

Bernd Jähne, Rudolf Mester, Erhardt Barth, and Hanno Scharf (Eds.)

# Complex Motion

1st International Workshop on Complex Motion  
Schloss Reisenburg, Günzburg/Germany  
October, 12 - 14, 2004  
Proceedings

# Divide-and-Conquer Strategies for Estimating Multiple Transparent Motions

Cicero Mota<sup>1</sup>, Ingo Stuke<sup>2</sup>, Til Aach<sup>2</sup>, and Erhardt Barth<sup>1</sup>

<sup>1</sup> Institute for Neuro- and Bioinformatics, University of Luebeck, Germany  
{mota, barth}@inb.uni-luebeck.de

<sup>2</sup> Institute for Signal Processing, University of Luebeck, Germany  
{aach, stuke}@isip.uni-luebeck.de

**Abstract** Transparent motions are additive or multiplicative superpositions of moving patterns and occur due to reflections, semi-transparencies, and partial occlusions. The estimation of transparent motions remained a challenging nonlinear problem. We here first linearize the problem in a way which makes it accessible to the known methods used for the estimation of single motions, such as structure tensor, regularization, block matching, Fourier methods, etc. We present the results for two motion layers but there is no limit to the number of layers. Finally, we present a way to categorize different transparent-motion patterns based on the rank of a generalized structure tensor.

## 1 Introduction

Motion estimation is a core problem in computer vision - see for example [1–3] in this volume. Most motion models used in standard applications are still rather simple and thus fail with more complex motion patterns. As a particular class of complex motion patterns, the multiple transparent motions treated here are additive or multiplicative superpositions of single moving patterns and occur due to reflections, semi-transparencies, and partial occlusions.

An algorithm for the estimation of two transparent motions was first proposed by Shizawa and Mase [4]. A layered representation of image sequences was presented in [5] and approaches based on nulling filters and velocity-tuned mechanisms have been proposed in [6,7]. A phase-based solution for the estimation of two transparent overlaid motions and the separation of the image layers was proposed by Vernon [8], and a solution for the separation of the image layers by using the constrained least-squares method was proposed in [9]. However, the estimation of transparent motions and the separation of the corresponding layers remained a challenging nonlinear problem [10]. Here we show how this problem can be naturally split into a linear and a nonlinear part. The linear part is then accessible to known methods used for the estimation of single motions, such as methods based on the structure tensor, regularization, block matching, Fourier analysis, etc. This becomes very useful since, in our approach, the nonlinear part has a closed-form solution. For simplicity, we restrict ourselves to the case of only two transparent motions at a given location in the image. This is certainly the most likely case in applications but, theoretically, our solutions are not limited by the number

of transparent layers. In fact, our results are presented such that the generalization to more than two layers is straightforward and the principle of how to generalize has been presented before [11–14]. Moreover, our approach provides confidence measures that allow for the categorization of different motion patterns and for the automatic detection of the number of moving layers. A related problem is the estimation of multiple global motions - see [15, 16] in this volume.

This paper is organized as follows. Section 2 introduces a differential constraint equation for transparent motions. The problem is then split into a linear and a nonlinear part. Two different algorithms to solve the linear part are presented and the nonlinear part is solved analytically. Section 3 introduces a Fourier domain constraint for transparent motions. The goal is to estimate the phase shifts corresponding to the motion vectors. The problem is, again, solved by splitting into linear and nonlinear parts. We show how to use the estimated phase shifts for separation of the image layers. Finally, the Fourier constraint is transformed back to the space domain to obtain a block matching constraint. Experimental results are presented for synthetic and real image sequences.

## 2 Differential Methods

Differential methods are based on the well-known constant brightness constraint equation [17], i.e., the motion field  $\mathbf{u} = (u_x, u_y)^T$  of an image sequence  $g(\mathbf{x}, t)$  is constrained by

$$u_x g_x + u_y g_y + g_t = 0 \quad (1)$$

where  $g_r = \partial g / \partial r$ ,  $r \in \{x, y, t\}$ . We write the above equation in short form as  $\alpha(\mathbf{u})g(\mathbf{x}, t) = 0$ , where  $\alpha(\mathbf{u}) = u_x \partial / \partial x + u_y \partial / \partial y + \partial / \partial t$ . Next, a similar constraint will be derived for an additive model of transparent motions.

**Constraint Equation for Transparent Motions.** We consider an additive superposition of two image sequences (layers)  $f(\mathbf{x}, t) = g_1(\mathbf{x}, t) + g_2(\mathbf{x}, t)$ . If the motion fields are sufficiently smooth to be considered ‘locally constant’, the layers can be modeled as  $g_1(\mathbf{x}, t) = \varphi_1(\mathbf{x} - t\mathbf{u})$  and  $g_2(\mathbf{x}, t) = \varphi_2(\mathbf{x} - t\mathbf{v})$  with constant motion fields  $\mathbf{u}$  and  $\mathbf{v}$  respectively. In this case, the operators  $\alpha(\mathbf{u})$  and  $\alpha(\mathbf{v})$  commute and we obtain the following constraint equation for the motion vectors [4]:

$$\alpha(\mathbf{u})\alpha(\mathbf{v})f(\mathbf{x}, t) = 0. \quad (2)$$

Since this transparent-motion constraint is nonlinear, the estimation of the motion vectors by the direct use of Equation (2) leads to non-convex problems. We overcome this difficulty by splitting the solution into a linear and a nonlinear part. By expanding Equation (2), we obtain

$$c_{xx}f_{xx} + c_{yy}f_{yy} + f_{tt} + c_{xy}f_{xy} + c_{xt}f_{xt} + c_{yt}f_{yt} = 0 \quad (3)$$

where  $f_{rs} = \partial^2 f / \partial r \partial s$ ,  $r, s \in \{x, y, t\}$ ; and

$$\begin{aligned} c_{xx} &= u_x v_x & c_{xt} &= u_x + v_x & c_{xy} &= u_x v_y + u_y v_x \\ c_{yy} &= u_y v_y & c_{yt} &= u_y + v_y \end{aligned} \quad (4)$$

are the so-called *mixed motion parameters*. In case of a multiplicative superposition  $f(\mathbf{x}, t) = g_1(\mathbf{x}, t)g_2(\mathbf{x}, t)$ , the constraint is the same except for  $f_{rs} = f \frac{\partial^2 f}{\partial r \partial s} - \frac{\partial f}{\partial r} \frac{\partial f}{\partial s}$  [18]. The introduction of the mixed motion parameters splits, in a natural way, the problem of transparent motion estimation in two parts: a linear part where we solve for the parameters  $c_{rs}$ ,  $r, s \in \{x, y, t\}$ ; and a nonlinear part where we solve Equation (4) for the motion vectors. Since Equation (3) is linear we can use different methods for the estimation of the mixed motion parameters. We will describe some of these methods in Section 2.1.

## 2.1 Linear Part: Estimation of the Mixed Motion Parameters

**The Structure Tensor.** Let time be parameterized such that Equation (1) reads

$$\tilde{u}_x g_x + \tilde{u}_y g_y + u_t g_t = 0 \quad (5)$$

with a unity parameter vector  $\mathbf{u}_e = (\tilde{u}_x, \tilde{u}_y, u_t)^T$ . If the variables  $g_x, g_y, g_t$  are independent with equal variances and  $\mathbf{u}_e$  is constant, the best fit  $\hat{\mathbf{u}}_e$ , in a *least-squares* sense, is the minimizer of the functional

$$E(\mathbf{u}_e) = \int |\mathbf{u}_e \cdot \nabla g(\mathbf{x}, t)|^2 \omega(\mathbf{x}, t) d\Omega, \quad (6)$$

where  $\Omega$  is a neighborhood of the point of interest and  $\omega(\mathbf{x}, t)$  is an weighting function. Therefore,  $\hat{\mathbf{u}}_e$  is the minimal eigenvector of the *structure tensor* [19]

$$\mathbf{J}_1 = \int \nabla g(\mathbf{x}, t) \otimes \nabla g(\mathbf{x}, t) \omega(\mathbf{x}, t) d\Omega. \quad (7)$$

The motion vector is then recovered from  $\hat{\mathbf{u}}_e / \hat{u}_t$ .

For the mixed motion parameters, we proceed in analogy and look for a unity minimizer  $\mathbf{c}_e = (c_{xx}, c_{yy}, c_{tt}, c_{xy}, c_{xt}, c_{yt})^T$  of the functional

$$E(\mathbf{c}_e) = \int |\mathbf{c}_e \cdot \mathbf{f}_{(2)}(\mathbf{x}, t)|^2 \omega(\mathbf{x}, t) d\Omega, \quad (8)$$

where  $\mathbf{f}_{(2)} = (f_{xx}, f_{yy}, f_{tt}, f_{xy}, f_{xt}, f_{yt})^T$ . Note that  $c_{tt}$  replaces 1 as the coefficient of  $f_{tt}$  in Equation (3). Again, the optimal estimator  $\hat{\mathbf{c}}_e$  is the minimal eigenvector of

$$\mathbf{J}_2 = \int \mathbf{f}_{(2)}(\mathbf{x}, t) \otimes \mathbf{f}_{(2)}(\mathbf{x}, t) \omega(\mathbf{x}, t) d\Omega \quad (9)$$

and the mixed motion parameters are recovered from  $\hat{\mathbf{c}}_e / \hat{c}_{tt}$  [11].

*Confidence Measures.* Clearly the estimator  $\hat{\mathbf{c}}_e$  ( $\hat{\mathbf{u}}_e$ ) is reliable only if the minimal eigenvalue of  $\mathbf{J}_2$  (or  $\mathbf{J}_1$ ) is small compared to the others (ideally, exactly one eigenvalue should be zero). Therefore, confidence for the quality of the estimation can be derived from the eigenvalues of  $\mathbf{J}_N$ ,  $N = 1, 2$ . however, it is useful to know the confidence before the estimation is performed. Let  $H_N, K_N, S_N$  represent the trace, the determinant, and the sum of the central minors of  $\mathbf{J}_N$  respectively. These numbers scale as  $K^{1/m} \leq (S/m)^{1/(m-1)} \leq H/m$  (with  $m = (N+1)(N+2)/2$ ). In the ideal (noise free) case of only one zero eigenvalue, we have  $K = 0$ ,  $S \neq 0$  and in practice the above scaling relation can be used to define confidence measures [11].

Moving Pattern	rank $\mathbf{J}_1$	rank $\mathbf{J}_2$
○	0	0
	1	1
+	2	2
●	2	3
● +	3	4
● + ●	3	5
others	3	6

**Table 1.** Different motion patterns (first column) and the ranks of the generalized structure tensors for 1 and 2 motions respectively (columns 2 and 3). Bars indicate motions of 1D (straight) patterns and filled circles motions of 2D patterns - see text for further details. The shown correspondences between the different motion patterns and the ranks of the two tensors can be used to identify the different motion patterns. In general, the rank of  $\mathbf{J}_N$ ,  $N = 1, 2, \dots$  induces a natural order of complexity for patterns consisting of  $N$  additive layers [12].

*Local Categorization of the Moving Patterns.* Besides allowing for motion estimation, the structure tensor allows for a local categorization of the moving pattern  $\varphi$ : rank  $\mathbf{J}_1 = 0$  corresponds to regions with constant intensity (○) and any motion vector is admissible in this region; rank  $\mathbf{J}_1 = 1$  corresponds to the motion of a straight pattern (|), in this case admissible motion vectors are constrained by a line; other moving patterns (●) correspond to the rank  $\mathbf{J}_1 = 2$ ; and non-coherent motion like noise, appearing and disappearing objects, etc. correspond to rank  $\mathbf{J}_1 = 3$ . Remarkably, in the case of transparent motions, the categorization of the moving patterns is again accessible through the rank  $\mathbf{J}_2$ . Table 1 summarizes these correspondences. For further details see [12].

**Regularization.** Here we show how to apply a Horn-Schunck-type regularization method for the estimation of the mixed motion parameters. To emphasize the dependency on  $\mathbf{c}$ , we rewrite Equation (3) as  $\mathbf{c} \cdot \mathbf{f}_{(2)r} + f_{tt} = 0$ , where  $\mathbf{f}_{(2)r} = (f_{xx}, f_{yy}, f_{xy}, f_{xt}, f_{yt})^T$ . At a given time, we then look for a field  $\mathbf{c} = (c_{xx}, c_{yy}, c_{xy}, c_{xt}, c_{yt})^T$  that minimizes the functional

$$\int \frac{1}{\lambda^2} |\mathbf{c} \cdot \mathbf{f}_{(2)r} + f_{tt}|^2 + |\nabla \mathbf{c}|^2 d\Omega, \quad (10)$$

where  $\lambda = \lambda(\mathbf{x})$ . The Euler-Lagrange equation is

$$(\mathbf{c} \cdot \mathbf{f}_{(2)r} + f_{tt})\mathbf{f}_{(2)r} = \lambda^2 \Delta \mathbf{c} \quad (11)$$

Using the approximation  $h^2 \Delta \mathbf{c} \approx \check{\mathbf{c}} - \mathbf{c}$ , where  $h$  is a normalization constant assimilated by  $\lambda$ , and solving for  $\mathbf{c}$ , we obtain a Gauss-Seidel iteration step defined by

$$\mathbf{c}^{k+1} = \check{\mathbf{c}}^k - \frac{\check{\mathbf{c}}^k \cdot \mathbf{f}_{(2)r} + f_{tt}}{\lambda^2 + |\mathbf{f}_{(2)r}|^2} \mathbf{f}_{(2)r}. \quad (12)$$

This iteration step defined by (12) can be implemented either directly as in [20], by simple methods like *successive over-relaxation* or by more sophisticated methods like

*multi-grid relaxation*. Next, we show how to solve for the motion vectors  $\mathbf{u}$  and  $\mathbf{v}$  given  $\mathbf{c}$ .

## 2.2 Nonlinear Part: Solving for the Motion Vectors

The key to our solution is the interpretation of the motion vectors as complex numbers [11], i.e.,  $u = u_x + ju_y$ , and  $v = v_x + jv_y$  and the observation that

$$uv = c_{xx} - c_{yy} + jc_{xy} = A_0, \quad u + v = c_{xt} + jc_{yt} = A_1. \quad (13)$$

In the above equations, the last equalities are just the definitions of  $A_0$  and  $A_1$ . Hence, the motion vectors can be recovered as the roots of the complex polynomial

$$Q_2(z) = (z - u)(z - v) = z^2 - A_1z + A_0 \quad (14)$$

since the coefficients of  $Q_2(z)$  depend only on the mixed motion parameters. However, Equation (4) is a over-determined system of equations for the motion vectors. Consequently, not all possible values for the mixed motion parameters vector  $\mathbf{c}$  correspond to motion vectors. To better understand this issue, we consider Equations (2) and (3) in the Fourier domain where they become

$$(u_x\xi_x + u_y\xi_y + \xi_t)(v_x\xi_x + v_y\xi_y + \xi_t)\mathbf{F}(\xi_x, \xi_y, \xi_t) = 0 \quad (15)$$

and

$$(c_{xx}\xi_x^2 + c_{yy}\xi_y^2 + c_{tt}\xi_t^2 + c_{xy}\xi_x\xi_y + c_{xt}\xi_x\xi_t + c_{yt}\xi_y\xi_t)\mathbf{F}(\xi_x, \xi_y, \xi_t) = 0 \quad (16)$$

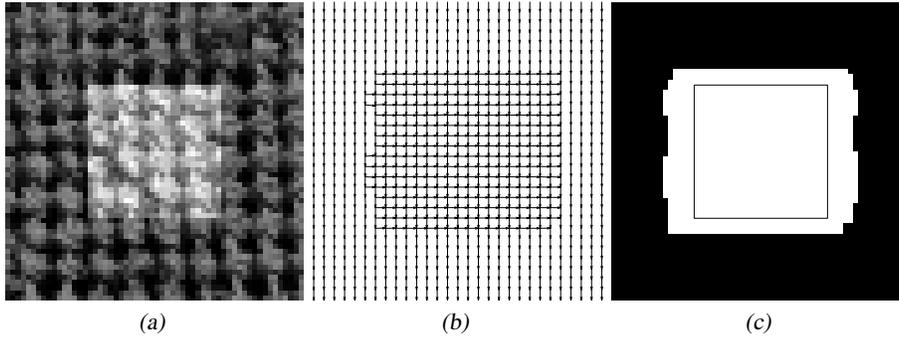
respectively.  $\mathbf{F}(\xi_x, \xi_y, \xi_t)$  represents the Fourier transform of  $\mathbf{f}(x, y, t)$ . Therefore, fitting the motion vectors  $\mathbf{u}, \mathbf{v}$  to Equation (2) is equivalent to fitting two planes to the support of  $\mathbf{F}(\xi_x, \xi_y, \xi_t)$  while fitting a parameter vector  $\mathbf{c}$  to Equation (3) is equivalent to fitting a quadric to the support of  $\mathbf{F}(\xi_x, \xi_y, \xi_t)$ . Such a quadric represents two planes if and only if its associated matrix has exactly two nonzero eigenvalues of opposite signs. Therefore, we conclude that a vector  $\mathbf{c}$  of mixed motion parameters corresponds to two motion vectors if and only if

$$\begin{vmatrix} c_{xx} & \frac{c_{xy}}{2} & \frac{c_{xt}}{2} \\ \frac{c_{xy}}{2} & c_{yy} & \frac{c_{yt}}{2} \\ \frac{c_{xt}}{2} & \frac{c_{yt}}{2} & c_{tt} \end{vmatrix} = 0 \text{ and } \left| \frac{c_{xx}}{2} \quad \frac{c_{xy}}{2} \right| + \left| \frac{c_{xx}}{2} \quad \frac{c_{xt}}{2} \right| + \left| \frac{c_{yy}}{2} \quad \frac{c_{yt}}{2} \right| < 0. \quad (17)$$

The role of the above conditions is to exclude the case when Equation (3) is valid but the Fourier transform of the motion signal is not restricted to two planes.

## 2.3 Experimental Results.

Figure 1 shows results for a synthetic image sequence with transparent motions. The algorithm first determines one motion using  $\mathbf{J}_1$  if the confidence for one motion is high. If the confidence test fails ( $H_1 > \epsilon_0$ ,  $K_1^{2/3} > \epsilon_1 S_1$ ), two motions are estimated by



**Figure 1.** Results for synthetic data: (a) the central frame of a synthetic input sequence to which dynamic noise with an SNR of 35 dB was added; (b) the estimated motion fields; (c) the segmentation defined by confidence for one (black) and two (white) motions. The background moved down and the foreground square to the right, the superposition of the two was additive. The means/standard-deviations for the components of the estimated motion fields are (0.0002/0.0029, 1.0001/0.0043) and (1.0021/0.0134, 0.0003/0.0129).

using  $\mathbf{J}_2$ . If confidence for two motions fails ( $K_2^{5/6} > \epsilon_2 S_2$ ) no motion is estimated (although this procedure could be extended for an arbitrary number of motions). The values  $\epsilon_0 = 0.001$ ,  $\epsilon_1 = 0.2$ ,  $\epsilon_2 = 0.3$  were used for the confidence parameters. We used  $[1, 0, -1]^T [1, 1, 1]$  as first order derivative filter, an integration window of  $5 \times 5 \times 5$  pixels and a weight function of  $\omega = 1$ . Second-order derivatives were obtained by applying the first order filter two times. Figure 2 shows results for more realistic image sequences. The Gauss-Seidel iteration (Equation 12) was applied to estimate the motion fields for both sequences. Gaussian derivatives with  $\sigma = 1$  and a kernel size of 7 pixels were used for first order derivatives. Again, second order derivatives were obtained by applying the first-order filter twice. The parameter  $\lambda = 1$  and 200 iterations were used. Even better results could be obtained with optimized filters - see [21] in this volume.

### 3 Extensions

#### 3.1 Phase-Based Approach

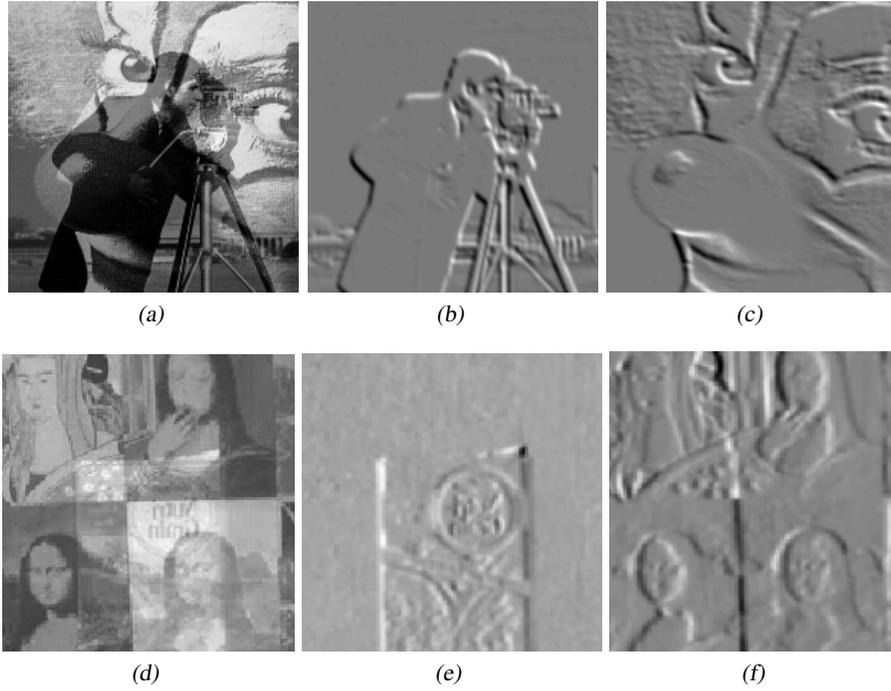
Frequency-domain based approaches to transparent motions are based on the observation that motion induces a phase shift [8, 14, 22]. For transparent motions, the multiple phase shifts lead to the following equations.

#### The Constraint Equations

$$F_{t_k}(\omega) = \phi_1^k G_1(\omega) + \phi_2^k G_2(\omega), \quad k = 0, \dots \quad (18)$$

To obtain the phase shifts from these constraints, we first simplify notation by setting  $\Phi_k = (\phi_1^k, \phi_2^k)$  and  $\mathbf{G} = (G_1, G_2)$ . We then obtain the following expressions:

$$F_{t_k} = \Phi_k \cdot \mathbf{G}, \quad k = 0, \dots \quad (19)$$



**Figure 2.** Results for natural images with synthetic, additive motions (up and to the right): (a) the central frame of the input sequence; (b) the result of applying  $\alpha(\hat{\mathbf{u}})$  to (a); (c) the result of applying  $\alpha(\hat{\mathbf{v}})$  to (a). The errors/standard-deviations of the estimated motion components are  $(0.9956/0.0106, -0.0032/0.0101)$  and  $(-0.0101/0.0129, 0.9868/0.0144)$ . Results for a real sequence: panels (d), (e) and (f) correspond to (a), (b), and (c) above. In this movie, the Mona-Lisa painting moves to the right and a right-moving box is transparently over-imposed due to reflections. The quality of the motion estimation is here demonstrated by showing how well the motion layers are separated.

Our goal now is to obtain the phase-components vector  $\Phi_1 = (\phi_1, \phi_2)$  by cancellation of the unknown Fourier-transforms vector  $\mathbf{G}$  of the image layers in the system above. First, we define the polynomial

$$p(z) = (z - \phi_1)(z - \phi_2) = z^2 + a_1z + a_2 \quad (20)$$

with unknown coefficients  $a_1 = -(\phi_1 + \phi_2)$ ,  $a_2 = \phi_1\phi_2$ . Now the phase terms  $\phi_1, \phi_2$  are the roots of  $p(z)$ , i.e.,  $p(\phi_n) = 0$ , for  $n = 1, 2$ . Second, we observe that

$$F_{t_{m+2}} + a_1F_{t_{m+1}} + a_2F_{t_m} = (\Phi_{m+2} + a_1\Phi_{m+1} + a_2\Phi_m) \cdot \mathbf{G} \quad (21)$$

$$= (\phi_1^m p(\phi_1), \phi_2^m p(\phi_2)) \cdot \mathbf{G} = 0 \quad (22)$$

and

$$F_{t_{m+2}} = -a_2F_{t_m} - a_1F_{t_{m+1}} \quad m = 0, \dots \quad (23)$$

**Solving for the Phase Shifts.** To solve for the phase shifts we apply again the strategy of splitting the problem into linear and nonlinear parts. First, we solve Equations (23) for  $a_1, a_2$  (linear part). Second, we obtain the unknown phase changes  $\phi_1, \phi_2$  as the roots of  $p(z)$  (nonlinear problem).

Since we have two unknowns, we need at least two equations for solving for  $a_1, a_2$ . Therefore we consider the first two Equations of (23), i.e.

$$\begin{pmatrix} F_{t_2} \\ F_{t_3} \end{pmatrix} = - \begin{pmatrix} F_{t_0} & F_{t_1} \\ F_{t_1} & F_{t_2} \end{pmatrix} \begin{pmatrix} a_2 \\ a_1 \end{pmatrix}. \quad (24)$$

Clearly, we can obtain  $a_1, a_2$  only if the matrix in the above equation is nonsingular. Nevertheless, in case of a singular but nonzero matrix, we can still obtain the phase shifts. To understand why, we will discuss all the cases in which  $\mathbf{A}$  is singular. First note that the matrix  $\mathbf{A}$  nicely factors as

$$\mathbf{A} = \begin{pmatrix} F_{t_0} & F_{t_1} \\ F_{t_1} & F_{t_2} \end{pmatrix} = \mathbf{B} \begin{pmatrix} G_1 & 0 \\ 0 & G_2 \end{pmatrix} \mathbf{B}^T \quad (25)$$

where

$$\mathbf{B} = \begin{pmatrix} 1 & 1 \\ \phi_1 & \phi_2 \end{pmatrix}. \quad (26)$$

Therefore,

$$\det \mathbf{A} = G_1 G_2 (\phi_1 - \phi_2)^2. \quad (27)$$

It follows that there are only two non-exclusive situations where the matrix  $\mathbf{A}$  can become singular: (i) the Fourier transform of at least one layer vanishes at the frequency  $\omega$ , and (ii) the phase shifts are equal. Therefore, we have

1. rank  $\mathbf{A} = 1$ : the possible cases are  $G_1 = 0, G_2 \neq 0$ ;  $G_1 \neq 0, G_2 = 0$  or  $\phi_1 = \phi_2, G_1 + G_2 \neq 0$  and we can compute the double phase or one of the two distinct phases from

$$F_{t_1} = F_{t_0} \phi. \quad (28)$$

2. rank  $\mathbf{A} = 0$ : in this case  $G_1 = G_2 = 0$  or  $\phi_1 = \phi_2, G_1 + G_2 = 0$  and all equations in (18) degenerate to

$$F_{t_k} = 0, \quad k = 0, \dots \quad (29)$$

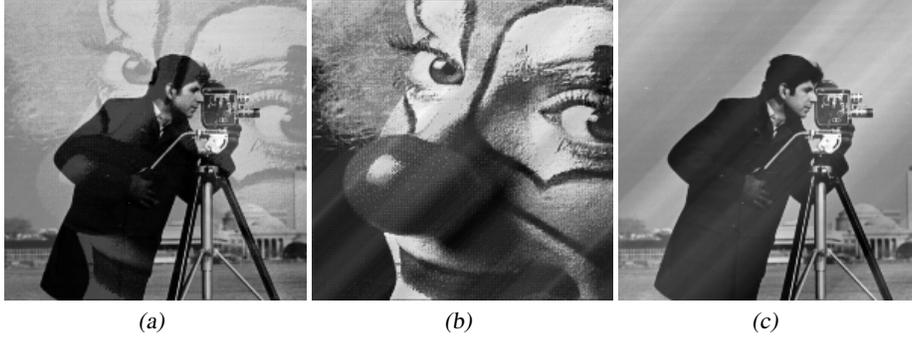
Finally, Equation (27) implies that rank  $\mathbf{A} \leq 1$  everywhere if and only if  $\phi_1 = \phi_2$  everywhere, i.e., the image sequence does not have any transparent layers.

### 3.2 Layer Separation.

Once the phase shifts are known, it is possible to obtain the transparent layers as follows:

$$\begin{pmatrix} F_{t_0} \\ F_{t_1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \phi_1 & \phi_2 \end{pmatrix} \begin{pmatrix} G_1 \\ G_2 \end{pmatrix}. \quad (30)$$

Note, however, that the separation is not possible at all frequencies. The problematic frequencies are those where two or more phase values are identical because the rank of



**Figure 3.** Results of layer separation. Shown are in (a) the same input as in Figure 2a, and in (b) and (c) the separated layers. The errors due to the still incomplete interpolation of missing frequencies are seen as oriented patterns.

the matrix  $\mathbf{B}$  is then reduced. This is an important observation because it defines the support where multiple phases can occur by the following equation:

$$\phi_1 = \phi_2 \iff e^{j(\mathbf{u}-\mathbf{v})\cdot\omega\Delta t} = 1 \iff (\mathbf{u} - \mathbf{v}) \cdot \omega = 2k\pi, \quad k = 0, \dots \quad (31)$$

On the above defined lines, the Fourier transforms at the transparent layers cannot be separated. A possible solution would be to interpolate the values on these lines from the neighboring frequency values of the separated layers.

### 3.3 Block Matching

**The Block Matching Constraint.** By transforming Equation (23) back to the space domain, we obtain the following block matching constraint equation for transparent motions [13]

$$e(f, \mathbf{x}, \mathbf{u}, \mathbf{v}) = f_0(\mathbf{x} - \mathbf{u} - \mathbf{v}) - f_1(\mathbf{x} - \mathbf{u}) - f_1(\mathbf{x} - \mathbf{v}) + f_2(\mathbf{x}) = 0. \quad (32)$$

From this constraint a number of different algorithms for the estimation of multiple motions can be derived. We here present a hierarchical algorithm which is based on a combination of statistical model discrimination and hierarchical decision making. First, a single-motion model is fitted to the sequence by exhaustive search. If the fit is poor, the single-motion hypothesis is rejected and the algorithm tries to fit two transparent motions. If the confidence for two motions is low, no motion is estimated.

**The stochastic image sequence model.** Apart from distortions and occlusions, the block matching constraint may differ from zero due to noise. Therefore additional information about the distribution of the noise can help to determine whether or not the difference between the best block matching fit and the true motion can be explained by the noise model. Different motion types lead to different noise distributions of the error signals. This can be used to select the most likely motion model.

We model the observed image intensity at each spatial location and time step as

$$f_k(\mathbf{x}) = \bar{f}_k(\mathbf{x}) + \epsilon_k(\mathbf{x}), \quad \epsilon_k(\mathbf{x}) \sim \mathcal{N}(0, \sigma^2), \quad k = 0, 1, \dots \quad (33)$$

Therefore, from Equation (32) and the noise model, we have

$$e(f, \mathbf{x}, \mathbf{u}, \mathbf{v}) = e(\bar{f}, \mathbf{x}, \mathbf{u}, \mathbf{v}) + \varepsilon(\mathbf{x}), \quad (34)$$

where  $\varepsilon(\mathbf{x}) = \epsilon_0(\mathbf{x} - \mathbf{u} - \mathbf{v}) - \epsilon_1(\mathbf{x} - \mathbf{u}) - \epsilon_1(\mathbf{x} - \mathbf{v}) - \epsilon_2(\mathbf{x})$ . Hence, in case of a perfect match of the transparent motion model, the motion-compensated residual can be modeled as

$$e(f, \mathbf{x}, \mathbf{u}, \mathbf{v}) = \varepsilon(\mathbf{x}) \sim \mathcal{N}(0, 4\sigma^2). \quad (35)$$

Consequently, the sum of squared differences over the block (denoted  $\text{BM}_2$ ) obeys the  $\chi^2$  distribution with  $|\Omega|$  degrees of freedom, i.e.,

$$\text{BM}_2(\mathbf{x}, \mathbf{u}, \mathbf{v}) = \frac{1}{4\sigma^2} \sum_{\mathbf{y} \in \Omega} e(f, \mathbf{y}, \mathbf{u}, \mathbf{v})^2 \sim \chi^2(|\Omega|), \quad (36)$$

where  $\Omega$  is the set of pixels in the block under consideration and  $|\Omega|$  is the number of elements in  $\Omega$ . A block matching algorithm can be obtained by minimization of the above expression.

If there is only one motion inside  $\Omega$ , i.e.  $f_1(\mathbf{x}) = f_0(\mathbf{x} - \mathbf{v})$ , the value of

$$\text{BM}_1(\mathbf{v}) = \frac{1}{|\Omega|} \sum_{\mathbf{x} \in \Omega} (f_1(\mathbf{x}) - f_0(\mathbf{x} - \mathbf{v}))^2 \quad (37)$$

will be small for the correct motion vector  $\mathbf{v}$ . If  $\Omega$  includes two motions, the value  $\text{BM}_1$  will be significantly different from zero for any single vector  $\mathbf{v}$ , because one vector cannot compensate for two motions.

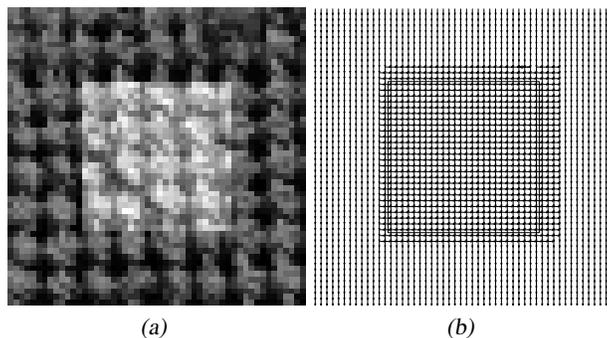
**Motion-Model Discrimination.** There are several possibilities to find the most likely motion model. To save computation time, we opt for a significance test which allows for a hierarchical estimation of the motion vectors. If we allow a percentage  $\alpha$  of misclassifications, we can derive a threshold  $T_N$  for  $\text{BM}_N$ ,  $N = 1, 2$  as follows [23]: let the null-hypothesis  $H_0$  mean that the model of  $N$  transparent motions is correct.  $T_N$  is then determined by

$$\text{prob}(\text{BM}_N > T_N | H_0) = \alpha. \quad (38)$$

$H_0$  is rejected if  $\text{BM}_N > T_N$ . The threshold can be obtained from tables of the  $\chi^2$  distribution.

### 3.4 Experimental Results

Figure 3 shows the separation of a synthetic additive overlaid image sequence. The missing phase shifts were interpolated by averaging the neighboring values. The interpolation errors are visible as oriented structures. A better interpolation method could help to reduce the errors. Figure 4 shows the results of motion estimation by block matching with a  $5 \times 5$  window. Full search has been performed to find the best match according to the confidence test described by Equation (38).



**Figure 4.** Block matching results. Shown are in (a) the central frame of the same input sequence as in 1a, and in (b) the estimated motion fields. The area corresponding to the transparent object has been depicted in (b) for better visualization.

## 4 Discussion

We have shown how to split the problem of estimating multiple transparent motions into a linear and a nonlinear part. This strategy has allowed us to extend classical but powerful algorithms for the estimation of motion to cases where standard single-motion models would fail. We have thereby reduced the difficulties in estimating multiple transparent motions to well-known difficulties in the standard, single-motion case: noisy images, aperture problem, occlusion, etc. The algorithms have been presented for two transparent motions, but are not limited to only two motions since extensions to more motions are straightforward.

The methods presented for solving the linear part of the problem have particular trade-offs. The structure-tensor method is fast and accurate but usually does not produce dense flows; the phase-based method suffers from windowing and fast Fourier transform artifacts; the regularization approach yields dense flow fields but is slow; and, finally, the block matching algorithm is very robust to noise but rather slow and does normally not yield sub-pixel accuracy.

The method proposed for solving the nonlinear part is the key which makes the overall approach so useful and lets us conclude that the difficulties in the estimation of transparent motion are, in essence, the same as for the estimation of single motions.

**Acknowledgment.** We gratefully acknowledge support from the *Deutsche Forschungsgemeinschaft* received under Ba 1176/7-2.

## References

1. Badino, H.: A robust approach for ego-motion estimation using a mobile stereo platform. In Jähne, B., Barth, E., Mester, R., Schar, H., eds.: *Complex Motion, Proceedings 1st Int. Workshop, Günzburg, Oct. 12–14. LNCS (2004)*

2. Sühling, M., Arigovindan, M., Jansen, C., Hunziker, P., Unser, M.: Myocardial motion and strain rate analysis from ultrasound sequences. In Jähne, B., Barth, E., Mester, R., Schar, H., eds.: *Complex Motion, Proceedings 1st Int. Workshop, Günzburg, Oct. 12–14. LNCS (2004)*
3. Sigal, L., Zhu, Y., Comaniciu, D., Black, M.: Graphical object models for detection and tracking. In Jähne, B., Barth, E., Mester, R., Schar, H., eds.: *Complex Motion, Proceedings 1st Int. Workshop, Günzburg, Oct. 12–14. LNCS (2004)*
4. Shizawa, M., Mase, K.: Simultaneous multiple optical flow estimation. In: *IEEE Conf. Computer Vision and Pattern Recognition. Volume I., Atlantic City, NJ, IEEE Computer Press (1990) 274–8*
5. Wang, J.Y.A., Adelson, E.H.: Representing moving images with layers. *IEEE Transactions on Image Processing* **3** (1994) 625–38
6. Darrell, T., Simoncelli, E.: Separation of transparent motion into layers using velocity-tuned mechanisms. Technical Report 244, MIT Media Laboratory (1993)
7. Darrell, T., Simoncelli, E.: Nulling filters and the separation of transparent motions. In: *IEEE Conf. Computer Vision and Pattern Recognition, New York, IEEE Computer Press (1993) 738–9*
8. Vernon, D.: Decoupling Fourier components of dynamic image sequences: a theory of signal separation, image segmentation and optical flow estimation. In Burkhardt, H., Neumann, B., eds.: *Computer Vision - ECCV'98. Volume 1407/II of LNCS., Springer Verlag (1998) 68–85*
9. Szeliski, R., Avidan, S., Anandan, P.: Layer extraction from multiple images containing reflections and transparency. In: *IEEE Conf. Computer Vision and Pattern Recognition. Volume 1., Hilton Head Island (2000) 246–253*
10. Anandan, P.: Motion estimation and optical flow computation: a personalized view. *Complex Motion, 1st Int. Workshop, Günzburg, Oct. 12–14 (2004) Invited talk.*
11. Mota, C., Stuke, I., Barth, E.: Analytic solutions for multiple motions. In: *Proc. IEEE Int. Conf. Image Processing. Volume II., Thessaloniki, Greece, IEEE Signal Processing Soc. (2001) 917–20*
12. Mota, C., Dorr, M., Stuke, I., Barth, E.: Categorization of transparent-motion patterns using the projective plane. *International Journal of Computer & Information Science* **5** (2004) 129–140
13. Stuke, I., Aach, T., Barth, E., Mota, C.: Multiple-motion estimation by block-matching using markov random fields. *International Journal of Computer & Information Science* **5** (2004) 141–152
14. Stuke, I., Aach, T., Mota, C., Barth, E.: Estimation of multiple motions: regularization and performance evaluation. In Vasudev, B., Hsing, T.R., Tescher, A.G., Ebrahimi, T., eds.: *Image and Video Communications and Processing 2003. Volume 5022 of Proceedings of SPIE. (2003) 75–86*
15. Cremers, D.: Bayesian approaches to motion-based image and video segmentation. In Jähne, B., Barth, E., Mester, R., Schar, H., eds.: *Complex Motion, Proceedings 1st Int. Workshop, Günzburg, Oct. 12–14. LNCS (2004)*
16. Nordberg, K., Vikstén, F.: Representation of motion fields and their boundaries. In Jähne, B., Barth, E., Mester, R., Schar, H., eds.: *Complex Motion, Proceedings 1st Int. Workshop, Günzburg, Oct. 12–14. LNCS (2004)*
17. Horn, B., Schunck, B.: Determining optical flow. *Artificial Intelligence* **17** (1981) 185–203
18. Langley, K., Fleet, D.J.: A logarithmic transformation applied to multiplicative transparencies and motion discontinuities. *Ophthalmology Physiology and Optics* **14** (1994) 441
19. Haußecker, H., Spies, H.: Motion. In Jähne, B., Haußecker, H., Geißler, P., eds.: *Handbook of Computer Vision and Applications. Volume 2. Academic Press (1999) 309–96*
20. Stuke, I., Aach, T., Mota, C., Barth, E.: Linear and regularized solutions for multiple motion. In: *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing. Volume III., Hong Kong, IEEE Signal Processing Soc. (2003) 157–60*

21. Scharr, H.: Optimal filters for extended optical flow. In Jähne, B., Barth, E., Mester, R., Scharr, H., eds.: *Complex Motion, Proceedings 1st Int. Workshop, Günzburg, Oct. 12–14. LNCS (2004)*
22. Fleet, D.J., Jepson, A.D.: Computation of component image velocity from local phase information. *International Journal of Computer Vision* **5** (1990) 77–104
23. Aach, T., Kaup, A.: Bayesian algorithms for adaptive change detection in image sequences using Markov random fields. *Signal Processing: Image Communication* **7** (1995) 147–160