

Image and Vision Computing 21 (2003) 383-399



A fast incremental approach for accurate measurement of the displacement field

Azeddine Beghdadi^{a,*}, Mostefa Mesbah^b, Jerôme Monteil^a

^aL2TI, Laboratoire de Traitement et Transport de l'Information, Institut Galilée-Université Paris Nord, 99 Avenue, J.B. Clément, 93 400 Villetaneuse, France ^bSignal Processing Research Centre, Queensland University of technology, GPO Box 2434, Brisbane QLD 4001, Australia

Received 23 November 2001; received in revised form 21 January 2003; accepted 23 January 2003

Abstract

A fast method for accurate measurement of displacement field that combines template matching and differential techniques is proposed. This method, which is able to deal with the problems of non-rigid object movement as well as large inter frame displacements, operates in three steps. First, a sparse displacement field is obtained by a classic template matching technique. Then, this information is propagated through the overall image to obtain a dense displacement field. Finally, this field is considered as approximate solution before the use of classical differential technique for optical flow estimation. To illustrate the efficiency of the proposed procedure, a number of experimental results using both synthetic and real images are presented and discussed. The results show significant improvement over those obtained using standard multigrid approaches, especially in the cases of textured, very structured images or large inter frame displacements. © 2003 Elsevier Science B.V. All rights reserved.

Keywords: Optical flow; Motion estimation; Template matching; Cross-correlation; Multiresolution analysis

1. Introduction

The Optical Flow (OF) is defined as the apparent movement or velocity distribution¹ of the 'brightness patterns' in the image plane. Its estimation from image sequences has been an active field of research since being introduced by Horn and Schunck in the early 80's [22]. OF techniques have been widely used in areas such as meteorology [50], biomedical imaging [8,15,26,33,41], material sciences [5,29,32] and other industrial problems. OF estimation methods can be grouped into four typical categories: differential methods, feature-based approaches, spatio-temporal (energy-based) approaches, and model-based methods [4,20].

The present contribution is a differential-based method that uses a special correlation strategy. The differential approach has been chosen for its simplicity and ease of computation. It is well known, however, that both differential and correlation methods suffer from some drawbacks. For instance, correlation methods are computationally very involved. The amount of computation time becomes prohibitive when accuracy in the estimation of a dense flow field is required. Whereas, differential methods become unreliable when large displacements are concerned. To overcome this problem, some multiresolution techniques with coarse-to-fine strategy were proposed [11,14]. In these techniques, a multiresolution decomposition, wherein a sequence of approximations at successively coarser resolutions is performed, is adopted. This decomposition is carried out by using low pass filters followed by subsampling operation. When these operations are iterated, a hierarchical structure is generated leading to what is called 'multiresolution pyramid'. The idea behind this decomposition is to perform a coarse estimation of the flow field in a reduced copy of the original image and conveys this estimated flow as an initial approximation of the field to be computed at the next finer resolution. This approach assumes that the image at coarser resolution of the pyramid still contains salient features of the signal. This requirement is, however, hard to fulfill in the case of structured images such as texture where the fine structures are lost at coarser resolutions of the pyramid. In this case, the flow field estimation error could propagate towards the finer levels of

^{*} Corresponding author. Tel.: +33-1-49-40-40-57; fax: +33-1-49-40-40-61.

E-mail addresses: beghdadi@galilee.univ-paris13.fr (A. Beghdadi), m. mesbah@qut.edu.au (M. Mesbah).

¹ Throughout the text we use 'velocity flow field' and 'displacement flow field' interchangeably since these two flows are equivalent with respect to the temporal sampling interval Δt .

the pyramid leading to a poor estimate of the flow field. In contrast to the differential approaches, features-based methods allow, under some conditions, to estimate large displacements with a good precision [29], but at the expense of a much longer computational time.

The incremental approach proposed in this paper combines a differential technique with a special correlation method for flow field estimation. It is shown that this strategy allows a fast and accurate (about 1/10 of a pixel) estimation of the dense flow field.

The paper is organized into five sections. We first review some existing OF estimation techniques and discuss their advantages and drawbacks (Section 2). We then expose the basic idea behind our method and give some details on its computer implementation (Section 3.1 and 3.2). In Section 4, the results of applying the proposed method on real life image sequences are presented and compared to the standard multigrid approach. Section 5 summarizes the contribution of this work and proposes possible future directions and improvements.

It should be noted that it is not the intention of the authors to compare the performance of the proposed method to the existing OF estimation techniques. A full and objective comparison of the large number of methods published in this field during the last two decades is a rather time consuming task. In this paper, we focus in the present work on answering the following questions:

Is it possible to use the computationally most attractive differential method of Horn and Schunk in the case of large displacement without introducing additional constraints as done in other known approaches (for a review see Mitiche and Bouthemy [31]). In the following, we show that this challenging question can be addressed and solved in a simple way. An approach allowing the computation of the dense flow field without additional constraint is proposed. The main attractive characteristics of this method are its simplicity, accuracy, and ease of implementation.

2. Existing methods

In the following we briefly summarize the most widely cited OF estimation approaches. Drawbacks as well as advantages of each approach are highlighted. OF estimation approaches can be broadly classified into four categories, namely, differential methods, correspondence (correlationbased) methods, spatio-temporal (energy-based) methods and the parametric (model-based) methods.

2.1. Differential methods

Differential methods [22,35,36,39,40,42,45,46] are based on the assumption that the observed brightness (intensity) *E* of any object point is constant over time. Consequently, any change in intensity at a point in the image is due to motion. Relative motion between the camera and the object will cause position (x, y) of a point M in the image at time t to change to a new position $(x + \delta x, y + \delta y)$ over the interval time δt . By the constant brightness assumption, the intensity of the object point will be the same in the two images. This assumption is mathematically stated as:

$$E(x, y, t) = E(x + \delta x, y + \delta y, t + \delta t)$$
(1)

By expanding the right side of the above equation as a Taylor's series about $(\mathbf{x}, t)^T$ and discarding terms higher than the first order, we can obtain the standard OF equation [22]:

$$E_t + \nabla E \mathbf{v} = 0 \tag{2}$$

where **v** is the velocity vector, E_t is the partial derivative of E with respect to t, and ∇ is the gradient operator. This equation is called 'Brightness Invariance Equation'. In practice, to derive such equation two fundamental assumptions are used, namely, Intensity conservation during the movement and small inter frame displacements. The validity of the latter requires fine temporal sampling of the data flow. Furthermore, the use of differential operators assumes small displacement, in general lower than the pixel. When the latter assumption cannot be guaranteed, other approaches such as multiresolution decomposition can be used [10,11,27,28]. Accurate measurement could then be obtained with such methods even for large displacements or for large and irregular intensity gradient deviations in the image plane [37]. The problem of these methods, as we mentioned before, is that they assume that the brightness constraint is still valid at coarser resolutions of the image decomposition [31].

It is well known that problem of OF estimation using Eq. (2) is ill-posed and consequently could not be solved without additional constraints to guarantee the uniqueness of the solution [6]. Depending on the choice of the additional constraint, various methods were derived [17,22,24,30,34,46,49]. In Section 3.1, the smoothness constraint used for regularization in Horn and Schunk method [22] is recalled and analyzed. This method serves as a basis of our new technique called Translated Optical Flow (TOF).

2.2. Correlation-based (matching or correspondence) methods

In contrast to differential approaches, correlation methods [2,25,29,44], can deal with large motions. The basic idea behind these methods consists in establishing an inter frame correspondence. This could be achieved by tracking a pattern or a set of local features in two successive frames. These techniques are known as block or pattern matching methods [18]. The easiest and the standard way of choosing such patterns is to consider only a set points of interest where intensity abruptly changes. Such features are for example corners, end line points or other salient features [43,51] which often guaranty the uniqueness of the pattern matching.

The correspondence procedure consists of searching the displacement vector of a given pattern from its initial position in the first frame into the second one. In general, this could be achieved using an exhaustive search for the displacement vector **s** that corresponds to the maximum of a local similarity measure such as the inter-block correlation coefficient. In practice, an analysis block is chosen on which a smaller sliding matching block is defined. Other distance measures such as Sum of Square Difference (SSM) defined by:

$$SSD(\mathbf{x}, \mathbf{s}) = \sum_{j=-n}^{n} \sum_{i=-n}^{n} W(i, j) [E_1(\mathbf{x} + (i, j)) - E_2(\mathbf{x} + \mathbf{s} + (i, j))]^2$$
(3)

can be used where W is a weighting function corresponding to the analysis window of size $(2n + 1) \times (2n + 1)$, and E_1 and E_2 are the intensity distributions in two successive frames.

The main difficulty encountered when using such methods is the validity of the uniqueness of the solution. Indeed, the 2D correlation function may not be unimodal and may exhibit a flat shape. In such cases, it becomes very difficult to localize the optimum solution. Moreover, even when the correlation is unimodal, its peak could be wide leading to the well-known *aperture problem* [2,50]. For these reasons, the methods based on correlation or matching give only an approximate solution of the displacement field. Furthermore, they often fail in estimating the displacement field in homogeneous region.

2.3. Energy-based methods

These methods are based on the multidecomposition of the image sequence using a bank of spatio-temporal filters, such as Gabor-like filters, tuned to a family of frequency bands [1,13,21,47]. Such techniques analyze the image sequence in the multi-dimensional space spanned by two spatial and one temporal frequencies (\mathbf{k}, ω). For this reason they are sometimes called frequency-based techniques.

The main idea behind the energy-based methods is the fact that a 2D pattern translating with a constant velocity **v** in the image plane has all of its power lies on a plane containing the origin in the spatio-temporal frequency space (**k**, ω). The plane equation is given by:

$$\omega = -\mathbf{v}'\mathbf{k}$$

Based on this fact, the image velocity can be computed in the Fourier domain by finding the plane in which most of the power is concentrated. This can be achieved by using filters that respond to oriented spatio-temporal energy.

Velocity estimation using these techniques is a two-stage process. In the first stage, also called image representation, sets of spatio-temporal frequency selective filters are used to sample the power spectrum of the input image. In the second stage, the filter outputs are resolved to measure the velocity of the image motion at each of the number of spatial locations and spatial frequencies.

The main drawback of these approaches is that they are computationally demanding and require a longer image sequence than other techniques, for example seven frames are used in Ref. [50]. Furthermore, the efficiency of these techniques is limited, as in differential methods, to small displacements.

2.4. Model-based methods

The basic idea of these methods is to propose a plausible model for motion and to estimate the model parameters fitting the model to the experimental data in order to compute the flow field. Common models of image flow in a local neighbourhood Ω include constant, affine, and quadratic [3,7,12,38,48]. For the case of affine model we have:

$$\mathbf{u}(\mathbf{x}, \mathbf{a}) = \begin{pmatrix} a_1 + a_3 x + a_5 y \\ a_2 + a_4 x + a_6 y \end{pmatrix}$$

while for the case of the case of quadratic model we have:

$$\mathbf{u}(\mathbf{x}, \mathbf{a}) = \begin{pmatrix} a_1 + a_3 x + a_5 y + a_7 x^2 + a_9 x y + a_{11} y^2 \\ a_2 + a_4 x + a_6 y + a_8 x^2 + a_{10} x y + a_{12} y^2 \end{pmatrix}$$

The main drawback of this approach is the fact that the accuracy is model-dependent. Furthermore, these approaches assume apriori knowledge of the nature of the motion. With more complicated motions model allowed, the problem of computational cost arises. So instead of solving for 2 parameters per pixel for the case of constant velocity, we will be required to solve for 6 in the affine model and 12 in the quadratic model. Also since these methods are local, they do not allow for propagation of the optical flow to uniform bright image regions from nearby points which results in a sparse flow field.

3. Incremental approach (translated optical flow)

Our method, called Translated Optical Flow (TOF), is based on both differential and matching methods. The differential side of the method is nothing but the well-known Horn and Schunk's technique (HST). It is worth noting here, that HST is the basis of several known optical flow methods. In the following, we briefly recall Horn and Schunck's technique.

3.1. The differential method of horn and Schunck

Horn and Schunck's method is based on the intensity distribution conservation assumption during the motion expressed in as shown in Eq. (2). To solve this ill-posed problem, that is finding the velocity vector $\mathbf{v} = (u, v)^T$, Horn and Schunck introduced an additional constraint called *smoothness constraint* (SC), which constraints the flow field to be smooth over the whole image. One way of obtaining the smooth flow field is by minimizing the following functional:

$$\Re(\mathbf{v}) = \iint [E_x u + E_y v + E_t] dx dy + \alpha^2 \iint [u_x^2 + u_x^2 + v_x^2 + v_y^2] dx dy$$

$$(4)$$

where the subscripts x, y and t stand for the partial derivations with respect to the spatial and temporal coordinates, respectively. The first term in the left-hand side of Eq. (4) represents a penalty on the deviation of the estimated velocity field from the brightness invariance equation (Eq. (2)), while the second term is a penalty on the deviation of the velocity filed components from smooth surfaces. The positive constant parameter α^2 controls the relative contribution of the latter penalty to the total cost function \aleph . This minimization problem is solved numerically using an iterative gradient search which leads to the following expression of u and v:

$$u^{(n+1)} = \bar{u}^{(n)} - \frac{E_x(E_x\bar{u}^{(n)} + E_y\bar{v}^{(n)} + E_t)}{\alpha^2 + E_x^2 + E_y^2}$$

$$v^{(n+1)} = \bar{v}^{(n)} - \frac{E_y(E_x\bar{u}^{(n)} + E_y\bar{v}^{(n)} + E_t)}{\alpha^2 + E_x^2 + E_y^2}$$
(5)

where *n* stands for the iteration number and $\bar{u}^{(n)}$ and $\bar{v}^{(n)}$ are the neighbourhood averages of $u^{(n)}$ and $v^{(n)}$. Horn and Schunk [22] approximated the derivative terms using finite difference technique. They used Eq. (6), where (i, j) stand for pixel coordinates and the subscripts (1, 2) are the indices related to two successive frames.

$$E_{x}(i,j) = \frac{1}{2} \left[E_{x,1}(i,j) + E_{x,2}(i,j) \right]$$

$$E_{y}(i,j) = \frac{1}{2} \left[E_{y,1}(i,j) + E_{y,2}(i,j) \right]$$

$$E_{t}(i,j) = \left[E_{2}(i,j) - E_{1}(i,j) \right]$$
(6)

k = i

j j+1

where

$$E_{x,k}(i,j) = 1/2[E(i+1,j,k) - E(i,j,k) + E(i+1,j+1,k) - E(i,j,k)] \text{ for } (k = 1,2)$$

$$E_{y,k}(i,j) = 1/2[E(i,j+1,k) - E(i,j,k) + E(i+1,j,k)] \text{ for } (k = 1,2)$$

$$E_k(i,j) = 1/4[E(i,j,k) + E(i+1,j,k)] \text{ for } (k = 1,2)$$

$$+ E(i,j+1,k) + E(i+1,j+1,k)] \text{ for } (k = 1,2)$$
(7)

Notice that in Eq. (7), the partial derivative terms E_x , E_y and E_t are computed on the same spatio-temporal point (see Fig. 1)

3.2. TOF technique

The basic idea of TOF is to combine HS method with a classical result derived from Helmolt's theorem [19]. This theorem states that any vector movement \mathbf{s} of a material surface element could be decomposed into elementary and small enough movements composed mainly of translation, rotation, and a uniform dilation of the two orthogonal components of a shear. Thus, in any small neighbourhood of a given image point, say O, one can write the displacement as follows:

$$\mathbf{s}(\mathbf{x}) = \mathbf{s}(\mathbf{x}_O) + J(\mathbf{x} - \mathbf{x}_O) \equiv \mathbf{s}_O + \delta \mathbf{s}$$
(8)

where \mathbf{s}_O is the displacement of O and J is the Jacobian given by

$$J = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}$$
(9)

The derivatives are computed at the point O. If J is bounded, which is the case when the partial derivative exist and are

j j+1

4



k = 2

Fig. 1. The location where the partial derivatives are estimated when using Horn and Schunck's method (k = 1, 2).

386



Fig. 2. The location where the partial derivatives are estimated when using TOF (k = 1, 2).

continuous [16], we can write:

$$\|J(\mathbf{x} - \mathbf{x}_O)\| \le \|J\| \|(\mathbf{x} - \mathbf{x}_O)\| \tag{10}$$

where $||J|| = \sup ||J\mathbf{x}||$ for ||x|| = 1.

In the following, we make use of this assumption since we consider slow transformations such that the partial derivatives expressed in Eq. (9) are quite small (such is the case when the temporal sampling rate is high enough). We can write:

$$\|\delta \mathbf{s}\| = \|J(\mathbf{x} - \mathbf{x}_O)\| \le \|J\| \|(\mathbf{x} - \mathbf{x}_O)\| \le M$$

$$\tag{11}$$

where M is a small enough constant. Thus, for such transformations, if one considers a small element of surface centred at O, we can interpret the movement as follows. If an observer is on a referential centered at O, he can only notice the local deformation δs . As a consequence, and due to the assumption of small deformation δs , differential techniques, such as HST, could be used for estimating the flow field relative to this local referential.

Now let us apply this idea to the analysis of image sequence. Consider two image frames with the two luminance distributions $E_1 = E(t_1)$ and $E_2 = E(t_2)$. Also consider a surface element Σ centered at an arbitrary point O, origin at the instant t_1 of a referential R_0 related to the initial frame F_1 , and a referential R which moves with the origin O (i.e. O is the origin of R at any instant t). Let $\mathbf{v}(P|R_0)$ be the velocity vector of a point P in the neighbourhood of O, measured in R_0 and similarly let $\mathbf{v}(R|R_0)$ be the translation vector of R with respect to R_0 , defined by:

$$\mathbf{v}(R|R_0) = \overrightarrow{O(t_1)O(t_2)}/\Delta t = \mathbf{s}_O/\Delta t \tag{12}$$

where $\Delta t = t_2 - t_1^2 \mathbf{v}(P|R_0)$ can be written as:

$$\mathbf{v}(P|R_0) = \mathbf{v}(R|R_0) + \mathbf{v}(P|R) \tag{13}$$

If now we use the intensity constraint in R₀ (i.e. we assume

that the luminance distribution E of a given element Σ is invariant under motion), then the same assumption applies in the R referential. Furthermore, in the case of slow transformation satisfying condition Eq. (11), the following relation holds:

$$\|\mathbf{v}(P|R)\| = \|\delta \mathbf{s}\| \le M \tag{14}$$

Using this simple analysis we showed that the conditions which need to be fulfilled in order to apply the differential technique (i.e. intensity invariance and small displacements) are all satisfied in the referential R_0 . In such referential, the numerical expressions of the partial derivative terms expressed in Eq. (6) could be estimated from two successive frames, numbered 1 and 2, (see Fig. 2) as follows:

$$E_{x}(i,j) = 1/2 [E_{x,1}(i,j) + E_{x,2}(i+u_{O},j+v_{O})]$$

$$E_{y}(i,j) = 1/2 [E_{y,1}(i,j) + E_{y,2}(i+u_{O},j+v_{O})]$$
(15)

$$E_t(i,j) = [E_2(i + u_O, j + v_O) + E_1(i,j)]$$

where) (u_O, v_O) are the components of the displacement vector \mathbf{s}_O . Furthermore, the following expressions, where index k stands for frame number, still hold.

$$\begin{split} E_{x,k}(i,j) &= 1/2[E(i+1,j,k) - E(i,j,k) + E(i+1,j+1,k) \\ &- E(i,j+1,k)] \text{ for } (k=1,2) \\ E_{y,k}(i,j) &= 1/2[E(i,j+1,k) - E(i,j,k) + E(i+1,j+1,k) \\ &- E(i+1,j,k)] \text{ for } (k=1,2) \\ E_k(i,j) &= 1/4[E(i,j,k) - E(i+1,j,k) + E(i,j+1+1,k) \\ &+ E(i+1,j+1,k)] \text{ for } (k=1,2) \end{split}$$

Since in general the components (u_O, v_O) of s_O are not integer quantities, an interpolation on the gray-levels may be necessary in order to compute the partial derivatives. In our case a cubic-spline interpolation described in Ref. [29] is used.

It is worth noting here that any differential-based optical flow estimation algorithm could be used with the help of

² In all the rest of the paper we use $\Delta t = 1$. This allows to simplify the analysis and to indiscriminately identify velocity vector and displacement vector.

the given expressions in order to compute the relative displacement field δs . The next step is to compute the overall displacement s which is simply the sum of the relative displacement δs and the translation vector s_0 . Because of the presence of the translation vector term s_O , this method is called *Translated Optical Flow*(TOF).

In summary, the use of TOF requires that we must have at our disposal an approximate solution with sufficient accuracy of the translation displacement vector \mathbf{s}_{O} at any point of the image. In Section 3.3, we show how to obtain this first approximate solution using a classical correlation method. Before going any further, let us make two important remarks. First, it should be mentioned that the use of correlation method is not the only way to compute the translation term and that any method allowing the determination of the approximate translation vector \mathbf{s}_{O} could be used. This gives our method a significant flexibility. The second remark relies on the fact that the introduction of the translation term will modify the way the partial derivatives are computed and could be combined with any of the known optical flow differential methods without affecting the basic idea of TOF.

3.3. Approximate solution determination

The main idea of the method used for estimating the approximate solution of the flow field is to use the classical correlation method to track some patterns but only at some strategic points. Obviously, correlation method is time consuming and the idea of limiting its use to some image features of interest will significantly reduce the computational load. Once these points are localized and the correspondence between the two frames is well established, this solution is then propagated to the other neighboring points using a constraint of global smoothness analogous to that use in HS algorithm. We call this constraint 'propagation constraint' and will be termed PC in the following.

In practice, the correspondence technique implemented in our method is based on a correlation search strategy applied only at regularly spaced nodes of a virtual grid deposited on the first frame. The grid node correspondence allows the determination of a sparse translation displacement field. In the next stage of the strategy, on each node a square analysis window of size proportional to the grid step and where the search correspondence method works is defined. Typical size of this window is $c = \alpha P$, where P is the grid step and α is a factor controlling the width (in this experiment it is chosen in the interval [1/4,1/2]). The correspondence procedure between the first and the second frame uses a template-matching method based on the Maximum Cross-Correlation (MCC) method. The MCC method consists of an exhaustive search over all possible values of the position vector \mathbf{s} for the displacement vector which would maximize the cross-correlation coefficient r(s)

defined by:

$$r(\mathbf{s}) = \frac{\int \int E_1(\mathbf{x}) E_2(\mathbf{x} + \mathbf{s}) d^2 x}{\left[\int \int E_1^2(x) d^2 \mathbf{x} \int \int E_2^2(\mathbf{x} + \mathbf{s}) d^2 \mathbf{x}\right]^{1/2}}$$
(16)

As already mentioned in Section 2.2, the validity of this method depends on the shape of the correlation function. It is well known that the correlation method works well and the solution is unique when the correlation curve is unimodal. In the case where the correlation curve is not unimodal, the uniqueness is not guaranteed. In the following, we will first assume that the correlation curve is unimodal. We will then extend this method to more general situations where such hypothesis no longer holds.

At the end of this matching procedure we have at our disposal a set $\{s_{m,n}\}$ of displacement vectors where the indices (m, n) stand for node spatial coordinates. This sparse field is to be transformed into a dense field with the use of the PC. As in Horn and Schunk's method, this constraint expresses the fact that the searched field would be regular and smooth (i.e. all neighboring points will be subjected to approximately the same amount of displacement). As stated in Section 3.1, this constraint could be treated as the minimization problem, where the function to be minimized is given by:

$$\iint [u_x^2 + u_y^2 + v_x^2 + v_y^2] \mathrm{d}x\mathrm{d}y$$

Using Lagrangian formalism one can easily derive the following Euler equations:

$$\Delta \mathbf{s} = \begin{bmatrix} u_{xx} + u_{yy} \\ v_{xx} + v_{yy} \end{bmatrix} = 0 \tag{17}$$

Using this constraint, the problem of estimating the optical flow becomes the displacement s that minimizes the following cost function:

$$\sum_{n} \sum_{m} \|\mathbf{s}(\mathbf{x}_{m,n}) - \mathbf{s}_{m,n}\|$$
(18)

subject to the $\Delta s = 0$.

This constrained minimization attempts to find a smooth field that matches the sparse field computed using the correlation method at the grid nodes $\mathbf{x}_{m.n.}$

The approximate solution for the dense field is obtained by performing an interpolation on the sparse field derived from a correlation method applied to a limited set of points of interest. Other algorithms for flow field estimation use this strategy to reduce the computation time involved in computing displacement field using correlation method [29]. Such approach, however, usually yields erroneous estimates when local deformations result in between the grid nodes. It is the case, for example, of the particular situations displayed in Fig. 3(a) and (b). Indeed, these figures represent a synthetic grid in its initial state and final deformed state, respectively. As it can be seen if, for convenience and ease of computation, we decide to estimate the displacement only



Fig. 3. An example of very localized deformations that may lead to an erroneous displacement field in the case of the intorpolation methods using estimated values at the nodes.

on the basis of grid nodes information it would not be possible to recover the motion information of, for example, the grid bars by simply interpolating in between the nodes. Whereas, it is the aim of TOF technique to solve this problem since differential approach based on the CI (Eq. (2)) uses the whole spatial information contained in the frame.

The diagram shown in Fig. 4 summarizes the different steps involved in the 'TOF' method.

4. Results and discussion

In the following, we present the results of applying TOF for estimating the flow field to four different image sequences. The proposed method is compared with a multiresolution implementation of HS method since it is usually this sort of approaches, which are used when large amplitude displacements are considered. The image sequences used in this study are the following:

- Synthetic images representing a grid at different stages of deformations (Figs. 5 and 6). This constitutes an excellent example well suited to TOF technique,
- another synthetic image representing a random binary pattern (Fig. 7),
- a real image sequence (from INSA de Lyon) representing a random speckle-like pattern (Fig. 8), obtained by spraying a paint over a material,
- a real image of town known as 'Goldhill' (Fig. 9), and.- a real image sequence downloaded from http:// www.cs.ubc.ca/nest/lci/vista/index-pix.html and called 'Hamburg Taxi' (Fig. 10)

In these experiments, the first three images are numerically transformed in order to obtain two successive frames per image. The main advantage of this experimental approach is that we have at our disposal the exact flow field since the images are deformed on the basis of a predetermined motion-based model. It would then be possible to objectively judge the quality of the flow field estimated by the different studied motion estimation methods. Whereas, in the last example of 'Hamburg taxi', the true flow field is unknown. To compare the computed optical flow fields we use some classical error measures.

4.1. Evaluation of a motion estimation method when the true flow field is known

One way of evaluating the efficiency of a given motion estimation method is to compute the discrepancy between the true flow field, when known, and the one derived from the method under consideration. Such discrepancy could be expressed in terms of a distance or a norm. Given two matrices A and B of size $M \times N$, the distance between them



Fig. 4. TOF diagram.

in terms of the Frobenius norm is given by:

$$d(A,B) = \left[\frac{1}{MN} \sum_{j=1}^{M} \sum_{i=1}^{N} (a_{ij} - b_{ij})^2\right]^{1/2}$$
(19)

The main shortcoming of this error measure is that it reflects a global deviation of A from B and consequently does not

心 医胆道	CO.	Signal	Nam	1023	儒
章 法法律	W282	1000	6000	2200	100
2 建酸	部間	842	133	部軍	8
a roome	2022002.0	And second	Market .		-
3 御祭	計画版	1000	地域	北部	靏
う 恐縮さ	設設	152	1800	國際	劉
NO MONTON	-	PER LONG	-	NOTES	
8 F. 16	R. CON	院领	122	1992	巖
國 新聞院	到福祉	和原	BRAK	1000	围
31 15178	税 (10)3	2552	1212256	62128	20
	1253	建設	國際	932	胞
OR BOOLER	199720		Renauti	oragen.	1009
新新教	領認	調調	1263	統設	蘣
國 副語	SK 52	新聞	1000	的外部	億
and the second	Sec. 1	distants.	Distantion of	No diverse	-
S 0280	COLUMN STATE	2014年	STORE OF	1000	22

(a) – Original image (magnification x 2)

give any indication on the error distribution. One way of taking into account this aspect is to incorporate some spatial information in the error distribution. Fleet and Jepson [13] have proposed such approach, where an angular measure of error is used in order to evaluate the deviation of the computed flow field from the correct one. This method is briefly described next.



(b) - Deformed image (magnification x 2)



(c) – Results using the correlation method (magnification x 2)



(d) – Difference between the image (b) and the image reconstructed from (a) and (f)



(e) – Difference between the image (b) and the image reconstructed from (a) and (i)

Fig. 5. Application of the 'Translated Optical flow' for the case of synthetic image.

390



Fig. 5 (continued)

A given velocity vector $\mathbf{v} = (u, v)$, or equivalently a displacement vector of (u, v) pixels per unit time, could be interpreted as a vector of components (u, v, 1) in the 3D spatio-temporal space (x - y - t). In such space, an angular deviation Ψ_E of the estimated displacement field $\mathbf{v}_e = (u_e, v_e, 1)^T$ from the orientation of the correct one

 $\mathbf{v}_c = (u_c, v_c, 1)^T$ is a more plausible and attractive objective measure for flow field estimation evaluation. This measure is defined by:

$$\Psi_E = \cos^{-1}(\mathbf{u}_c, \mathbf{u}_e) \tag{20}$$



(f) – Reconstructed field (/10) using TOF (α = 40, Iter.=250) (g) – Histogram of the differences between (c) and (f)

Fig. 6. Comparison between HS and TOF methods for the case of synthetic image.

where

$$\mathbf{u}_{c} = \frac{(u_{c}, v_{c}, 1)^{T}}{\sqrt{u_{c}^{2} + v_{c}^{2} + 1}} \qquad \mathbf{u}_{e} = \frac{(u_{e}, v_{e}, 1)^{T}}{\sqrt{u_{e}^{2} + v_{e}^{2} + 1}}$$
(21)

One drawback of such measure is that the angular errors may be relatively small in the case of large displacements. Moreover, the estimated deviation corresponding to two symmetrical estimated vectors (symmetry with



(d) – Reconstructed field (/10) using HSM (α = 40, Iter.= 250) (e) – Histogram of the differences between (c) and (d)



(f) – Reconstructed field (/10) using TOF (α = 40, Iter.= 250) (g) – Histogram of the differences between (c) and (f) Fig. 7. Comparison between the HS and TOF methods for the case of binary image.

respect to the correct vector) might yield very different angular deviations [38]. Keeping these in mind, it would be better to use a deviation histogram, which is nothing but the distribution of the magnitude differences between the estimated vector and the correct one. This histogram may help in evaluating the efficiency of the flow filed estimation method. Indeed, for instance, a narrow shaped histogram is an indication of an efficient estimation.



(f) – Reconstructed field using TOF ($\alpha = 40$, Iter. = 250) (g) – Histogram of the differences between (c) and (f)

Fig. 8. Comparison of the methods HSM and TOF in the case of an image of the 'speckle' type.

4.2. Evaluation of a motion estimation method when the flow field is unknown

In the case of an unknown flow field, which is the most encountered situation in real life, we make use of the Inverse Reconstruction Method (IRM) in evaluating the flow field estimation method. Indeed, let I_1 and I_2 be

the two frames of the image sequence under study and let s(r) be the estimated flow. Thus, given I_1 and $\mathbf{s}(\mathbf{r})$, it would be possible to recover an estimated frame, say \hat{I}_2 . This reconstructed frame could then be compared to the actual given frame I_2 . One can, for example, use the Mean Square Error (MSE) to quantify the deviation of the estimated frame from the correct one. This approach has been tested on synthetic images and its performance compared to the angular deviation method of Fleet and Jepson (which is only useful when the true flow field is known) [23]. It was found that the two error measures are correlated [9]. This motivated us to use IRM as an evaluation criterion of flow field estimation technique. It is worth keeping in mind, however, that this approach should not be considered as a rigorous and universal method for evaluation. Indeed, consider for example a homogeneous region in the initial frame I_1 one can easily derive various flow field $\mathbf{s}(\mathbf{r})$ yielding the same final frame \hat{I}_2 . However, we feel that the introduction of additional criteria (objective criteria such as dissimilarity measure or smoothness constraint, or subjective criterion



(f) – Reconstructed field (/20) using FOT (α = 40, Iter.=250) (g) – Histogram of the differences between (c) and (f)

Fig. 9. Comparison of the methods HSM and TOF in the case of a real image 'Goldhill'.



(h) – Reconstructed field (/20) using TOF from the exact values of the field at the nodes of the network.



4.3. Results

To show how TOF works, a first experiment is conducted on a synthetic image (Fig. 5) composed of a grid with a thickness of 22 pixels and with a regular step in both horizontal and vertical directions embedded in a textured homogeneous background (Fig. 5(a)). A local deformation of amplitude 3 pixels is applied to a circular region 24 pixels wide. This local deformation is firstly applied to the image center then followed by a uniform dilation of 1% strength and centered on the upper left corner of the image (this point is supposed to be fixed during all the transformation). The result of the transformation is shown in Fig. 5(b). The local deformation is clearly visible at the center of the image.

The estimated flow field using the correlation method described in Section 3.3 is shown in Fig. 5(c). This sparse vector field is propagated to the rest of the image using the Propagation Constraint (PC) expressed in Eq. (21). The result is shown in Fig. 5(f). By comparing this estimated flow field to the correct one (shown in Fig. 5(j)), one can see that the local deformation of the image is not taken into account when using the correlation method. This is not surprising since the extent of the local deformation magnitude is of order size of the grid step. This is more noticeable in the histogram of the deviations between the estimated flow field and the correct one as shown in Fig. 5(g). Notice that in all the conducted experiments the error measures are computed in the quadrilateral delimited by grid nodes because, in the scope of our method, the flow field has no meaning outside this region. This is confirmed in Fig. 5(d), where the difference between the correct deformed image (Fig. 5(b)) and the one estimated by IRM



(i) – Histogram of the differences between (c) and (f)

Fig. 9 (continued)

(using the original image of Fig. 5(a) and the estimated flow field of Fig. 5(f)) is shown.

Fig. 5(h) shows the flow field δs estimated by TOF method. This flow field is clearly high in the central zone where the local deformation was applied. When summing the flow fields shown in Fig. 5(f) and (h), the total displacement field is obtained as displayed in Fig. 5(i). Using TOF, a good precision in flow field estimation was achieved. The maximum distance error obtained was 0.7 pixels compared to the 2.3 pixels of the correlation method alone. The new mean error of 0.06 pixels is better than the 0.26 pixels due to the correlation method. If one compares the flow field obtained using TOF with the correct flow field shown in Fig. 5(j), one can clearly see that the local deformation has been well taken into account as illustrated in Fig. 5(e).

In the following experiments (see Figs. 6-10), the results obtained using our method are compared with those obtained with the Horn and Schunck based mutiresolution method. The optimum parameters used in the two studied methods are mentioned in the figure captions. As it can be noticed, TOF technique clearly outperforms the multi-resolution-based approach when dealing with large displacements. This is confirmed by the example shown in Fig. 6 where a punch effect is applied to the synthetic grid image. The displacement field, which attains in some areas a magnitude of ten pixels, is estimated with an average precision of order of 1/10 pixels with the proposed approach whereas multiresolution method attains an average value of 1/4. This superiority is also well illustrated in the shape of the error histogram, which exhibits a very narrow mode.

This gain in precision is visible in the image sequence shown in Fig. 7 where a binary image is subjected to a translation of 9 pixels towards the right bottom direction. Indeed, on this textured image, the coarse analysis inherent to the multiresolution approach image yields inevitably a loss of information which in turn is propagated to lower



(a) - Initial image (magnification x 2)



(c) – Difference between the image (b) and the image reconstructed from (a) and (e)





(b) - Final image (magnification x 2)



(d) – Difference bewteen the image (b) and the image reconstructed from (a) and (f)



er. = 250) (f) – Field reconstructed by TOF (α = 40, Iter. = 250)

Fig. 10. Comparison between HSM and TOF methods in the case of the real sequence 'Hamburg Taxi'.

levels of the pyramid leading to a poor estimation of the flow field compared to that of TOF method.

Another test sample shown in Fig. 8, corresponding to a speckle like image similar to that used in the previous example, is subjected to a small rotation (1°) resulting in

a maximum displacement of 4 pixels. In this case, the TOF method, with an average precision of about 1/25 pixels, outperforms once more the HS method.

To test the efficiency of the proposed method, a real image, called 'Goldhill', shown in Fig. 9 and composed of

different structures is considered. A displacement field composed of a uniform dilation of 20% followed by a circular deformation attaining 50 pixels at the central zone are simulated on this image. Once more, the TOF method is shown to outperform HS technique. Notice here that the attained precision is less important than that obtained with the previous experiments. This is mainly due to the high dilation amplitude used in this case, which result in a poor displacement field on the virtual grid nodes.

It is worth noticing here that the grid step is a decisive parameter in the proposed method. Indeed, a small grid step will result in a better estimation of the displacement field but with an increase in computation time especially in the correlation computation step. Furthermore, in the case of a small step, a small size analysis window for matching is used. This does not guarantee the existence of a unimodal bell-shaped correlation curve. When such situation occurs the correlation method could yield erroneous displacement vectors which are propagated via PC to other parts of the image leading to a poor estimated flow field when using TOF. The importance, for the second step of TOF technique, of the displacement vectors obtained with the correlation step is demonstrated in Fig. 9(h) and (i). The dense flow field reconstructed when the TOF method is applied with the a priori knowledge of the actual displacement vectors of the nodes of a virtual grid is shown in Fig. 9(h). In this case an improvement is observed as shown in Fig. 9(i). Through these results, it is shown that the use of TOF requires a good precision in the estimation of the sparse flow field computed on regularly spaced nodes. One possible improvement is to select other points of interest not necessarily on a regular grid but randomly distributed on the image plane or better belonging to some salient features of the image signal (contrasted points, corners, end of line, etc.). Once these points of interest are selected, any correlation method such that described in Ref. [50] would be used to efficiently estimating the sparse flow field which is then propagated to the other points of the image using Eq. (21).

The last example is a real image sequence called 'Hamburg Taxi', shown in Fig. 10 whose true displacement field is unknown. To evaluate the quality of the estimated flow field, the IRM is used. Fig. 10(c) displays the difference between the last image of the real image sequence and the one reconstructed when using the flow field obtained with HS method (Fig. 9). Fig. 10(d) corresponds to the difference between the last image of the real image sequence and that of the one reconstructed using the flow field obtained with TOF method (Fig. 10(f)). In both figures the difference in magnitude have been multiplied by a factor 2 for ease of visual appearance. It could be noticed that the image reconstructed using TOF better fits the real original one than does the reconstructed one using HSM. Moreover, Fig. 10(f) shows that the displacement vectors of the two vehicles moving in opposite direction are better discriminated when using TOF method than when using the HSM. Using the HMS results in an oversmoothed flow field between the two

mobiles as already reported in Ref. [50]. This drawback is avoided thanks to the PC which guarantees the stability of the nodes displacement vectors, supposed to be correctly computed, and allows to control the extent of smoothing effect.

5. Conclusions

In this paper, we have shown that TOF method can efficiently deal with large displacements in the case of continuous and slow deformations, where the classical multiresolution approaches fail. The proposed approach offers two advantages, namely time saving in the computation and good precision (1/10 pixel) when compared to similar approaches. The obtained results show that this method can be potentially used in different scientific areas such as material deformation estimation, medical investigations, and meteorological analysis. This method, however, fails to give reliable flow field in the presence of discontinuities and occlusions. These issues are being investigated by the authors.

References

- E.H. Adelson, J.R. Bergen, Spatiotemporal energy models for the perception of motion, J. Opt. Soc. Am. A2 (1985) 284–299.
- [2] P. Anadan, A Computational framework and an algorithm for the measurement of visual motion, Int. J. Comput. Vision 2 (1989) 283–310.
- [3] J.P. Anandan, J.R. Bergen, K.J. Hanna, R. Hingorani, Hierarchical model-based motion estimation, in: M.I. Sezan, R.L. Lagendijk (Eds.), Motion Analysis and Image Sequence Processing, Kluwer Academic Publishers, Dordecht, 1993, pp. 1–22.
- [4] J.L. Barron, D.J. Fleet, S.S. Beauchemin, Performance of optical flow techniques, Int. J. Comput. Vision 12 (1) (1994) 43–77.
- [5] N. Ben Amar, A. Beghdadi, P. Viaris de Lesegno, An all-digital method for accurate measurements of mechanical deformations, Scanning 18 (1996) 327–330.
- [6] A. Del Bimbo, P. Nesi, J.L.C. Sanz, Analysis of optical flow contraints, IEEE Trans. Image Process. 4 (4) (1995) 460–469.
- [7] M.J. Black, A. Jepson, Estimating optical flow in segmented images using variable-order parametric models with local deformations, IEEE Trans. Pattern. Anal. Machine Intell. 18 (10) (1996) 972–986.
- [8] N. Cornelius, T. Kanade, Adapting Optical Flow to Measure Object Motion in Reflectance and X-ray Image Sequences, Siggraph/Sigart Interdisciplinary Workshop on Motion: Representation and Perception, Toronto, Ont., Canada, 1983, pp. 50–58.
- [9] Z. Diab, Champ de mouvement et flux optique, Ecole Polytechnique de Montréal, 1995, http://www.ai.polymtl.ca/CdL/Zafer.
- [10] W. Enkelmann, Investigations of Multigrid Algorithms for the Estimation of Optical Flow Fields in Image Sequences, Workshop on Motion: Representation and Analysis, Kiawah Island Resort, Charleston/SC, May 7–9, IEEE Computer Society Press, New York, 1986, pp. 81–87.
- [11] W. Enkelmann, Investigations of multigrid algorithms for the estimation of optical flow fields in image sequences, Comput. Vision, Graphics Image Process. 43 (1988) 150–177.
- [12] C. Fennema, W. Thompson, Velocity determination in scenes containing several moving objects, Comput. Graphics Image Process. 9 (1979) 301–315.

- [13] D.J. Fleet, A.D. Jepson, Computation of component image velocity from local phase information, Int. J. Comput. Vision 5 (1) (1990) 77-104.
- [14] F. Glazer, Multilevel relaxation in low-level computer vision, in: A. Rosenfeld (Ed.), Multiresolution Image Processing and Analysis, Springer, Berlin, 1984, pp. 312–330.
- [15] J.D. Gorce, D. Friboulet, I.E. Magnin, Méthode d'estimation du mouvement des parois cardiaques à partir d'images 3D, Innov. Techn. Biol. Med. 15 (5) (1994).
- [16] G.H. Golub, C.F. Van Loan, Matrix Computations, John Hopkins University Press, Baltimore, USA, 1996.
- [17] S. Ghosal, P. Vanek, A fast scalable algorithm for discontinuous optical flow estimation, IEEE Trans. Pattern Anal. Mach. Intell. 18 (2) (1996) 181–194.
- [18] E.L. Hall, Computer Image Processing and Recognition, Academic Press, New York, 1979.
- [19] H. Helmholz, Über die Integrale der hydrodynamischen Gleichungen, welche den Wirbelbewegungen entsprechen, Crelles J. 55 (1858) 25.
- [20] D.J. Heeger, Model for the extraction of the image flow, J. Opt. Soc. Am. A4 (1987) 1455–1471.
- [21] D.J. Heeger, Optical flow from spatiotemporal filters, Int. J. Comput. Vision 1 (1988) 279–302.
- [22] K.P. Horn, B.G. Schunck, Determining optical flow, Artificial Intell. 17 (1981) 185–203.
- [23] T. Lin, J.L. Barron, Image reconstruction error for optical flow, Technical Report, University of Western Ontario, 1994.
- [24] B. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, Proceedings DARPA Image Understanding Workshop (1981) 121–130.
- [25] M.W. Mak, W.G. Allen, A lip-tracking system based on morphological processing and block matching techniques, Signal Process.: Image Commun. 6 (1993) 335–348.
- [26] G.E. Mailloux, F. Langlois, P.Y. Simard, M. Bertrand, Restoration of the velocity field of the heart from two-dimensional echocardiograms, IEEE Trans. Med. Imaging 8 (2) (1989) 143–153.
- [27] E. Mémin, P. Pérez, Adaptative multigrid and variable parameterization for optical-flow estimation, Internal Publication, IRISA, No. 1057, Oct. 1996, pp. 1–19.
- [28] E. Mémin, P. Pérez, Dense estimation and object-based segmentation of the optical flow with robust techniques, IEEE Trans. Image Process. 7 (5) (1998) 703–719.
- [29] S. Mguil-Touchal, F. Morestin, M. Brunet, Mesure de champs de déformations par une méthode optique de corrélation directe d'images digitales, Colloque National Mécamat (Aussois) (1996) 179–182.
- [30] A. Mitiche, Y. Wang, J.K. Aggarwal, Experiments in computing optical flow with the gradient-based, multiconstraint method, Pattern Recogn. 20 (2) (1987) 173–179.
- [31] A. Mitiche, P. Bouthemy, Computation and analysis of image motion: a synopsis of current problems and methods, Int. J. Comput. Vision 19 (1) (1996) 29–55.
- [32] J. Monteil, A. Beghdadi, P. Viaris de Lesegno, Mesures de champ de déformation par une technique incrementale de flux optique, Revue de la Métallurgie—CIT 97 (2) (2000) 229–238.

- [33] R. Mongrain, Estimation de paramètres hémodynamiques à l'aide de séquences angiographiques, PhD Thesis, Université de Montréal, p. 117, 1993.
- [34] H.-H. Nagel, W. Enkelmann, An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences, IEEE Trans. Pattern Anal. Machine Intell. Pami-8 (5) (1986).
- [35] H.-H. Nagel, On the estimation of optical flow: relations between different approaches and some new results, Artificial Intell. 33 (1987) 299–324.
- [36] P. Nesi, A. Del Bimbo, D. Ben-Tzvi, A robust algorithm for optical flow estimation, Comput. Vision, Graphics Image Process. 62 (1995) 59–68.
- [37] J.M. Odobez, P. Bouthemy, Estimation robuste multiéchelle de modèles paramétrés de mouvement sur des scènes complexes, Traitement du Signal 12 (2) (1995) 113–128.
- [38] M. Otte, H.-H. Nagel, in: J.-O. Eklundh (Ed.), Optical Flow Estimation: Advances and Comparisons, Proceedings of the Third European Conference on Computer Vision-ECCV'94, Stockholm Sweden, 1994, pp. 1–19.
- [39] R. Paquin, E. Dubois, A spatio-temporel gradient method for estimating the displacement field in time-varying imagery, Comput. Vision, Graphics Image Process. 21 (1983) 205–221.
- [40] J.A. Perrone, Simple technique for optical flow estimation, J. Opt. Soc. Am. 7 (2) (1990) 264–278.
- [41] S. Ruan, A. Bruno, R. Collorec, J.L. Coatrieux, Estimation de mouvement 3D en coronarographie, Innov. Techn. Biol. Med. 14 (2) (1993).
- [42] B.G. Schunck, Image flow segmentation and estimation by constraint line clustering, IEEE Trans. Pattern Anal. Machine Intell. 11 (10) (1989) 1010–1027.
- [43] C. Schmid, R. Mohr, C. Bauckhage, Evaluation of interest point detectors, Int. J. Comput. Vision 37 (2) (2000) 151–172.
- [44] A. Singh, An estimation-theoric framework for image-flow computation, Proceedings of the Third International Conference in Computer Vision, Osaka, 1990, pp. 168–177.
- [45] O. Tretiak, L. Pastor, Velocity estimation from image sequences with second order differential operators, Proceedings International Conference on Pattern Recognition, Montréal, Quebec, 1984, pp. 16–19.
- [46] A. Verri, F. Girosi, V. Torre, Differential techniques for optical flow, J. Opt. Soc. Am. 7 (5) (1990) 912–922.
- [47] A.B. Watson, A.J. Ahumada, Model of human visual-motion sensing, J. Opt. Soc. Am. A2 (1985) 322–342.
- [48] A.M. Waxman, K. Whon, Contour evolution, neighbourhood deformation and global image flow: planar surfaces in motion, Int. J. Robotics Res. 4 (1985) 95–108.
- [49] K. Wohn, L.S. Davis, P. Thrift, Motion estimation based on multiple local constraints and nonlinear smoothing, Pattern Recogn. 16 (6) (1983) 563–570.
- [50] W. Qing X, A correlation-relaxation-labeling framework for computing optical flow-template matching from a new perspective, IEEE Trans. Pattern Anal. Mach. Intell. Pami-17 (8) (1995).
- [51] B. Zitova, J. Flusser, J. Kautsky, G. Peters, Feature point detection in multiframe images, Proceedings of Czech Pattern Recognition Workshop, 2000.