In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03), Hong Kong, April 2003, vol. **III**, pp. 157–160.

LINEAR AND REGULARIZED SOLUTIONS FOR MULTIPLE MOTIONS

Ingo Stuke, Til Aach

Institute for Signal Processing
University of Lübeck
Ratzeburger Allee 160, 23538 Lübeck, Germany

*Cicero Mota, Erhardt Barth**

Institute for Applied Physics
J. W. Goethe-University Frankfurt
*Institute for Neuro- and Bioinformatics
University of Lübeck

ABSTRACT

We extend a novel framework for the estimation of multiple transparent motions to include regularization. We use mixed-motion parameters to obtain linear Euler-Lagrange equations with a regularization term. The equations are solved iteratively for the mixed-motion parameters based on an update rule that is similar to the case of only one motion. The motion parameters are then obtained as the roots of a complex polynomial of a degree that is equal to the number of overlaid motions. An experimental error analysis is performed and reported.

1. INTRODUCTION

This paper addresses the problem of estimating multiple transparent motions that can occur in computer-vision applications, e.g. in case of semi-transparencies and occlusions, and also in medical x-ray projections imaging, when different layers of tissue move independently. An overview of the problem of multiple motions has been given in [2]. To our knowledge, the problem of two motions has been first solved in [5] by the use of spatio-temporal Gabor filters and fourth-order moments derived from these filters. An alternative solution that is also based on the frequency domain is given in [6], where a nonlinear system of four equations is solved to estimate the phase change and from there two transparent motions. A recent analysis of the spectral properties of two motions can be found in [8]. In general, frequency-based methods suffer from requiring large local windows. Others have introduced the useful and intuitive notion of 'layers' [7]. As an important extension of the methods mentioned above, we have provided analytic solutions for up to four transparent motions [4]. Our approach also delivers numerical solutions for more than four motions. Here we extend the solution to include regularization.

*Work supported in parts by the *Deutsche Forschungsgemeinschaft* under Ba 1176/7-1.

2. THE CASE OF ONLY ONE MOTION

We start by recalling a classical optical-flow method. We consider image sequences defined by intensity $f(x, y, t)$. The classical constant-brightness constraint for the motion vector $\boldsymbol{v} = (v_x, v_y)$ is

$$\alpha(\boldsymbol{v})f = 0 \tag{1}$$

where $\alpha(\boldsymbol{v}) = v_x \frac{\partial}{\partial x} + v_y \frac{\partial}{\partial y} + \frac{\partial}{\partial t}$ is the derivative operator along $\mathbf{V} = (v_x, v_y, 1)$. The usefulness of this notation will become clear later on.

We thus have only one equation for two unknown components of the motion vector. To solve this problem, in [3] a regularization term was used that should minimize changes over space in the motion vector. This leads to the following functional, which needs to be minimized:

$$\iint \big(\alpha(\mathbf{v})f\big)^2 + \lambda^2\Big((\partial_x v_x)^2 + (\partial_y v_x)^2 + (\partial_x v_y)^2 + (\partial_y v_y)^2\Big)\mathbf{d}\Omega \tag{2}$$

Ω is the support region, here the whole image plane, and λ the regularization parameter. Based on the calculus of variations the following two Euler-Lagrange equations are obtained:

$$\begin{aligned} f_x^2 v_x + f_x f_y v_y + f_x f_t &= \lambda^2 \triangle v_x \\ f_x f_y v_x + f_y^2 v_y + f_y f_t &= \lambda^2 \triangle v_y, \end{aligned} \tag{3}$$

where Δ is the Laplace operator.

Note that the above system is linear in the motion components v_x and v_y. Using $\Delta v = \hat{v} - v$, where $\hat{v}$ is the weighted average over the eight direct neighbours of v, and solving the system (3) for v_x and v_y we obtain

$$\begin{aligned} v_x^{l+1} &= \hat{v}_x^l - f_x \frac{P}{D} \\ v_y^{l+1} &= \hat{v}_y^l - f_y \frac{P}{D} \end{aligned} \tag{4}$$

with

$$P = f_x \hat{v}_x^l + f_y \hat{v}_y^l + f_t$$
$$D = \lambda^2 + f_x^2 + f_y^2 \quad (5)$$

where l denotes the iteration step.

Due to the large size and the sparseness of the matrix involved, the system needs to be solved iteratively for example by using the Gauss-Seidel method.

3. THE CASE OF TWO MOTIONS

Here we extend the above method to the case of two transparent motions. We shall see that the resulting update rule is very similar to the case of one motion, the difference being that we have to deal with a higher number of equations.

3.1. Basic equations

An image sequence $f(x, y, t)$ with two transparent motions $\mathbf{v}$ and $\mathbf{u}$ is described as:

$$f(x, y, t) = f_1(x - v_x t, y - v_y t) + f_2(x - u_x t, y - u_y t). \quad (6)$$

The task is to determine $\mathbf{v} = (v_x, v_y)^T$ and $\mathbf{u} = (u_x, u_y)^T$ given f. To do so, we use the optical-flow equation introduced by Shizawa and Mase [5]:

$$\alpha(\mathbf{u})\alpha(\mathbf{v})f = 0 \quad (7)$$

Note that the equation involves the concatenated directional derivatives along $\mathbf{u}$ and $\mathbf{v}$. After expanding the above equation we obtain the following expression:

$$\alpha(\mathbf{u})\alpha(\mathbf{v})f = f_{xx} u_x v_x + f_{yy} u_y v_y + f_{xy}(u_x v_y + u_y v_x) + f_{xt}(u_x + v_x) + f_{yt}(u_y + v_y) + f_{tt} = 0 \quad (8)$$

As in [4] we use the following notation:

$$c_{xx} = u_x v_x \qquad c_{yy} = u_y v_y$$
$$c_{xy} = u_x v_y + u_y v_x \qquad c_{xt} = u_x + v_x \quad (9)$$
$$c_{yt} = u_y + v_y \qquad c_{tt} = 1$$

Eq. (7) then becomes:

$$\alpha(\mathbf{u})\alpha(\mathbf{v})f = f_{xx} c_{xx} + f_{yy} c_{yy} + f_{xy} c_{xy} + f_{xt} c_{xt} + f_{yt} c_{yt} + f_{tt} c_{tt} = 0. \quad (10)$$

As we shall see, this notation leads to a linear formulation of the problem.

3.2. Regularization

As with one motion we still have only one equation but now the number of unknowns is five. We therefore need four more equations. We employ again the calculus of variation and define, in analogy to the method used in [3], the following regularization term:

$$N = (\partial_x c_{xx})^2 + (\partial_y c_{xx})^2 + (\partial_x c_{yy})^2 + (\partial_y c_{yy})^2 + (\partial_x c_{xy})^2 + (\partial_y c_{xy})^2 + (\partial_x c_{xt})^2 + (\partial_y c_{xt})^2 + (\partial_x c_{yt})^2 + (\partial_y c_{yt})^2. \quad (11)$$

We thus obtain the parameters c as the values that minimize the above term together with the squared optical-flow term (7), i.e.:

$$\iint (\alpha(\mathbf{u})\alpha(\mathbf{v})f)^2 + \lambda^2 N \, \mathbf{d}\Omega.$$

λ is the regularization parameter and Ω the whole image plane over which we integrate. Note that, at this stage, we work on finding the c's and not the motion vector components. This has the great advantage that we obtain an Euler-Lagrange system of differential equations that is linear! As we shall see, this would not be the case, when working directly on the motion vectors themselves. Note that if the velocities $\mathbf{u}$ and $\mathbf{v}$ are smooth, the parameters c will also be smooth. The five Euler-Lagrange equations that we obtain are the following:

$$\begin{aligned}
& f_{xx}^2 c_{xx} + f_{xx} f_{yy} c_{yy} + f_{xx} f_{xy} c_{xy} \\
& \quad + f_{xx} f_{xt} c_{xt} + f_{xx} f_{yt} c_{yt} + f_{xx} f_{tt} = \lambda^2 \triangle c_{xx} \\
& f_{yy} f_{xx} c_{xx} + f_{yy}^2 c_{yy} + f_{yy} f_{xy} c_{xy} \\
& \quad + f_{yy} f_{xt} c_{xt} + f_{yy} f_{yt} c_{yt} + f_{yy} f_{tt} = \lambda^2 \triangle c_{yy} \\
& f_{xy} f_{xx} c_{xx} + f_{xy} f_{yy} c_{yy} + f_{xy}^2 c_{xy} \\
& \quad + f_{xy} f_{xt} c_{xt} + f_{xy} f_{yt} c_{yt} + f_{xy} f_{tt} = \lambda^2 \triangle c_{xy} \\
& f_{xt} f_{xx} c_{xx} + f_{xt} f_{yy} c_{yy} + f_{xt} f_{xy} c_{xy} \\
& \quad + f_{xt}^2 c_{xt} + f_{xt} f_{yt} c_{yt} + f_{xt} f_{tt} = \lambda^2 \triangle c_{xt} \\
& f_{yt} f_{xx} c_{xx} + f_{yt} f_{yy} c_{yy} + f_{yt} f_{xy} c_{xy} \\
& \quad + f_{yt} f_{xt} c_{xt} + f_{yt}^2 c_{yt} + f_{yt} f_{tt} = \lambda^2 \triangle c_{yt}.
\end{aligned} \quad (12)$$

3.3. Linear and nonlinear formulations of the problem

As noted above, the system (12) is linear in the c's. Let us consider the following example to illustrate that this would not be the case with the motion parameters themselves. To obtain the Euler-Lagrange differential equations, one must differentiate the functional to be minimized with respect to the unknown variables. If we differentiate only the expression of the squared optical-flow equation with respect to v_x, we obtain

$$\frac{\partial}{\partial v_x}(\alpha(\mathbf{u})\alpha(\mathbf{v})f)^2 = 2\Big(f_{xx}u_xv_x + f_{xy}(u_xv_y + u_yv_x) + f_{xt}(u_x+v_x) + f_{yy}u_yv_y + f_{yt}(u_y+v_y) + f_{tt}\Big)(f_{xx}u_x + f_{xy}u_y + f_{xt}) \quad (13)$$

i.e., we obtain an equation that is nonlinear in v_x, v_y, u_x and u_y. Therefore we note that it is indeed the introduction of the mixed motion parameters c that leads to a linear formulation of the problem.

3.4. Update rule

In analogy to the case of one motion we obtain the following update rules for the system (12):

$$\begin{aligned} c_{xx}^{l+1} &= \hat{c}_{xx}^l - f_{xx}\frac{P}{D} \\ c_{yy}^{l+1} &= \hat{c}_{yy}^l - f_{yy}\frac{P}{D} \\ c_{xy}^{l+1} &= \hat{c}_{xy}^l - f_{xy}\frac{P}{D} \\ c_{xt}^{l+1} &= \hat{c}_{xt}^l - f_{xt}\frac{P}{D} \\ c_{yt}^{l+1} &= \hat{c}_{yt}^l - f_{yt}\frac{P}{D} \end{aligned} \quad (14)$$

with

$$\begin{aligned} P &= f_{xx}\hat{c}_{xx}^l + f_{yy}\hat{c}_{yy}^l + f_{xy}\hat{c}_{xy}^l + f_{xt}\hat{c}_{xt}^l + f_{yt}\hat{c}_{yt}^l + f_{tt}, \\ D &= \lambda^2 + f_{xx}^2 + f_{yy}^2 + f_{xy}^2 + f_{xt}^2 + f_{yt}^2. \end{aligned} \quad (15)$$

The iteration will deliver the values of the mixed motion parameters c. From these parameters we still need to extract the velocity vectors $\mathbf{u}$ and $\mathbf{v}$. We accomplish this by using the novel method described in [4]. The velocity vectors are thereby treated as complex numbers:

$$\mathbf{u} = u_x + iu_y,\ \mathbf{v} = v_x + iv_y.$$

These complex numbers are related to the mixed motion parameters c by the following equations:

$$\begin{aligned} \mathbf{uv} &= A_0 = c_{xx} - c_{yy} + ic_{xy} \\ \mathbf{u}+\mathbf{v} &= A_1 = c_{xt} + ic_{yt} \end{aligned} \quad (16)$$

Note that A_0 and A_1 are homogeneous symmetric functions in $\mathbf{u}$ and $\mathbf{v}$ and, by Vieta's theorem, the coefficients of the complex polynomial

$$Q(z) = (z-\mathbf{u})(z-\mathbf{v}) = z^2 - A_1z + A_0 \quad (17)$$

that has the complex roots $\mathbf{u}$ and $\mathbf{v}$. These roots can be obtained analytically (even for the case of up to four motions [4]). Thus the main steps are (i) solve the linear system for the c's (ii) find the roots of the complex polynomial $Q(z)$, and (iii) take the real parts of the z's as x and the imaginary parts as y components of the motion vectors.

3.5. Boundary treatment

We have chosen to extend the image by copying the boundary pixels into an extended margin of size one. Such, the first-order derivatives will be zero outside the boundary of the original image thus minimizing boundary effects.

4. GENERALIZATION TO MORE THAN TWO MOTIONS

In this section we will show that similar update rules can be obtained for the case of more than two motions. In case of n transparent motions, the optical-flow equation is given by

$$\alpha(\mathbf{v_1})\cdots\alpha(\mathbf{v_n})f = \sum_I c_I f_I = 0 \quad (18)$$

with the notation $I = I_1, \cdots, I_m$, $m = (n+1)(n+2)/2$. I_i are ordered sequences with elements (x, y, t). For example in Eq. (9) $I_1 = xx$, $I_4 = xt$, and $I_6 = tt$. The functional to be minimized is

$$\iint \Big(\sum_I c_I f\Big)^2 + \lambda^2 \sum_{I\setminus I_m} \Big((\partial_x c_I)^2 + (\partial_y c_I)^2\Big)\mathbf{d}\Omega \quad (19)$$

and the Euler-Lagrange differential equations are

$$\Big(\sum_I c_I f_I\Big) f_{I_i} = \lambda^2 \triangle c_{I_i},\ i = 1, \cdots, m-1. \quad (20)$$

Discretization with $\triangle c_{I_i} = \hat{c}_{I_i} - c_{I_i}$ leads to

$$\begin{aligned} \Big(\sum_{I\setminus I_m} c_I f_I\Big) f_{I_i} + \lambda^2 c_{I_i} &= \lambda^2 \hat{c}_{I_i} - f_{I_i} f_{I_m} \\ i &= 1, \cdots, m-1 \end{aligned} \quad (21)$$

It is now straightforward to show that equation (21) leads to the following update rule:

$$c_{I_i}^{l+1} = \hat{c}_{I_i}^l - f_{I_i}\frac{P}{D} \quad i = 1, \cdots, m-1 \quad (22)$$

with

$$\begin{aligned} P &= \sum_I \hat{c}_I^l f_I \\ D &= \lambda^2 + \sum_{I\setminus I_m} f_I^2 \end{aligned} \quad (23)$$

Since the system is positive definite, we know that even for more than two motions equation (22) is the only possible solution of the system (21). For up to four motions, the motion parameters can be obtained from the mixed-motion parameters analytically as described above since the increased number of motions will just increase the order of the polinomial. For more than four motions, the roots of the polynomial $Q(z)$ can be found by numerical methods.

5. RESULTS

5.1. Derivatives and filters

The derivatives were computed by multiplication in the frequency domain with a filter function that corresponds to blurred second order derivatives, e.g. $-\omega_x^2 \exp(-(\omega_x^2+\omega_y^2+\omega_t^2)/(2 \times 0.3^2))$ for estimating the derivative f_{xx}, with ω beeing the transform variables, such that, e.g., ω_x corresponds to x. The value of 0.3 for the scale parameter σ has been found to be optimal for minimizing the errors reported below. Obviously, this parameter can and should be used to approach the problem on multiple scales.

5.2. Error performance

We have performed an error analysis on synthetic and natural image sequences. Results on natural images show that we obtain a good segmentation of overlaid transparent objects. Results on natural images with known ground truth have not been computed yet. The results on synthetic images are summarized below. We used spatial noise patterns that were filtered to obtain a $1/\omega$ frequency spectrum that is considered to be typical for natural images. Two such patterns that move with different velocities in different directions were then superimposed. Depending on the directions and velocities the mean squared errors in estimating the velocity components are in the range of 4×10^{-6} to 1.34×10^{-3} and the standard deviations in the range of 0.002 to 0.03. With a dynamical noise that is added to the overlaid image sequences and is uniformly distributed in the range 0 to 1 percent of the maximum image-intensity value, the mean squared errors are in the range 4×10^{-4} to 3.4×10^{-3} and the standard deviations in the range 0.02 to 0.05. When the noise is in the range of 0 to 5 percent the mean squared errors range from 5.3×10^{-3} to 5.8×10^{-2} and the standard deviations from 0.07 to 0.19. The velocity vectors had the components (0,1) and (1,0), (-1,1) and (1,1), (1,0) and (1,1), (2,0) and (0,2). For all results, iterations have been stopped after 400 steps but similar results are obtained after about 100 iterations. The value of the parameter λ was 0.1.

6. DISCUSSION

We have presented a novel method for the estimation of multiple transparent motions that is based on an iterative solution of a linear system of equations. The system is obtained by introducing a regularization term for the mixed-motion parameters. The motion-vector components are then obtained from the mixed-motion parameters by solving for the roots of a complex polynomial. Alternative regularization procedures could be used, since we have succeeded to linearize the problem of determining multiple overlaid motions. By doing so we can easily incorporate regularization and deal with more that two motions. We have obtained good simulation results on synthetic sequences. However, the way we have computed the partial derivatives still needs to be optimized and filters other than derivatives could be used, as outlined in [4], to increase robustness. Possible extensions to cope with occluded motions have been proposed in [1].

7. REFERENCES

[1] E. Barth, I. Stuke, and C. Mota. Analysis of motion and curvature in image sequences. In *Proc. IEEE Southwest Symp. Image Analysis and Interpretation*, pages 206–10, Santa Fe, NM, Apr. 7–9, 2002. IEEE Computer Press.

[2] M. J. Black and P. Anandan. The robust estimation of multiple motions: parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1):75–104, Jan. 1996.

[3] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17(1–3):185–203, Aug. 1981.

[4] C. Mota, I. Stuke, and E. Barth. Analytic solutions for multiple motions. In *Proc. IEEE Int. Conf. Image Processing*, volume II, pages 917–20, Thessaloniki, Greece, Oct. 7–10, 2001. IEEE Signal Processing Soc.

[5] M. Shizawa and K. Mase. Simultaneous multiple optical flow estimation. In *IEEE Conf. Computer Vision and Pattern Recognition*, volume I, pages 274–8, Atlantic City, NJ, June 1990. IEEE Computer Press.

[6] D. Vernon. Decoupling Fourier components of dynamic image sequences: a theory of signal separation, image segmentation and optical flow estimation. In H. Burkhardt and B. Neumann, editors, *Computer Vision - ECCV'98*, volume 1407/II of *LNCS*, pages 68–85. Springer Verlag, Jan. 1998.

[7] J. Y. A. Wang and E. H. Adelson. Representing moving images with layers. *IEEE Transactions on Image Processing*, 3(5):625–38, 1994.

[8] W. Yu, K. Daniilidis, S. Beauchemin, and G. Sommer. Detection and characterization of multiple motion points. In *IEEE Conf. Computer Vision and Pattern Recognition*, volume I, pages 171–7, Fort Collins, CO, June 23–25, 1999. IEEE Computer Press.