Divide-and-Conquer Strategies for Estimating Multiple Transparent Motions

Cicero Mota[1], Ingo Stuke[2], Til Aach[2], and Erhardt Barth[1]

[1] Institute for Neuro- and Bioinformatics, University of Luebeck, Germany
{mota, barth}@inb.uni-luebeck.de
[2] Institute for Signal Processing, University of Luebeck, Germany
{aach, stuke}@isip.uni-luebeck.de

Abstract. Motion estimation is essential in a variety of image processing and computer vision tasks, like video coding, tracking, directional filtering and denoising, scene analysis, etc. 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. Theoretically, our solution does not limit the number of transparent 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 essential in a variety of image processing and computer vision tasks, like video coding, tracking, directional filtering and denoising, scene analysis, etc. Transparent motions are additive or multiplicative superpositions of 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 [10]. A layered representation of image sequences was presented in [16] and approaches based on nulling filters and velocity-tuned mechanisms have been proposed in [3, 2]. A phase-based solution for the estimation of two transparent overlaid motions and separation of the image layers was proposed by Vernon [15] and a solution for the separation of the image layers using constrained least square was proposed in [14]. However, the estimation of transparent motions remained a challenging nonlinear problem. Here we show how the problem can be naturally split into linear and nonlinear parts. The linear part is accessible to known methods used for the estimation of single motions, such as structure tensor, regularization, block matching, Fourier methods, etc. The nonlinear part has fortunately closed form solution. For simplicity, we restrict ourselves to the double transparent motion case, i.e., every point of the image sequence presents at most two motion vectors. This is certainly the most

interesting case but, theoretically, our solutions are not limited by the number of transparent layers and they are presented in such a way that is easy to derive the corresponding formulas to the general case [9, 8, 11, 12]. Even more, the solutions allow for the categorization of different motion patterns and for automatic detection of the number of moving layers. In summary, our approach for the estimation of transparent motion allows us to conclude that the difficulties in the estimation of transparent motion are, in essence, the same as for the estimation of single motion itself.

This paper is organized as follows. Section 2 introduces a differential constraint equation for transparent motion. The problem is then split into linear and nonlinear parts. Two different algorithms to solve the linear part are presented and the nonliner part is solved analiticaly. Section 3 introduces a Fourier domain constraint for transparent motion. 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 space domain to obtain a block-matching constraint. Results are presented for synthetic and real iamge sequences.

2 Differential Methods

Differentials methods are based on the well known constant brightness constraint equation [6], i.e., the motion field $\mathbf{u} = (u_x, u_y)^T$ of an image sequence $g(\mathbf{x}, t)$ is constrained by

$$u_x g_x + u_y g_y + g_t = 0 \tag{1}$$

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 an additive model for transparent motion is introduced and a similar constraint is derived.

Constraint Equation for Transparent Motion. 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 smooth enough 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 $\alpha(\mathbf{u})$, $\alpha(\mathbf{v})$ commute and we obtain a constraint equation for the motion vectors [10]

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

Since the transparent motion constraint is nonlinear, trying to estimate the motion vectors directly by using Equation (2) leads to non-convex problems. Instead, we overcome this difficult by splitting the solution into a linear and a nonlinear part. Expanding Equation (2), we obtain

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

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

$$c_{xx} = u_x v_x \qquad c_{xt} = u_x + v_x \qquad c_{xy} = u_x v_y + u_y v_x$$
$$c_{yy} = u_y v_y \qquad c_{yt} = u_y + v_y \tag{4}$$

are the so called *mixed motion parameters.* In case of a multiplicative super-position $f(\mathbf{x}, t) = g_1(\mathbf{x}, t)g_2(\mathbf{x}, t)$, the constraint is the same except for $f_{rs} = f\partial^2 f/\partial r \partial s - \partial f/\partial r\, \partial f/\partial s$ [7]. 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 look 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 such methods in Section 2.1 of this work.

2.1 Linear Part: Estimation of the Mixed Motion Parameters

The Structure Tensor. This method consists in first supposing that time is parameterized such that Equation (1) reads

$$\tilde{u}_x g_x + \tilde{u}_y g_y + u_t g_t = 0. \tag{5}$$

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

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

where Ω 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* [5]

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

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

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

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

$$\mathbf{J_2} = \int \mathbf{f}_{(2)}(\mathbf{x}, t) \otimes \mathbf{f}_{(2)}(\mathbf{x}, t)\, \omega(\mathbf{x}, t)\, \mathrm{d}\Omega \tag{9}$$

and the mixed motion parameters are recovered from $\hat{\mathbf{c}}_e/\hat{c}_{tt}$ [9].

Moving Pattern	rank $\boldsymbol{J}_1$	rank $\boldsymbol{J}_2$
○	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, 2 motions (table rows). The correspondence between the different motion patterns and the tensor ranks that can be used to estimate the confidence for a particular pattern, i.e., a proper motion model. The rank of $\boldsymbol{J}_N$, $N = 1, 2$ induces a natural order of complexity for patterns consisting of N additive layers.

Confidence Measures. Clearly the estimator $\hat{\mathbf{c}}_{\mathrm{e}}$ ($\hat{\mathbf{u}}_{\mathrm{e}}$) is reliable only if the minimal eigenvalue of $\mathbf{J}_2$ ($\mathbf{J}_1$) is small compared to the others (ideally, exactly one zero eigenvalue). Therefore, confidence on the estimators can be accessed from the eigenvalues of $\mathbf{J}_n$, $n = 1, 2$. Alternatively, it is interesting to have confidence even before the estimation be carried out. 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 $K^{1/m} \leq (S/m)^{1/(m-1)} \leq H/m$ and, for a ideal model, we had $K = 0$, $S \neq 0$. Hence, they can be used as confidence measures [9].

Local Categorization of the Moving Patterns. Besides allowing for motion estimation, the structure tensor allows for a local categorization of the moving pattern φ : rank $\boldsymbol{J}_1 = 0$ corresponds to the motion of regions with constant intensity (○) and any motion vector is admissible in this region; rank $\boldsymbol{J}_1 = 1$ 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 $\boldsymbol{J}_1 = 2$; and non-coherent motion like noise, popping up objects, etc. correspond to rank $\boldsymbol{J}_1 = 3$. Surprisingly, in the case of transparent motion, the categorization of the moving patterns is again accessible trhough the rank $\boldsymbol{J}_2$. Table 1 summarizes these correspondences. For more details see [8].

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 $\mathbf{c} \cdot \mathbf{f}_{(2)\mathrm{r}} + f_{tt} = 0$, where $\mathbf{f}_{(2)\mathrm{r}} = (f_{xx}, f_{yy}, f_{xy}, f_{xt}, f_{yt})^T$. At a given time, we the look for a field $\mathbf{c} = (c_{xx}, c_{yy}, c_{xy}, c_{xt}, c_{yt})^T$ that minimizes the functional

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

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

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

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

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

This iteration step can be implemented either directly [13], in a simple method like *successive overrelaxation* or in more sophisticated methods like *multigrid*. Next, we show how to solve for the motions vectors given $\mathbf{c}$.

2.2 Solving for the Motion Vectors.

The key to our solution is the interpretation of the motion vectors as complex numbers [9], i.e., $u = u_x + \mathrm{j}u_y$, and $v = v_x + \mathrm{j}v_y$ and the observation that

$$uv = c_{xx} - c_{yy} + \mathrm{j}c_{xy} = A_0, \quad u + v = c_{xt} + \mathrm{j}c_{yt} = A_1. \tag{13}$$

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

$$Q_2(z) = (z - u)(z - v) = z^2 - A_1 z + A_0 \tag{14}$$

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

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

and

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

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

$$\begin{vmatrix} c_{xx} & \frac{c_{xy}}{2} & \frac{c_{xt}}{2} \\ \frac{c_{xy}}{2} & c_{yy} & \frac{c_{yt}}{2} \\ \frac{c_{xt}}{2} & \frac{c_{yt}}{2} & c_{tt} \end{vmatrix} = 0 \text{ and } \begin{vmatrix} c_{xx} & \frac{c_{xy}}{2} \\ \frac{c_{xy}}{2} & c_{yy} \end{vmatrix} + \begin{vmatrix} c_{xx} & \frac{c_{xt}}{2} \\ \frac{c_{xt}}{2} & c_{tt} \end{vmatrix} + \begin{vmatrix} c_{yy} & \frac{c_{yt}}{2} \\ \frac{c_{yt}}{2} & c_{tt} \end{vmatrix} < 0. \tag{17}$$

The role of the above Equations 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 shown transparence. We first look for one motion using $\mathbf{J}_1$ if confidence fails ($H_1 > \epsilon_0$, $K_1^{2/3} > \epsilon_1 S_1$), we look for two motions using $\mathbf{J}_2$. If confidence for two motion fails ($K_2^{5/6} > \epsilon_2 S_2$)

(a)	(b)	(c)
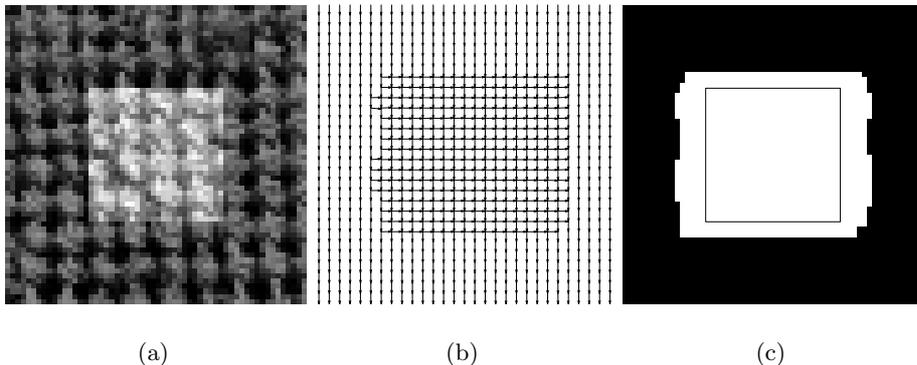
Fig. 1. Shown are: (a) the central frame of a synthetic sequence shown additive transparence (SNR 35 dB); (b) the estimated motion fields; (c) the segmentation corresponding to single and double motion fields. The mean/standard-deviation for the components of the estimated motion fields are $(0.0002/0.0029, 1.0001/0.0043)$ and $(1.0021/0.0134, 0.0003/0.0129)$.

zero is assigned to the motion vector. 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 weight function $\omega = 1$. Second order derivatives are obtained by applying the first order filter two times. Figure 2 show results for synthetic and 'real' sequences. The Gauss-Seidel iteration (Equation 12) was applied to estimate the motion fields for both sequences. Gaussian derivatives with $\sigma = 1$ and size of 7 pixel were used for first order derivatives. As before, second order derivatives were obtained by applying the first order filter twice. The parameter $\lambda = 1$ and 200 iterations were used.

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 [4, 15, 12].

The Constraint Equation For transparent motions, this translates to

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

To obtain the phase shifts, 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 for the above system :

$$F_{t_k} = \Phi_k \cdot \mathbf{G}, \qquad k = 0, \ldots \qquad (19)$$
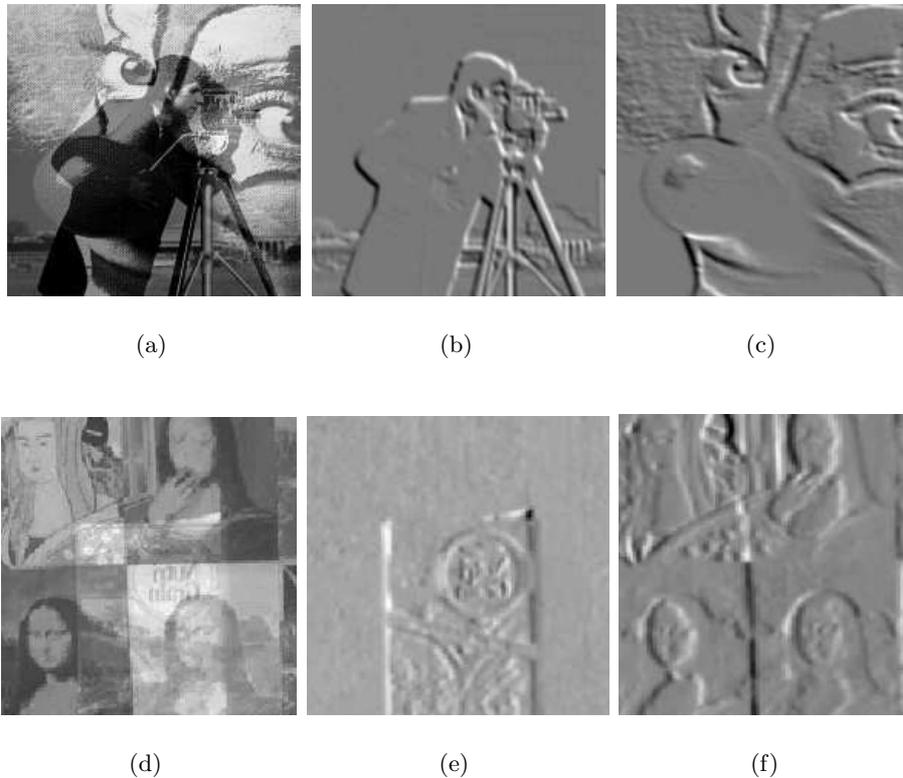
Fig. 2. Shown are: (a) the central frame of a synthetic sequence shown transparence; (b) the result of applying $\alpha(\hat{\mathbf{u}})$ to (a); (c) the result of applying $\alpha(\hat{\mathbf{v}})$ to (a); (d), (e) and (f) as in before but for a 'real' sequence. For the sequence in the first row, the error/standard-deviation of the estimated motion components after are $(0.9956/0.0106, -0.0032/0.0101)$ and $(-0.0101/0.0129, 0.9868/0.0144)$.

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

$$p(z) = (z - \phi_1)(z - \phi_2) = z^2 + a_1 z + a_2 \qquad (20)$$

with unknown coefficients $a_1 = -(\phi_1 + \phi_2)$, $a_2 = \phi_1 \phi_2$. Now the phase terms ϕ_1, ϕ_2 are the roots of $p(z)$, i.e., $p(\phi_n) = 0$, for $n = 1, 2$. Second, we observe that

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

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

and

$$F_{t_{m+2}} = -a_2 F_{t_m} - a_1 F_{t_{m+1}} \qquad m = 0, \dots \qquad (23)$$

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 ϕ_1, ϕ_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

$$\begin{pmatrix} F_{t_2} \\ F_{t_3} \end{pmatrix} = - \begin{pmatrix} F_{t_0} & F_{t_1} \\ F_{t_1} & F_{t_2} \end{pmatrix} \begin{pmatrix} a_2 \\ a_1 \end{pmatrix} \tag{24}$$

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

$$\boldsymbol{A} = \begin{pmatrix} F_{t_0} & F_{t_1} \\ F_{t_1} & F_{t_2} \end{pmatrix} = \boldsymbol{B} \begin{pmatrix} G_1 & 0 \\ 0 & G_2 \end{pmatrix} \boldsymbol{B}^T \tag{25}$$

where

$$\boldsymbol{B} = \begin{pmatrix} 1 & 1 \\ \phi_1 & \phi_2 \end{pmatrix}. \tag{26}$$

Therefore,

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

It follows that there are only two non-exclusive situations where the matrix $\boldsymbol{A}$ can become singular: The Fourier transform of at least one layer vanishes at the frequency ω; The phase shifts are equal. Therefore, we have

1. rank $\boldsymbol{A} = 1$: 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

$$F_{t_1} = F_{t_0} \phi. \tag{28}$$

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

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

Finally, Equation 27 implies that rank $\boldsymbol{A} \leq 1$ everywhere if and only $\phi_1 = \phi_2$ everywhre, i.e., the image sequence does not show transparence.

3.2 Layer Separation.

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

$$\begin{pmatrix} F_{t_0} \\ F_{t_1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \phi_1 & \phi_2 \end{pmatrix} \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} \tag{30}$$

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

(a) (b) (c)
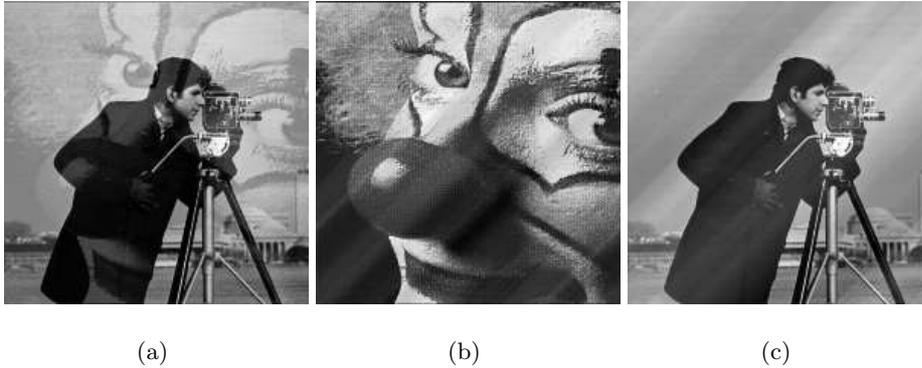
Fig. 3. Layer separation. Shown are: (a) the central frame of a synthetic image sequence shown transparence; (b) and (c) the corresponding layers. The errors due to the interpolation of missing frequencies are seen as oriented patterns.

because the rank of 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:

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

On the above defined lines, the Fourier transforms of 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. Transforming Equation (23) back to the space domain, we obtain the block-matching constraint equation for transparent motion [11]

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

From this constraint a number of different algorithms for the estimation of multiple motions could be derived. We here present a hierarchical algorithm 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.

The stochastic image sequence model. Apart from distortions and occlusions the non zero results of the block-matching constraint may be caused by noise. Additional information about the distribution of the noise hence helps to

determine whether or not the observed error signals after the block-matching process is explainable by the noise model. Different motion types lead to different noise distributions of the error signals which is helpful for selecting the most likely motion model.

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

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

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

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

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

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

Consequently, the sum BM_2 of squared differences over the block obeys the χ^2 distribution with $|\Omega|$ degrees of freedom, i.e.,

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

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

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

$$\mathrm{BM}_1(\boldsymbol{v}) = \frac{1}{|\Omega|} \sum_{\boldsymbol{x} \in \Omega} (f_1(\boldsymbol{x}) - f_0(\boldsymbol{x} - \boldsymbol{v}))^2 \tag{37}$$

will be small for the correct motion vector $\boldsymbol{v}$. On the other hand, if Ω includes two motions, the value BM_1 will tend to be far from zero for any 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 a hierarchical estimation of the motion vectors. If we allow a percentage α of misclassifications, we can derive a threshold T_N for BM_N $N = 1, 2$ as follows ([1]): let the null-hypothesis H_0 mean that the model of N transparent motion is correct. T_N is determined by

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

H_0 is rejected if $BM_N > T_N$. The threshold can be obtained from tables for the χ^2 distribution.

<div align="center">(a)   (b)</div>
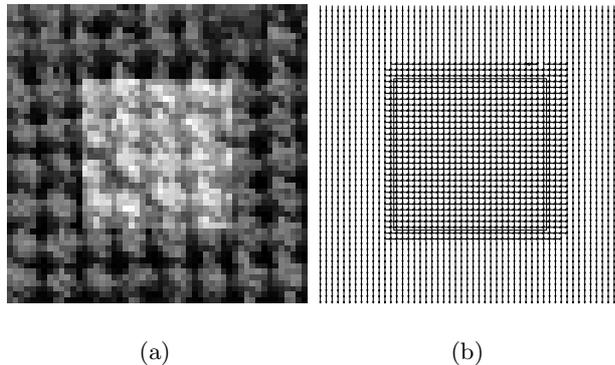
Fig. 4. Shown are: (a) the central frame of an image sequence presenting regions of single and transparent motion. (b) estimated motion fields. The area corresponding to the transparent object has been depicted in (b) for better visualization.

3.4 Experimental Results

Figure 3 shows the separation of a synthetically additive overlaid image sequence. The missing phase shifts were interpolated by averaging the neighbors values. The interpolation errors are visible as oriented structures. A better interpolation method can help to reduce the errors. Figure 4 shows the results of a block-matching search using 5×5 blocks. Full search has been performed to find the best match according to the confidence test described by Equation (38).

4 Discussion

We show how to split the problem of estimating motion in sequences shown transparencies into a linear and a nonlinear part. This strategy allow us to extend classical but powerful algorithms for the estimation of motion in such sequences were standard single motion estimation methods fail. By doing so, we reduced the difficulties in estimating transparent motion to well known difficulties in the standard single motion case: noisy images, aperture problem, occlusion, etc. The algorithms presented here are in no way limited to two transparent motions. The methods were presented in a way that is easy to derive corresponding $N-$motions formulas. The presented methods have, as always, vantages and disadvantages. The structure tensor is fast but usually does not produce dense flows; The phase based suffer under windowing and fast Fourier transform artifacts; The regularization approach has dense flows as its main advantage by is usually slow; The block-matching algorithm is very robust to noisy but slow and does not allow for sub-pixel accuracy.

References

1. T. Aach and A. Kaup. Bayesian algorithms for adaptive change detection in image sequences using Markov random fields. *Signal Processing: Image Communication*, 7(2):147–160, 1995.
2. T. Darrell and E. Simoncelli. Nulling filters and the separation of transparent motions. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 738–9, New York, June 14–17, 1993. IEEE Computer Press.
3. T. Darrell and E. Simoncelli. Separation of transparent motion into layers using velocity-tuned mechanisms. Technical Report 244, MIT Media Laboratory, 1993.
4. D. J. Fleet and A. D. Jepson. Computation of component image velocity from local phase information. *Int. J. of Computer Vision*, 5(1):77–104, 1990.
5. H. Haußecker and H. Spies. Motion. In B. Jähne, H. Haußecker, and P. Geißler, editors, *Handbook of Computer Vision and Applications*, volume 2, pages 309–96. Academic Press, 1999.
6. B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17(1–3):185–203, Aug. 1981.
7. K. Langley and D. J. Fleet. A logarithmic transformation applied to multiplicative transparencies and motion discontinuities. *Opthalmology Physiology and Optics*, 14:441, 1994.
8. C. Mota, M. Door, I. Stuke, and E. Barth. Categorization of transparent-motion patterns using the projective plane. *International Journal of Computer and Information Science*, 5(2):119–140, June 2004.
9. 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.
10. 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.
11. I. Stuke, T. Aach, E. Barth, and C. Mota. Multiple-motion estimation by block-matching using markov random fields. *International Journal of Computer and Information Science*, 5(2):141–152, June 2004.
12. I. Stuke, T. Aach, C. Mota, and E. Barth. Estimation of multiple motions: regularization and performance evaluation. In *Image and Video Communications and Processing 2003*, v. 5022 of *Proc. SPIE*, pages 75–86, May 2003.
13. I. Stuke, T. Aach, C. Mota, and E. Barth. Linear and regularized solutions for multiple motion. In *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, volume III, pages 157–60, Hong Kong, Apr. 2003. IEEE Signal Processing Soc.
14. R. Szeliski, S. Avidan, and P. Anandan. Layer extraction from multiple images containing reflections and transparency. In *IEEE Conf. Computer Vision and Pattern Recognition*, volume 1, pages 246–253, Hilton Head Island, 2000.
15. D. Vernon. Decoupling Fourier components of dynamic image sequences: a theory of signal separation, image segmentation and optical flow estimation. In *Computer Vision - ECCV'98*, volume 1407/II of *LNCS*, pages 68–85, 1998.
16. J. Y. A. Wang and E. H. Adelson. Representing moving images with layers. *IEEE Transactions on Image Processing*, 3(5):625–38, 1994.