Multi-Valued Motion Fields Estimation for Transparent Sequences with a Variational Approach

Alonso Ramirez-Manzanares, Mariano Rivera, Pierre Kornprobst, François Lauze

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Multi-Valued Motion Fields Estimation for Transparent Sequences with a Variational Approach

Alonso Ramirez-Manzanares — Mariano Rivera — Pierre Kornprobst — François Lauze

N° 5920

Mai 2006

Thème BIO

ISSN 0249-6399

Multi-Valued Motion Fields Estimation for Transparent Sequences with a Variational Approach

Alonso Ramirez-Manzanares* , Mariano Rivera† , Pierre Kornprobst‡ , François Lauze§

Thème BIO — Systèmes biologiques
Projets Odyssée

Rapport de recherche n° 5920 — Mai 2006 — 25 pages

Abstract: Most optical flow algorithms provide flow fields as single valued functions of the image sequence domains. Only a very few of them attempt to recover multiple motion vectors at given location, which is necessary when some transparent layers are moving independently. In this report we introduce a novel framework for modeling multivalued motion fields, and propose an energy minimization formulation with smoothing terms and terms implementing velocity model competition. We illustrate the capabilities of this approach on synthetic and real sequences.

Key-words: approches variationnelles, mouvement transparent, champs de vecteurs multi-valués, compétition de modéls

* Centro de Investigacion en Matematicas, Apdo Postal 402, 36000 Guanajuato Gto. Mexico
† Centro de Investigacion en Matematicas, Apdo Postal 402, 36000 Guanajuato Gto. Mexico
‡ INRIA, Odyssée Lab., 2004 route des Lucioles BP 93, 06902 Sophia Antipolis, France
§ IT University of Copenhagen, Rued Langgaardsvej 7, DK-2300 Copenhagen S, Denmark

Estimation de champs de déplacements multi-valués pour les mouvements transparents avec une approche variationnelle

Résumé : La majorité des algorithmes pour le flot optique permettent d'obtenir une estimation d'un mouvement unique en chaque position. Peu d'approches permettent en revanche d'estimer plusieurs vitesses en une même position, ce qui se produit quand des couches transparentes se déplacent de manière indépendantes. Nous introduisons dans ce rapport un nouveau cadre pour modéliser des champs de vecteurs multivalués, et nous proposons une formulation énergétique comportant des termes de régularité et de compétition entre les différentes vitesses. Les performances de cette approche sont évaluées sur un ensemble de séquences synthétiques et réelles.

Mots-clés : variational approaches, transparent motion, multi-valued motion fields, model competition

1 Introduction

There exists a very wide literature on apparent motion estimation, also called optical flow. Such a craze for optical flow is notably due by the number of applications that require some motion estimation to perform their tasks. We refer the reader to [24, 12, 1, 3] for some reviews on this topic. Although less models are proposed concerning multiple motions, it is our conviction that considering more complex stimuli will also bring some new solutions and ideas for simple optical flow estimation.

In this paper we propose a framework based on a finite sampling of the space of velocities. Having chosen a finite set of admissible velocities, our goal is to recover a coherent spatio-temporal field that encodes at each location the presence of one or more velocities from our sample set. To recover such a field, we start with local velocity measurements, and then we minimize an energy function that encodes our prior knowledge about the optical flow smoothness and the expected number of motions (relatively small, says one or two) at a particular site of the image.

The paper is organized as follows. Section 2 reviews and comments on some related works on multiple motion estimations. Then Section 3 describes the proposed framework and related notations are introduced. Section 4 states a discrete variational model to handle multiple motions, and the role of each term in the resulting energy is discussed. The performance of the approach is illustrated in Section 5, on synthetic, synthesized realistic and real sequences. We conclude and present furture work in Section 6.

2 Related Work On Multiple Motion Estimation for Transparent Sequences

Motion estimation methods rely on a form for data conservation along motion trajectories and some spatial or spatiotemporal regularity. Regularity in that context corresponds to some local smoothness assumption of the motion field. The most elementary form of data conservation and probably the most used is the Lambertian assumption, or brightness constancy, which states that intensities remain constant. Given a sequence $f(\mathbf{x},t) = f(x_1, x_2, t)$, then the conservation can be stated as the *Displayed Frame Difference Equation* (DFD)

$$f(\mathbf{x} - u, t+1) = f(\mathbf{x}, t). \tag{1}$$

or the linearization of it, the *Optical Flow Constraint Equation* (OFC)

$$\left(u^1 \frac{\partial}{\partial x_1} + u^2 \frac{\partial}{\partial x_2} + \frac{\partial}{\partial t}\right) f(\mathbf{x}, r) = (\nabla f(\mathbf{x}, t))^T \begin{pmatrix} u^1 \\ u^2 \\ 1 \end{pmatrix} = 0 \tag{2}$$

where $\nabla f = (f_{x_1}, f_{x_2}, f_t)^T$ and (u^1, u^2) are the spatial components of the velocity vector u. The gradient ∇f provides an affine constraint on the velocity space and is sometimes refereed to as a "motion constraint vector."

Although widely used, this model has a well known limited validity, intensities do not always remain constant due to, among others, changing lightning conditions, specularities and clearly it cannot cope with multiple motions, especially in the case of transparency.

Transparency can be modeled as a superposition of moving layers, a linear superposition meaning addition of layer intensities, or a generalized one [23] where intensity addition is replaced by an operation with similar formal algebraic properties such as multiplication in reflection. A simple superposition model was introduced by Burt *et al.* in [4] for the case of two motions. The observed image sequence f is assumed to come from the combination $f = P_1 \oplus P_2$ of two moving patterns P_1 and P_2 with respective motions u_1 and u_2, such that brightness constancy hold for each (P_i, u_i):

$$P_i(\mathbf{x} - u_i, t+1) = P_i(\mathbf{x}, t)$$

or

$$(\nabla P_i(\mathbf{x}, t))^T \begin{pmatrix} u_i^1 \\ u_i^2 \\ 1 \end{pmatrix} = 0.$$

In the mere case of linear superposition, i.e. when the combination operation $\oplus$ is just a pixelwise addition $f = P_1 + P_2$, then one gets

$$f(\mathbf{x} - u_i, t-1) - f(\mathbf{x}, t) = P_j(\mathbf{x} - u_i, t-1) - P_j(\mathbf{x}, t) =: D_j(\mathbf{x}, t) \tag{3}$$

where $(i, j) = (1, 2)$ or $(2, 1)$ and the displaced frame difference is non zeros, but one of the patterns has been eliminated. In the case that the motion of each pattern P_i is constant on at least three frames at times $t-2, t-2$ and t then the "difference pattern" D_j satisfies the DFD e $D_j(\mathbf{x} - u_j, t-1) - D_j(\mathbf{x}, t) = 0$ and assuming that u_i is known, u_j can be computed by a single motion estimation technique. Burt *et al.* then derive from this fact a three frames algorithm for estimating u_1 and u_2. They start, in a multiresolution setting, with a coarse estimate of u_1 (for instance), and use a single motion algorithm on the resulting difference pattern D_2 in order to compute an estimate of u_2. This estimate is then used to form the difference pattern D_1 and get a new estimate of u_1 from it. This process is iterated until convergence.

A more thorough study and extension of this idea is proposed in a subsequent paper [9], where a frequency domain interpretation, including multiresolution pyramid effects, is provided. In particular a "dominant velocity extraction" mechanism is explained, and the latter is used by Irani and Peleg in [14] (see also [15]).

Starting from the linear superposition principle, Shizawa and Mase explore in a series of papers [29, 30, 31] a frequency domain, total least squares formulation of the multiple motion problems. They start from the single motion case, the OFC constraint equation (2) is replaced by the spatiotemporal *linear* homogeneous one

$$(\nabla f(\mathbf{x}, t))^T \begin{pmatrix} u^1 \\ u^2 \\ u^3 \end{pmatrix} = 0, \quad \vec{u} = (u^1, u^2, u^3) \neq 0 \tag{4}$$

or his frequency domain counterpart

$$\vec{u}^T \omega \hat{f}(\omega) = 0$$

where $\omega = (\omega_1, \omega_2, \omega_3)$ are the spatial and temporal frequencies and $\hat{f}$ is the Fourier transform of f (the $2\pi\sqrt{-1}$ multiplicative constant has been dropped). In that case, assuming constant motion, the best $\vec{u}$ can be retrieved as the minimizer of the energy

$$E_{single}(\vec{v}) = \frac{\vec{v}^t \left(\int \omega\omega^t |\hat{f}(\omega)^2| d\omega \right) \vec{v}}{\vec{v}^t \vec{v} \int |\hat{f}(\omega)^2| d\omega}.$$

This is a total least squares problem whose solution is given as the (an) eigenvector corresponding to the smallest eigenvalue of the 3×3 symmetric, positive (semi-)definite matrix

$$A = \int \omega\omega^t |\hat{f}(\omega)^2| d\omega$$

which is a *Structure Tensor* (see [11, 5, 36] for instance). For the recovery of n motions at a given location, the linear, first order, constraint (4) is replaced by a n-th order, n-th multilinear one obtained by "cascading" the linear first order ones. For example, in the two motions case, the pair $(\vec{u_1}, \vec{u_2})$ would be a zero of the bilinear symmetric map

$$(\vec{v_1}, \vec{v_2}) \mapsto \vec{v_2}^T \mathcal{H}(f) \vec{v_1} = 0 \tag{5}$$

where $\mathcal{H}$ is the spatiotemporal Hessian operator. Multilinear maps can be factored through linear ones using the Tensor Product construction, and this leads them to a two stages formulation of the multiple motion recovery as: first a total least squares computation on the n-th tensor power of the velocity space, which singles out one element (in fact a line) of that space, and secondly a decomposition of this element into a tensor product of n velocities. A closed-form formula for that decomposition is provided in the case $n = 2$, but becomes rapidly more complicated for higher orders. Very recently, Mota *et al.* have extended these ideas in [21] and Mühlich and Aach have proposed an algebraic framework based on homogeneous parts of symmetric algebras in [22].

The very algebraic nature of the motion constraint in frequency domain has lead Vernon to propose in [35] an algorithm for the decoupling of moving patterns, for both transparency and occlusion models. An algorithm for the specific problem of reflections is proposed by Zou and Kambhamettu in [38].

The non homogeneous form (with $u_i^3 = 1$) of equation (5) provides the *2-folds optical flow constraint equation* as introduced by Shizawa and Mase in [30]:

$$\left(v_1^1 \frac{\partial}{\partial x_1} + v_1^2 \frac{\partial}{\partial x_2} + \frac{\partial}{\partial t} \right) \left(v_2^1 \frac{\partial}{\partial x_1} + v_2^2 \frac{\partial}{\partial x_2} + \frac{\partial}{\partial t} \right) f(\mathbf{x}, t) = 0 \tag{6}$$

This form is used by Liu *et al.* in [19] with Hermite polynomial based differentiation filters and specific checks for the presence of single or multiple motions. Darell and Simoncelly

"dualize" this constraint in [10] in order to construct some Fourier "donuts" used to respond to one or more velocities. The 2-folds optical flow constraint is used in the present work, in order to build local multiple motions "probes".

The nonlinear form of this constraint then provides what one may call the *2-folds displaced frame difference equation*, and also used in this present work,

$$f(\mathbf{x}+\vec{u_1}+\vec{u_2},t+2)-f(\mathbf{x}+\vec{u_1},t+1)-f(\mathbf{x}+\vec{u_2},t+1)+f(\mathbf{x},t+2)=0. \qquad (7)$$

It can be extended to more than two motions and has been used as starting point by several authors. For instance, Stuke *et al.* use it in [32] to derive a block-matching approach to the multiple motion problem. In their subsequent work [33], the authors regularized spatially the block matching solution by promoting smooth solutions with a Markov Random Field (MRF) framework, improving the noise robustness of the method. However, finding a solution results in a *computationally heavy minimization* (because of the use of a field of binary indicator variables) and complex (due to a statistical confidence test used to discern the number of motions at each pixel).

Starting with equation (7), Pingault *et al.* in [25] perform a N-th order Taylor expansion around velocity values. A multi-resolution non linear least squares estimation is performed, using a Levenberg-Marquardt algorithm. Recently, Auvray *et al.* proposed in [2] an algorithm based on equation (7). The method is also multi-resolution, uses a simplex algorithm for its initialization and adds a postprocessing step, especially efficient when the two velocities are "close".

The approaches described above are based on a single higher order constraint designed to "react" to multiple motions. In the other hand, a series of methods have been developed by incorporating several single, low order, motion constraints. When dealing with transparency, they all use an essentially unmentioned idea of a local dominance of one of the layers in some spatiotemporal neighborhood of the image sequence. These local dominances are scattered in the image plane/volume and are associated with different layers at different positions. We will now describe a few of these approaches.

In the robust statistics approach of Black and Anandan [6], the transparency is treated through a segmentation approach. The image plane is assumed to be partitioned into regions, each but one corresponding to a parametric motion model $u = u(a)$, a being the parameter vector for the region. This is done by iteratively estimating on a region $\mathcal{O}_i$ a dominant motion $u(a_i)$, the inlier pixel region $\mathcal{R}_i$ for that motion and the outlier region $\mathcal{O}_{i+1} = \mathcal{O}_i - \mathcal{R}_i$, providing, after n iterations the decomposition

$$\mathcal{R}_1 \cup \mathcal{R}_2 \cup \ldots \mathcal{R}_n \cup \mathcal{O}_n$$

where $\mathcal{R}_i$ moves with velocity $u(a_i)$, $i = 1 \ldots n$ and $\mathcal{O}_n$ is the final outlier region. In their paper, they apply the strategy to a image pair (I_1, I_2) with $n = 2$. The motion parameters a_1 and a_2 are then assumed to represent the motion of two layers that cover the entire image

plane. These layers are recovered by a nulling process

$$L_2 = I_2(\mathbf{x} - u(a_1)) - I_1$$
$$L_1 = I_2(\mathbf{x} - u(a_2)) - I_1.$$

The authors describe the process as a "no-model" one (with respect to transparency) and as a limited one. It is however clear that authors assume that transparency is due to superposition of moving patterns, through the ways these patterns are recovered, via a nulling process similar to the one of [4].

Mixture models for multiple motion computation have been introduced by Jepson and Black in [16, 17]. A parametric layered flow model is considered. One assumes that the motion can be explained by up to N parametric motion fields with parameter vectors $\vec{a}_n$. As input, one uses motion vector constraints as mention above, and the probability of observing constraint $\vec{c}_r$ at location r, assuming velocity layer n is given by $p_n(\vec{c}_r|r, \vec{a}_n)$. In order to take outlier measurements into account, a model for it is added in a "zero-th layer" $p_0(\vec{c}_r)$. Assume then that each layer is given the probability α_i of being selected (including $i = 0$ and thus requesting that $\sum_{i=0}^{N} \alpha_i = 1$), then one can write the mixture model for a constraint $\vec{c}_k$

$$p\left(\vec{c}_r|r, (\vec{a}_i)_{i=1\ldots N}, (\alpha_i)_{i=0\ldots N}\right) = \sum_{i=0}^{N} \alpha_i p_i(\vec{c}_r|r, \vec{a}_i).$$

The problem is then to compute the best mixture and motion parameters $(\alpha_i, \vec{a_i})_i$. This is usually done using EM-like algorithms.

Ju, Black *et al* [18] proposed the "Skin and Bones" model, in which multi-layered affine models are defined on small rectangular patches of the image (bones), then an ownership field defines the likelihood that each pixel comes from a particular layer. The goal is to solve for the affine model parameters and the ownership field. This is done within a robust estimation framework using an EM-algorithm. An inter-patch regularization (skin) term introduces a regularization effect in the model parameters estimation. Another layered representation is proposed by Black *et al.* in [7]. In that approach, they consider that multiple motions may appear due to occlusions and limited forms of transparency. The method introduces models for illumination changes and specular reflections, and allows one to eliminate them, improving the computation of the optical flow of the scene. In this formulation, a set of membership weights are computed in order to indicate which layer is more likely to belong to a region. Although the method captures the changes in illumination, it does not allow to compute the optical flow of moving transparencies. Weiss and Adelson [37] and recently Rivera et al. [27] proposed EM-based approaches for computing different layered motion models in an image sequence and its segmentation based on these models. They use as prior knowledge the smooth feature of the velocities. The solution in such cases is given by a field of probabilities measures that indicates layers ownership. Last methods produce pixel-wise unimodal solutions (single motions) because of the use of a distance measure for single motions as well as their entropy controls.

This section has focused on optical flow recovery and does not include some of the related questions, and the especially important one of layer recovery, at the exception of [35, 38]. We mention here the work of Toro *et al.* [34] where the knowlegde of motion is fundamental, as opposed to to work of Sarel and Irani [28], where such a separation is performed by optimizing some correlation measures. Also non mentioned here are the perceptual/neurophysiological aspects of transparency.

3 Problem Statement: From Local to Global

Let us assume that we have an estimation of the likelihood of a set of velocities at each spatio-temporal position. Our goal is to propose an approach which integrate this local velocity information in order to get a more global and robust velocity information. This integration is necessary for dealing with complex motions (such as transparent motion sequences) and with the noise, as we will see in the sequel.

First, let is define a finite sampling of the velocity space, i.e. we consider N vectors

$$\{u_1, \ldots, u_N\},$$

describing the set of possible velocities (such a predefined finite sampling of the velocity space is inspired in the human visual cortex where the different cells are tuned to a specific velocity).

Given a gray-scale image sequence $f : (x,t) \in \Omega \times [0,T] \rightarrow R$, the input is set of functions $d(f, r, u_i) \in \mathbb{R}^+|_{i=1\ldots N}$, where $r = (x,t)$ stands for the spatio-temporal coordinates, describing at each position if the velocity u_i is likely, at a **local** scale, at the position r. We show in Figure 1 an illustration of a the velocity space representation and we refer to Section 3.1 for the estimation of $d(f, r, u_i)$.

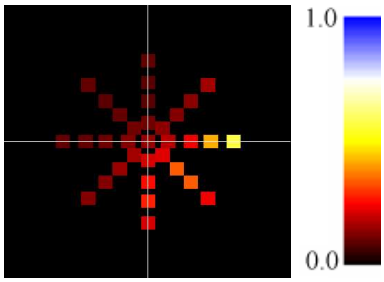


Figure 1: Example of a velocity space composed by 33 velocity vectors, specified through their magnitudes and orientations, respectively $\{0, 1, 2, 3, 4\}$ pixels and $\{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3}{4}\pi, \pi, \frac{5}{4}\pi, \frac{3}{2}\pi, \frac{7}{4}\pi\}$ radians. The color indicates the likelihood of each velocity.

The goal is to compute the velocities (one or more) at each spatio-temporal position r. For this, we will associate to each possible motion u_i a variable $\alpha_i(r)$ which indicates if

such a model is present or not in r. In order to obtain a robust solution, we perform an integration process over the local information provided by d and introduce prior constrains about the possible number of simultaneous motions in r. This will be explained in next subsections.

3.1 Computing Local Velocity Information

We first introduce here the general mechanism we have used in order to select the local velocity descriptors from multiple motion operators. We assume that we are given a family of such operators

$$\mathbf{M} = \{M^{(k)}(u_{i_1}, \ldots, u_{i_k}), \quad k = 1, \ldots N, \quad i_1 < \cdots < i_k\}$$

where $M^{(k)}(u_{i_1}, \ldots, u_{i_k})f(r) \approx 0$ if the velocity vectors $u_{i_1}, \ldots, u_{i_k}$ explain the motion of the image sequence f at the r position. In the case we concern here $k = 1, 2$. For each vector u_i in the base of velocities , we consider the subset $\mathbf{M}_{u_i}$ of all the operators involving u_i and define

$$d(f, r, u_i) = \min_{M \in \mathbf{M}_{u_i}} \frac{1}{k} \|Mf(s)\|^2_{W_r}, \tag{8}$$

where $\|Mf(s)\|^2_{W_r}$ denotes the sum of the L2-norm of $Mf(s)$ for all s in the 3×3 spatial window center at r.

When dealing with one motion ($k = 1$), two well known filters satisfy these requirements, the non-linear correlation:

$$M^{(1)}_{\mathcal{C}}(u_i)f(x, t) \stackrel{def}{=} f(x, t) - f(x - u_i, t - 1), \tag{9}$$

and its differential counterpart

$$M^{(1)}_{\mathcal{D}}(u_i)f(x_1, x_2, t) \stackrel{def}{=} \left(u_{i1}\frac{\partial}{\partial x_1} + u_{i2}\frac{\partial}{\partial x_2} + \frac{\partial}{\partial t}\right) f(x_1, x_2, t).$$

Shizawa and Mase proposed in [31] to build multiple motion operators for velocities $v_1, \ldots, v_k$ as

$$M^{(k)}_{sm}(v_1, \ldots, v_k) = M^{(1)}_{\mathcal{D}}(v_1) \ldots M^{(1)}_{\mathcal{D}}(v_k)$$

where products of the form $\frac{\partial}{\partial x}\frac{\partial}{\partial y}$ are expanded as $\frac{\partial^2}{\partial x \partial y}$.

On the other hand, cascading instead the nonlinear correlation filters $M^{(1)}_{\mathcal{C}}$ provides the operator $M^{(k)}_s$, for instance for $k = 2$ we have

$$M^{(2)}_s(v_1, v_2)f(x, t) = f(y, \tau) - f(y - v_1, \tau - 1) - f(y - v_2, \tau - 1) + f(y - v_1 - v_2, \tau - 2),$$

that corresponds to distance reported in Stuke et al. [32].

The local measurements $d(f, r, u_i)$ used in this work have been derived from $M^{(1)}_{\mathcal{D}}$, $M^{(1)}_{\mathcal{C}}$, $M^{(2)}_{sm}$ and $M^{(2)}_s$, and we, therefore, place us in the additive framework model of [4].

3.2 Objective: Motion Detection Variables

Based on the previous discussion, we define the problem unknowns as the vector valued field α such that a vector in the r position is

$$\alpha(r) = (\alpha_1(r), \ldots, \alpha_N(r)), \quad \alpha_i(r) \in [0,1] \quad \forall r \in \Omega \times [0,T],$$

therefore $\alpha_i(r)$ can be interpreted as the probability of observe the velocity u_i at the position r. Note that although the entries of $\alpha(r)$ are probabilities, $\alpha(r)$ is not a probability measure (as in [37, 27]) in the sense that it is not constrained to sum one. This is, if two motions u_i and u_j are present at a particular pixel position, r, then we expect that both associated probabilities $\alpha_i(r)$ and $\alpha_j(r)$ will be close to one.

In Section 4, we propose an approach for computing the α vector field by means of a variational integration process of the local information $d(f, r, u_i)$ (eqn. (8)).

4 Global Motion Integration via a Variational Approach

Let $d(f, r, u_i)$ defined as in Section 3.1, we look for the velocity distribution minimizing the energy

$$\mathcal{E}(\alpha) = \sum_r \Big\{ \sum_i d(f, r, u_i)\alpha_i^2(r) \tag{10}$$

$$+ \frac{\lambda_s}{2} \sum_{s:s\in\mathcal{N}_r} \sum_i w_i(r,s)[\alpha_i(r) - \alpha_i(s)]^2 \tag{11}$$

$$+\lambda_c \Big[c\bar{\alpha}^2(r) - \sum_i \alpha_i^2(r) \Big] \Big\}, \tag{12}$$

$$\text{subject to } \alpha_i(r) \in [0,1], \forall i;$$

with $\bar{\alpha}(r) \stackrel{def}{=} \frac{1}{N}\sum_i \alpha_i(r)$, where c is a positive scalar, λ_s and λ_c are some positive constants, the weights $w_i(r,s)$ will be defined in the sequel and $\mathcal{N}_r \stackrel{def}{=} \{s : r, s \in \Omega \times [0,T], \|r-s\| < 2\}$ is the spatio-temporal neighborhood of the r position.

Before going more into details, let us give a general idea on the meaning of each term. The first term (10) is called the attach term since it links the input (the functions d's) to the unknown α (see Section 4.1). The second term (11), see Section 4.2, is a smoothing term and its role is to integrate local to global motion estimation. The last term (12), see Section 4.3, gives a prior that controls the number of active motion layers. The compromise between the last term and the attach term introduces a motion model competition mechanism.

4.1 Attach Term

In order to compute the presence of the i-th model, we use an approach related with the outlier rejection method [8] and with the EM formulation [7, 37, 27, 18]. Remind that

function $d(f, r, u_i)$ is close to zero when the velocity u_i explains correctly the motion at position r, and is a positive large value otherwise. Minimizing term (10) with respect to $\alpha_i(r)$ produces $\alpha_i(r)$ close to 0 for high $d(f, r, u_i)$ values, indicating in this way that such a motion model is not likely at position r. Otherwise, the $\alpha_i(r)$ is free and its value is established by the next terms and the bound constraint.

4.2 Spatial Regularization

Term (11) allows us to integrate the local information by regularization, in order to obtain a more global estimation. In the previous work of Stuke *et al.*[33] was noted the necessity of a spatial regularization process. However, given that their approach is based on the computation of categorical variables, hard (combinatorial) optimization methods are required, for instance the computationally-expensive Gibbs Sampler algorithm.

Differently, in our approach, spatial–temporal smoothness means that we want to diminish the difference between the real valued vector $\alpha(r)$ and the ones $\alpha(s)$ in its neighborhood, $\mathcal{N}_r$. Given that our indicator variables are real valued, we can use differentiable potentials with the well-known algorithmic advantages. We use the approach presented in [26] for achieving such a regularization, i.e. a directional one. Therefore the smoothing process is controlled by directional fixed weights,

$$w_i(r, s) = \frac{(s - r)^T \bar{\mathbf{I}}_i (s - r)}{\|s - r\|^4},$$

generated from the i^{th} tensor associated to the i^{th} velocity model: $\bar{\mathbf{I}}_i = \gamma I_d + U_i {U_i}^T$, where I_d is the identity matrix, $\gamma = 0.1$ and $U_i = [u_{i1}, u_{i2}, 1]^T / \|[u_{i1}, u_{i2}, 1]\|$ is a homogeneous-coordinate unitary vector. For a small γ values these weights, $w_i(r, s)$, promote a strong smoothness along the i^{th} velocity direction, see [26]. This is illustrated in Figure 2. As consequence piece-wise smooth optical flows are recovered and the boundaries are well-defined along the velocity model (see results in Figure 6).

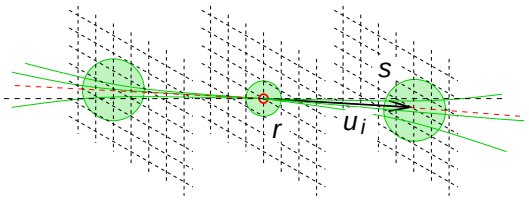


Figure 2: The diffusion coefficients w_i. The diffusion process is performed in the spatio-temporal neighborhood of a given point r, according to the associated velocity u_i. The domain of influence is schematically represented by the circles. The strength of the influence of the point r to the point s then depends on the spatio-temporal distance between r and s, taking into account trajectories leaving r with speed u_i.

4.3 Intra-Model Competition

To introduce the intra-model competition prior, fundamental in our approach, we first remind the expected behavior of the attach term (see Section 4.1): if velocity u_i explains locally the motion at position r, then $d(f, r, u_i)$ is small and consequently the corresponding α_i value is not penalized. Since our aim is to detect multiple simultaneously motions (transparent motions) and therefore we may have several α's switched-on at a given position. Thus we may have problems at sites where multiple spurious matches are locally detected, for example in homogeneous regions, where $d(f, r, u_i)$ is small for many (maybe all) the velocities. For this reason we need a mechanism for eliminating spurious models (to switch-off α's) and to promote the valid ones, i.e., to recover almost binary solutions. So that, our intra-model competition term should behave similarly to entropy-control potentials (as the Shannon's or Gini's used respectively in [37, 27]) in the sense of remove spurious models. Although, in our case we need a suitable term for a no measure of probabilities and multi-modal solutions (see subsection 3.2).

Thus, we use the contrast potential (12) that depends on the $\alpha(r)$ mean value and the parameter c. The c parameter is very useful for controlling the number of switched-on models as will be explained bellow.

To understand the potential's behavior, one can see that the first term penalizes the number of switched-on models while the second term promotes to switch-on models and avoids the trivial solution: zero. Therefore for a fixed mean value (controlled by the first term) the second term prefers high contrasted solutions. It is important to note that our potential (12) can be tuned such that for a given c value a multimodal solution (with two or more detected motions) has lower energy than a unimodal one or conversely. That makes an important distinction with respect to entropy based measured that always have lower energy for unimodal solutions [37, 27]. Additionally, our proposed potential, based on quadratic terms, is easily differentiable and therefore simple minimization algorithms can be used, for instance a Gauss-Seidel scheme.

5 Experiments

5.1 Algorithmic Details

Cost function $\mathcal{E}(\alpha)$, defined by (10)–(12), is quadratic so that it can be minimized by solving the linear system

$$\frac{\partial \mathcal{E}(\alpha)}{\partial \alpha_i(r)} = 0, \quad \forall i,$$

with the constraint $\alpha_i(r) \in [0, 1]$. This is achieved with a Gauss–Seidel iterative scheme,

$$\alpha_i(r) = \frac{\lambda_s \sum_{s \in \mathcal{N}_r} w_i(r, s)\alpha_i(s) - c\lambda_c \bar{\alpha}^{(prev)}(r)}{d(f, r, u_i) + \lambda_s \sum_{s \in \mathcal{N}_r} w_i(r, s) - \lambda_c}, \tag{13}$$

that has the advantage of low memory requirements. The bound constraints on $\alpha_i(r)$ are enforced by projecting non-feasible values to bounds at each iteration. We noted that for obtaining a smooth algorithm convergence, was important to keep fixed the mean of the previous iteration, $\bar{\alpha}^{(prev)}(r)$, for updating the current $\alpha(r)$ vector. This can be seen as an over-relaxation strategy. We initially set $\alpha_i(r) = 0.5, \forall i, r$.

Additionally, a Deterministic Annealing strategy in the λ_c parameter introduces the *intra-model competition* once an approximate solution with valid representative models have predominant $\alpha_j(r)$ values. For each iteration $k = 1, 2, \ldots, n$, we used $\lambda_c^{(k)} = \lambda_c a_k$, where λ_c is the chosen contrast level and $a_k = 1 - 0.95^{(100k/n)}$ is a factor that increases to 1 in the approximately 90% of the total number of iterations, n. This deterministic annealing process ensures that the intra-model regularization term (12) is fully active only when a preliminary solution is available. We note that the computed results are sensible to the annealing speed of λ_c: a premature increment could lead us to an incorrect solution. Nevertheless, we used the same annealing scheduling in all our experiments.

The tuning for the spatial regularization parameter is relatively easy: the large λ_s value eliminates noise but a too large value over-smooth the solution, i.e. the motion boundaries are blurred. We found that $\lambda_s \in [50, 100]$ produced an adequate noise reduction in all the experiments.

According to our experiments the parameter $c = 1$ performs well for most noise-free synthetic sequences. When one is processing noise-contaminated or real sequences, several spurious models may be present in the final solution because of false matches. So that, the prominent models are obtained by increasing this parameter within the small interval $1 \leq c \leq 4$.

One example of the used velocity basis is the one composed by 33 velocity vectors, specified through their magnitudes and orientations, respectively $\{0, 1, 2, 3, 4\}$ pixels and $\{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3}{4}\pi, \pi, \frac{5}{4}\pi, \frac{3}{2}\pi, \frac{7}{4}\pi\}$ radians, see Figure 3. We choose it according to the present displacements in our test sequences, but a different basis can be chosen depending on the problem. This change does not affect significantly a previous parameter selection.

5.2 Global Coherent Motion Estimation for Non Transparent Motion Sequences

In this subsection, we experiment with non transparent motion sequences. The first example deals with the aperture problem and motion integration, while the second illustrates the performances of our algorithm on a real sequence.

5.2.1 Minimization Procedure Performs Motion Integration

The first experiment concerns single motion sequences and it shows how a correct global estimation is obtained based on local velocity estimations. This integration is illustrated with a synthetic sequence that consists of an oblique bar translating in the horizontal direction (Figure 3).

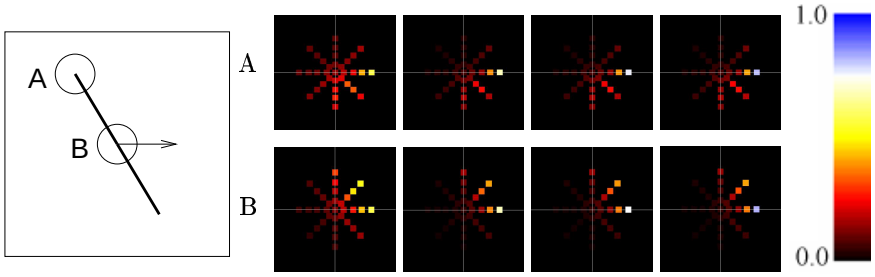


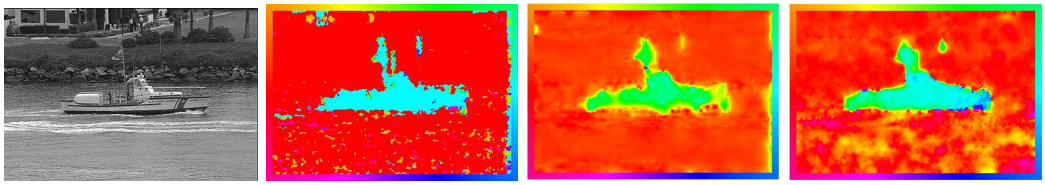
Figure 3: Translating bar example. We show the α's evolution (small squares denote the associated velocities) for the iteration number 1, 3, 5 and 10, for the 2 points marked in the figure on the left. The pseudocolor scale for the range in the alpha values [0,1] is shown to the right-hand side.

Interestingly, some psychophysics studies show how motion integration is performed in tracking tasks. In [20] is shown that the eyes will first follow the normal direction of the bar, according to 1D motion detectors. Then, after few milliseconds, there is a correction of the pursuit toward the horizontal direction, once 2D cues from ending points are integrated (That is illustrated by Figure 3 region B). These kind of experiments suggest that there is a parallel processing between 1D and 2D motion signal with different temporal dynamics and that some time is needed to extract from them a stable response.

Figure 3 shows the evolution of the probabilities α at two given spatio-temporal location, depending on the convergence of the energy minimization. The integration phenomenon can be observed, so that the iterations of the optimization procedure can be interpreted as time evolution in real experiments.

5.2.2 Non Transparent Multiple Motion Sequence.

Figure 4 shows an example of the computed result with our algorithm on a real sequence, called `coastguards`. The background moves roughly horizontally to the left, while the foreground object, a coastguard boat, moves roughly to the right. For comparison purposes some results computed with standard variational approaches [13, 1] are displayed. Figure 4 (b) shows the flow corresponding to the most probable velocity at every position. The orientation of the solution is color-coded: the color associated to a each orientation is shown in the border of the image, for instance, a red pixel in the image indicates a motion to the left. In this case, as expected, unimodal solutions are obtained.



(a) (b) (c) (d)

Figure 4: Optical flow in real scene (single motion case). (a) One frame of the coastguard sequence. Results computed with (b) our approach, (c) Horn and Schunck method [13] and (d) Aubert et al. approach [1].

5.3 Global Coherent Motion Estimation for Transparent Motion Sequences

In this subsection, we present several experiments for transparent motions. The first sequence used is a synthetic one, the three next ones were artificially created from real photographs while the last one is a natural sequence with transparency and occlusions..

5.3.1 Illustration of the Local Measurements.

The Figure 5 (a) shows a synthetic sequence with transparent motion similar to the one used in [33]. The sequence dimensions are $54 \times 54 \times 16$ and it is composed by a moving background (with velocity $\hat{u} = [0, -1]$) and an overlapped moving transparent square (with velocity $\hat{v} = [1, 0]$). Now we demonstrate how local measures could be highly disturbed by noise in the acquisition process. By looking only in the distance measures, several incorrect movements could be considered as valid candidates in each position. For instance, Figures 5(b) and 5(c) shows the present movements associated to the minimum distance value, for the Stuke distance [32] and the Shizawa distances [31], see Section 3.1. The data were noise corrupted with a Signal to Noise Ratio (SNR) equal to 30dB. As one can see, the quality of this first approximation is poor i.e., several incorrect movements are detected and so that a regularization process is required, which is discussed in the next paragraph.

5.3.2 Regularization of Local Measurements.

As mention above, consider the synthetic sequence shown in Figure 5(a). Gaussian noise has been added in order to evaluate the robustness of our proposal. Figures 5(d)-(i) shows the results for a frame. Note that the method can deal with a strong noise corruption, as a SNR=10dB, and shows better performance than the approach reported in [33] (see Figures 5(j) and 5(k)). Note that our method produces relatively good results even for a extreme corrupted sequence, as the one shown in Figures 5(f) and 5(i). For comparison purposes, Figures 5(j) and 5(k) show the computed optical flow with the method reported

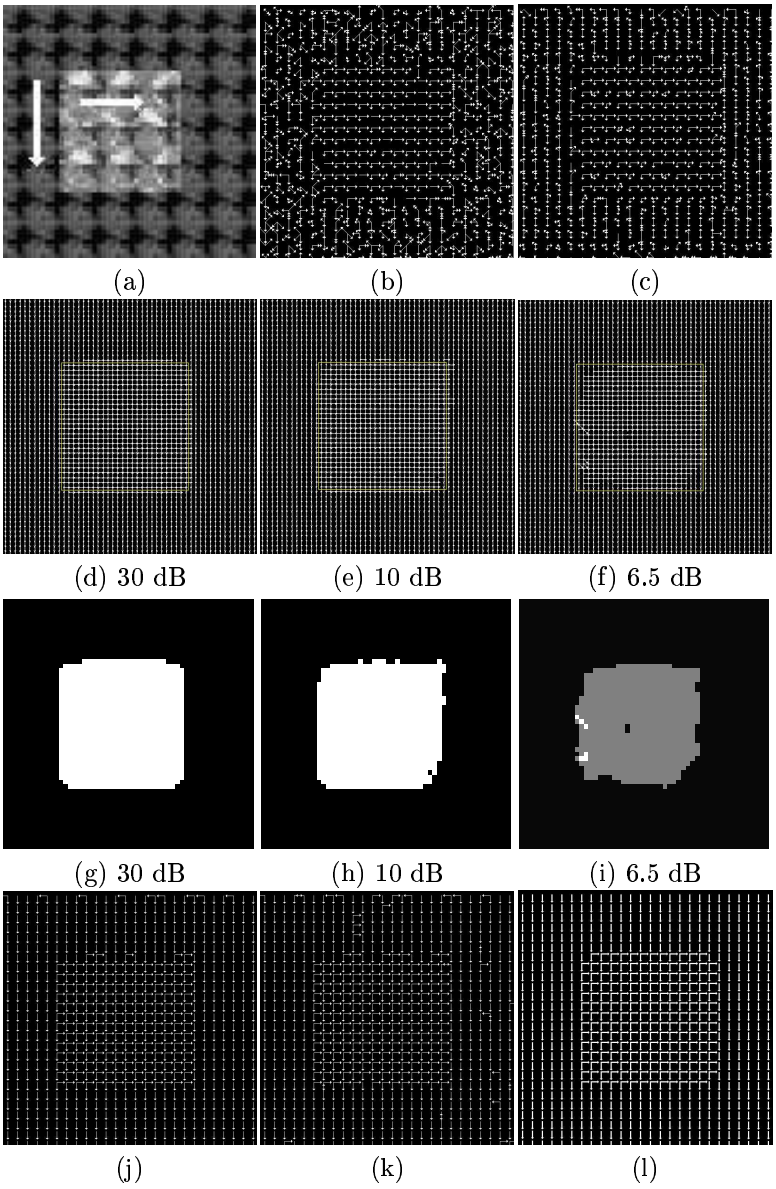


Figure 5: First row: performance of the two local motion estimators. (a) noise-free frame. (b) Non-regularized velocities for [32] distance (SNR=30dB) and (c) for [31] distance. Second and third rows: Our results for different amounts of Gaussian noise, as input we used [32] distance. Figures (d)-(f) velocity fields and Figures (g)-(i) number of motions per pixel: in (g),(h) white = 2 velocities, and black = 1 velocity. In Figure (i) white = 3 velocities, gray = 2 and black = 1. Last row: the Gibbs Sampler (GS) scheme [33]: (j) results in noise-free sequence and (k) in a noise-corrupted one (SNR=30dB). (l) same than (d), but sampled for comparison with (j) and (k).

in [33] for the fixed velocities basis defined in Section 3.2. The noise-free case is shown in Figure 5(j), and the case SNR=30dB is in Figure 5(k). In that case, a hard optimization is performed by using the computationally expensive Gibbs Sampler algorithm. The shown results correspond to the computed solution after 150,000 iterations (about 2.5 hours, in a PC Pentium IV, 3.0 GHz) and represent 150 times than the computational time required by our approach. Furthermore, we observe a bad performance of the method in [33] when the sequences are noise corrupted [see result in Figure 5(k) and compare with the one computed with the proposed method in about 1 minute shown in Figure 5(l)].

We can verify that our spatial regularization, jointly with our intra-model competition, develops well in order to separate the velocities that are present in a region. Figure 6 shows the evolution of the layer associated with velocity [1,0]. Note that the layer takes a significantly large value [by growing from small values (red-yellow) to 1 (blue)] in the square region and that the contribution of this layer is completely eliminated in the background region.

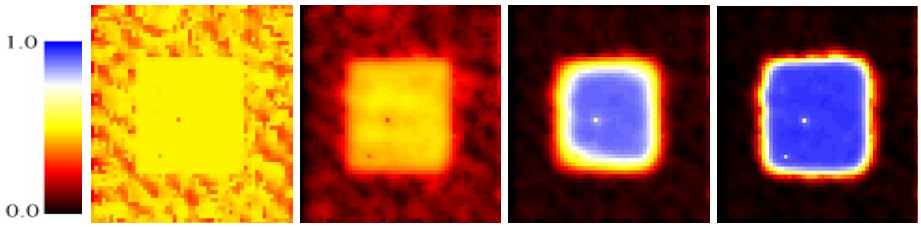


Figure 6: Evolution for the layer associated with velocity [1,0]. For this experiment the sequence was strongly corrupted noise (SNR=15dB). We show the layer values in the pseudo–color scale shown in the left. Note that the presence of the movement [1,0] in the background is pushed to zero because the spatial regularization and the intra-model competition mechanism of the algorithm.

5.3.3 Realistic Texture Sequences

It is important to note that high textured sequences are relatively easy to solve using local motion measures. The real performance of a method for transparent motion should be evaluated in realistic textured scenes: recovering transparent motion in sequences with homogeneous regions presents difficulties because several models may locally explain the data. We have tested our approach using a series of experiments, where we use both synthetic and real image sequences.

For the aim of comparison, we tested the method in the sequence shown in figure 7, which is similar to the one presented in [10]. In this sequence, the left image is moving with velocity $\hat{u} = [1, 0]$ and the right image is moving with velocity $\hat{v} = [-1, 0]$. The transparent region corresponds to the area where the two images overlap. The sequence have dimensions

$64 \times 64 \times 20$ The results for the frame 8 are shown in figure 7. Figure 7a shows the recovered map for 1 or 2 movements, the white regions indicates the presence of 2 movements and the black ones indicates the presence of 1 movement. Figure 7b shows the recovered multi-velocity field. For the sake of clarity, we show separately the recovered field for the velocity [1,0] in Figure 7c and the recovered field for the velocity [-1,0] in Figure 7d.

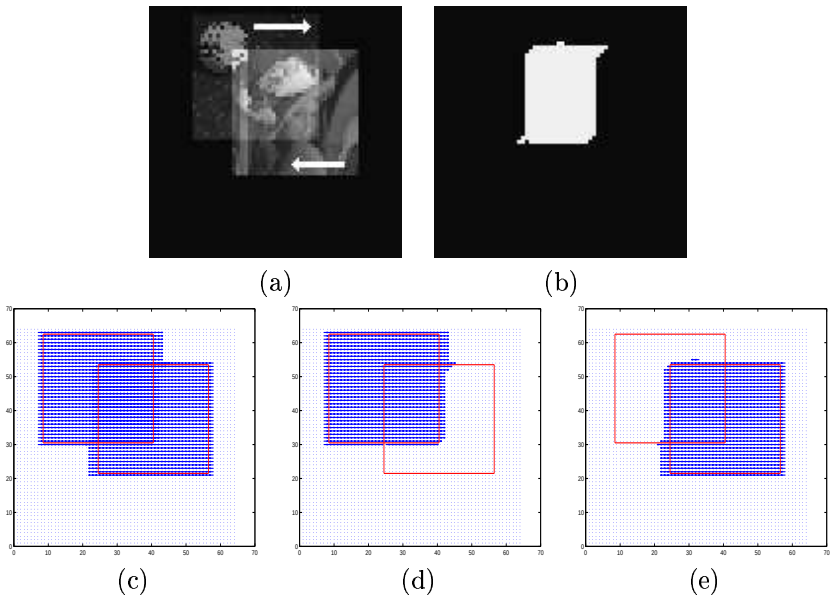


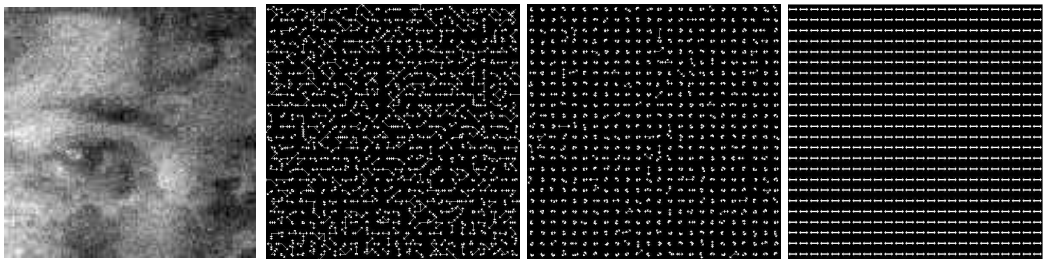
Figure 7: Results for the transparent sequence (a)

The second one presented here is composed of two photographs: a face (limited textured scene) and a Mars landscape, see Figure 8. Figures 9(a) and 9(b) shown the computed optical flow associated to the minimum distance value for the Stuke and Shizawa distances, respectively. We corrupt the sequence with a strong noise (SNR=8dB) (Figure 9 (c)), and the computed velocities field is shown in Figure 9 (d). For this experiment the distance of Shizawa's work was used in the attach term. Note that the right optical flow is recovered in all the pixels regardless the high amount of noise. Figure 10 represents the evolution of one of the two active layers, in this case the one associated with the velocity [-1,0]. One can observe that gaps corresponding to non-textured regions are correctly filled.

Figure 11 shows the methods performance for the case when the transparent region is composed by different combinations of velocities across the time and when the image layers contains non-textured regions (realistic ones). Figure 11(a) shows a frame of the transparent sequence. The changing velocities are schemed in Figure 11(b): the background is moving with velocity [-1,0] and the airplane is moving with velocity [1,-1], [1,0] and [2,2] in equal

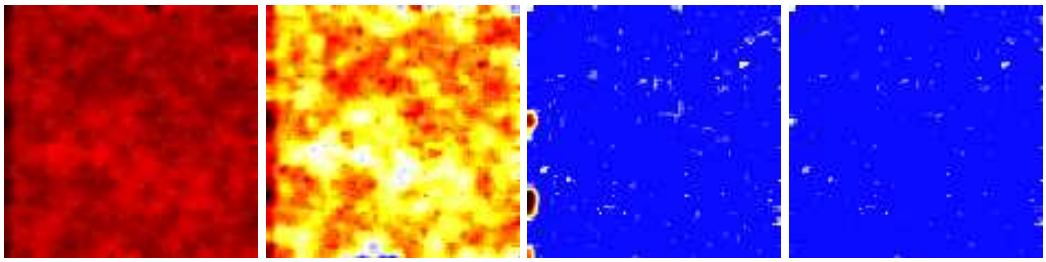
(a) (b) (c)

Figure 8: Realistic synthetic sequence. (a) limited-texture image, I_1, with motion $u = [1, 0]$. (b) Rocky Martian landscape, I_2, with motion $v = [-1, 0]$. (c) transparent generated sequence with $f = \frac{3}{5}I_1 + \frac{2}{5}I_2$.



(a) (b) (c) (d)

Figure 9: Transparent motion estimation on a realistic sequence corrupted with Gaussian noise. (a) Central frame highly noise corrupted (SNR=8dB). Velocities associated with the minimum distance for (b) Stuke and (c) Shizawa measures (SNR=30dB). In (d) the result obtained with the proposed method for the high corrupted sequence in (a) (SNR =8dB), note that we recovered the right velocities in all positions.



Iteration 11 Iteration 21 Iteration 31 Iteration 100

Figure 10: Evolution in the values for the layer associated with the velocity [-1,0].

time intervals[1]. The obtained multi-velocity vector fields are shown in Figures 11(c), 11(d) and 11(e). For this experiment we used the Stukes's distance [32] in the attach term.

5.3.4 Transparency and Occlusion in a Real Sequence

In order to show the performance of the proposed method in a real situation, we shown the results obtained for a real sequence in Figure 12. The sequence is composed by two robots moving slope down, see Figure 12(a), 12(b) and 12(c)[2]. The upper-left robot is located behind a glass, the lower-right one is located in front of the camera and the reflex of the second one is located into the upper-central part. The associated resultant vector fields are shown in Figures 12(d), 12(e) and 12(f) for the 3^{th}, 12^{th} and 22^{th} frame respectively). The recovered velocities were [1.5, 0.4] pixels for upper-left robot and [-1.5,0.5] for both the lower-right robot and its reflex. Note that despite the fact that the lower right-robot is moving a little faster than its reflex (easy to deduct form the projection geometry), both were associated to the same velocity model, because the discrete nature of the velocity basis.

In this experiment, we perform a spatio-temporal Gaussian smoothing process ($\sigma = 0.5$) of the input sequence and we processed only the regions that contain displacements as is explained below. The static background was removed automatically by thresholding the difference between consecutive frames, and then applying opening-closing morphologic operators. By using this pre-process, we obtain an activity-mask that indicates the pixels where a change in time occurred, i.e. the regions where the optical needs to be computed.

In all previous experiments, we computed a dense optical flow in at most 200 minimization iterations.

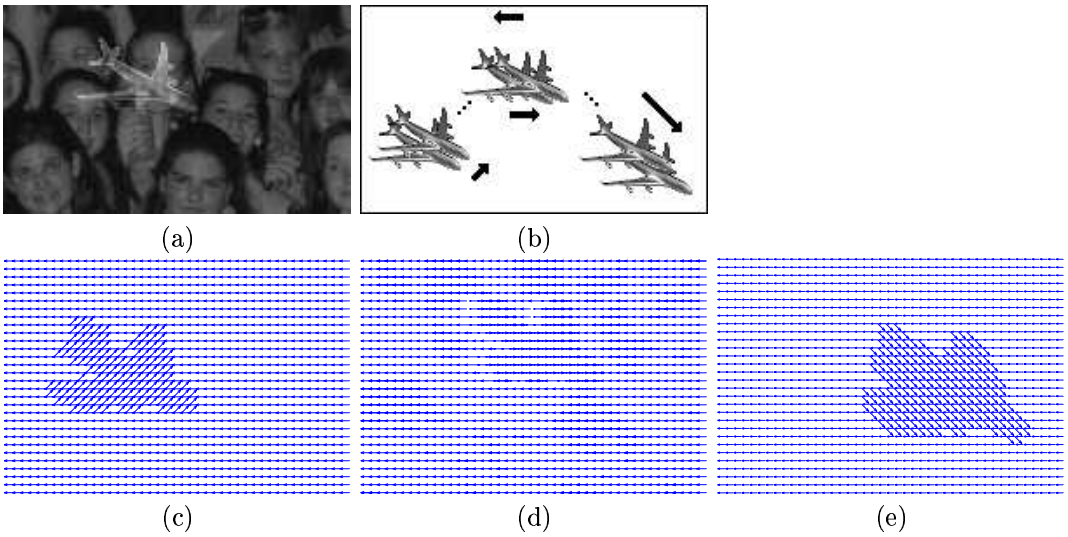
6 Conclusion

In this report we have presented an energy cost formulation in order to estimate multiple motions. The unknown is a vector valued field that indicates the present motions in a particular spatio–temporal position. Our formulation extends previous works based on layered optical flow computation, by using a distance measure suitable for transparent motions and proposing an intra–model competition mechanism proper for multi–valued solutions. The proposed intra–model competition mechanism behave for the multi–motion case, as those used for entropy–control in probability measure approaches for single motion, this term is by itself a novel contribution of this work.

Our formulation allows us to tackle sequences having single or multiple layers moving. The optimization process makes the integration of local velocities information by using suitable diffusion terms. The performance of the presented approach is demonstrated by synthetic experiments in textured and non-textured sequences as well as real sequences.

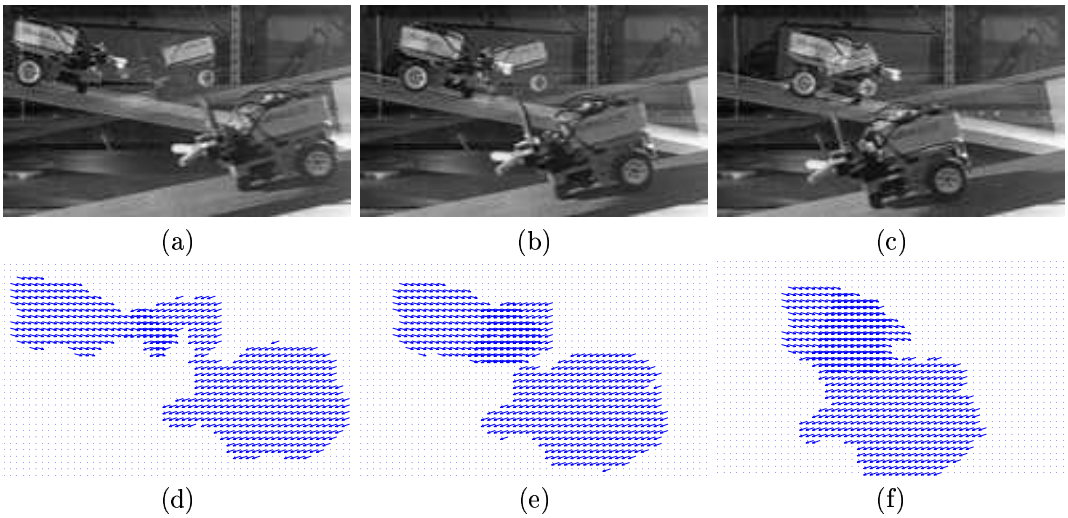
[1] The data sequence and the results can be downloaded at the public web site `http://www.geocities.com/transparent_of_sequences/index.html`.

[2] the AVI file and the computed flows can be downloaded at the public web site `http://www.geocities.com/transparent_of_sequences/index.html`.



(a) (b) (c) (d) (e)

Figure 11: Results for a synthetic transparent sequence in which both the velocity of the background and the velocity of the object changes across the time. (a) A frame taken from the sequence. (b) Scheme of velocities: the airplane experiment velocities [1,-1],[1,0] and [2,2], and the background experiment the velocity [-1,0]. (c), (d) and (e) Sampled recovered multi-velocity fields for frames 5, 23 and 39 respectively.



(a) (b) (c) (d) (e) (f)

Figure 12: Experiment with a real transparent sequence. (a)(b)(c) Frames 3^{th}, 12^{th} and 22^{th} of the real sequence: the upper-left robot is moving slope down behind a glass, the lower-right on is moving slope down in front of camera and its reflex is captured in the upper–central part. (d), (e), (f) Sampled recovered multi-velocity fields for the respective frames.

In future work it is planned to focus on the diffusion terms and investigate how the different velocity maps may interact together. We wish also to evaluate our approach on test sequences used in psychophysics, which will certainly suggest some improvement of the current model.

References

[1] G. Aubert, R. Deriche, and P. Kornprobst. Computing optical flow via variational techniques. *SIAM Journal of Applied Mathematics*, 60(1):156–182, 1999.

[2] V. Auvray, P. Bouthemy, and J. Lienard. Multiresolution parametric estimation of transparent motions. In *Proceedings of the International Conference on Image Processing*, 2005.

[3] J.L. Barron, D.J. Fleet, and S.S. Beauchemin. Performance of optical flow techniques. *The International Journal of Computer Vision*, 12(1):43–77, 1994.

[4] J.R. Bergen, P.J. Burt, R. Hingorani, and S. Peleg. Computing Two Motions From Three Frames. In *Proceedings of the third International Conference on Computer Vision*, pages 27–32, Osaka, Japan, December 1990.

[5] J. Bigun, G. H. Granlund, and J. Wiklund. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(8):775–790, August 1991. Report LiTH-ISY-I-0828 1986 and Report LiTH-ISY-I-1148 1990, both at Computer Vision Laboratory, Linköping University, Sweden.

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

[7] M.J. Black, D.J. Fleet, and Y. Yacoob. Robustly estimating changes in image appearance. *Computer Vision and Image Understanding*, 78:8–31, 2000.

[8] M.J. Black and P. Rangarajan. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *The International Journal of Computer Vision*, 19(1):57–91, 1996.

[9] P.J. Burt, R. Hingorani, and R.J. Kolczynski. Mechanisms for Isolating Component Patterns in the Sequential Analysis of Multiple Motion. In *Proceedings of the IEEE Workshop on Visual Motion*, pages 187–193, Princeton, NJ, October 1991.

[10] T. Darrell and E. Simoncelli. Separation of transparent motion into layers using velocity-tuned mechanisms. In *MIT Media Laboratory Vision and Modeling Group Technical Report*, number 244, 1993.

[11] W. Förstner. A Feature Based Corresponding Algorithm for Image Matching. *International Archives of Photogrammetry and Remote Sensing*, 26:150–166, 1986.

[12] B. Galvin, B. McCane, K. Novins, D. Mason, and S. Mills. Recovering motion fields: an evaluation of eight optical flow algorithms. *British Machine Vision Conference*, pages 195–204, 1998.

[13] B.K. Horn and B.G. Schunck. Determining Optical Flow. *Artificial Intelligence*, 17:185–203, 1981.

[14] M. Irani and S. Peleg. Motion analysis for image enhancement: resolution, occlusion, and transparency. *Journal on Visual Communications and Image Representation*, 4(4):324–335, 1993.

[15] M. Irani, B. Rousso, and S. Peleg. Computing Occluding and Transparent motions. *The International Journal of Computer Vision*, 12(1):5–16, January 1994.

[16] A. Jepson and M.J. Black. Mixture Models for Optical Flow Computation. In *CVPR93*, pages 760–761, 1993.

[17] A. Jepson and M.J. Black. Mixture Models for Optical Flow Computation. Technical Report RBCV-TR-93-44, University of Toronto, Department of Computer Science, April 1993.

[18] S.X. Ju, M.J. Black, and A.D. Jepson. Skin and Bones: Multi-layer, Locally Affine, Optical Flow and Regularization with Transparency. In *Proceedings of CVPR 96*, pages 307–314, San Francisco, CA, June 1996.

[19] H.C. Liu, T.H. Hong, M. Herman, and R. Chellappa. Spatio-temporal filters for transparent motion segmentation. In *Proceedings of the International Conference on Image Processing*, pages 464–468, Washington, USA, 1995.

[20] G.S. Masson, Y. Rybarzcyk, E. Castet, and D.R. Mestre. Temporal dynamics of motion integration for the initiation of tracking responses at ultra-short latencies. *Visual Neuroscience*, 17(5):754–767, 2000.

[21] C. Mota, I. Stuke, T. Aach, and E. Barth. Divide-and-Conquer Strategies for Estimating Multiple Transparent Motions. In *Proceedings of the 1st International Workshop on Complex Motion, Schloss Reisensburg, Germany. Lecture Notes on Computer Science, LNCS Vol. 3417*, 2005.

[22] M. Mühlich and T. Aach. A Theory of Multiple Orientation Estimation. In *European Conference on Computer Vision*, volume 2, pages 69–82, 2006.

[23] A. V. Oppenheim. Superpostion in a Class of Nonlinear Systems. In *IEEE International Convention*, pages 171–177, New York, USA, 1964.

[24] M. Orkisz and P. Clarysse. Estimation du flot optique en présence de discontinuités: une revue. *Traitement du Signal*, 13(5):489–513, 1996.

[25] M. Pingault, E. Bruno, and D. Pellerin. A Robust Multiscale B-Spline Function Decomposition for Estimating Motion Transparency. *IEEE Transactions on Image Processing*, 12(11):1416–1426, November 2003.

[26] A. Ramirez-Manzanares and M. Rivera. Brain nerve boundless estimation by restoring and filtering intra-voxel information in DT-MRI. In *Second Workshop on Variational and Level Sets Methods*, pages 71–80, October 2003.

[27] M. Rivera, O. Ocegeda, and J. L. Marroquin. Entropy controlled gauss-markov random measure field models for early vision. In *LNCS 3752, Springer-Verlag VLSM 2005*, pages 137–148, 2005.

[28] B. Sarel and M. Irani. Separating transparent layers through layer information exchange. In T. Pajdla and J. Matas, editors, *Proceedings of the 8th European Conference on Computer Vision*, pages 328–341, Prague, Czech Republic, 2004. Springer–Verlag.

[29] M. Shizawa and K. Mase. Simultaneous Multiple Optical Flow Estimation. In *International Conference on Pattern Recognition*, volume 1, pages 274–278, 1990.

[30] M. Shizawa and K. Mase. Principle of Superposition: a Common Computational Framework for Analysis of Multiple Motion. In *IEEE Workshop on Visual Motion*, pages 164–172, 1991.

[31] M. Shizawa and K. Mase. A unified computational theory for motion transparency and motion boundaries based on eigenergy analysis. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pages 289–295, Lahaina, Hawai, June 1991. IEEE.

[32] I. Stuke, T. Aach, E. Barth, and C. Mota. Estimation of multiple motions by block matching. In *4th ACIS International Conference onSoftware Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2003)*, pages 358–362, 2003.

[33] I. Stuke, T. Aach, E. Barth, and C. Mota. Multiple-motion-estimation by block matching using MRF. In *ACIS, International Journal of Computer and Information Science*, 2004.

[34] J. Toro, F. Owens, and R. Medina. Using Known Motion Fields for Image Separation in Transparency. *Pattern Rcognition Letters*, 24:597–605, 2003.

[35] D. Vernon. Decoupling Fourier Components of Dynamic Image Sequences: a Theory of Signal Separation, Image Segmentation and Optical Flow Estimation. In *European Conference on Computer Vision*, volume 2, pages 69–85, June 1998.

[36] J. Weickert and H. Hagen, editors. *Visualization and Processing of Tensor Fields.* Springer Verlag, 2006.

[37] Y. Weiss and E.H. Adelson. A unified mixture framework for motion segmentation: incorporating spatial coherence and estimating the number of models. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pages 321–326, San Francisco, CA, June 1996. IEEE.

[38] W. Zhou and C. Kambhamettu. Separation of Reflection by Fourier Decoupling. In *Asian Conference on Computer Vision*, Jeju Island, Korea, January 2004.

Contents

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399