Boundary Preserving Joint Estimation of Optical Flow and Disparity in a Sequence of Stereoscopic Images

H. Weiler, A. Mitiche, A. Mansouri
INRS-EMT
Place Bonaventure 800 rue de la Gauchetiere O.
Suite 6900 Montreal, Quebec H5A 1K6
{weiler,mitiche,mansouri}@inrs-telecom.uquebec.ca

Abstract

The purpose of this study is to investigate a new method for the joint estimation of optical flow and disparity in a sequence of stereoscopic images. For a sequence of two pairs of stereoscopic images, the problem involves the estimation of four fields, two fields of image motion and two of disparity. Stereoscopy relates the four fields, and there is no relation between any three of these fields except through the fourth. We present a new method where three of the fields are estimated independently by anisotropic diffusion to preserve their boundaries, and the fourth deduced using the integrability constraint. Neither spatial nor temporal correspondence is explicitly referenced. Extention of the scheme over time requires the estimation of only two fields and an application of the integrability constraint at each time step. When motion or disparity in any of the fields is of large extent, the scheme is activated via multiresolution. Several experiments confirm the advantage of the method both in terms of accuracy and speed of execution.

Keywords: Machine vision, motion, disparity, estimation

1 Introduction

Optical flow and disparity play a central role in serveral image processing and computer vision problems such as coding, image segmentation, motion tracking, and image three-dimensional interpretation. These are fundamental problems in useful applications such as robotics, surveillance, image database retrieval, and video analysis.

Optical flow and disparity have each been subject to numerous studies [1][3][2]. Because the two fields are related in stereoscopy [4][5], computational benefits may accrue from their joint estimation [6] [7][8][9][10]. Implementation of joint estimation involves determining four fields: two motion fields and two disparity fields. The stereokinematic constraint [4] [5] relates three of these fields to the fourth. There is no relation between any three of these fields except though the fourth. Therefore, joint estimation can be done according to one of two basic paradigms: (1) The four fields are estimated simultaneously using, for instance, a variational formulation to minimize an objective function containing terms of conformity to data, regularization terms of smoothness, and a term to include the stereokinematic constraint, (2) three of the fields are estimated independently, each according, for instance, to a variational formulation as in [1][2], and the fourth field is estimated using the stereokinematic constraint.

Estimation along the first paradigm involves solving a very large system of equations, nonlinear because of the stereokinematic constraint term. For instance, there are 32×10^4 variables to estimate with images of size 200×200. Extending the process over time requires solving each time such a large scale system of equations. Estimation along the second paradigm involves solving independently three much smaller problems to estimate three fields, followed by an application of the stereokinematic constraint to estimate the fourth field. Extending the estimation over time requires estimating only two fields and an application of the stereokinematic constraint. This, therefore, affords significant computational savings.

Previous studies have adopted hybrid paradigms with no obvious formal justification, and have in common that they reference either temporal or stereoscopic correspondence explicitly, requiring the intervention of external processes such as search or interpolation [6] [7][8][9][10]. In this study, we state the problem according to the the second paradigm and in a way that forgoes the need for explicit reference to correspondence. Motion and disparity of small extent is treated directly. In the presence of motion or disparity of large extent, joint motion-disparity multiresolution estimation is done by a coarse-to-fine algorithm we present in Section 3.

The remainder of this paper is organized as follow. Section 2 states the problem and describes its resolution. Section 3 explains our implementation. Section 4 gives experimental results. Section 5 contains a conclusion.

2 Formulation

We define an image sequence as a real function over a domain $\Omega \times [0, S] \times [0, T]$, where Ω is an open subset of $\mathbf{R}^2$ representing the spatial coordinates domain, $[0, T]$ is an interval of time, and $[0, S]$ an interval of $\mathbf{R}$. An image is,

therefore, a map:

$$\begin{aligned} I : \Omega \times [0,S] \times [0,T] &\mapsto \mathbf{R} \\ \mathbf{x}, s, t &\mapsto I(\mathbf{x}, s, t) \end{aligned}$$

For a fixed value of $s \in [0,S]$ we obtain a temporal sequence of images and for two distinct fixed values we obtain a stereoscopic image sequence. Variable s can be thought of as the parameter of the trajectory of a sequence of image planes in these planes coordinate domain. This definition of an image sequence generalizes the notion of a stereoscopic image sequence.

Let $(\mathbf{x}, s, t) \in \Omega \times [0,S] \times [0,T]$ and $\mathbf{x} + D(\mathbf{x}, s + ds, t + dt)$ its correspondant at $(s + ds, t + dt)$, where D designates displacement. The assumption that I does not change at corresponding points,

$$I(\mathbf{x}+D(\mathbf{x},s+ds,t+dt),s+ds,t+dt) = I(\mathbf{x},s,t)$$

yields the following motion and disparity constraints:

$$\begin{aligned} \nabla I \cdot \mathbf{w} + I_t &= 0 \\ \nabla I \cdot \mathbf{d} + I_s &= 0 \end{aligned} \quad (1)$$

where $\mathbf{w}$ is the optical velocity vector, $\mathbf{d}$ the disparity vector, ∇I the spatial gradient of I, I_t and I_s the partial derivatives of I with respect to t and s. Because

$$\begin{aligned} \mathbf{w} &= \frac{\partial D}{\partial t} \\ \mathbf{d} &= \frac{\partial D}{\partial s} \end{aligned}$$

we also have the *integrability constraint*:

$$\frac{\partial \mathbf{w}}{\partial s} = \frac{\partial \mathbf{d}}{\partial t} \quad (2)$$

Constraint (2) is the continuous form of the stereokinematic constraint introduced in [4][5] for stereoscopic image sequences. Its particularity is that it references neither temporal nor disparity correspondance explicitly.

Consider two consecutive stereoscopic images of a stereoscopic sequence: $I^{l,t}, I^{r,t}$, the left and right images at time t, and $I^{l,t'}, I^{r,t'}$, the left and right images at subsequent time t'. Let $\mathbf{w}^{l,t} = (u^{l,t}, v^{l,t})$ and $\mathbf{w}^{r,t} = (u^{r,t}, v^{r,t})$ designate left and right optical velocity at time t, and $\mathbf{d}^t = (\delta_1^t, \delta_2^t)$, $\mathbf{d}^{t'} = (\delta_1^{t'}, \delta_2^{t'})$ designate disparity at t and t'.

Constraints (1) and (2) suggest that the left and right motion fields at time t and the disparity field at time t can be estimated independently before estimating the disparity field at time t' using the integrability-stereokinematic constraint. The left and right motion fields and the disparity field are estimated at time t as:

$$\begin{aligned} \mathbf{w}^{l,t} &= \arg\min_{\mathbf{w}} \{ \int_\Omega \alpha (\nabla I^{l,t} \cdot \mathbf{w} + I_t^{l,t})^2 \\ &\quad + \phi(\|\nabla u\|) + \phi(\|\nabla v\|) \\ \mathbf{w}^{r,t} &= \arg\min_{\mathbf{w}} \{ \int_\Omega \alpha (\nabla I^{r,t} \cdot \mathbf{w} + I_t^{r,t})^2 \\ &\quad + \phi(\|\nabla u\|) + \phi(\|\nabla v\|) \\ \mathbf{d}^t &= \arg\min_{\mathbf{d}} \{ \int_\Omega \alpha (\nabla I^{l,t} \cdot \mathbf{d} + I_s^{l,t})^2 \\ &\quad + \phi(\|\nabla \delta_1\|) + \phi(\|\nabla \delta_2\|) \end{aligned} \quad (3)$$

where α is a positive constant to weigh the contribution of the first term of the integrand (conformity to data) against the other two terms (regularization), and ϕ is a quadratic function as in [1] or a boundary preserving function as in [2]. The Euler-Lagrange equations of the first problem in (3) are:

$$\begin{aligned} 2\alpha I_x^{l,t}(I_x^{l,t}u^{l,t} + I_y^{l,t}v^{l,t} + I_t^{l,t}) &= div(\frac{\phi'(\|\nabla u^{l,t}\|)}{\|\nabla u^{l,t}\|}\nabla u^{l,t}) \\ 2\alpha I_y^{l,t}(I_x^{l,t}u^{l,t} + I_y^{l,t}v^{l,t} + I_t^{l,t}) &= div(\frac{\phi'(\|\nabla v^{l,t}\|)}{\|\nabla v^{l,t}\|}\nabla v^{l,t}) \end{aligned}$$

with the boundary conditions on $\partial\Omega$:

$$\begin{aligned} \frac{\phi'(\|\nabla u^{l,t}\|)}{\|\nabla u^{l,t}\|}\frac{\partial u^{l,t}}{\partial \mathbf{n}} &= 0 \\ \frac{\phi'(\|\nabla v^{l,t}\|)}{\|\nabla v^{l,t}\|}\frac{\partial v^{l,t}}{\partial \mathbf{n}} &= 0 \end{aligned}$$

where $\mathbf{n}$ is the unit normal to $\partial\Omega$. Similar equations can be written for the other two problems in (3).

The Euler-Lagrange equations are discretized using the left images $I^{l,t}, I^{l,t'}$ and solved iteratively [1] [2]. A similar procedure applies for the estimation of the right motion field using the right images $I^{r,t}$ and $I^{r,t'}$, and for the estimation of the disparity field at time t using the images $I^{l,t}$ and $I^{r,t}$.

When the left and right motion fields and the disparity field are estimated at time t, the disparity field at time t' is deduced using the integrability-stereokinematic contraint. For motion and disparity of small extent, this done simply by finite difference approximation of the derivatives in (2):

$$\mathbf{d}^{t'} = \mathbf{w}^{r,t} - \mathbf{w}^{l,t} + \mathbf{d}^t \quad (4)$$

It is important to note that extending the estimation over time requires estimating only two fields and an application of the integrability constraint at each time step. This, therefore, affords significant computational savings. It is also important to note that the variables in (4) are all evaluated at the same spatial coordinates. There is explicit reference to neither temporal nor disparity correspondence. Explicit reference to correspondence would call for the intervention of complex external processes such as interpolation to bring evaluation on grid points of the images positional array. However, direct iterative resolution of a discretized version of the Euler-Lagrange equations, and the use of the integrability constraint in the form (4) is valid for motion and disparity of small extent. In the presence of either motion or disparity of large extent, the estimation of the fields in (3), and the application of the integrability constraint (4)

are performed at each level of a multiresolution pyramid according to a scheme described in Section 3.

For the estimation problems in (3), we used the Aubert function, $\phi(z) = 2\sqrt{1+z^2} - 2$. With this function, the regularization term of the functionals in (3) has the following effect: In regions where motion (disparity) is homogeneous, smoothing is done by isotropic diffusion, as with the Horn and Schunk method. At motion (disparity) boundaries, however, and unlike with the Horn and Schunk method, smoothing is done by anisotropic diffusion to preserve these boundaries because it is favored along motion boundaries and inhibited across [2]. When discretized, the Euler-Lagrange equations corresponding to problems (3) yield large scale sparse systems of nonlinear equations that can be solved efficiently via the half-quadratic algorithm [2]. The algorithm can be shown to converge to a unique solution.

3 Multiresolution joint estimation

All multiresolution approaches have in common the use of pyramids of images. These pyramids are obtained by image filtering and subsampling. We built such a pyramid using the anisotropic diffusion Perona-Malik [11] filter to preserve image boundaries, and subsampling by 2. Joint estimation is then performed on the pyramid using a coarse-to-fine strategy. Estimation starts at the coarsest level where motion and disparity have an extent small enough to validate a variational, gradient-based independent estimation of three of the fields, and the deduction of the fourth by the integrabilty constraint. At any subsequent finer level, a correction to the estimates of the previous level are calculated using the integrability constraint and the same motion (disparity) variational gradient-based estimation algorithms. This coarse-to-fine algorithm is given in Figure 1.

/ Variables and symbols */*	
$w_r(l,t,p)$	: velocity of point **p** at time **t** on level **l** in the right image
$w_l(l,t,p)$	: velocity in the left image
$d(l,t,p)$	: disparity between both image
$I_r, I_l(l,t,p)$	: brightness in the right or left image
dw_r, dw_l	: correction of the velocity
$dd(l,t,p)$	: correction of the disparity
l_{min}, l_{max}	: lowest and highest level in the pyramid
dt	: sampling rate between two images
$I'w_r, I'w_l$	: image at time **t+dt**, created by moving image $I(t)$ with flow field **w(t)**
$I'd$	: image at time **t+dt**, created by moving image $I_r(t)$ with disparity field **d(t)**
Aub	: Aubert/Kornprobst/Deriche algorithm [2]
/ Estimation on lowest level */*	
w_r, w_l:	
$w(l_{min}, t)$	$= Aub(I(l_{min}, t), I(l_{min}, t+dt))$
$d(l_{min}, t)$	$= Aub(I_l(l_{min}, t), I_r(l_{min}, t))$
$d(l_{min}, t+dt)$	$= w_r(l_{min}, t) - w_l(l_{min}, t)$ $+d(l_{min}, t)$
/ Correction of flow and disparity */*	
for ($l = l_{min}$ TO $l = l_{max}$)	
w_l, w_r :	
$w(l+1, t, p)$	$= 2w(l, t, \frac{p}{2})$
$d(l+1, t, p)$	$= 2d(l, t, \frac{p}{2})$
I_l, I_r :	
$I'w(l+1, t+dt)$	$= I(l+1, t) + w(l+1, t)$
$I'd(l+1, t)$	$= I_r(l+1, t) + d(l+1, t)$
dw_l, dw_r :	
$dw(l+1, t)$	$= Aub(I'w(l+1, t+dt)$ $, I(l+1, t+dt))$
$dd(l+1, t)$	$= Aub(I'd(l+1, t), I_l(l+1, t))$
/ Joint estimation of disparity and flow */*	
$dd(l+1, t+dt)$	$= dw_r(l+1, t) - dw_l(l+1, t)$ $+dd(l+1, t)$
$d(t+dt)$	$= d(t+dt) + dd(t+dt)$
/**************************************/	
w_l, w_r :	
$w(l+1, t)$	$= w(l+1, t) + dw(l+1, t)$
$d(l+1, t)$	$= d(l+1, t) + dd(l+1, t)$

Table 1. Coarse-to-fine joint estimation of optical flow and disparity

4 Experimental verification

To verify our method, we ran several experiments two of which we describe here. The scene in these expriments consists of a background and two objects which are cut out of real images. Motions and disparities are synthetic to provide ground truth for the purpose of quantitative evaluation. The sequence in the first experiment consists of two pairs of binocular images, the secong displayed in Figure 2. The background and the objects are given disparities in the first pair of binocular images and are made to move such that disparities in the second pair of binocular images are (-1,0) for the background, and (1,1) for the two objects. This first experiment is to test the method with motions and disparities of small extent. In the second experiment, the sequence consists of ten pairs of binocular images, the last one displayed in Figure 3. We used disparities and motions to produce disparities of large extent in the last (tenth) pair of binocular images. These are (-10,0) for the background, (-17,-12) for the object on the left, and (-2,12) for the other. Although disparities in these experiments may not correspond to any kind of real life scene or camera, the generated sequences do provide adequate test data, with ground truth against which we can contrast the estimates of the method. We show the results of the estimation of disparities for the last binocular pair of images. These results, shown graphically in Figures 2 and 3, and quantitatively in Table 2, indicate the efficiency of joint estimation by the method. The top box of Table 2, labelled ”Actual”, shows the ground

truth of disparities for the background (labelled "BG") and the left and right objects (labelled "Obj 1" and "Obj 2", for both experiments (labelled "Exprmnt 1" and "Exprmnt 2"). respectively). The middle box of the table contains the results obtained with the *joint estimation* method, applied incrementally via the integrability constraint. The lower box (labelled "Aubert") shows the results obtained by direct estimation. Note that for the second experiment, the disparities are just too large to be estimated directly, even with the assistance of multiresolution. As can be seen from these examples, the joint estimation method performs well, and better than independent, direct estimation, particularly in the case of large displacements.

Actual	BG	Obj 1	Obj 2.
Exprmnt 1 - x	-1.0	1.0	1.0
Exprmnt 1 - y	0.0	1.0	1.0
Exprmnt 2 - x	-10.0	-17.0	-2.0
Exprmnt 2 - y	0.0	-12.0	12.0
Joint estimation	BG	Obj 1	Obj 2.
Exprmnt 1 - x	-0.8	0.96	0.92
Exprmnt 1 - y	0.02	0.81	0.87
Exprmnt 2 - x	-7.65	-14.42	-2.61
Exprmnt 2 - y	0.16	-8.55	8.46
Aubert	BG	Obj 1	Obj 2.
Exprmnt 1 - x	-0.81	0.87	0.90
Exprmnt 1 - y	0.02	0.78	0.87
Exprmnt 2 - x	.	.	.
Exprmnt 2 - y	.	.	.

Table 2. Experimental results

5 Conclusion

The purpose of this study was to investigate a new method for the joint estimation of optical flow and disparity in a sequence of stereoscopic images. This method, which uses the integrability constraint, and multiresolution in the presence of large displacements, has given promising results in a number of experiments.

References

[1] B. K. P Horn and B. G. Schunk, Determining Optical Flow. *Artificial Intelligence*, 17, 1981, 185-203.

[2] G. Aubert, R. Deriche, and P. Kornprobst, Computing optical flow via variational techniques. *SIAM Journal of Applied Mathematics*, 60(1), 199, 156-182.

[3] A. Mitiche and P. Bouthemy, Computation and Analysis of Image Motion: A Synopsis of Current Problems and Methods, *International Journal of Computer Vision* 19(1), 1996, 29-55.

[4] A. Mitiche, On Combining Stereopsis and Kineopsis for Space Perception, *First International Conference on Artificial Intelligence Applications*, Denver, CO, 1984, pp. 156-160

[5] A. Mitiche, A Computational Approach to the Fusion of Stereopsis and Kineopsis, in W.N. Martin and J.K. Aggarwal (Eds.) *Motion Understanding: Robot and Human Vision*, (Kluwer Academic Publishers, 1988), 81-99.

[6] A. Tamtaoui and C. Labit, Constrained disparity and motion estimation for 3DTV image sequence coding, *Signal Processing: Image Communication*, 4, 1991, 45-54.

[7] J. Liu and R. Skerjanc, Stereo and motion correspondence in a sequence of stereo images, *Signal Processing: Image Communication*, 5, 1993, 305-318.

[8] Y. Altunbasak, A. M. Tekalp, G. Bozdagi, Simultaneous motion-disparity estimation and segmentation from stereo, *First International Conference on Image Processing*, Austin, Texas, 1994, pp. 73-77.

[9] I. Patras, N. Alvertos, and G. Tzirita, Joint Disparity and Motion Field Estimation in Stereoscopic Image Sequences, Computer Science Dept, University of Greece, FORTH-ICS TR-157, 1995.

[10] Laganière, R., *Analyse stéréocinétique d'une séquence d'images: Estimation des champs de mouvement et de disparité*, Ph.D. thesis, INRS-Télécommunications, 1995.

[11] P. Perona and J. Malik: Scale-Space and Edge Detection Using Anisotropic Diffusion, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7), 1990, 629-639.

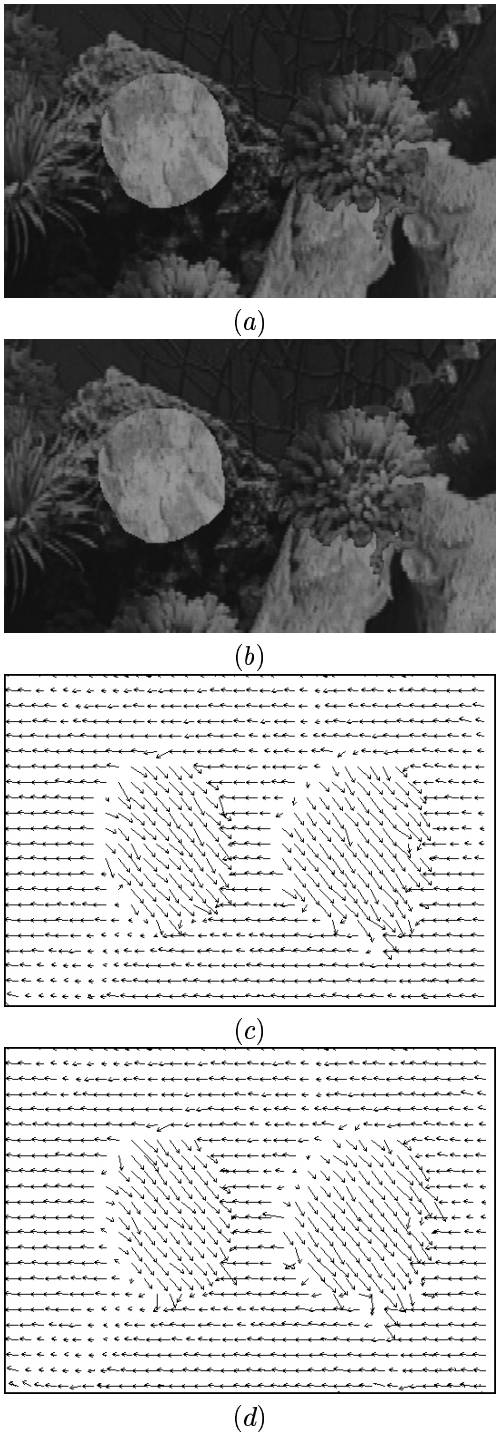

Figure 1. Experimental results, small motions: a) left side image, b) right side image, c) resulting disparity field with *Joint estimation*, d) resulting disparity field with independent estimation

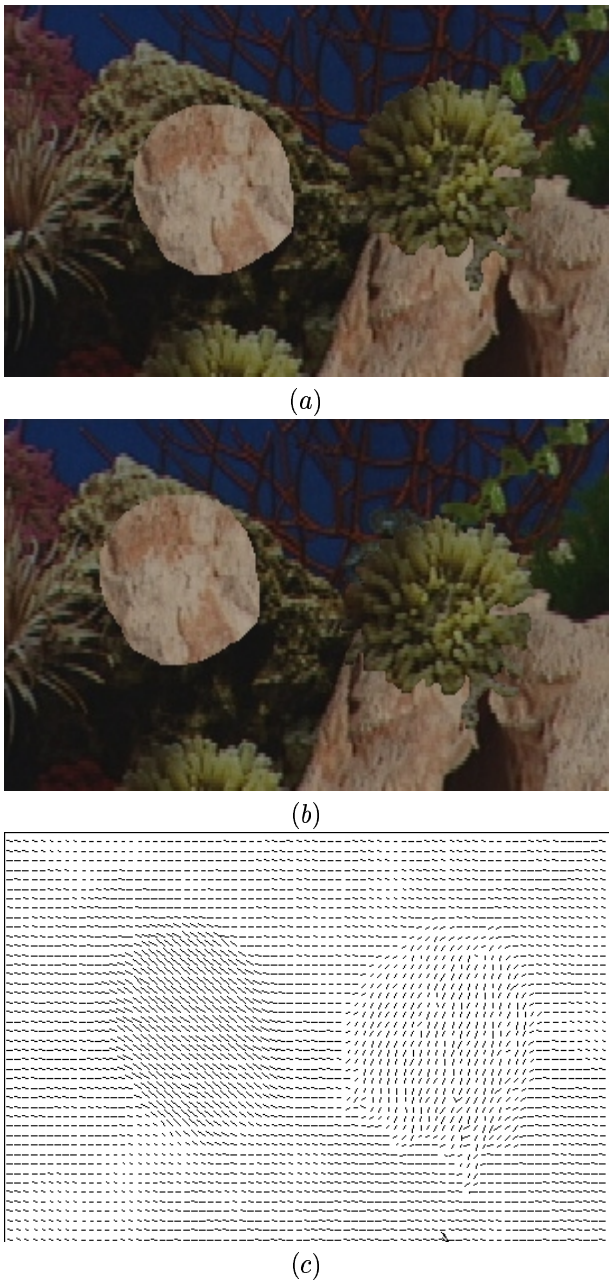

Figure 2. Experimental results, large motions: a) left side image, b) right side image, c) resulting disparity field with *Joint estimation*