'une estimation des paramètres par moindres carrés [8, 17]. Deuxièmement, le critère \mathcal{C}_{VCO} n'est pas robuste aux correspondances de points erronées. Enfin, les valeurs calculées avec le critère \mathcal{C}_{VCO} ne sont pas fiables lorsque le nombre de correspondances de points est faible [16].

V-fold Cross-Validation (VCV). Le principe de la VCV est de diviser l'ensemble des correspondances de points en V sous-ensembles disjoints de tailles à peu près identiques (avec V souvent choisi comme $V = \min(\sqrt{n}, 10)$). Une étude détaillée de la VCV est donnée dans [14]. Soit $\mathbf{P}_{\boldsymbol{\theta}}^{[v]}$ la matrice des paramètres de la déformation obtenue en laissant le v -ème sous-ensemble de côté et soit m_v le nombre de points de ce même sous-ensemble. Le critère de la VCV est donné par :

$$\mathcal{C}_{VCV}(\boldsymbol{\theta}) = \sum_{v=1}^V \frac{m_v}{n} \sum_{k=1}^{m_v} \frac{1}{m_v} \left\| \mathbf{q}_k - \mathcal{W}(\mathbf{p}_k; \mathbf{P}_{\boldsymbol{\theta}}^{[v]}) \right\|^2. \quad (5)$$

Comme pour la VCO, ce critère n'est pas robuste aux données erronées. Il est cependant possible de le rendre robuste [14] en remplaçant la moyenne $\sum_{k=1}^{m_v} \frac{1}{m_v} \left\| \mathbf{q}_k - \mathcal{W}(\mathbf{p}_k; \mathbf{P}_{\boldsymbol{\theta}}^{[v]}) \right\|^2$ dans l'équation (5) par une mesure plus robuste comme la moyenne tronquée.

2.3 Autres approches

Il existe d'autres approches permettant de sélectionner automatiquement des hyperparamètres. En voici quelques exemples : le critère informatif d'Akaike (*Akaike Information Criterion*) [12]), le critère C_P de Mallou (*Mallow's C_P*) [13] ou le critère de longueur minimale de description (*Minimum Description Length*). Il existe aussi des variantes robustes de ces critères ; par exemple, une version robuste du critère C_P de Mallou est proposée dans [13]. Le principal inconvénient de ces approches est qu'elles ont été conçues pour sélectionner un modèle parmi un ensemble donné de modèles [8]. Elles sont donc mal adaptées pour le réglage d'hyperparamètres continus comme l'échelle d'un M-estimateur. De plus, ces méthodes sont peu fiables lorsque le nombre de correspondances de points est faible.

3 Notre nouveau critère

Les approches présentées en §2 ont la caractéristique commune de n'utiliser que les correspondances de points, aussi bien pour l'estimation des paramètres que pour l'estimation des hyperparamètres. Nous proposons ici une nouvelle approche qui utilise toute l'information disponible c.-à-d.

les correspondances de points mais aussi les données photométriques. Pour cela, nous combinons les approches géométriques et photométriques de la manière suivante :

- étant donné un jeu d’hyperparamètres θ , l’approche géométrique est utilisée pour estimer les paramètres naturels de la déformation ;
- la fonction de coût typiquement optimisée dans l’approche photométrique est utilisée pour évaluer la qualité des hyperparamètres θ en se basant sur l’idée suivante : de bons hyperparamètres doivent conduire à une estimation des paramètres induisant une déformation qui minimise la différence photométrique entre l’image source et l’image cible transformée.

En d’autres termes, nous proposons d’utiliser les correspondances de points comme un jeu d’apprentissage et les pixels des images (c.-à-d. l’information photométrique) comme un jeu de test. Étant donnés des paramètres naturels P_θ issus de l’estimation de la déformation à partir de correspondances en utilisant un ensemble donné d’hyperparamètres θ , notre critère est ainsi défini :

$$\mathcal{C}_*(\theta) = \frac{1}{|\mathfrak{R}|} \sum_{\mathbf{p} \in \mathfrak{R}} \|\mathcal{S}(\mathbf{p}) - \mathcal{T}(\mathcal{W}(\mathbf{p}; P_\theta))\|^2, \quad (6)$$

où \mathcal{S} et \mathcal{T} sont respectivement les images source et cible et où $|\mathfrak{R}|$ est la région d’intérêt (de taille $|\mathfrak{R}|$), souvent définie par un rectangle obtenu en supprimant une marge au domaine de l’image source.

Notons que le critère de l’équation (6) correspond à la fonction de coût typiquement minimisée dans l’approche photométrique [10, 1]. La différence avec l’approche photométrique est que notre critère est considéré comme une fonction des hyperparamètres θ et non des paramètres naturels P de la déformation.

Robustesse. Les occultations et les spécularités sont deux exemples de phénomènes pouvant être considérés comme des données aberrantes pour notre critère. Les problèmes liés à ce genre d’erreurs peuvent être évités en utilisant une mesure plus robuste que celle utilisée dans l’équation (6), comme une moyenne tronquée :

$$\mathcal{C}'_*(\theta) = \frac{1}{\frac{100-\alpha}{100}|\mathfrak{R}|} \sum_{\mathbf{p} \in \mathfrak{R}_\alpha} \|\mathcal{S}(\mathbf{p}) - \mathcal{T}(\mathcal{W}(\mathbf{p}; P_\theta))\|^2, \quad (7)$$

où \mathfrak{R}_α est un sous-ensemble de \mathfrak{R} obtenu en enlevant les $\alpha\%$ de pixels qui produisent les valeurs les plus élevées de $\|\mathcal{S}(\mathbf{p}) - \mathcal{T}(\mathcal{W}(\mathbf{p}; P_\theta))\|^2$.

4 Résultats expérimentaux

4.1 Détails techniques

Nous spécialisons ici notre contribution générique afin de mener quelques expériences.

Modèle de déformation. Comme dans [5], nous utilisons des déformations libres à base de B-splines. Cette déformation est paramétrée par un ensemble de $\frac{l}{2}$ points de

contrôle 2D rangés dans une matrice $P \in \mathbb{R}^{\frac{l}{2} \times 2}$. Elle est définie par :

$$\mathcal{W}(\mathbf{q}; P) = \mathbf{w}(\mathbf{p})^\top P, \quad (8)$$

où $\mathbf{w} : \mathbb{R}^2 \rightarrow \mathbb{R}^l$ est la fonction définie par :

$$\mathbf{w}(\mathbf{p}) = [N_1(x)N_1(y) \quad \dots \quad N_{l_x}(x)N_{l_y}(y)]^\top. \quad (9)$$

où $\mathbf{p} = (x, y)$ et où l_x et l_y sont deux hyperparamètres donnant le nombre de points de contrôle selon l’axe des abscisses et des ordonnées respectivement (avec $l = l_x l_y$). Les fonctions $N_i : \mathbb{R} \rightarrow \mathbb{R}$ constituent la base de l’espace vectoriel des B-splines à une dimension [18].

Terme de régularisation. Nous utilisons l’énergie de torsion comme terme de régularisation :

$$\mathcal{R}(P) = \sum_{i=1}^2 \int_{\Omega} \left\| \frac{\partial^2 \mathcal{W}^i}{\partial \mathbf{p}^2}(\mathbf{p}; P) \right\|_{\mathcal{F}}^2 d\mathbf{p}, \quad (10)$$

où Ω est le domaine de définition de la déformation, \mathcal{W}^i est la i -ème coordonnée de \mathcal{W} et $\|\cdot\|_{\mathcal{F}}$ est la norme de Frobenius de la matrice hessienne de la déformation. Avec les déformations B-splines, il est possible de montrer qu’il existe une matrice B telle que $\mathcal{R}(P) = \|BP\|_{\mathcal{F}}^2$.

M-estimateur. Nous utilisons le M-estimateur de Cauchy défini par la fonction ρ suivante :

$$\rho(x; \gamma) = \log \left(1 + \frac{x^2}{\gamma^2} \right), \quad (11)$$

où $\gamma \in \mathbb{R}_+^*$ est un hyperparamètre contrôlant l’échelle du M-estimateur (c.-à-d. sa sensibilité aux données erronées). Nous omettons ici les détails mais il peut être montré qu’utiliser cet M-estimateur est raisonnable au vu des erreurs rencontrées lorsque l’on utilise SIFT ou SURF.

Optimisation. Tous les critères étudiés dans les expériences à venir sont optimisés en utilisant une recherche exhaustive. Cette approche est potentiellement coûteuse en temps de calcul mais a l’avantage de produire des résultats fiables car non soumis aux problèmes de minima locaux.

4.2 Données synthétiques

Génération des données. Une image source est générée en extrayant une zone rectangulaire d’une image texturée (choisie aléatoirement parmi 15 images). L’image cible est construite en déformant une autre partie de la même image texturée avec une déformation B-spline connue \mathcal{W}^* dont les 5×5 points de contrôle sont choisis aléatoirement de manière à ce que l’amplitude moyenne de la déformation soit de 20 pixels. Les tailles des images source et cible sont respectivement 160×160 et 320×240 pixels. Un bruit gaussien centré d’écart type égal à 5% de la valeur maximale d’un pixel est ajouté aux images. Un ensemble $\mathcal{P} = \{\mathbf{p}_i \leftrightarrow \mathbf{q}_i\}_{i=1}^n$ de correspondances de points est construit en choisissant aléatoirement des points dans l’image source et en calculant leurs correspondants dans l’image cible avec \mathcal{W}^* . Ces points sont ensuite perturbés en ajoutant des erreurs qui suivent une distribution de Cauchy avec un paramètre d’échelle de 1 pixel.

Oracle. Nous appelons *oracle* la déformation estimée à partir des points de \mathcal{P} la plus proche possible de la vérité terrain \mathcal{W}^* ³. Il s’agit de la déformation induite par les paramètres P_o et les hyperparamètres θ_o solutions de ce problème :

$$\min_{(P, \theta)} \iint_{\Omega_{\mathcal{W}^*}} \|\mathcal{W}^*(\mathbf{p}) - \mathcal{W}(\mathbf{p}; P)\| d\mathbf{p}. \quad (12)$$

Erreur Géométrique Relative (EGR). L’EGR mesure la différence entre une déformation estimée et l’oracle. Soit θ_\bullet les hyperparamètres minimisant le critère C_\bullet et soit P_{θ_\bullet} les paramètres estimés à partir des points pour les hyperparamètres θ_\bullet . L’EGR est définie par :

$$\iint_{\mathbf{p} \in \Omega_S} \frac{\|\mathcal{W}(\mathbf{p}; P_o) - \mathcal{W}(\mathbf{p}; P_{\theta_\bullet})\|}{\|\mathcal{W}(\mathbf{p}; P_o)\|} d\mathbf{p}. \quad (13)$$

Nous comparons dans la figure 2 les EGR obtenues avec différentes approches pour sélectionner l’échelle γ du M-estimateur et le paramètre de régularisation λ :

- notre critère (Photo) et ses variantes robustes pour des seuils α de 25% (Photo25) et de 50% (Photo50) ;
- la VCV (VFold) et ses variantes robustes pour des seuils α de 20% (VFold20) et de 40% (VFold40).

Le nombre de points de contrôle des déformations est fixé à 8×8 et 100 correspondances de points sont utilisées. Les valeurs reportées sont obtenues sur 100 essais différents. Nous observons sur la figure 2 que les EGR les plus faibles sont obtenues avec notre critère photométrique. Nous remarquons aussi que la VCV non robuste est très sensible à la présence de correspondances de points erronées. Ces valeurs erronées n’ont que très peu d’impact sur la version non robuste de notre critère puisque celui-ci repose principalement sur les données photométriques.

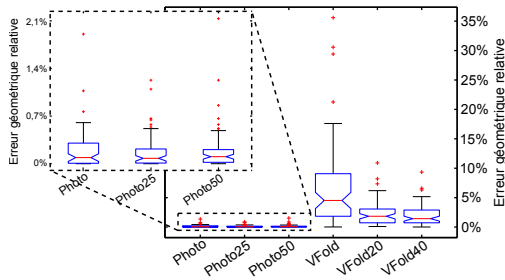


Figure 2 – « Boîtes à moustaches » des ERG obtenues pour différents critères de sélection des hyperparamètres. De manière générale, notre critère (Photo) et ses variantes robustes (Photo25, Photo50) donnent de meilleurs résultats que ceux reposant sur la VCV (VFold, VFold20 et VFold40).

4.3 Données réelles

Ici, les images sources sont des images numériques et les images cibles sont obtenues en imprimant les images

3. En fonction des modèles de déformation et des correspondances de points, l’oracle et la vérité terrain ne sont pas nécessairement identiques.

sources puis en les photographiant. Une vérité terrain est établie en cliquant plusieurs centaines de points à la main. Notons que l’illustration de la figure 1 est un exemple de recalage sur de telles données.

Image cubiste. La figure 3 montre les recalages obtenus avec différents critères pour la sélection de ces hyperparamètres : paramètre de régularisation, échelle du M-estimateur et nombre de points de contrôle. L’algorithme SIFT [3] a été utilisé pour extraire 314 correspondances de point, dont approximativement 8% de fausses. De manière générale, nous observons sur la figure 3 que les résultats obtenus avec notre critère sont meilleurs que ceux obtenus avec le critère de la VCV. Le tableau 1 confirme cette constatation. Cela provient principalement du fait qu’il y a un manque de correspondances de points dans la partie inférieure droite des images (notre approche permet de combler ce manque).

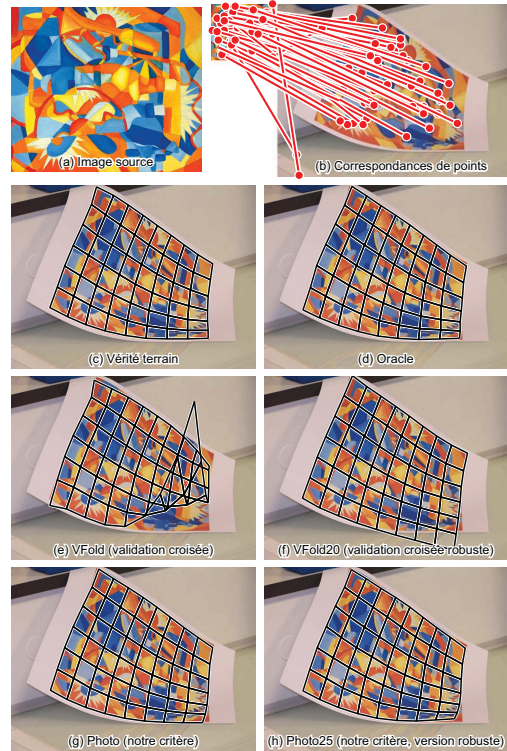


Figure 3 – Recalage d’images avec 3 hyperparamètres réglés par des critères différents. Dans ce test, les 2 variantes de notre approche donnent les meilleurs résultats.

Critère	EGR
VCV	1.852%
VCV (variante robuste)	0.675%
Notre critère	0.190%
Notre critère (variante robuste)	0.197%

Tableau 1 – Erreur géométrique relative (EGR) pour l’expérience de la figure 3.

« **Waterfall** » de Maurits Escher. Nous reportons dans la figure 4 les résultats obtenus pour des tests similaires au test précédent. Ne sont considérés ici que les hyperparamètres de régularisation et d'échelle du M-estimateur. Une occultation a été artificiellement ajoutée dans l'image cible. L'algorithme SURF [9] a été utilisé pour extraire 621 correspondances, dont approximativement 12% de fausses. Comme dans le cas précédent, les hyperparamètres estimés avec nos critères donnent de meilleurs résultats que ceux estimés avec le critère de la VCV. Pour les deux critères, les versions robustes sont meilleures que les versions non robustes. Bien que le critère de la VCV ne repose pas sur l'information photométrique, la présence d'une occultation dans l'image cible influence quand même ce critère puisqu'elle introduit de fausses correspondances de points.

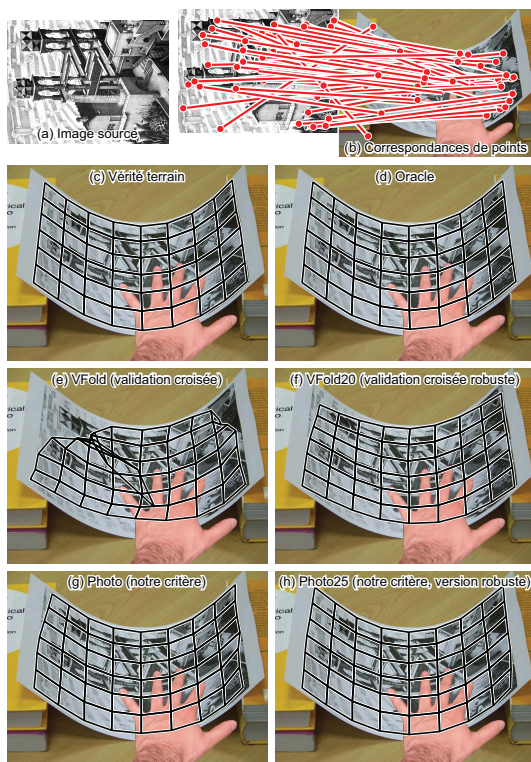


Figure 4 – Recalage d'images avec 2 hyperparamètres sélectionnés par des critères différents. De manière générale, les variantes robustes aboutissent à de bons résultats. Les meilleurs résultats sont obtenus avec notre critère robuste.

5 Conclusion

Nous avons proposé un nouveau critère permettant de sélectionner automatiquement des hyperparamètres en recalage d'images. Nous avons montré qu'en général notre critère aboutit à des hyperparamètres meilleurs que ceux obtenus avec d'autres méthodes. Cela est rendu possible par la combinaison des approches géométriques et photométriques au recalage d'images. Bien que pour des raisons pratiques nous ayons limité nos expériences à certains hyperparamètres particuliers, notre approche est générique et

pourrait donc être utilisée avec d'autres hyperparamètres ou d'autres modèles de déformations.

Références

- [1] R. Szeliski. Image alignment and stitching: A tutorial. *Foundations and Trends in Computer Graphics and Vision*, 2:1–104, 2006.
- [2] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60:91–110, 2004.
- [3] A. Vedaldi et B. Fulkerson. VLFeat: An open and portable library of computer vision algorithms. <http://www.vlfeat.org/>, 2008.
- [4] R. Hartley et A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004.
- [5] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, et D. Hawkes. Nonrigid registration using free-form deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging*, 18:712–721, 1999.
- [6] F. Bookstein. Principal warps: Thin-Plate Splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:567–585, 1989.
- [7] J. Modersitzki. *Numerical Methods for Image Registration*. Oxford Science, 2004.
- [8] A. Bartoli. Maximizing the predictivity of smooth deformable image warps through cross-validation. *Journal of Mathematical Imaging and Vision*, 31(2-3):133–145, 2008.
- [9] H. Bay, A. Ess, T. Tuytelaars, et L. Van Gool. SURF: Speeded up robust features. *Computer Vision and Image Understanding*, 110:346–359, 2008.
- [10] M. Irani et P. Anandan. About direct methods. Dans *Workshop on Vision Algorithms*, 1999.
- [11] S. Baker et I. Matthews. Lucas-Kanade 20 years on: A unifying framework. *International Journal of Computer Vision*, 56:221–255, 2004.
- [12] M. Cetin et A. Erar. Variable selection with Akaike information criteria: a comparative study. *Hacettepe Journal of Mathematics and Statistics*, 31:89–97, 2002.
- [13] E. Ronchetti et R. Staudte. A robust version of Mallows's C_p . *Journal of the American Statistical Association*, 89:550–559, 1994.
- [14] J. De Brabanter, K. Pelckmans, J. Suykens, J. Vandewalle, et B. De Moor. Robust cross-validation score functions with application to weighted least squares support vector machine function estimation. Rapport technique, Katholieke Universiteit Leuven, 2003.
- [15] G. Wahba et S. Wold. A completely automatic French curve: fitting spline functions by cross-validation. *Communications in Statistics*, 4:1–17, 1975.
- [16] G. Wahba. *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, 1990.
- [17] M. Farenzena, A. Bartoli, et Y. Mezouar. Efficient camera smoothing in sequential structure-from-motion using approximate cross-validation. Dans *European Conference on Computer Vision*, 2008.
- [18] P. Dierckx. *Curve and Surface Fitting with Splines*. Oxford University Press, 1993.