

A variational approach for correspondence finding between three rectified images

Jeremy Yrmeyahu Kaminski^{1,*},
Michael Belhdeb², and Mina Teicher²

¹*Computer Science Department, Holon Institute of Technology,
52 Golomb St., Holon 58102, Israel*

²*Department of Mathematics, Bar-Ilan University,
Ramat-Gan 52900, Israel*

Corresponding author: kaminsj@hit.ac.il

Received 10 May 2007, accepted 9 July 2007

Abstract

We present a variational approach for the computation of dense correspondence between three rectified images. Using rectification allows reducing the problem to a single parameter functional. Classically, the functional is composed of a data term and a regularization term. We introduce an improved model, where the pixel wise data term is combined with a template matching approach, which is well adapted to the case of rectified images. This modification increases the accuracy significantly.

The minimization is practically handled by solving the Euler-Lagrange equation. In our setting, the Euler-Lagrange equation is replaced by a parabolic evolution equation. Under general assumptions, we prove that this equation has a single asymptotic state, that is obtained by any initial guess.

1 Introduction

Recent works used a variational approach to recover a dense disparity map from a set of two weakly calibrated stereoscopic images [2, 3]. In this context, the computation of this apparent motion, the optical flow, uses the *fundamental matrix* [4], to relate corresponding pixels in the two views. An energy term is minimized, taking into account the geometric constraint defined by the fundamental matrix, as well as the intensity information with an appropriate regularization term. The associated Euler-Lagrange equations allow to solve this minimization problem, and so to find the speed vector.

More images bring more information and so, more accuracy in the results. In this work we shall use the generalization of the fundamental matrix to three views, the trilinear tensor [6]. Moreover, the rectification embeds the geometric information provided by the trifocal tensor. We shall investigate more in depth this approach by using a triplet of rectified images, produced by the method of Zhang [7] and his team, after the work of Hartley [8]. In the framework of the variational approach, it allows us to compute only one Euler-Lagrange equation to solve the problem of minimization.

We shall first in Section 2 present the rectification method we used. Then in Section 3, the minimization problem is presented, with the mathematical foundation. Finally, some experiments are shown in Section 4.

2 Rectification

The use of the process of rectification is of major importance on practical applications. It consists in re-sampling pairs of stereo images taken from different viewpoints in order to produce a pair of "matched epipolar projections". In the case of three images, a conventional disposition is a reference bottom image, a second right image, and finally a top image. Concerning the bottom and right images, there are projections in which the epipolar lines run parallel to the x-axis and consequently, disparities between the images are in the x direction only. In practice, the epipolar lines are to be transformed to lines parallel to the x-axis, then the epipoles of each image should be mapped to points at infinity (with a third coordinate equals to 0). Which can be done with the x-axis works obviously with the y-axis, and so, concerning the bottom and top images, we can map the epipoles to the points at infinity in the same way. The question is to find a good projective transformation H to map the epipoles to points at infinity.

We use here the results of Zhang and his team [7], by posing three con-

straints on the trinocular rectification. The rectified images are denoted as: \bar{b} (bottom, the bar means rectified image), \bar{r} (right) and \bar{t} (top), respectively. The three constraints can be described as follows: images \bar{b} , \bar{r} and \bar{t} are rectified if:

1. All the epipolar lines of images \bar{b} and \bar{r} are horizontal and the corresponding points have the same y-coordinate.
2. All the epipolar lines of images \bar{b} and \bar{t} are vertical and the corresponding points have the same x-coordinate.
3. For any triplet of corresponding points (\bar{p}_b , \bar{p}_r and \bar{p}_t), the disparity u between \bar{p}_b, \bar{p}_r and \bar{p}_b, \bar{p}_t is equal.

We have so, for $\bar{p}_b = (x, y)$ in \bar{b} , the corresponding point $\bar{p}_r = (x + u, y)$ in \bar{r} , and the corresponding point $\bar{p}_t = (x, y + u)$ in \bar{t} .

For convenience, the canonical fundamental matrices of the rectified images are used to represent the rectification constraints. Considering the three rectified image pairs $\bar{b}\bar{r}$, $\bar{b}\bar{t}$ and $\bar{r}\bar{t}$, the fundamental matrices to obtain are:

$$\bar{F}_{br} \simeq \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \bar{F}_{bt} \simeq \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \bar{F}_{rt} \simeq \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix} \quad (1)$$

where \simeq means "equal up to a non-zero scale".

We need three rectification homographies H_b , H_r , H_t that make the fundamental matrices have the form (1) and can be parametrized by a set of independent free parameters over which it is possible to minimize the distortion. Following [7], we therefore have

$$H_r^T \bar{F}_{br} H_b = \lambda_1 F_{br}, \quad H_t^T \bar{F}_{bt} H_b = \lambda_2 F_{bt}, \quad H_t^T \bar{F}_{rt} H_r = \lambda_3 F_{rt}, \quad (2)$$

for some non-zero $\lambda_1, \lambda_2, \lambda_3$. Furthermore, the homographies have to map the epipoles to infinity. This leads to direct expression for the last row of each homography matrix. Thus, the equation set (2) is a linear under-constrained problem with 6 degrees of freedom. One can exploit the fact that the set of solutions is infinite in order to enforce additional geometric constraints, that minimizes the distortion generally caused by the rectification (for more details see [5]).

3 Energy minimization and anisotropic diffusion

3.1 Correspondence finding

Finding dense correspondences between images is a difficult problem. One has to find for each pixel in the first image, where the corresponding pixel is in the second (and third image). Therefore a measure of similarity has to be used in order to maximize the similarity between corresponding pixels. If we denote by $I_j(x, y)$ the intensity of the image j at pixel (x, y) , the optimization problem can be formulated as the minimization of $\sum_{i,j} (I_j(x', y') - I_i(x, y))^2$, where (x', y') is the corresponding pixel of (x, y) .

This minimization problem can be handled either a local approach [9] or a global approach [10]. It is well known that global methods produce a more consistent motion between nearby pixels, but are less efficient for discontinuities preservation. Therefore in order to improve results obtained by global methods at image discontinuities, Nagel and Enkelmann proposed to replace the isotropic regularization term found in [10], by an anisotropic operator [13]. More details are given in Section 3.3.

In the case, two rectified images are used, these ingredients are combined in [2, 3]. In the sequel, we show how the framework introduced in [2] can be generalized to three rectified images with some further improvement. Since our approach consists in minimizing a global functional on the image, we shall first recall some basic facts.

3.2 Energy minimization

Following [11], we summarize the modeling assumptions used in the sequel and define the matching problem in the context of the calculus of variations. We consider two images intensities $I_1^\sigma = I_1 * G_\sigma$ and $I_2^\sigma = I_2 * G_\sigma$ at a given scale σ , i.e. resulting from the convolution of two square-integrable functions $I_1 : \mathbb{R}^2 \rightarrow \mathbb{R}$ and $I_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ with a Gaussian kernel of standard deviation σ . Given a region of interest Ω , a bounded region of \mathbb{R}^2 , (we may require its boundary $\partial\Omega$ to fulfill some regularity constraints, e.g. that of being of class \mathcal{C}^2), we look for a function $u : \Omega \rightarrow \mathbb{R}^2$ assigning to each point x in Ω a displacement vector $u(x) \in \mathbb{R}^2$. This function is searched for in a set \mathbf{F} of admissible functions such that it minimizes an energy functional $\mathbf{E} : \mathbf{F} \rightarrow \mathbb{R}$ of the form :

$$\mathbf{E}(u) = \mathbf{J}(u) + \mathbf{R}(u) \quad (3)$$

where $\mathbf{J}(u)$ measures the "similarity" between I_1^σ and $I_2^\sigma \circ (Id + u)$ and $\mathbf{R}(u)$ is a measure of the "irregularity" of u (Id is the identity mapping of \mathbb{R}^2).

The similarity term will be defined in terms of global or local statistical measures on the intensities of I_1^g and $I_2^g o(Id + u)$, and the irregularity term will generally be a measure of the variations of u in Ω . In summary, the matching problem is defined as the solution of the following minimization problem:

$$u^* = \arg \min_{u \in \mathbf{F}} \mathbf{E}(u) = \arg \min_{u \in \mathbf{F}} (\mathbf{J}(u) + \mathbf{R}(u)) \quad (4)$$

Assuming that \mathbf{E} is sufficiently regular, its first variation at $u \in \mathbf{F}$ in the direction of $h \in \mathbf{F}$ is defined by :

$$\delta \mathbf{E}(u, h) = \lim_{\epsilon \rightarrow 0} \frac{\mathbf{E}(u + \epsilon h) - \mathbf{E}(u)}{\epsilon} = \left. \frac{d\mathbf{E}(u + \epsilon h)}{d\epsilon} \right|_{\epsilon=0}$$

If a minimizer u^* of \mathbf{E} exists, then $\delta \mathbf{E}(u^*, h) = 0$ must hold for every $h \in \mathbf{F}$. Assuming that \mathbf{F} is a linear subspace of a Hilbert space \mathbf{H} , endowed with a scalar product $(\cdot, \cdot)_{\mathbf{H}}$, we define the gradient $\nabla_{\mathbf{H}} \mathbf{E}(u)$ of \mathbf{E} by requiring that

$$\forall h \in \mathbf{F}, \quad \left. \frac{d\mathbf{E}(u + \epsilon h)}{d\epsilon} \right|_{\epsilon=0} = (\nabla_{\mathbf{H}} \mathbf{E}(u), h)_{\mathbf{H}}$$

The Euler equations are then equivalent to $\nabla_{\mathbf{H}} \mathbf{E}(u^*) = 0$. Rather than solving them directly, the search for a minimizer of \mathbf{E} is done using a "gradient descent" strategy. Given an initial estimate $u_0 \in \mathbf{H}$, we introduce time and a differentiable function, also noted u from the interval $[0, T]$ into \mathbf{H} (we say that $u \in \mathcal{C}_1([0, T]; \mathbf{H})$) and we solve the following initial value problem:

$$\begin{cases} \frac{du}{dt} = -\nabla_{\mathbf{H}} \mathbf{E}(u) = -(\nabla_{\mathbf{H}} \mathbf{J}(u) + \nabla_{\mathbf{H}} \mathbf{R}(u)) \\ u(0)(\cdot) = u_0(\cdot) \end{cases} \quad (5)$$

That is, we start from the initial field u_0 and follow the gradient of the functional \mathbf{E} (the minus sign is because we are minimizing). The solution of the matching problem is then taken as the asymptotic state (i.e. when $t \rightarrow \infty$) of $u(t)$.

Consequently, equation (5) may be viewed as a first-order ordinary differential equation with values in \mathbf{H} . It turns out that studying it from such an abstract viewpoint allows to prove the existence and uniqueness of several types of solutions (mild, strong, classical) of (5), by borrowing tools from functional analysis and the theory of semi-groups of linear operators. In the present Section, we study the generic minimization flow within this abstract framework. The linear operator $-\nabla_{\mathbf{H}} \mathbf{R}(u)$ defined by the regularization term will be simply noted A and the non-linear matching term

$-\nabla_{\mathbf{H}}\mathbf{J}$ will be generically noted F . The unknown of the problem is an \mathbf{H} valued function $u : [0, +\infty[\rightarrow \mathbf{H}$ defined on \mathbb{R}^+ .

We want here to establish the properties required by A and F in order for equation (5), which is now written as a semi-linear abstract initial value problem of the form :

$$\begin{cases} \frac{du}{dt} - Au(t) = F(u(t)), & t > 0 \\ u(0) = u_0 \in \mathbf{H} \end{cases} \quad (6)$$

to have a unique solution. For this purpose, we shall use the following notations. Let $D(A)$ be the domain of A and S_A be the corresponding semi-group (see [1]).

A continuous solution u of the integral equation

$$u(t) = S_A(t)u_0 + \int_0^t S_A(t-s)F(u(s))ds$$

is called a *mild solution* of the initial value problem (6). The last definition is motivated by the following argument. If (6) has a classical solution then the \mathbf{H} valued function $k(s) = S_A(t-s)u(s)$ is differentiable for $0 < s < t$ and, thanks to the theorem claiming that for all $u \in D(A)$, we have $S_A(t)u \in D(A)$ and $\frac{d}{dt}S_A(t)u = AS_A(t)u = S_A(t)Au$ (Theorem 1.2.4 in [12]):

$$\begin{aligned} \frac{dk}{ds} &= -AS_A(t-s)u(s) + S_A(t-s)u'(s) \\ &= -AS_A(t-s)u(s) + S_A(t-s)Au(s) + S_A(t-s)F(u(s)) = S_A(t-s)F(u(s)) \end{aligned}$$

If $F \circ u \in \mathcal{L}^1([0, T[, \mathbf{H})$ then $S_A(t-s)F(u(s))$ is integrable and integrating (3.2) from 0 to t yields

$$k(t) - k(0) = u(t) - S_A(t)u_0 = \int_0^t S_A(t-s)F(u(s))ds$$

hence

$$u(t) = S_A(t)u_0 + \int_0^t S_A(t-s)F(u(s))ds$$

The definition of the mild solution is thus natural. Sufficient conditions on A and F for (6) to have a unique mild solution are given by the following theorem.

Theorem 1. *Let $F : H \rightarrow H$ be uniformly Lipschitz continuous on H and let $-A$ be a maximal monotone operator. Then the initial value problem (6) has a unique mild solution $u \in \mathcal{C}([0, T[, \mathbf{H})$. Moreover, the mapping $u_0 \rightarrow u$ is Lipschitz continuous from H into $\mathcal{C}([0, T[, \mathbf{H})$.*

The proof can be found for example in Theorem 6.1.2 of [12]. Since H is a Hilbert space, taking an initial value $u_0 \in D(A)$ suffices to obtain existence and uniqueness of a strong solution.

3.3 Anisotropic Diffusion

The diffusion allows to homogenize a picture. In image processing, the diffusion eliminates the local perturbations of the signal. This is why we use in previous section a convolution with a Gaussian. The inconvenient is that the iterations make the contours more and more blurred. The diffusion has to be uniform far from the contours, and perpendicular to the gradient on the contours, which is meant by "anisotropic". The key idea is so to forbid regularizing and smoothing across the discontinuities. We consider the term $\mathbf{R}(u)$ from equation (3).

The regularization operator we will use was introduced by Nagel and Enkelmann [13], and is expressed as :

$$\mathbf{R}(u) = \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \right) \cdot D(\nabla I_1) \cdot \begin{pmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \end{pmatrix} \quad (7)$$

with I_1 the function describing the first image intensity, $D(\nabla I_1)$ a regularized projection matrix perpendicular to ∇I_1 :

$$D(\nabla I_1) = \frac{1}{|\nabla I_1|^2 + 2\nu^2} \left[\begin{pmatrix} \frac{\partial I_1}{\partial y} \\ -\frac{\partial I_1}{\partial x} \end{pmatrix} \begin{pmatrix} \frac{\partial I_1}{\partial y} \\ -\frac{\partial I_1}{\partial x} \end{pmatrix}^t + \nu^2 Id \right] \quad (8)$$

Therefore we have the following limit:

$$\lim_{\|\nabla I_1\| \rightarrow 0} D(\nabla I_1) = \frac{1}{2} Id$$

On the other hand, when $\|\nabla I_1\| \rightarrow \infty$, then $D(\nabla I_1)$ become the projection operator over the direction orthogonal to the gradient.

The resulting effect is that the variation of the displacement is minimized in the direction orthogonal to the gradient of the first image at each pixel. This allows a smoothing effect along the objects boundaries, while preserving these boundaries.

3.4 Euler-Lagrange equation and parabolic equation

We are now in a position to introduce a first instance of the functional that we shall minimize as in the general problem given by (4). An improved model will be introduced in Section 3.5. We consider three rectified images. The rectification assumption is equivalent to assume that the three cameras are organized in a perfect L shape with same distances within the horizontal and vertical pairs. In order to make notation even more concrete and following notations used in Section 2, we shall denote the *rectified* image intensities functions by \mathbf{b} , \mathbf{r} and \mathbf{t} , which stands for *bottom*, *right* and *top*, the three camera locations in our setting. Since rectification is used, the pixel motion field is described by a single parameter u which a real-valued function of two parameters, as explained in Section 2. Hence the Euler-Lagrange equation will also contain only one unknown. Finally, the energy term to minimize is:

$$\begin{aligned}
E(u) = \iint_{\bar{b}} L(u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}) dx dy = & \iint_{\bar{b}} \left(\|\mathbf{r}(x + u(x, y), y) - \mathbf{b}(x, y)\|^2 \right. \\
& + \|\mathbf{t}(x, y + u(x, y)) - \mathbf{b}(x, y)\|^2 \\
& + \|\mathbf{r}(x + u(x, y), y) - \mathbf{t}(x, y + u(x, y))\|^2 \left. \right) dx dy \\
& + C \iint_{\bar{b}} R(u) dx dy,
\end{aligned} \tag{9}$$

where C is a constant that and the associated Euler-Lagrange equation is $\frac{\partial L}{\partial u} - \frac{\partial}{\partial x}(\frac{\partial L}{\partial u_x}) - \frac{\partial}{\partial y}(\frac{\partial L}{\partial u_y}) = 0$.

The part term $-\frac{\partial}{\partial x}(\frac{\partial L}{\partial u_x}) - \frac{\partial}{\partial y}(\frac{\partial L}{\partial u_y})$ is related to the regularization operator. After computation, it is given by $C \cdot \text{div}(D(\nabla \mathbf{b}) \nabla u) = 0$. Thus, the following equation has to be solved :

$$\begin{aligned}
& 2 \frac{\partial \mathbf{r}}{\partial x}(x + u, y) [2\mathbf{r}(x + u, y) - \mathbf{t}(x, y + u) - \mathbf{b}(x, y)] \\
& + 2 \frac{\partial \mathbf{t}}{\partial y}(x, y + u) [2\mathbf{t}(x, y + u) - \mathbf{r}(x + u, y) - \mathbf{b}(x, y)] \\
& - C \text{div}(D(\nabla \mathbf{b}) \nabla u) = 0
\end{aligned} \tag{10}$$

We have now to compute the asymptotic state when $t \rightarrow \infty$ of the

corresponding parabolic equation:

$$\begin{aligned} \frac{\partial u}{\partial t} = & -2\frac{\partial \mathbf{r}}{\partial x}(x+u, y) [2\mathbf{r}(x+u, y) - \mathbf{t}(x, y+u) - \mathbf{b}(x, y)] \\ & -2\frac{\partial \mathbf{t}}{\partial y}(x, y+u) [2\mathbf{t}(x, y+u) - \mathbf{r}(x+u, y) - \mathbf{b}(x, y)] \\ & + C \operatorname{div}(D(\nabla \mathbf{b}) \nabla u) \end{aligned} \quad (11)$$

We will prove here the convergence to this asymptotic state and its uniqueness, by using the theorem of Section 3.2. We shall look for a solution u in $H = L^2(\mathbb{R}^2)$. Let $F : H \rightarrow H$ be defined as follows:

$$\begin{aligned} F(u) = & -2\frac{\partial \mathbf{r}^\sigma}{\partial x}(x+u, y) [2\mathbf{r}^\sigma(x+u, y) - \mathbf{t}^\sigma(x, y+u) - \mathbf{b}^\sigma(x, y)] \\ & -2\frac{\partial \mathbf{t}^\sigma}{\partial y}(x, y+u) [2\mathbf{t}^\sigma(x, y+u) - \mathbf{r}^\sigma(x+u, y) - \mathbf{b}^\sigma(x, y)] \end{aligned}$$

Moreover let $A : \mathcal{D}(A) \subset H \rightarrow H$ be the differential operator defined as follows:

$$A(u) = -C \operatorname{div}(D(\nabla \mathbf{b}^\sigma) \nabla u)$$

Then the system obtained by blurring (11) can be written in a compact form, together with the initial guess, and has actually the form of equation (6):

$$\begin{cases} \frac{\partial u}{\partial t} + A(u) = F(u) \\ u(0) = u_0 \end{cases}$$

The image functions are defined in a bounded domain. For simplicity, we shall consider them as defined on the whole plane. More, we shall consider that $\mathbf{b}, \mathbf{r}, \mathbf{t} \in L^2(\mathbb{R}^2)$, which is physically amply justified.

Theorem 2. *Under the assumption that $\mathbf{b}, \mathbf{r}, \mathbf{t} \in L^2(\mathbb{R}^2)$, the function F is Lipschitz-continuous, and the Lipschitz constant L depends on the functions $\mathbf{b}, \mathbf{r}, \mathbf{t}$ and σ .*

Proof. Since $\mathbf{b}, \mathbf{r}, \mathbf{t} \in L^2(\mathbb{R}^2)$, we have $\mathbf{r}^\sigma, \mathbf{t}^\sigma \in W^{1,2}(\mathbb{R}^2)$ and $\mathbf{b}^\sigma, \mathbf{t}^\sigma \in L^\infty(\mathbb{R}^2)$, where $W^{1,2}(\mathbb{R}^2)$ denotes the Sobolev space:

$$W^{1,2}(\mathbb{R}^2) = \left\{ u \in L^2(\mathbb{R}^2) \left| \begin{array}{l} \exists g_1, g_2 \in L^2(\mathbb{R}^2), \forall \phi \in C_c^\infty(\mathbb{R}^2), \\ \int_{\mathbb{R}^2} \frac{\partial u}{\partial x} \phi' = - \int_{\mathbb{R}^2} g_1 \phi, \\ \int_{\mathbb{R}^2} \frac{\partial u}{\partial y} \phi' = - \int_{\mathbb{R}^2} g_2 \phi \end{array} \right. \right\},$$

where $C_c^\infty(\mathbb{R}^2)$ denotes the space of infinitely derivable points at infinity functions that are zero outside a compact domain.

Thus it clear that $\mathbf{r}^\sigma, \mathbf{t}^\sigma \in W^{1,2}(\mathbb{R}^2)$, because after the smoothing with a Gaussian filter, we obtain infinity derivable functions. Smoothing the images, which are basically bounded, also results in bounded functions. Therefore $\mathbf{b}^\sigma, \mathbf{t}^\sigma \in L^\infty(\mathbb{R}^2)$.

In the sequel, we shall denote $F_1(u) = -2\frac{\partial \mathbf{r}^\sigma}{\partial x}(x+u, y)[2\mathbf{r}^\sigma(x+u, y) - \mathbf{t}^\sigma(x, y+u) - \mathbf{b}^\sigma(x, y)]$ and $F_2(u) = -2\frac{\partial \mathbf{t}^\sigma}{\partial x}(x, y+u)[2\mathbf{t}^\sigma(x, y+u) - \mathbf{r}^\sigma(x+u, y) - \mathbf{b}^\sigma(x, y)]$, so that $F(u) = F_1(u) + F_2(u)$. Now, consider $u_1, u_2 \in H$. We have the following estimate (the norm $\| \cdot \|$ is $\| \cdot \|_{L^2(\mathbb{R}^2)}$) $\| F(u_1) - F(u_2) \| \leq \| F_1(u_1) - F_1(u_2) \| + \| F_2(u_1) - F_2(u_2) \|$. Therefore let us first exhibit an upper bound of $\| F_1(u_1) - F_1(u_2) \|$.

$$\begin{aligned}
& \| F_1(u_1) - F_2(u_2) \| = \\
& \left\| 2\frac{\partial \mathbf{r}^\sigma}{\partial x}(x+u_1, y)[2\mathbf{r}^\sigma(x+u_1, y) - \mathbf{t}^\sigma(x, y+u_1) - \mathbf{b}^\sigma(x, y)] - \right. \\
& \left. 2\frac{\partial \mathbf{r}^\sigma}{\partial x}(x+u_2, y)[2\mathbf{r}^\sigma(x+u_2, y) - \mathbf{t}^\sigma(x, y+u_2) - \mathbf{b}^\sigma(x, y)] \right\| \\
& \leq \left\| 4\frac{\partial \mathbf{r}^\sigma}{\partial x}(x+u_1, y)\mathbf{r}^\sigma(x+u_1, y) - 4\frac{\partial \mathbf{r}^\sigma}{\partial x}(x+u_2, y)\mathbf{r}^\sigma(x+u_2, y) \right\| \\
& + \left\| \mathbf{b}^\sigma(x, y)[2\frac{\partial \mathbf{r}^\sigma}{\partial x}(x+u_1, y) - 2\frac{\partial \mathbf{r}^\sigma}{\partial x}(x+u_2, y)] \right\| \\
& + \left\| 2\frac{\partial \mathbf{r}^\sigma}{\partial x}(x+u_1, y)\mathbf{t}^\sigma(x, y+u_1) - 2\frac{\partial \mathbf{r}^\sigma}{\partial x}(x+u_2, y)\mathbf{t}^\sigma(x, y+u_2) \right\| \\
& \leq 2 \left\| \frac{\partial(\mathbf{r}^\sigma)^2}{\partial x}(x+u_1, y) - \frac{\partial(\mathbf{r}^\sigma)^2}{\partial x}(x+u_2, y) \right\| \\
& + 2 \left\| \mathbf{b}^\sigma \right\|_\infty \left\| \frac{\partial \mathbf{r}^\sigma}{\partial x}(x+u_1, y) - \frac{\partial \mathbf{r}^\sigma}{\partial x}(x+u_2, y) \right\| \\
& + 2 \left\| \mathbf{t}^\sigma \right\|_\infty \left\| \frac{\partial \mathbf{r}^\sigma}{\partial x}(x+u_1, y) - \frac{\partial \mathbf{r}^\sigma}{\partial x}(x+u_2, y) \right\| \\
& \leq 2C_{lip}\left(\frac{\partial(\mathbf{r}^\sigma)^2}{\partial x}\right) \| u_1 - u_2 \| + 2 \left\| \mathbf{b}^\sigma \right\|_\infty C_{lip}\left(\frac{\partial \mathbf{r}^\sigma}{\partial x}\right) \| u_1 - u_2 \| \\
& + 2 \left\| \mathbf{t}^\sigma \right\|_\infty C_{lip}\left(\frac{\partial \mathbf{r}^\sigma}{\partial x}\right) \| u_1 - u_2 \| \\
& \leq 2(C_{lip}\left(\frac{\partial(\mathbf{r}^\sigma)^2}{\partial x}\right) + \left\| \mathbf{b}^\sigma \right\|_\infty C_{lip}\left(\frac{\partial \mathbf{r}^\sigma}{\partial x}\right) + \left\| \mathbf{t}^\sigma \right\|_\infty C_{lip}\left(\frac{\partial \mathbf{r}^\sigma}{\partial x}\right)) \| u_1 - u_2 \|,
\end{aligned}$$

where $C_{lip}(f)$ is the Lipschitz constant of the function f . It is clear that $C_{lip}\left(\frac{\partial(\mathbf{r}^\sigma)^2}{\partial x}\right)$ and $C_{lip}\left(\frac{\partial \mathbf{r}^\sigma}{\partial x}\right)$ exists since $\frac{\partial(\mathbf{r}^\sigma)^2}{\partial x}, \frac{\partial \mathbf{r}^\sigma}{\partial x} \in W^{1,2}(\mathbb{R}^2)$.

Eventually, a Lipschitz constant of F_1 is given by :

$$C_{lip}(F) = 2(C_{lip}\left(\frac{\partial(\mathbf{r}^\sigma)^2}{\partial x}\right) + \left\| \mathbf{b}^\sigma \right\|_\infty C_{lip}\left(\frac{\partial \mathbf{r}^\sigma}{\partial x}\right) + \left\| \mathbf{t}^\sigma \right\|_\infty C_{lip}\left(\frac{\partial \mathbf{r}^\sigma}{\partial x}\right)).$$

Similarly, we could find a Lipschitz constant of F_2 : points at infinity

$$C_{lip}(F) = 2(C_{lip}(\frac{\partial(\mathbf{t}^\sigma)^2}{\partial x}) + \|\mathbf{b}^\sigma\|_\infty C_{lip}(\frac{\partial\mathbf{t}^\sigma}{\partial x}) + \|\mathbf{r}^\sigma\|_\infty C_{lip}(\frac{\partial\mathbf{t}^\sigma}{\partial x})).$$

A Lipschitz constant for F is simply the sum of the two previous constants. \square

Since $\mathbf{b} \in L^2(\mathbb{R}^2)$, then $\mathbf{b}^\sigma \in W^{1,\infty}(\mathbb{R}^2)$. Thus $\nabla\mathbf{b}^\sigma$ is bounded and the eigenvalues of $D(\nabla\mathbf{b}^\sigma)$ are strictly positive. Therefore since $C > 0$, the operator A is a maximal monotone operator. Then using, theorem 1, we can conclude by the following result.

Theorem 3. *The system (6) has a unique mild solution for any initial guess.*

3.5 An improved model

The energy defined by expression (9) uses a mere difference between the intensities of corresponding pixels. In order to make the correspondence finding more robust, the use of a local region around each pixel is necessary. In [14], this is handled by requiring not only the intensities constancy but also the gradient constancy. The drawback of this approach is that it requires the computation of second order image derivatives in the Euler-Lagrange equation. Here we adopt another approach, the data term

$$\iint_{\bar{b}} \left(\|\mathbf{r}(x + u(x, y), y) - \mathbf{b}(x, y)\|^2 + \|\mathbf{t}(x, y + u(x, y)) - \mathbf{b}(x, y)\|^2 + \|\mathbf{r}(x + u(x, y), y) - \mathbf{t}(x, y + u(x, y))\|^2 \right) dx dy$$

is replaced by

$$\iint_{\bar{b}} \sum_{i=-n}^{i=n} \sum_{j=-n}^{j=n} \left(\|\mathbf{r}(x + u(x, y) + i, y + j) - \mathbf{b}(x + i, y + j)\|^2 + \|\mathbf{t}(x + i, y + u(x, y) + j) - \mathbf{b}(x + i, y + j)\|^2 + \|\mathbf{r}(x + u(x, y) + i, y + j) - \mathbf{t}(x + i, y + u(x, y) + j)\|^2 \right) dx dy,$$

where $2n + 1$ is the size of the square neighborhood of each pixel. We found that using $2n + 1 = 3$ yields optimal results. It is straight forward to see that the proof in theorem 2 also holds for this improved model. Therefore theorem 3 is also valid.

3.6 Numeric scheme

For the numerical solution for the system (6), we shall use a finite difference scheme. In this section, we shall denote by (i, j) the grid vertexes on which we compute the flow. The derivative with respect to time is given by:

$$\frac{u_{ij}^{k+1} - u_{ij}^k}{\tau},$$

where τ is the time step. Following [2], we use the following scheme for the regularization term:

$$\begin{aligned} \operatorname{div}(D(\nabla \mathbf{b}^\sigma) \nabla u) \approx & \frac{a_{i+1,j} + a_{i,j}}{2} \frac{u_{i+1,j}^{k+1} - u_{i,j}^{k+1}}{h_1^2} + \frac{a_{i-1,j} + a_{i,j}}{2} \frac{u_{i-1,j}^{k+1} - u_{i,j}^{k+1}}{h_1^2} \\ & + \frac{c_{i,j+1} + c_{i,j}}{2} \frac{u_{i,j+1}^{k+1} - u_{i,j}^{k+1}}{h_2^2} + \frac{c_{i,j-1} + c_{i,j}}{2} \frac{u_{i,j-1}^{k+1} - u_{i,j}^{k+1}}{h_2^2} \\ & + \frac{b_{i+1,j+1} + b_{i,j}}{2} \frac{u_{i+1,j+1}^{k+1} - u_{i,j}^{k+1}}{2h_1 h_2} + \frac{b_{i-1,j-1} + b_{i,j}}{2} \frac{u_{i-1,j-1}^{k+1} - u_{i,j}^{k+1}}{2h_1 h_2} \\ & - \frac{b_{i+1,j-1} + b_{i,j}}{2} \frac{u_{i+1,j-1}^{k+1} - u_{i,j}^{k+1}}{2h_1 h_2} - \frac{b_{i-1,j+1} + b_{i,j}}{2} \frac{u_{i-1,j+1}^{k+1} - u_{i,j}^{k+1}}{2h_1 h_2}, \end{aligned}$$

where $D(\nabla \mathbf{b}^\sigma) = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$.

Each component of the data similarity term in the first component term is approximated as follows:

$$\begin{aligned} & -2 \frac{\partial \mathbf{r}^\sigma}{\partial x}(x+u, y) [2\mathbf{r}^\sigma(x+u, y) - t^\sigma(x, y+u) - \mathbf{b}^\sigma(x, y)] \approx \\ & -2 \frac{\partial \mathbf{r}^\sigma}{\partial x}(i+u_{ij}^k, j) [2\mathbf{r}^\sigma(i+u_{ij}^k, j) + 2(u_{ij}^{k+1} - u_{ij}^k) \frac{\partial \mathbf{r}^\sigma}{\partial x}(i+u_{ij}^k, j) \\ & - \mathbf{t}(i, j+u_{ij}^k) - (u_{ij}^{k+1} - u_{ij}^k) \frac{\partial \mathbf{t}}{\partial y}(i, j+u_{ij}^k) - \mathbf{b}^\sigma(i, j)] \end{aligned}$$

The values for the images and their derivatives outside of the grid points are computed by interpolation. Eventually, we get an implicit scheme that leads to solve a large sparse linear system: $Ax = b$. For solving this system, since the matrix A is typically very large and sparse, an iterative method is used. We use the Gauss-Seidel iteration. If $A = L + D + U$, where L (respectively U) is a strictly lower (respectively upper) triangular matrix, while D is a diagonal matrix, then the iteration is performed as follows:

$$x^{n+1} = -(D - L)^{-1} U x^n + (D - L)^{-1} b$$

Moreover in order to speed up the computations and to handle large displacement, we embed the scheme into a multiresolution approach [10].

4 Experiments

We present here the results obtained with the improved model over three rectified images. Fig. 1 presents the three rectified images.

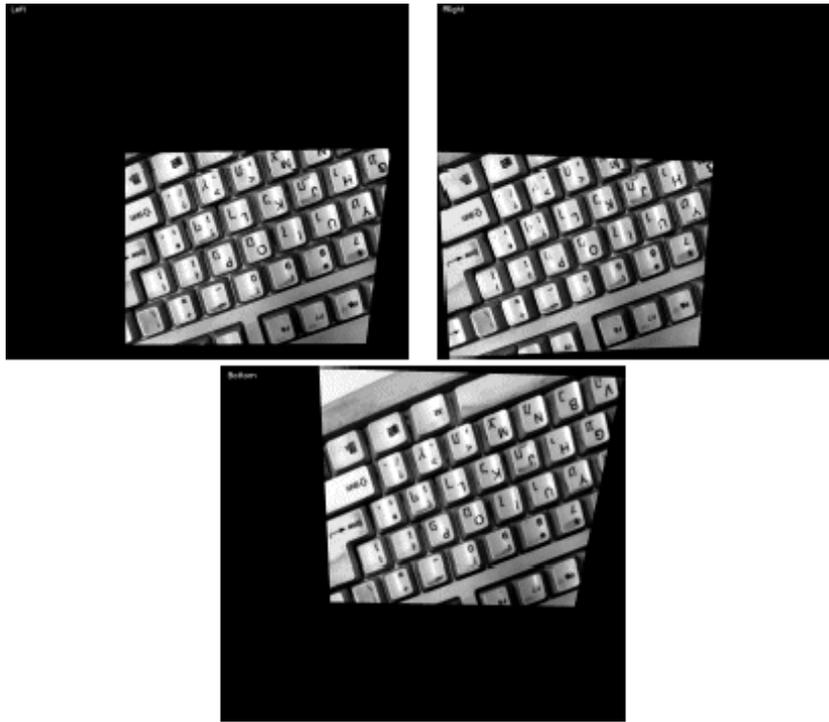


Figure 1: Three rectified images.

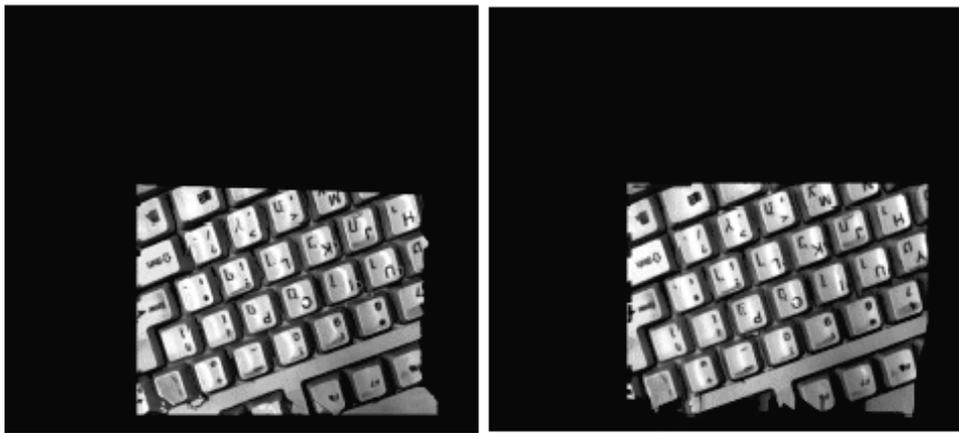
In Fig. 2 we show the results. Fig. 2a, the disparity is rendered as a gray-level images, while Figs. 2b and 2c shows the respectively the second and the third image after backward warping toward the first image, using the computed disparity.

The quality of these results is more precisely handled by some statistical data. In the following measures, the images are normalized so that the gray-levels values are between 0 and 1. Thus when we come to compare the original image 1 and the re-sampling of images 2 and 3 based the computed disparity, the two difference images have values between -1 and 1 at most. For the difference computed with the first and the second image, we found the following values: the minimum is -0.992157, the maximum

0.996078, while the mean value and the standard deviation are respectively $\mu_{12} = 0.0119701$ and $\sigma_{12} = 0.121034$. For the difference computed using the first and third images, we found similar values: the minimum and maximum are strictly identical, while the mean value and the standard deviation are respectively $\mu_{13} = 0.0138003$ and $\sigma_{13} = 0.102947$. In both cases, the statistical measures show that the computed disparity yields a high quality resampling.



(a)



(b)

(c)

Figure 2: The results computed with the improved model. (a) The rendered disparity; (b) the warping of the second image; (c) the warping of the third image.

5 Future work and discussion

We have present a complete and mathematical consistent way to handle the problem of correspondence finding between three images. Future work will incorporate more efficient numeric techniques like multigrid [15]. We shall also investigate the way to deal with more than three images at once.

References

- [1] H. Brezis, *Analyse Fonctionnelle*, Dunod, 1999.
- [2] L. Alvarez, R. Deriche, J. Sánchez, and J. Weickert, *Dense Disparity Map Estimation Respecting Image Discontinuities*, Technical Report INRIA 3874 (Sophia-Antipolis, France, 2000).
- [3] N. Slesareva, A. Bruhn, and J. Weickert, *Optic flow goes stereo: A variational method for estimating discontinuity-preserving dense disparity maps*, In: W. Kropatsch, R. Sablatnig, A. Hanbury (Eds.): *Pattern Recognition*. Lecture Notes in Computer Science, Vol. **3663**, 33-40 (Springer, Berlin, 2005).
- [4] O. Faugeras, *What can be seen in three dimensions with an uncalibrated stereo rig*, Computer Vision ECCV92, Lectures Notes in Computer Science, vol. **588** (Springer-Verlag, 1992).
- [5] M. Heinrichs and V. Rodehorst, *Trinocular Rectification for Various Camera Setups*, In Photogrammetric Computer Vision, 2006.
- [6] A. Shashua, *Trilinear Tensor: The Fundamental Construct of Multiple View Geometry and its Applications* (Hebrew University of Jerusalem, 1997).
- [7] Huaifeng Zhang, Jan Cech, Radim Sara, Fuchao Wu, Zhanyi Hu, *A Linear Trinocular Rectification Method for Accurate Stereoscopic Matching* (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Science, 2003).
- [8] R.I. Hartley, *Theory and practice of projective rectification*, International Journal of Computer Vision, **35**, 115-127 (1999).
- [9] B. Lucas and T. Kanade, *An iterative image registration technique with an application to stereo vision*, Proc. of Imaging Understanding Workshop, Washington, DC, pp. 121–130, 1981.

- [10] B.K.P. Horn and B.G. Schunk, *Determining optical flow*, Artificial Intelligence, **17**, 85–203 (1981).
- [11] G.H. Valadez, *Variational Methods for Multimodal Image Matching*, Doctoral Thesis, University of Nice - Sophia Antipolis, 2002.
- [12] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations* (Springer-Verlag, 1983).
- [13] H.-H. Nagel and W. Enkelmann *An investigation of smoothness constraints for the estimation of displacement vector fields from images sequences*, IEEE Trans. Pattern H Anal. Mach. Intell. **8**, 565-593 (1986).
- [14] A. Bruhn, J. Weickert, T. Kohlberger, and C. Schnorr, *A multigrid platform for real-time motion computation with discontinuity-preserving variational methods*, International Journal of Computer Vision **70**, 257-277 (2006).
- [15] W. Briggs, V. Henson, and S. McCormick, *A Multigrid Tutorial* (2-nd ed., SIAM, Philadelphia, 2000).