Available online at www.sciencedirect.com

ScienceDirect

C. R. Acad. Sci. Paris, Ser. I 346 (2008) 593–598

Comptes Rendus Mathematique

http://france.elsevier.com/direct/CRASS1/

Theory of Signals/Numerical Analysis

A fast algorithm for image registration

Jérôme Fehrenbach, Mohamed Masmoudi

Institut de mathématiques de Toulouse, Université de Toulouse, 31062 Toulouse cedex 09, France

Received 21 November 2007; accepted after revision 12 March 2008

Presented by Yves Meyer

Abstract

An algorithm is presented here to estimate a smooth motion at a high frame rate. It is derived from the non-linear constant brightness assumption. A hierarchical approach reduces the dimension of the space of admissible displacements, hence the number of unknown parameters is small compared to the size of the data. The optimal displacement is estimated by a Gauss–Newton method, and the matrix required at each step is assembled rapidly using a finite-element method. ***To cite this article: J. Fehrenbach, M. Masmoudi, C. R. Acad. Sci. Paris, Ser. I 346 (2008).***

Résumé

Un algorithme de recalage rapide d'images. On propose un algorithme permettant d'estimer un mouvement régulier de façon rapide. Cet algorithme est basé sur l'hypothèse de préservation du niveau de gris. Une approche hiérarchique permet de réduire l'espace des déplacements admissibles, et le nombre de paramètres est faible devant la taille des données. La fonction coût non-linéaire donnant le déplacement optimal est minimisée par la méthode de Gauss–Newton, et la matrice nécessaire à chaque pas est assemblée efficacement à l'aide d'une méthode d'éléments finis. ***Pour citer cet article : J. Fehrenbach, M. Masmoudi, C. R. Acad. Sci. Paris, Ser. I 346 (2008).***

Version française abrégée

Lorsque l'on observe un milieu qui se déforme, les images dépendent du temps t. Notons $I(x,t)$ l'intensité lumineuse du point x à l'instant t et $\mathbf{v}(x,t)$ la vitesse instantanée. L'hypothèse de préservation du niveau de gris est donnée par l'équation (1). Nous proposons ici d'utiliser la version intégrée de (1) entre deux instants où l'on observe les images I_0 et I_1, qui est l'équation (2). Il s'agit d'une équation non-linéaire que l'on cherche à résoudre par une méthode de type moindre carrés. Il s'agit de trouver un champ de déplacement admissible qui va minimiser la fonction-coût donnée par (3).

Nous proposons de réduire l'espace des inconnues, en considérant des déplacements qui sont linéaires par morceaux par rapport à chaque variable d'espace, selon un maillage qui est défini par une grille de contrôle $\mathcal{V}$, voir Fig. 1. Nous pouvons ainsi minimiser la fonction-coût non-linéaire (3) sur un espace de dimension relativement faible.

E-mail addresses: jerome.fehrenbach@math.univ-toulouse.fr (J. Fehrenbach), mohamed.masmoudi@math.univ-toulouse.fr (M. Masmoudi).

1631-073X/$ – see front matter
doi:10.1016/j.crma.2008.03.019

La minimisation de la fonction-coût se fait à l'aide de la méthode de Gauss–Newton. C'est une méthode itérative qui nécessite à chaque pas de résoudre l'équation (4). L'introduction d'un produit scalaire adapté sur l'espace des champs de déplacements admissibles permet d'assembler la matrice nécessaire à chaque pas de Gauss–Newton rapidement, avec une méthode d'éléments finis. Le second membre de (4) est aussi assemblé rapidement.

Un exemple est présenté sur des images synthétiques (Fig. 2) où de grands déplacements sont retrouvés grâce à une implémentation multigrille de l'algorithme. Un autre exemple présente une image réelle déformée à l'aide d'un champ de déplacement connu (Fig. 3). Pour traiter ces exemples, un terme de régularisation spatiale a été rajouté à la fonction-coût, il s'agit de la norme L^2 du gradient du champ de déplacement.

1. Introduction

When a medium moving according to a velocity field is observed, the images depend on the time variable t. Let $I(x,t)$ be the intensity of the point x at time t and $\mathbf{v}(x,t)$ the instantaneous velocity. The constant brightness assumption asserts that every point of the medium has a fixed grey level. This can be written as

$$\frac{\mathrm{d}I(x,t)}{\mathrm{d}t} = \nabla_x I(x,t)\cdot\mathbf{v}(x,t) + \frac{\partial I(x,t)}{\partial t} = 0. \tag{1}$$

Numerous methods use this model, in various ways. The vast family of optical-flow methods [2] consists of an approximation of this equation by finite differences both in time and in space (real images are pixelized and the sequence is sampled at a finite time step). This amounts to linearize the constant brightness assumption. Several models are considered for the deformation, the pioneer work is [5] where a local affine transformation is considered. Global models add a spatial constraint, see [4] where a displacement vector is estimated at each point and a fixed point algorithm is used to minimize the global cost-function. Multigrid methods have been developed [1,6].

Large deformations are considered in numerous papers, see for instance [7] where a combination of segmentation and intensity-based registration is described, or [3] where an optimal diffeomorphism between I_0 and I_1 is retrieved. To our knowledge, no real-time results are reported.

We propose here a fast non-linear registration method, where the velocity field belongs to a small dimensional vector space, namely piecewise bilinear vector fields. We use the non-linear version of the constant brightness equation (and not the linear version considered in [4]), and observe that it is possible to minimize rapidly the cost-function. More precisely, our contribution is to propose a Gauss–Newton algorithm to register two images, and show that the product of the Jacobian matrix of the cost function by its transpose can be assembled rapidly using a finite-element method. This is possible since an adequate inner product is defined on the set of admissible displacements. The convergence properties of Gauss–Newton algorithm provide an accurate estimate of the motion, and a fine spatial resolution can be achieved by a hierarchical estimation of the displacement.

The application we have in mind is the estimation of the deformations of a medium that is imaged (in particular medical or geophysical images). The type of motion to consider is different from the case where (rigid) objects move in a scene. It makes sense for these applications to look for a smooth deformation field, our principal mathematical assumption is that the deformation field is Lipschitz.

2. The continuous problem

We assume that images belong to $C^1(\Omega,\mathbf{R})$, where Ω is a rectangular domain. Let $\mathcal{L}$ be the set of Lipschitz vector fields on Ω. The space $\mathcal{L}_{\mathrm{ad}}$ of admissible displacement vector fields is the set of vector fields $\mathbf{V}\in\mathcal{L}$ that satisfy the following conditions:

(i) $\mathrm{Lip}(\mathbf{V}) < 1$ where $\mathrm{Lip}(\mathbf{V})$ is the Lipschitz constant of $\mathbf{V}$, this ensures that $\mathrm{Id}+\mathbf{V}$ is a homeomorphism,
(ii) $\|\mathbf{V}\|_\infty < K$ where K is an a-priori bound on the displacement between I_0 and I_1.

Let $\mathbf{V}$ be the motion between two successive images I_0 and I_1. The grey-level of the points is assumed to be constant during time (constant brightness assumption). We do not linearize the constant brightness equation. It can be stated as follows: find a displacement field $\mathbf{V}\in\mathcal{L}_{\mathrm{ad}}$ such that

$$I_1(\mathrm{Id}+\mathbf{V}) = I_0. \tag{2}$$

The equality (2) is not likely to be exactly satisfied, due to measurement errors and to the approximation induced by the constant brightness assumption. This equality is to be satisfied in the least squares sense. The optimal displacement field $\mathbf{V}$ is given by

$$\min_{\mathbf{V}\in\mathcal{L}_{\mathrm{ad}}} \frac{1}{2}\int_{\Omega} F(\mathbf{V})^2, \tag{3}$$

where $F:\mathcal{L}_{\mathrm{ad}} \to C^0(\Omega,\mathbf{R})$ is defined by $F(\mathbf{V}) = I_1(\mathrm{Id}+\mathbf{V}) - I_0$.

Note. In practise, spatial regularization terms are added to the cost function. This does not affect the calculations below.

Routine calculations show that the map F is differentiable in $\mathcal{L}_{\mathrm{ad}}$ and for $\mathbf{V}\in\mathcal{L}_{\mathrm{ad}}$ and $\mathbf{d}\in\mathcal{L}$, we have

$$DF(\mathbf{V})\cdot\mathbf{d} = G, \quad \text{where } G(x) = \nabla I_1\big(x+\mathbf{V}(x)\big)\cdot\mathbf{d}(x).$$

3. Discretization

We consider displacement fields $\mathbf{V}$ that are piecewise affine with respect to each space variable, on rectangles of size $h_x \times h_y$. A grid of control points is defined, with coordinates in $X\times Y$, where $X = (0, h_x, 2h_x, \ldots, Nh_x)$ and $Y = (0, h_y, 2h_y, \ldots, Mh_y)$, see Fig. 1. Let $\mathcal{V} = X\times Y$ be the set of all control points, and $\mathcal{R}$ the set of all rectangles of the grid. The vector space $\mathcal{C}$ of vector fields that are bilinear on each rectangle of $\mathcal{R}$ is of dimension $2(N+1)(M+1)$, since an element of $\mathcal{C}$ is determined by its values at the points of $\mathcal{V}$.

The space $\mathcal{C}$ is equipped with the following inner product:

$$\big(\mathbf{V}^1 \mid \mathbf{V}^2\big)_{\mathcal{C}} = \sum_{x\in\mathcal{V}} \big(\mathbf{V}^1(x) \mid \mathbf{V}^2(x)\big),$$

where $(\mathbf{V}^1(x) \mid \mathbf{V}^2(x))$ denotes the usual inner product in $\mathbf{R}^2$. Note that it differs from the classical inner product in $L^2(\Omega,\mathbf{R}^2)$. In the following, we still denote by F the restriction of $F:\mathcal{C}\to C^0(\Omega,\mathbf{R})$.

We will also use the following bilinear map (point-wise inner product):

$$\mathcal{C}\times\mathcal{C} \longrightarrow C^0(\Omega,\mathbf{R}),$$
$$(\mathbf{V}^1,\mathbf{V}^2) \longmapsto \mathbf{V}^1\odot\mathbf{V}^2,$$

such that for every point $x\in\Omega$, $\mathbf{V}^1\odot\mathbf{V}^2(x) = (\mathbf{V}^1(x) \mid \mathbf{V}^2(x))$.

4. Minimization

The minimization of the cost function defined in (3) is performed with a Gauss–Newton algorithm. Let $\mathbf{V}_0$ be the initial guess. We define a sequence $(\mathbf{V}_k)_k$ of iterates, with the Gauss–Newton update: $\mathbf{V}_{k+1} = \mathbf{V}_k + \mathbf{d}_k$, where $\mathbf{d}_k$ is the solution of

$$DF^T DF\mathbf{d}_k = -DF^T F, \tag{4}$$

where DF denotes $DF(\mathbf{V}_k)$. Eq. (4) can be solved using any linear solver, since the size of the unknown is small. We use the conjugate gradient method without preconditioning.

5. Assembling the matrix $DF^T DF$ and the vector $DF^T F$

Let $\mathcal{E}$ be the orthonormal basis of $\mathcal{C}$ defined as follows: $\mathcal{E} = (\mathbf{e}_i^x)_{i\in\mathcal{V}} \cup (\mathbf{e}_i^y)_{i\in\mathcal{V}}$, where $\mathbf{e}_i^x$ is the vector field that is zero at every control point of $\mathcal{V}$ except at the vertex i where it is directed along the horizontal axis, and $\mathbf{e}_i^y$ is the vector field that is zero at every control point of $\mathcal{V}$ except at the vertex i where it is directed along the vertical axis, see Fig. 1 for an example.

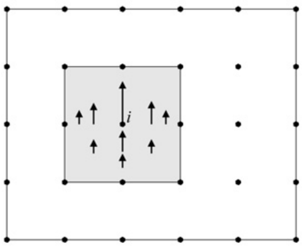

Fig. 1. The points of $\mathcal{V}$ are indicated by dots. An elementary displacement vector field $\mathbf{e}_i^y$ is shown: $\mathbf{e}_i^y$ is zero at the points of the non-shaded area, it is directed along the vertical with norm 1 at the vertex i, and is piecewise linear with respect to each space variable (for reason of clarity, the values at only some points are indicated).

If $\mathbf{V} \in \mathcal{C} \cap \mathcal{L}_{\mathrm{ad}}$ and $\mathbf{d} \in \mathcal{C}$,

$$DF(\mathbf{V}) \cdot \mathbf{d} = \nabla I_1(\mathrm{Id}+\mathbf{V}) \odot \mathbf{d}.$$

Let $\mathbf{V} \in \mathcal{C} \cap \mathcal{L}_{\mathrm{ad}}$ and $DF = DF(\mathbf{V})$. We are interested in evaluating the matrix $DF^T DF$. The coefficient of place (k,l) in the matrix $DF^T DF$ is

$$\left(DF^T DF\right)_{k,l} = \left(DF^T DF\mathbf{e}_k \mid \mathbf{e}_l\right)_{\mathcal{C}} = (DF\mathbf{e}_k \mid DF\mathbf{e}_l)_{L^2(\Omega)}.$$

Since the elementary displacements $\mathbf{e}_i^x$ and $\mathbf{e}_i^y$ are non-zero in the rectangles adjacent to the vertex i, the coefficient of place (k,l) in the matrix $DF^T DF$ is non-zero only for displacements $\mathbf{e}_k$ and $\mathbf{e}_l$ that are associated with vertices that are corners of one common rectangle. The matrix $DF^T DF$ has thus a sparse structure, and it can be assembled like a finite-element matrix, as we explain now.

$$\begin{aligned} DF^T DF &= \sum_{k,l}\left(DF^T DF\right)_{k,l}\mathbf{e}_k \otimes \mathbf{e}_l = \sum_{k,l}\int_\Omega \left(\nabla I_1(x') \mid \mathbf{e}_k(x)\right)\left(\nabla I_1(x') \mid \mathbf{e}_l(x)\right)\mathbf{e}_k \otimes \mathbf{e}_l \,\mathrm{d}x \\ &= \sum_{k,l}\sum_{R\in\mathcal{R}}\int_R \left(\nabla I_1(x') \mid \mathbf{e}_k(x)\right)\left(\nabla I_1(x') \mid \mathbf{e}_l(x)\right)\mathbf{e}_k \otimes \mathbf{e}_l \,\mathrm{d}x \\ &= \sum_{R\in\mathcal{R}}\sum_{k,l}\int_R \left(\nabla I_1(x') \mid \mathbf{e}_k(x)\right)\left(\nabla I_1(x') \mid \mathbf{e}_l(x)\right)\mathbf{e}_k \otimes \mathbf{e}_l \,\mathrm{d}x, \end{aligned}$$

where we write $x' = x + \mathbf{V}(x)$ for the sake of concision. For a given rectangle $R \in \mathcal{R}$, the set of $\mathbf{e}_k, \mathbf{e}_l$ to be considered are the elementary vector fields attached to the 4 corners of the rectangle R. There are 8 such vector fields (one in each direction for each of the 4 corners), and 8 quantities of the form $(\nabla I_1(x + \mathbf{V}(x)) \mid \mathbf{e}_k(x))$ must be computed. A loop is performed over the rectangles, at each step 64 coefficients of the matrix $DF^T DF$ are updated, but for symmetry reasons, only 36 different quantities are evaluated and they are straightforward to compute once $\nabla I_1(x + \mathbf{V}(x))$ is known.

The vector field $DF^T F \in \mathcal{C}$ is assembled rapidly in a similar way, we do not give the details here.

6. The algorithm

As a summary, we indicate here the algorithm that is proposed:

input: *images* I_0, I_1, *initial guess for the displacement* $\mathbf{V}_0$

1. set $k := 0$,
2. assemble the matrix $DF^T(\mathbf{V}_k)DF(\mathbf{V}_k)$ and the vector $DF^T(\mathbf{V}_k)F(\mathbf{V}_k)$,
3. solve the equation $DF^T DF\mathbf{d}_k = -DF^T F$ for $\mathbf{d}_k$, using conjugate gradient without preconditioning,
4. set $\mathbf{V}_{k+1} := \mathbf{V}_k + \mathbf{d}_k$, set $k := k+1$,
5. if the stopping criterion is not met, go back to step 2.

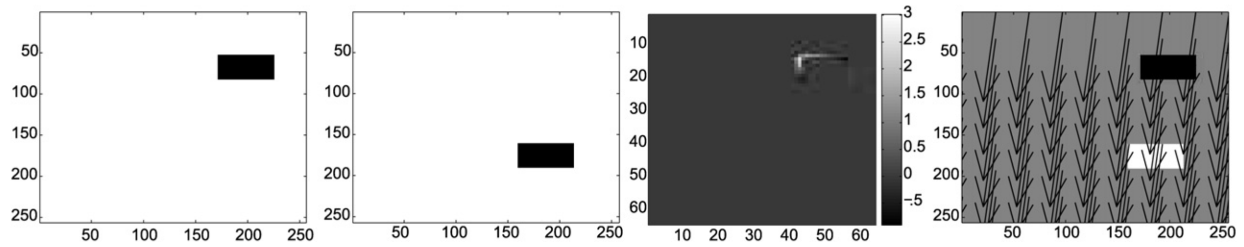

Fig. 2. The original image I_0; the target image I_1; the difference after registration at a coarse level of the multigrid algorithm; the estimated vector field.

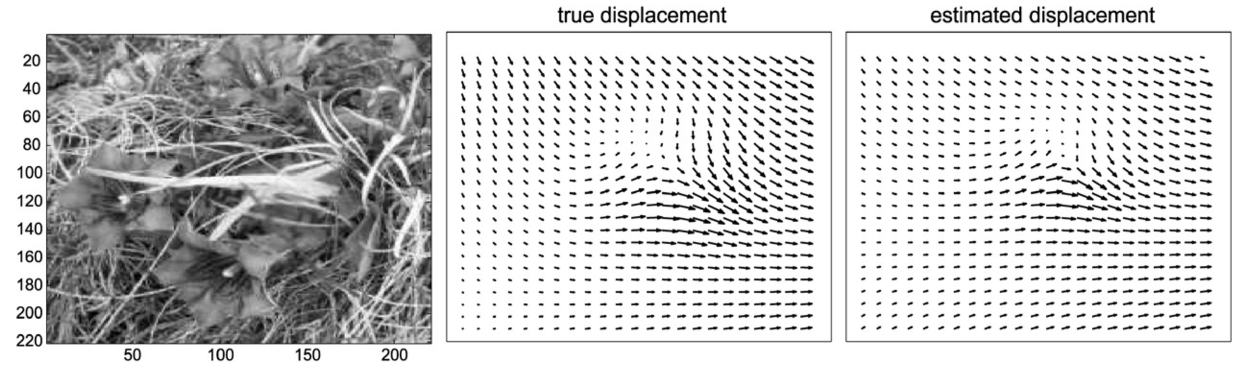

Fig. 3. The original image I_0; the true displacement vector field (the vector fields are undersampled and dilated for a sake of clarity); the estimated vector field.

The stopping criterion in step 5 is to fix in advance the number of steps.

Note. at the beginning, when no initial guess is available, we take $\mathbf{V}_0 \equiv 0$ as initial guess.

A hierarchical estimation of the displacement consists of applying this algorithm with a small number of control points (N and M small) and use the estimated displacement as initial guess for a finer grid (increase N and M, in practice by a factor 2). The process can be iterated to achieve a desired spatial resolution for the displacement.

7. Examples

To detect large displacements, a multigrid variant of the method was implemented. The initial and final images are coarsened, an estimation of the displacement is obtained on the coarse images and used as initial guess to register the finer images.

The first example is presented in Fig. 2, it consists of two synthetic images of size 256×256. The displacement was estimated in 10 Gauss–Newton steps at each resolution level and a spatial regularization term was added to the cost function (Tikhonov regularization of the norm of the gradient of the displacement). The total processing time is 4.3 s using a MatLab(R) implementation.

A second example is presented in Fig. 3, where a real image I_0 is used. A known displacement field was applied to this image, it is a sum of rotation, translation and local divergence. This forms an image I_1. The hierarchical algorithm described in Section 6 is applied to retrieve the deformation between the images I_0 and I_1. In this example, the motion is far from linear, but the estimation is reasonable.

References

[1] P. Ananden, A computational framework and an algorithm for the measurement of visual motion, Int. J. Comput. Vis. 2 (1989) 283–310.
[2] S. Beauchemin, J. Barron, The computation of optical flow, ACM Computing Surveys 27 (1995) 433–467.
[3] M. Beg, M. Miller, A. Trouvé, L. Younes, Computing large deformations metric mappings via geodesic flows of diffeomorphisms, Int. J. Comput. Vis. 61 (2) (2005) 139–157.

[4] B. Horn, B. Schunck, Determining optical flow, Artificial Intelligence 17 (1981) 185–204.
[5] B. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, in: Proc. Seventh International Joint Conference on Artificial Intelligence, Vancouver, Canada, 1981, pp. 674–679.
[6] E. Memin, P. Perez, Dense estimation and object-based segmentation of the optical flow with robust techniques, IEEE Trans. Image Proc. 7 (1998) 703–719.
[7] F. Richard, L. Cohen, A new image registration technique with free boundary constraints: application to mammography, Computer Vision and Image Understanding 89 (2003) 166–196.