Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®
Pattern Recognition Letters 25 (2004) 449–457
Pattern Recognition Letters
www.elsevier.com/locate/patrec

Optical flow using textures ☆

M.A. Arredondo *, K. Lebart, D. Lane

School of Engineering and Physical Sciences, Heriot Watt University, Mountbatten Building, Riccarton EH14 4AS Edinburgh, Scotland, UK

Received 12 June 2003; received in revised form 3 November 2003

Abstract

Motion estimation is a key problem in the analysis of image sequences. From a sequence of images we can only estimate an approximation of the image motion field called optical flow. We propose to improve optical flow estimation by including information from images of textural features. We compute the optical flow from intensity and textural images from first-order derivatives, then combine estimates using the spatial gradient as confidence measure. Experimental results with images for which the ground-truth optical flow is known show clearly that the estimate improves by including estimates from textural images. Experiments with several underwater images also show a qualitative improvement.

Keywords: Optical flow; Differential approach; Brightness constancy; Assumption; Texture; Underwater images

1. Introduction

Motion estimation is a key problem in the analysis of image sequences. The motion field is the 2D vector field which is the perspective projection on the image plane of the 3D velocity field of a moving scene. From the information available from a sequence of images (the spatial and temporal variation of the brightness pattern) it is only possible to derive an estimate of the motion field, called optical flow (OF) (Horn and Schunck, 1981). Motion information can play different roles in machine vision systems (e.g. scene motion detection, object segmentation, tracking, measurement of depth, 3D reconstruction) in different domains (e.g. television, mobile robotics, military applications, surveillance).

During the past two decades many methods for the estimation of optical flow have been proposed (Weber and Malik, 1995; Tsai et al., 1999; Ming et al., 2002; Zhang and Lu, 2000; Fleet and Jepson, 1990; Bruno and Pellerin, 2000). These methods can be classified in three groups:

- *Differential techniques:* Compute velocity from spatiotemporal derivatives of image intensity

☆ This work has been partially supported by the 5th Framework Programme of research of the European Community through the project AMASON (EVK3-CT-2001-00059).

* Corresponding author. Tel.: +44-131-4514-133; fax: +44-131-4514-155.

E-mail addresses: m.arredondo@hw.ac.uk (M.A. Arredondo), k.lebart@hw.ac.uk (K. Lebart), d.m.lane@hw.ac.uk (D. Lane).

0167-8655/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2003.11.007

or filtered versions of the image (Weber and Malik, 1995; Tsai et al., 1999).

- *Region-based matching:* Define the velocity as the shift that yields the best fit between image regions at different times (Ming et al., 2002; Zhang and Lu, 2000).
- *Frequency-based methods:* Use frequency and phase information by means of tools like the Fourier transform to estimate the velocity between frames (Fleet and Jepson, 1990; Bruno and Pellerin, 2000).

We present a differential method that includes optical flow estimates from images of textural features to improve the overall flow estimate. The rest of this article is organized as follows: Section 2 describes a method to estimate OF from first-order derivatives, how we compute textures and combine their information. In Section 3 we explain the error metrics we used. Some experimental results are on Section 4, we justify our use of the spatial gradient as a confidence measure and exhibit the complementarity of information that different textures can provide. The conclusions are on Section 5.

2. Optical flow from first-order derivatives

Let $E(\boldsymbol{x},t)$ denote the usual, continuous space-time intensity function, where $\boldsymbol{x}=(x,y)^{\top}$. If the intensity remains practically constant along a motion trajectory, we have: $\frac{\mathrm{d}E(\boldsymbol{x},t)}{\mathrm{d}t}=0$, where $\boldsymbol{x}$ varies by t according to the motion trajectory. This is a total derivative expression and denotes the rate of change of intensity along the motion trajectory. Using the chain rule of differentiation, it can be expressed as:

$$\frac{\partial E(\boldsymbol{x},t)}{\partial x}u(\boldsymbol{x},t)+\frac{\partial E(\boldsymbol{x},t)}{\partial y}v(\boldsymbol{x},t)+\frac{\partial E(\boldsymbol{x},t)}{\partial t}=0 \quad (1)$$

where $u(\boldsymbol{x},t)=\mathrm{d}x/\mathrm{d}t$ and $v(\boldsymbol{x},t)=\mathrm{d}y/\mathrm{d}t$ denote the components of the image velocity vector in terms of the continuous image coordinates, and the partial spatial derivatives of the image brightness are the components of the spatial gradient ∇E. Expression (1) can be rewritten as the image brightness constancy equation:

$$(\nabla E)^{\top}\boldsymbol{v}+E_{\mathrm{t}}=0 \quad (2)$$

where E_{t} denotes partial differentiation with respect to time. Eq. (2) provides one linear equation for two unknown components of the velocity vector (this is known as the aperture problem); hence, further constraints are necessary to solve for $\boldsymbol{v}$. Perhaps the simplest constraint is to assume that the motion is the same on a small spatial neighborhood: the optical flow can be estimated (Lucas and Kanade, 1981) within a patch Q (of size $N\times N$) as the vector, $\boldsymbol{v}$, that minimizes:

$$\Psi[\boldsymbol{v}]=\sum_{\boldsymbol{p}_i\in Q}[(\nabla E)^{\top}\boldsymbol{v}+E_{\mathrm{t}}]^2 \quad (3)$$

The solution to this least squares problem is:

$$\boldsymbol{v}=(A^{\top}A)^{-1}A^{\top}\boldsymbol{b} \quad (4)$$

where the ith row of the $N^2\times 2$ matrix A is the spatial image gradient evaluated at point $\boldsymbol{p}_i$, $\boldsymbol{b}$ is the N^2-dimensional vector of the partial temporal derivatives of the image brightness evaluated at $\boldsymbol{p}_1,\boldsymbol{p}_2,\ldots,\boldsymbol{p}_{N^2}$, after a sign change, and $\boldsymbol{v}$ is the optical flow at the center of patch Q.

Although the gradient equation was originally applied to the brightness function, we can apply the equation to other functions, assuming that Eq. (2) holds true at all or most pixels. Mitiche et al. (1987) identified other sources of such functions:

(1) *Multispectral images:* A signal recorded in several bands of the electromagnetic spectrum. For example, Markandey and Flinchbaugh (1990) and Barron and Klette (2002) estimated optical flow from color images.
(2) *Operators:* Spatial operators are applied to the original image to obtain new images (Bruno and Pellerin, 2000; Weber and Malik, 1995). The operators may compute properties such as local variance, contrast, entropy, etc.
(3) *Constraints on image motion:* Constraints are introduced on the kind of motion that causes the temporal changes. For example, the optical flow is smooth, meaning that neighboring points have similar velocities (Horn and Schunck, 1981; Arredondo et al., 2003).

We propose an approach that includes information from textural images to improve optical flow estimates. The textural images that we use are estimated using matrices that are designed to act as matched filters for certain types of quasiperiodic variations commonly found in textured images (Hsiao and Sawchuk, 1989). Hence, they contain information of the motion of prominent features. We would expect two advantages of estimating the optical flow on textural images: first, texture properties may vary less with illumination changes. If there is a strong illumination change, the optical flow equation does not hold for intensity. However, as the textures extract prominent features from the image, we would expect these features to remain present even after an illumination change, [1] and therefore Expression (2) remains true. The second advantage of applying the optical flow equation to textural images is that, by estimating the motion of a pixel from several images, we address the aperture problem as we can build an overconstrained system to solve for the two unknowns. Our motivation is to estimate accurate OF from underwater images, where illumination changes can be abrupt.

Previous researchers have used Gabor-energy filters (Bruno and Pellerin, 2000) and spatiotemporal filters based on derivatives of normalized Gaussian functions (Weber and Malik, 1995) as operators to obtain new images and estimate more accurate OF. In this article we use nine filters constructed from three basic vectors: $[1,2,1]^\top$, $[-1,0,1]^\top$, $[-1,2,-1]^\top$ as suggested by Laws (1980) (see Theodoridis and Koutroumbas, 1998):

$$\begin{bmatrix}1&2&1\\2&4&2\\1&2&1\end{bmatrix}\begin{bmatrix}-1&0&1\\-2&0&2\\-1&0&1\end{bmatrix}\begin{bmatrix}-1&2&-1\\-2&4&-2\\-1&2&-1\end{bmatrix}$$

$$\begin{bmatrix}-1&-2&-1\\0&0&0\\1&2&1\end{bmatrix}\begin{bmatrix}1&0&-1\\0&0&0\\-1&0&1\end{bmatrix}\begin{bmatrix}1&-2&1\\0&0&0\\-1&2&-1\end{bmatrix}$$

$$\begin{bmatrix}-1&-2&-1\\2&4&2\\-1&-2&-1\end{bmatrix}\begin{bmatrix}1&0&-1\\-2&0&2\\1&0&-1\end{bmatrix}\begin{bmatrix}1&-2&1\\-2&4&-2\\1&-2&1\end{bmatrix}$$

We generate the textural images by convolving the intensity image with the constructed filters and then estimating the standard deviation on local patches of the convolved images (Theodoridis and Koutroumbas, 1998). We have chosen to use this method to estimate the textural images because its computational simplicity and because, as we will see in Section 4.2, the information they provide is complementary.

To include information from the textures we estimate the optical flow in the intensity and textural images independently; if we estimate all nine textures we obtain $\boldsymbol{v}_i = (u_i, v_i)$ where $i = 1, 2, \ldots, 10$. Then we add these estimates, weighting each one according to the strength of the gradient $\nabla I_i = [\, G_{x,i} \quad G_{y,i} \,]^\top$ in the neighborhood Q used to estimate the flow:

$$\hat{u} = \sum_{i=1\ldots10} \frac{G_{x,i}}{\sum_{n=1\ldots10} G_{x,n}} u_i \tag{5}$$

$$\hat{v} = \sum_{i=1\ldots10} \frac{G_{y,i}}{\sum_{n=1\ldots10} G_{y,n}} v_i \tag{6}$$

3. Evaluation protocol

We compared the results of estimating the optical flow using images only and using images and textural images.

3.1. Error measurement

For the images where the true optical flow is known we use two different error metrics. We follow other authors (Fleet and Jepson, 1990; Barron et al., 1994) and measure the angular deviation between the estimated velocity $\boldsymbol{v}_e$ and the correct one $\boldsymbol{v}_c$. Let the velocities $\boldsymbol{v} = (u, v)^\top$ be represented as 3D direction vectors, $\boldsymbol{v} = \frac{1}{\sqrt{u^2+v^2+1}}(u, v, 1)^\top$. The angular error is defined as

$$\psi_E = \arccos(\boldsymbol{v}_c \cdot \boldsymbol{v}_e)$$

[1] Assuming that the illumination change is homogeneous at least in the region where the textural feature is estimated. This is a valid assumption, as the matrices to estimate the textural features are of size 3×3 pixels.

Table 1
Angular error on synthetic images

	Sinusoid2		Yosemite		Translating tree	
	Average error	Standard deviation	Average error	Standard deviation	Average error	Standard deviation
Image	6.42	2.83	24.17	26.01	4.48	4.83
Tex 1	2.18	2.64	18.95	21.39	5.45	8.21
Tex 2	85.45	36.65	29.02	28.23	7.21	11.85
Tex 3	69.01	37.75	35.07	29.31	16.33	23.10
Tex 4	74.10	38.22	26.28	28.82	4.57	6.52
Tex 5	81.38	36.20	34.88	32.10	7.82	11.45
Tex 6	77.68	36.13	50.58	33.62	15.44	20.15
Tex 7	81.20	36.89	34.51	33.72	7.55	10.66
Tex 8	82.14	35.01	51.28	37.99	11.02	16.17
Tex 9	79.67	37.80	67.57	39.19	18.79	23.48

The second error metric is the absolute magnitude of the differences between the components of the estimated velocity (u_e, v_e) and the correct one (u_c, v_c).

$$E_x = |u_c - u_e| \tag{7}$$

$$E_y = |v_c - v_e| \tag{8}$$

3.2. The images

The images used for the experiments are from the well known, ground-truthed, synthetic sequences: *Sinusoid2*, *Yosemite*, and *Translating tree* used by several other authors (Weber and Malik, 1995; Fleet and Jepson, 1990; Barron et al., 1994) to benchmark the performance of algorithms. [2] We also used images from underwater sequences to test the proposed approach on real images. Image derivatives were estimated using the approach proposed by Horn (1986), a Gaussian prefilter with a standard deviation of 1.5 pixels was used.

4. Experimental study

4.1. Suitability of textures

We want to show that at some points textural images can provide better OF estimates than those computed from intensity images alone. We estimated the OF on the intensity and textural images and calculated the angular error for the sequences where the true optical flow was known (see Table 1). The rather larger angular errors on some of the textural images are because while the estimate of one of the velocity components improves, the estimate of the other component is not as good because of low spatial gradient. Table 2 shows the percentage of points on each textural image that have a smaller angular error than their corresponding estimate on the intensity image. As we can see, textures can help us obtain better velocity estimates. In some cases the number of points that are better estimated on the textures is over 50% of the number of points on the image.

Table 2
Percentage of estimated points on each texture where the velocity has a smaller angular error than their corresponding estimate in the intensity image

	Sinusoid2 (%)	Yosemite (%)	Translating tree (%)
Tex 1	91.2	70.6	46.1
Tex 2	0.3	55.0	38.5
Tex 3	1.3	37.3	18.5
Tex 4	1.2	57.2	55.4
Tex 5	0.2	41.9	37.3
Tex 6	0.2	27.2	16.3
Tex 7	0.3	34.0	39.3
Tex 8	0.3	28.0	27.9
Tex 9	0.3	19.3	12.1

[2] The complete sequences can be obtained by anonymous ftp from `ftp.csd.uwo.ca` in the directory `/pub/vision`.

4.2. Complementarity of textures

The reason for using several textural images to improve the OF estimate is because each texture extracts different properties from the image. Thus, each texture will provide best estimates for different regions of the image. Fig. 1 exhibits this complementarity by showing in which intensity or textural image the optical flow is best estimated. After estimating the OF on the intensity and textural images, we calculated the angular error of the estimates at every point and determined on which (intensity or textural) image the error was smallest. On Fig. 1 white pixels show the intensity/textural image where the optical flow was best estimated for the Yosemite sequence. It is worth to note that in the three sequences for which the true OF is known, over 70% of the points are best estimated in textures 1, 2 and 4.

4.3. Relation between gradient and error

Verri and Poggio (1989) demonstrated that the bigger the spatial image gradient the more reliable the estimate of the motion will be. To verify this experimentally in the sequences we considered, we plotted spatial gradient against the error magnitude. Fig. 2 shows scatter plots of the spatial gradient on the y direction against the error magnitude for the Yosemite sequence. We can observe how estimates with higher error correspond to points with smaller spatial gradient. In the right plot, note how by considering the spatial gradient of a point over intensity and textural images the

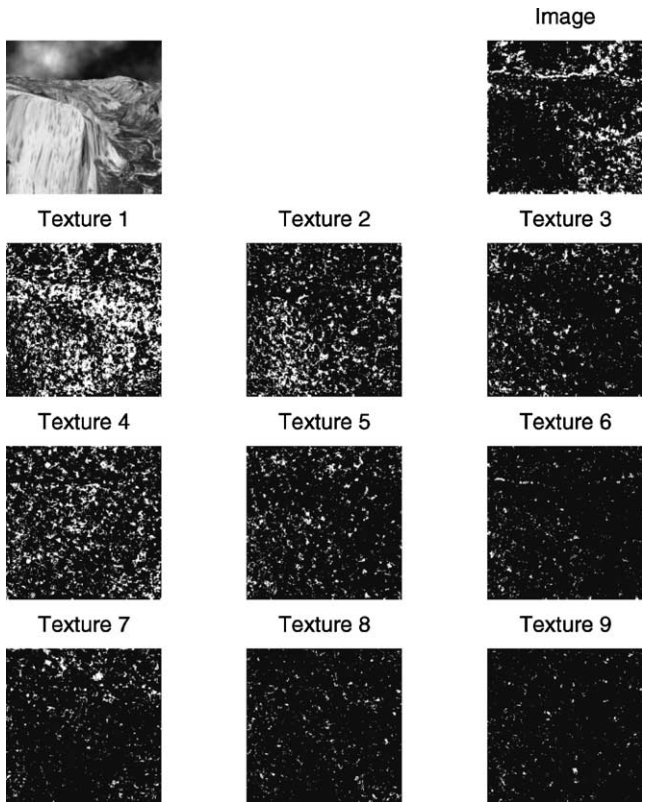

Fig. 1. On the top-left there is a frame from the Yosemite sequence. For the other images, pixels are white if the angular error was the smallest on that intensity/textural image.

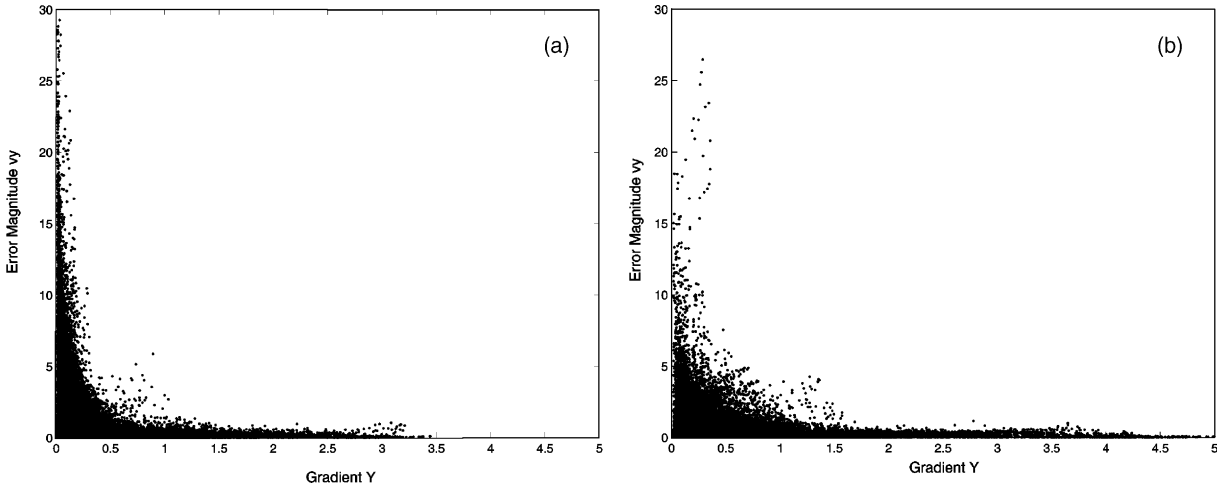

Fig. 2. (a) Scatter plot of spatial gradient in the y direction against error magnitude for the original images of the Yosemite sequence. (b) Scatter plot of the highest gradient (y direction) of a point over intensity and textural images against error magnitude.

range of error was reduced and how there are more points with higher spatial gradient and small error. The plots also show how in practice the relation spatial gradient error does not always hold: we have points with low gradient and small error and points with high gradient and a considerable error.

4.4. Optical flow results

After observing the improvement each texture can bring, we estimated the optical flow including information from textures 1, 2 and 4 (the filters used to estimate these textures are equivalent to a spot detector and the sobel operator). We chose to combine the information from these textures because of the quantitative and qualitative results of their optical flow estimates. In Table 3 we compare the mean magnitude of the error estimates using the intensity image only versus using intensity and textural images. We can observe how by including the information from the textures we obtain more reliable estimates. In the case of the Yosemite sequence the mean error decreases considerably.

In Table 4 we compare the mean and standard deviation of the angular error using only the intensity image versus using intensity and textural images. We can observe again how the error is reduced by including information from textural images.

Table 3
Error magnitude using only intensity images and using intensity plus texture images

	Intensity		Int + Txtrs	
	Average error	Standard deviation	Average error	Standard deviation
Sinusoid2 Vx	0.15	0.12	0.10	0.09
Sinusoid2 Vy	0.14	0.11	0.09	0.07
Yosemite Vx	0.79	1.06	0.40	0.52
Yosemite Vy	1.26	2.25	0.47	0.72
Translating tree Vx	0.23	0.21	0.20	0.24
Translating tree Vy	0.13	0.22	0.11	0.20

Table 4
Angular error using only intensity images and using intensity plus texture images

	Intensity		Int + Txtrs	
	Average error	Standard deviation	Average error	Standard deviation
Yosemite	24.17	26.01	12.87	15.87
Translating tree	4.48	4.83	4.24	7.29
Sinusoid2	6.42	2.83	4.07	2.76

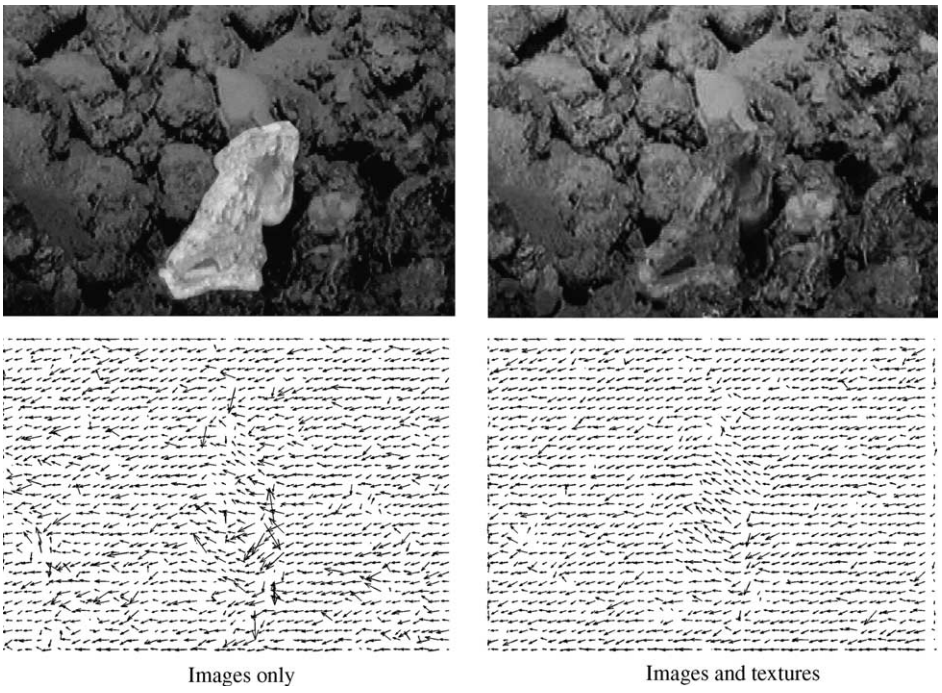

Fig. 3. Results of estimating the optical flow on underwater images. Images on the top are from the original sequence; the contrast of the top-left image has been modified to highlight the position of an octopus. In the bottom-right image, we can see how the direction of the flow in the octopus' region is slightly different, as it is lifting its head.

In our experiments our aim was to investigate if the inclusion of information from textures would improve the optical flow estimates. Further reduction of the error can be achieved; for example, authors like Weber and Malik (1995) reduce the error at the cost of a high number of filters (30) being used and a larger (10 frames) temporal support (suitable for parallel implementations but expensive otherwise) or obtaining a lower density of estimates (Fleet and Jepson, 1990). Barron et al. (1994) pointed that results for methods like Weber and Malik (1995) and Fleet and Jepson (1990) are expected to be good only when the input frequencies match those in the pass-band to which the filters are tuned.

Figs. 3 and 4 show the results of the proposed approach over sequences of underwater images. Since there is no ground-truth optical flow for these sequences, we can only evaluate the results qualitatively. On the results we can observe a smoother OF with less outliers.

5. Conclusions

We have shown that optical flow estimates can be improved by including information from textures. We have shown how OF is best estimated at some points by applying the image brightness constancy equation on textural images. In our experiments, a high percentage of the points of the images proved to be best estimated on textures. Since each texture extracts different characteristics from the image, the information they provide is complementary; by combining estimates from different textures we can improve estimates over the image. Using only three textures and two frames from the sequences we obtained an improvement on the overall OF without reducing the density of estimates. Empirically we have verified that points with high spatial image gradient are the locations where optical flow estimates are more reliable; we have used this property to combine OF estimates from the image and textures.

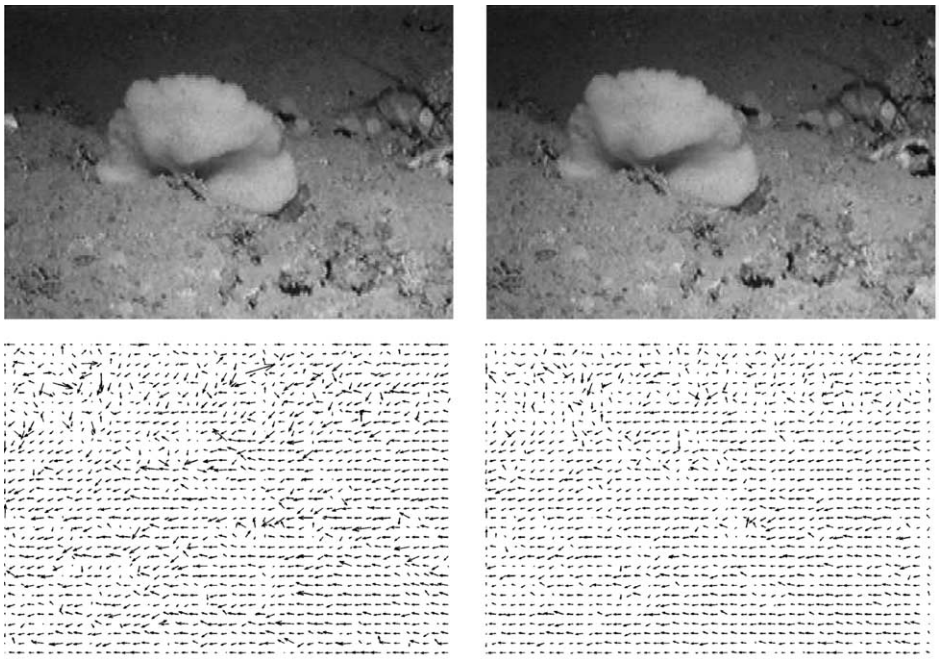

Fig. 4. Results of estimating the optical flow on underwater images. Images on the top show how the scene is moving to the left.

Acknowledgements

We would like to thank Emanuele Trucco for many useful discussions and comments. We also thank the comments of the referees to improve the quality of this paper.

References

Arredondo, M., Lebart, K., Lane, D.M., Serrurier, A., 2003. Recovery of 3D information from underwater images. In: 13th Int. Symp. on Unmanned Untethered Submersible Technol. (UUST).

Barron, J.L., Fleet, D.J., Beauchemin, S.S., Burkitt, T.A., 1994. Performance of optical flow techniques. IJCV 12 (1), 43–77.

Barron, J., Klette, R., 2002. Quantitative color optical flow. In: Proc. Int. Conf. on Pattern Recognition, vol. 4, pp. 251–255.

Bruno, E., Pellerin, D., 2000. Robust motion estimation using spatial gabor filters. In: X European Conf. Signal Process.

Fleet, D.J., Jepson, A.D., 1990. Computation of component image velocity from local phase information. IJCV 5 (1), 77–104.

Horn, B.K.P., 1986. Robot Vision, The MIT Electrical Engineering and Computing Series. MIT Press, Cambridge, Mass. USA.

Horn, B.K.P., Schunck, B.G., 1981. Determining optical flow. A.I. 17, 185–203.

Hsiao, J.Y., Sawchuk, A.A., 1989. Supervised textured image segmentation using feature smoothing and probabilistic relaxation technique. IEEE Trans. Pattern Anal. Machine Intell. 11 (12), 1279–1292.

Laws, K.I., 1980. Rapid texture identification. Proc. SPIE 238, 376–380.

Lucas, B.D., Kanade, T., 1981. An iterative image registration technique with an application to stereo vision. In: 7th Int. Joint Conf. on Artificial Intelligence, pp. 674–679.

Ming, Y., Haralick, R.M., Shapiro, L.G., 2002. Estimating optical flow using a global matching formulation and graduated optimization. In: Proc. Int. Conf. on Image Process., pp. 289–292.

Markandey, V., Flinchbaugh, B.E., 1990. Multispectral constraints for optical flow computation. In: Proc. 3rd Int. Conf. on Computer Vision, pp. 38–41.

Mitiche, A., Wang, Y.F., Aggarwal, J.K., 1987. Experiments in computing optical flow with the gradient-based multiconstraint method. Pattern Recognition Lett. 20 (2), 173–179.

Theodoridis, S., Koutroumbas, K., 1998. Pattern Recognition Lett. Elsevier Science and Technology.

Tsai, C.J., Galatsanos, N.P., Katsaggelos, A.K., 1999. Optical flow estimation from noisy data using differential techniques. In: Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Process., vol. 6, pp. 3393–3396.

Verri, A., Poggio, T., 1989. Motion field and optical flow: Qualitative properties. IEEE Trans. Pattern Anal. Machine Intell. 11 (5), 490–498.

Weber, J., Malik, J., 1995. Robust computation of optical-flow in a multiscale differential framework. IJCV 14 (1), 67–81.

Zhang, D., Lu, G., 2000. An edge and color oriented optical flow estimation using block matching. In: Proc. 5th Int. Conf. on Signal Process., pp. 1026–1032.