SPATIAL/JOINT SPACE-TIME MOTION SEGMENTATION OF IMAGE SEQUENCES BY LEVEL SET PURSUIT

Abdol-Reza Mansouri and Amar Mitiche

INRS-Télécommunications, Institut National de la Recherche Scientifique
Place Bonaventure, P.O. Box 644, Montréal, Québec, Canada, H5A 1C6
[mansouri,mitiche]@@inrs-telecom.uquebec.ca

1. PROBLEM STATEMENT

Segmentation of image sequences based on motion is an important problem with numerous applications in video processing (object-based frame conversion) and video compression (MPEG-4) [1]. Motion-based segmentation can be carried out either on a frame by frame basis, resulting in a 2D *spatial region segmentation* in each frame of the sequence, or it can be carried out on a multi-frame basis, whereby the image sequence is treated as a function defined on a spatio-temporal volume, resulting in a spatio-temporal 3D *spatio-temporal volume segmentation* problem [2].

The basic problem we address in this paper is the following: Given a set of sparse point correspondences, how does one obtain a dense motion field and a motion-based segmentation of the image sequence ? We propose a novel solution to this problem by formulating it as pursuit in segmentation space. This segmentation is defined by level set evolution equations, allowing changes in segmentation topology [4].

The main novelty of our proposed algorithm is that the number of distinct motion regions and their parameters need not be known prior to segmentation and are determined by the segmentation itself. Furthermore, the algorithm we propose applies equally to (frame-by-frame) spatial motion-based segmentation and to (multi-frame) joint space-time motion-based segmentation. This extends our prior work on motion-based image segmentation with level sets in both the spatial [3] and spatio-temporal domains [2] where the number of distinct motion regions as well as their precise motion parameters need to be computed through a complex clustering operation prior to segmentation.

We formulate motion-based image segmentation as pursuit in the space of image segmentations, in analogy to classical matching pursuit, the basic idea consisting of iteratively segmenting the image by focusing on residual regions. The segmentation obtained by our algorithm is based on motion alone and does not use intensity boundaries as an auxiliary. We formulate our algorithm for the case of spatial motion segmentation and illustrate it on a real image sequence with natural motion; our proposed algorithm applies verbatim to the case of joint space-time motion segmentation with level sets, and we refer the reader to [2] for details of the latter.

This work was supported by the Natural Sciences and Engineering Research Council of Canada under Strategic Grant STR224122.

2. MOTION-BASED IMAGE SEGMENTATION

2.1. Basic Models

Let $\{I^0, I^1\}$ be images at time instants t_0 and t_1. Let the domain of both images be $\Omega \subset \mathbb{R}^2$. We define motion segmentation as the partitioning $\{\Omega_i\}_{i=1}^N$ of the domain Ω such that each region Ω_i is characterized by a distinct motion $\boldsymbol{T}_i$. The goal of motion-based segmentation is the estimation of $\{(\Omega_i, \boldsymbol{T}_i)\}_{i=1}^N$ from the image pair $\{I^0, I^1\}$. The following observation model relates an element Ω_i of the partition and the corresponding motion transformation $\boldsymbol{T}_i$:

$$I^0(\boldsymbol{x}) = I^1(\boldsymbol{T}_i\boldsymbol{x}) + \eta_i(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in \Omega_i, \forall i \in \{1, \ldots, N\},$$

where the η_i are assumed to be independent zero-mean stationary Gaussian white noise processes with variance σ^2. We assume that the transformations $\boldsymbol{T}_i$ all belong to a parametric family Θ of transformations (e.g. translational, Euclidean, affine, ...). Let $\mathcal{C} = \{(\boldsymbol{x}_j, \vec{v}_j)\}_{j=1}^p$ be a given family of point correspondences, where $\boldsymbol{x}_j \in \Omega$ is a point in I^0 and $\boldsymbol{x}_j + \vec{v}_j$ the corresponding point in I^1. The only assumption we make on the family $\mathcal{C}$ is that for each $\boldsymbol{T}$ in $\{\boldsymbol{T}_i\}_{i=1}^N$ there exist a subset of $\mathcal{C}$ that defines it. This means that the family $\mathcal{C}$ should not be too sparse so as to miss possible image motions. We define $\Theta_{\mathcal{C}}$ to be the subset of Θ consisting of the transformations obtained from $\mathcal{C}$. For each element $(\boldsymbol{x}_j, \vec{v}_j)$ of $\mathcal{C}$, we construct the quadruple $(\boldsymbol{x}_j, \boldsymbol{U}_j, \epsilon_j, \widehat{\boldsymbol{T}}_j)$ where $\boldsymbol{U}_j \subset \Omega$ is a neighborhood of $\boldsymbol{x}_j$ (typically a square block centered at $\boldsymbol{x}_j$), ϵ_j the associated error measure defined by:

$$\epsilon_j = \inf_{\boldsymbol{T} \in \Theta_{\mathcal{C}}} \frac{1}{\int_{\boldsymbol{U}_j} d\boldsymbol{x}} \int_{\boldsymbol{U}_j} (I^1(\boldsymbol{T}\boldsymbol{x}) - I^0(\boldsymbol{x}))^2 d\boldsymbol{x},$$

and $\widehat{\boldsymbol{T}}_j$ the transformation in $\Theta_{\mathcal{C}}$ for which this infimum is reached. We retain only the quadruples for which $\epsilon_j < \sigma^2$, obtaining a family $\mathcal{C}_\Theta = \{(\boldsymbol{x}_j, \boldsymbol{U}_j, \epsilon_j, \widehat{\boldsymbol{T}}_j) | \epsilon_j \leq \sigma^2\}$ of cardinality less than or equal to that of $\mathcal{C}$.

2.2. MAP Formulation

The maximum a posteriori motion-based segmentation $\{\Omega_i^\star, \boldsymbol{T}_i^\star\}_{i=1}^N$ is given by

$$\begin{aligned}\{\Omega_i^\star, \boldsymbol{T}_i^\star\}_{i=1}^N &= \arg \max_{\{\hat{\Omega}_i, \widehat{\boldsymbol{T}}_i\}_{i=1}^N} P(\{\hat{\Omega}_i, \widehat{\boldsymbol{T}}_i\}_{i=1}^N | I^0, I^1, \mathcal{C}_\Theta) \\ &= \arg \max_{\{\hat{\Omega}_i, \widehat{\boldsymbol{T}}_i\}_{i=1}^N} \{P(I^1 | I^0, \{\hat{\Omega}_i, \widehat{\boldsymbol{T}}_i\}_{i=1}^N, \mathcal{C}_\Theta) \\ &\quad \cdot P(\{\hat{\Omega}_i, \widehat{\boldsymbol{T}}_i\}_{i=1}^N | I^0, \mathcal{C}_\Theta)\}\end{aligned}$$

Since I^1 is related to I^0 through the transformations $\boldsymbol{T}_i$, we can express the likelihood $P(I^1|I^0, \{\hat{\Omega}_i, \widehat{\boldsymbol{T}}_i\}_{i=1}^N, \mathcal{C}_\Theta)$ as follows:

$$P(I^1|I^0, \{\hat{\Omega}_i, \widehat{\boldsymbol{T}}_i\}_{i=1}^N, \mathcal{C}_\Theta) = P(I^1|I^0, \{\hat{\Omega}_i, \widehat{\boldsymbol{T}}_i\}_{i=1}^N) = \prod_{i=1}^N \prod_{\boldsymbol{x}\in\hat{\Omega}_i} \mathcal{N}_i(\xi(\widehat{\boldsymbol{T}}_i, \boldsymbol{x})),$$

where $\xi(\boldsymbol{T}, \boldsymbol{x}) = I^1(\boldsymbol{T}\boldsymbol{x}) - I^0(\boldsymbol{x})$ is the motion-compensated intensity residual, and $\mathcal{N}_i$ is the density function of the noise process η_i. We make the simplifying assumption that Ω_i and $\boldsymbol{T}_j$ are independent conditioned on I^0 and $\mathcal{C}_\Theta$, and, assuming independence of the motion regions on I^0 and $\mathcal{C}_\Theta$, we obtain:

$$P(\{\hat{\Omega}_i, \widehat{\boldsymbol{T}}_i\}_{i=1}^N|I^0, \mathcal{C}_\Theta) = \prod_{i=1}^N P(\hat{\Omega}_i) \prod_{j=1}^N P(\widehat{\boldsymbol{T}}_j|I^0, \mathcal{C}_\Theta).$$

We shall also take a simple prior for the Ω_i, one that favors shorter boundaries $\partial\Omega_i$. With this choice of prior, the MAP estimation problem is equivalent to the following energy minimization problem:

$$\{\Omega_i^\star, \boldsymbol{T}_i^\star\}_{i=1}^N = \arg \min_{\{\hat{\Omega}_i, \widehat{\boldsymbol{T}}_i\}_{i=1}^N} E(\{\hat{\Omega}_i, \widehat{\boldsymbol{T}}_i\}_{i=1}^N|I^0, I^1, \mathcal{C}_\Theta)$$

$$\begin{aligned} E(\{\hat{\Omega}_i, \widehat{\boldsymbol{T}}_i\}_{i=1}^N|I^0, I^1, \mathcal{C}_\Theta) &= \frac{1}{2\sigma^2}\sum_{i=1}^N \int_{\hat{\Omega}_i} \xi^2(\widehat{\boldsymbol{T}}_i, \boldsymbol{x}) d\boldsymbol{x} \\ &+ \frac{\lambda}{2\sigma^2}\sum_{i=1}^N \oint_{\partial\hat{\Omega}_i} ds \\ &- \sum_{i=1}^N \log P(\widehat{\boldsymbol{T}}_i|I^0, \mathcal{C}_\Theta) \end{aligned}$$

where ds is the differential of arclength and λ a positive constant. Assume now that $\widehat{\boldsymbol{T}}_i = \boldsymbol{T}_i$ for all i and $\hat{\Omega}_i = \Omega_i$ for $i \neq 1$. We can then view the above functional as a function of $\hat{\Omega}_1$ only, and we can write

$$\begin{aligned} E(\{\hat{\Omega}_i, \widehat{\boldsymbol{T}}_i\}_{i=1}^N|I^0, I^1, \mathcal{C}_\Theta) &= \frac{1}{2\sigma^2}\int_{\hat{\Omega}_1} \xi^2(\boldsymbol{T}_1, \boldsymbol{x}) d\boldsymbol{x} \\ &+ \frac{1}{2\sigma^2}\int_{\hat{\Omega}_1^c} \psi(\{\boldsymbol{T}_i\}_{i=2}^N, \boldsymbol{x}) d\boldsymbol{x} + \frac{\lambda}{2\sigma^2}\oint_{\partial\hat{\Omega}_1} ds \\ &+ \frac{\lambda}{2\sigma^2}\sum_{i=2}\oint_{\partial\Omega_i} ds - \sum_{i=1}^N \log P(\boldsymbol{T}_i|I^0, \mathcal{C}_\Theta) \end{aligned}$$

where $\psi(\{\boldsymbol{T}_i\}_{i=2}^N, \boldsymbol{x}) = \sum_{i=2}^N \chi_{\Omega_i}(\boldsymbol{x})\xi^2(\boldsymbol{T}_i, \boldsymbol{x})$, χ_{Ω_i} being the indicator function of the set Ω_i. Approximating ψ by its expected value $E(\psi)$ yields $\int_{\hat{\Omega}_1^c} \psi(\{\boldsymbol{T}_i\}_{i=2}^N, \boldsymbol{x}) d\boldsymbol{x} = \sigma^2 \int_{\hat{\Omega}_1^c} d\boldsymbol{x}$. With this approximation, the Euler-Lagrange descent equation of E with respect to $\hat{\Omega}_1$ yields the following level set evolution equation:

$$\frac{\partial u(\boldsymbol{x}, t)}{\partial t} = -(\xi_1^2(\boldsymbol{x}) - \sigma^2 + \lambda\kappa)\|\vec{\nabla} u\|$$

where $u : \mathbb{R}^2 \times \mathbb{R}^+ \to \mathbb{R}$ evolves according to the above equation in such a way that the zero-level set $\{u(., \infty) = 0\}$ obtained at convergence minimizes the functional E with respect to $\partial\hat{\Omega}_1$, and κ is the curvature of the zero-level set of u. An estimate of Ω_1 is then given by $\hat{\Omega}_1 = \{u(., \infty) > 0\}$.

2.3. Motion Segmentation by Basic Level Set Pursuit

The above approximation and derivation suggest the following sequential algorithm: Let $(\boldsymbol{x}_1, \boldsymbol{U}_1, \epsilon_1, \widehat{\boldsymbol{T}}_1)$ be the (or a) quadruple in $\mathcal{C}_\Theta$ with minimum error ϵ_1. Consider the level set evolution equation

$$\frac{\partial u_1(\boldsymbol{x}, t)}{\partial t} = -(\xi_1^2(\boldsymbol{x}) - \sigma^2 + \lambda\kappa_1)\|\vec{\nabla} u_1\|$$

and define $\hat{\Omega}_1 = \{u_1(., \infty) > 0\}$. This yields the motion-based segmentation $\{\hat{\Omega}_1, \hat{\Omega}_1^c\}$ of the image domain Ω into one region and its complement. Let now $\mathcal{C}_\Theta^{(1)}$ be the subset of $\mathcal{C}_\Theta$ obtained by removing from $\mathcal{C}_\Theta$ all quadruples $(\boldsymbol{x}, \boldsymbol{U}, \epsilon, \boldsymbol{T})$ for which $\boldsymbol{U} \subset \hat{\Omega}_1$. Let $(\boldsymbol{x}_2, \boldsymbol{U}_2, \epsilon_2, \widehat{\boldsymbol{T}}_2)$ be the (or a) quadruple in $\mathcal{C}_\Theta^{(1)}$ with minimum error ϵ_2. Consider the level set evolution equation

$$\frac{\partial u_2(\boldsymbol{x}, t)}{\partial t} = -(\xi_2^2(\boldsymbol{x}) - \sigma^2 + \lambda\kappa_2)\|\vec{\nabla} u_2\|$$

and define $\hat{\Omega}_2 = \{u_2(., \infty) > 0\} \cap \hat{\Omega}_1^c$. This yields the motion-based segmentation of Ω into the three regions $\hat{\Omega}_1, \hat{\Omega}_2$, and $(\hat{\Omega}_1 \cup \hat{\Omega}_2)^c$. Assume we are now at step l of the pursuit, that is, a motion-based segmentation $\{\hat{\Omega}_1, \cdots, \hat{\Omega}_l, (\hat{\Omega}_1 \cup \cdots \cup \hat{\Omega}_l)^c\}$ of Ω has been obtained. We then define $\mathcal{C}_\Theta^{(l)}$ as the subset of $\mathcal{C}_\Theta$ obtained by removing from $\mathcal{C}_\Theta$ all quadruples $(\boldsymbol{x}, \boldsymbol{U}, \epsilon, \boldsymbol{T})$ for which $\boldsymbol{U} \subset \bigcup_{i=1}^l \hat{\Omega}_i$. Letting $\widehat{\boldsymbol{T}}_{l+1}$ be the transformation corresponding to the quadruple with smallest error in $\mathcal{C}_\Theta^{(l)}$, we solve the level set partial differential equation

$$\frac{\partial u_{l+1}(\boldsymbol{x}, t)}{\partial t} = -(\xi_{l+1}^2(\boldsymbol{x}) - \sigma^2 + \lambda\kappa_{l+1})\|\vec{\nabla} u_{l+1}\|$$

and we define $\hat{\Omega}_{l+1} = \{u_{l+1}(., \infty) > 0\} \cap (\bigcup_{i=1}^l \hat{\Omega}_i)^c$. Thus, at each step in the pursuit, only the residual of the motion regions obtained at the previous step, i.e. the complement of the union of the regions obtained, is segmented. As in classical matching pursuit, where the best approximating function is included in the linear combination at each step, we consider the best approximating transformation at each step (in the sense of minimum error associated to the quadruple of that transformation) and perform segmentation according to that transformation. If at step l of the pursuit, we obtain $\mathcal{C}_\Theta^{(l)} = \emptyset$ or $\mathcal{C}_\Theta^{(l)} = \mathcal{C}_\Theta^{(l-1)}$, then the pursuit terminates. The number of distinct motion regions is then defined to be l. Note that the number of motion regions follows directly from the segmentation itself and therefore need not be known prior to segmentation as is the case with [3]. Note also that a form of clustering is implicit in this pursuit; indeed, correspondence points which are englobed by a region during segmentation are discarded and do not give rise to additional segmentations. Yet, in contrast to [3], this clustering operation is a natural by-product of the segmentation itself, and does not involve additional thresholds or parameters.

2.4. Motion Segmentation by Level Set Pursuit

The basic level set pursuit algorithm described in section 2.3 suffers from a drawback: The selection of the transformation on the basis of which segmentation is performed is guided by an error function which is computed only in a neighborhood of a correspondence point. Thus, the motion-based segmentation of the image domain which has been computed at some stage of the pursuit may need to be modified in light of the segmentation of its residual. This suggests that at each stage of the pursuit, all of the existing segmentations need be considered for readjustment. Such an approach then yields the following level set pursuit algorithm: Let $(\boldsymbol{x}_1, \boldsymbol{U}_1, \epsilon_1, \widehat{\boldsymbol{T}}_1)$ be the (or a) quadruple in $\mathcal{C}_\Theta$ with minimum error ϵ_1. Consider the level set evolution equation

$$\frac{\partial u_1(\boldsymbol{x},t)}{\partial t} = -(\xi_1^2(\boldsymbol{x}) - \sigma^2 + \lambda\kappa_1)\|\vec{\nabla} u_1\|$$

and define $\hat{\Omega}_1 = \{u_1(.,\infty) > 0\}$. This yields the motion-based segmentation $\{\hat{\Omega}_1, \hat{\Omega}_1^c\}$ of the image domain Ω into one region and its complement. Let now $\mathcal{C}_\Theta^{(1)}$ be the subset of $\mathcal{C}_\Theta$ obtained by removing from $\mathcal{C}_\Theta$ all quadruples $(\boldsymbol{x}, \boldsymbol{U}, \epsilon, \widehat{\boldsymbol{T}})$ for which $\boldsymbol{U} \subset \hat{\Omega}_1$. Let $(\boldsymbol{x}_2, \boldsymbol{U}_2, \epsilon_2, \widehat{\boldsymbol{T}}_2)$ be the (or a) quadruple in $\mathcal{C}_\Theta^{(1)}$ with minimum error ϵ_2. Consider now the *system* of level set evolution equations

$$\begin{cases} \frac{\partial u_1(\boldsymbol{x},t)}{\partial t} = -(\xi_1^2(\boldsymbol{x}) - \zeta_1^2(\boldsymbol{x}) + \lambda\kappa_1)\|\vec{\nabla} u_1\| \\ \frac{\partial u_2(\boldsymbol{x},t)}{\partial t} = -(\xi_2^2(\boldsymbol{x}) - \zeta_2^2(\boldsymbol{x}) + \lambda\kappa_2)\|\vec{\nabla} u_2\| \end{cases}$$

where

$$\begin{aligned} \zeta_1^2(\boldsymbol{x}) &= \chi_{\{u_2 \le 0\}}(\boldsymbol{x})\sigma^2 + \chi_{\{u_2 > 0\}}(\boldsymbol{x})\xi_2^2(\boldsymbol{x}), \\ \zeta_2^2(\boldsymbol{x}) &= \chi_{\{u_1 \le 0\}}(\boldsymbol{x})\sigma^2 + \chi_{\{u_1 > 0\}}(\boldsymbol{x})\xi_1^2(\boldsymbol{x}), \end{aligned}$$

and define $\hat{\Omega}_2 = \{u_2(.,\infty) > 0\} \cap \hat{\Omega}_1^c$. Similarly, assume we are now at step l of the pursuit, that is, a motion-based segmentation $\{\hat{\Omega}_1, \cdots, \hat{\Omega}_l, (\hat{\Omega}_1 \cup \cdots \cup \hat{\Omega}_l)^c\}$ of Ω has been obtained. We then define $\mathcal{C}_\Theta^{(l)}$ as the subset of $\mathcal{C}_\Theta$ obtained by removing from $\mathcal{C}_\Theta$ all quadruples $(\boldsymbol{x}, \boldsymbol{U}, \epsilon, \widehat{\boldsymbol{T}})$ for which $\boldsymbol{U} \subset \bigcup_{i=1}^l \hat{\Omega}_i$. Letting $\widehat{\boldsymbol{T}}_{l+1}$ be the transformation in $\mathcal{C}_\Theta^{(l)}$ corresponding to the quadruple with smallest error, we are led to the following system of level set partial differential equations

$$\begin{cases} \frac{\partial u_1(\boldsymbol{x},t)}{\partial t} = -(\xi_1^2(\boldsymbol{x}) - \zeta_1^2(\boldsymbol{x}) + \lambda\kappa_1)\|\vec{\nabla} u_1\| \\ \cdots \\ \frac{\partial u_l(\boldsymbol{x},t)}{\partial t} = -(\xi_l^2(\boldsymbol{x}) - \zeta_l^2(\boldsymbol{x}) + \lambda\kappa_l)\|\vec{\nabla} u_l\| \\ \frac{\partial u_{l+1}(\boldsymbol{x},t)}{\partial t} = -(\xi_{l+1}^2(\boldsymbol{x}) - \zeta_{l+1}^2(\boldsymbol{x}) + \lambda\kappa_{l+1})\|\vec{\nabla} u_{l+1}\| \end{cases}$$

with $\zeta_k^2(\boldsymbol{x}) = \min_{i \neq k, u_i > 0} \xi_i^2(\boldsymbol{x})$, the min operator equating to σ^2 whenever the defining condition is void, and we define $\hat{\Omega}_{l+1} = \{u_{l+1}(.,\infty) > 0\} \cap (\bigcup_{i=1}^l \hat{\Omega}_i)^c$. Here again, the number of motion regions follows directly from the segmentation itself and if at step l of the pursuit, $\mathcal{C}_\Theta^{(l)} = \emptyset$ or $\mathcal{C}_\Theta^{(l)} = \mathcal{C}_\Theta^{(l-1)}$, then the level set pursuit terminates and the number of distinct motion regions is defined to be l. Although computationally more costly than basic level set pursuit, level set pursuit allows regions previously defined to change if necessary; a fundamental stability property of level set pursuit is that for a small enough value of λ, these regions will not disappear. Thus, motion regions previously found persist (with modifications) throughout the pursuit. In this sense therefore, the regions created by pursuit in this particular order are "stable". For simplicity, we prove this result for $\lambda = 0$, the proof for $0 < \lambda < \delta$ being similar albeit longer.

Proposition 1 *Assume $\lambda = 0$. Assume that at step l of level set pursuit, the image domain Ω has been segmented into the regions $\{\hat{\Omega}_1^{(l)}, \cdots, \hat{\Omega}_l^{(l)}, (\hat{\Omega}_1^{(l)} \cup \cdots \cup \hat{\Omega}_l^{(l)})^c\}$, $\hat{\Omega}_i^{(l)}$ being the estimate, at step l, of motion region Ω_i, $i = 1, \cdots, l$. Then, for $i = 1, \cdots, l$, the estimate $\hat{\Omega}_i^{(l+1)}$ of the region Ω_i obtained at step $l+1$ of level set pursuit is non-empty.*

Proof. Consider the i^{th} region, $1 \le i \le l$. Then, by construction of the algorithm, there exists a quadruple $(\boldsymbol{x}_i, \boldsymbol{U}_i, \epsilon_i, \widehat{\boldsymbol{T}}_i) \in \mathcal{C}_\Theta^{(l)}$ with ϵ_i the minimum error over all transformations in $\Theta_\mathcal{C}$ and $\boldsymbol{U}_i \subset \hat{\Omega}_i^{(k)}$, for $k \le l$. Thus there exists a non-empty open set $\boldsymbol{U} \subset \boldsymbol{U}_i$ such that $\forall \boldsymbol{x} \in \boldsymbol{U}, \xi(\widehat{\boldsymbol{T}}_i, \boldsymbol{x}) \le \xi(\boldsymbol{T}, \boldsymbol{x})$ for all $\boldsymbol{T} \in \Theta_\mathcal{C}$. It follows from the level set evolution equations that $\frac{\partial u_i(\boldsymbol{x},t)}{\partial t} \ge 0, \forall \boldsymbol{x} \in \boldsymbol{U}$, and hence the function $t \mapsto u_i(\boldsymbol{x}, t)$ is non-decreasing, $\forall \boldsymbol{x} \in \boldsymbol{U}$. Since $\boldsymbol{U} \subset \hat{\Omega}_i^{(l)}$ by assumption, we therefore have that $\boldsymbol{U} \subset \hat{\Omega}_i^{(l+1)}$ as well. This proves that $\hat{\Omega}_i^{(l+1)}$ is non-empty.

3. EXPERIMENTAL RESULTS

We illustrate our level set pursuit algorithm on the *pingpong* sequence, with $\sigma^2 = 1000$ and $\lambda = 1500$. Figs. (1a) and (1b) show frames 30 and 40, respectively, of this sequence. Fig. (1c) shows the given correspondence vectors. Figs. (1a), (1b), and (1c) thus constitute the data from which level set pursuit is made and a motion-based segmentation of Fig. (1a) is obtained. The family $\mathcal{C}$ consists here of 12 correspondence points (4 of which are not seen in Fig. (1c) for the corresponding vectors are null). Letting Θ to be the group of affine plane transformations, the subset $\Theta_\mathcal{C}$ of Θ is constructed from $\mathcal{C}$ and contains 495 transformations. The subsets $\mathcal{C}_\Theta^{(1)}$ and $\mathcal{C}_\Theta^{(2)}$ of $\mathcal{C}$ computed during pursuit are non-empty, while $\mathcal{C}_\Theta^{(3)} = \emptyset$; the number of distinct motion regions is then 3, and Fig. (1d) shows the zero level set curves obtained at termination of level set pursuit. Fig. (2a) shows the first motion region found by the algorithm; this corresponds to the background, which happens to correspond to the identical transformation. The second motion region found by the algorithm is shown in Fig. (2b); this corresponds to the ball, and has mainly upward translational motion. Finally, the third motion region found by the algorithm is the player's arm, which has mainly a counter-clockwise rotational motion around the elbow, and is shown in Fig. (2c). It is important to note that all these distinct motion regions have been precisely delineated by the algorithm despite the fact that no use was made of intensity boundaries and that only motion information was used in the segmentation. It is also important to note that apart from occlusion regions, all of the image domain has been covered by the pursuit, as was expected. Fig. (2d) shows the dense motion field obtained (with a subsampling of 8 for better visibility) by level set pursuit: Note that the ball's translation and the arm's rotation are evident in this motion field.

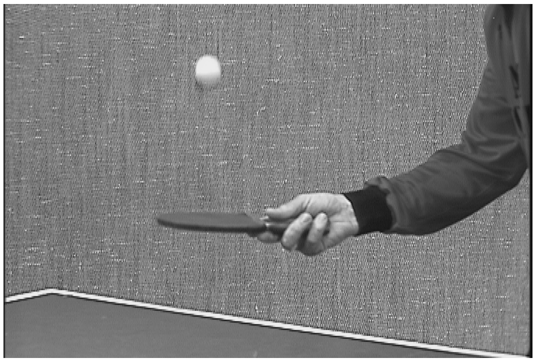
4. REFERENCES

[1] N. Diehl, "Object-oriented motion estimation and segmentation in image sequences," *Signal Process. Image Commun.,* vol. 3, no. 1, pp. 23-56, 1991.

[2] A.-R. Mansouri, A. Mitiche, R. Feghali, "Spatio-Temporal Motion Segmentation Via Level Set Partial Differential Equations," *Proc. of the* 5^{th} *IEEE Southwest Symposium on Image Analysis and Interpretation,* Santa Fe, April 2002.

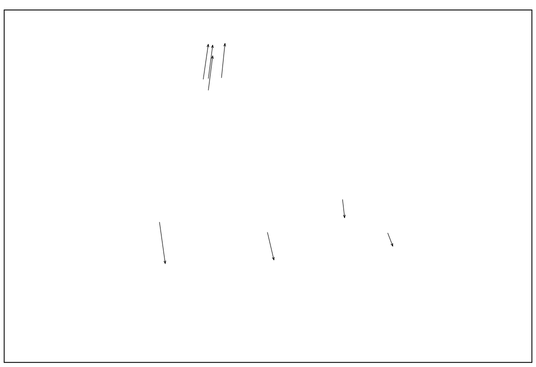
[3] A.-R. Mansouri, J. Konrad, "Multiple motion segmentation with level sets," submitted to *IEEE Transactions on Image Processing,* April 2000 (to appear).

[4] J.A. Sethian, *Level Set Methods,* Cambridge University Press, 1996.

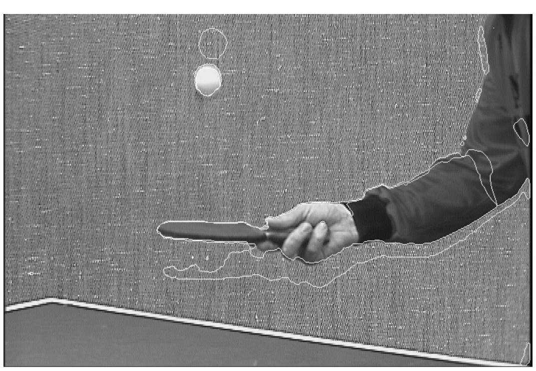


(a) frame 30 of *pingpong* sequence (I^0)

(b) frame 40 of *pingpong* sequence (I^1)

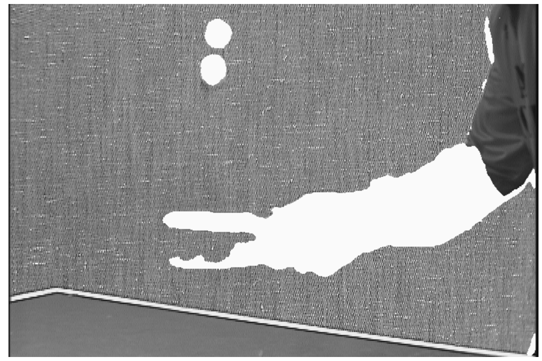


(c) point correspondences ($\mathcal{C}$)



(d) zero level sets after pursuit ($\{u_i(.,\infty) = 0\}_{i=1}^{3}$)

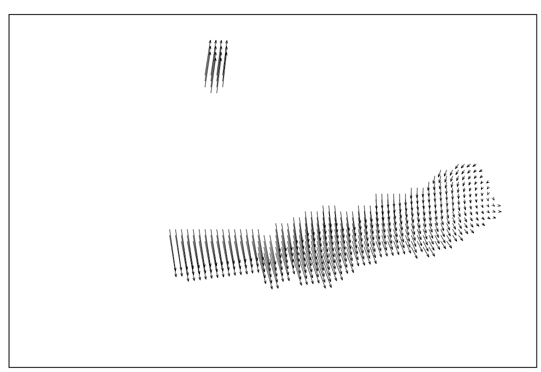
Fig. 1.



(a) background ($\hat{\Omega}_1$)

(b) ball ($\hat{\Omega}_2$)

(c) arm ($\hat{\Omega}_3$)



(d) motion field corresponding to $\{(\hat{\Omega}_i, \widehat{\boldsymbol{T}}_i)\}_{i=1}^{3}$

Fig. 2.