

STATISTICAL DEFORMABLE MODEL-BASED SEGMENTATION OF IMAGE MOTION

Charles KERVRANN † and Fabrice HEITZ ‡

† IRISA/INRIA
Campus Universitaire de Beaulieu
35042 Rennes Cedex, France
kervrann@irisa.fr

‡ ENSPS/LSIIT UPRES-A CNRS 7005
Boulevard Sébastien Brant
67400 Illkirch, France
Fabrice.Heitz@ensps.u-strasbg.fr

Abstract

We present a statistical method for the motion-based segmentation of deformable structures undergoing non-rigid movements. The proposed approach relies on two models describing the shape of interest, its variability and its movement. The first model corresponds to a statistical deformable template that constrains the shape and its deformations. The second model is introduced to represent the optical flow field inside the deformable template. These two models are combined within a single probability distribution which enables to derive optimal shape and motion estimates using a Maximum Likelihood approach. The method requires no manual initialization and is demonstrated on synthetic data and on a medical X-ray image sequence.

EDICS number : IP1 IMAGE PROCESSING

- IP1.11 MOTION DETECTION AND ESTIMATION
- IP1.12 IMAGE SEQUENCE PROCESSING

Author to contact :

First (Personal) Name: Charles
Last (Family) Name: KERVRANN
Address: IRISA/INRIA
TEMIS project
Campus universitaire de Beaulieu
35042 Rennes Cedex, France
Telephone number: (33 +) 2 99 84 74 29
Fax number: (33 +) 2 99 84 71 71
E-mail address: kervrann@irisa.fr

STATISTICAL DEFORMABLE MODEL-BASED SEGMENTATION OF IMAGE MOTION

Abstract

We present a statistical method for the motion-based segmentation of deformable structures undergoing non-rigid movements. The proposed approach relies on two models describing the shape of interest, its variability and its movement. The first model corresponds to a statistical deformable template that constrains the shape and its deformations. The second model is introduced to represent the optical flow field inside the deformable template. These two models are combined within a single probability distribution which enables to derive optimal shape and motion estimates using a Maximum Likelihood approach. The method requires no manual initialization and is demonstrated on synthetic data and on a medical X-ray image sequence.

1 Introduction

The segmentation of motion information from visual input is an important preliminary task in most dynamic image analysis problems. In an increasing number of application fields – biomedical image analysis for instance – the structures to be modeled undergo deformations which have to be analyzed and characterized. Deformable models have been introduced to incorporate geometric information about shapes and their variability [4, 7, 12, 17]. Deformable model-based segmentation schemes usually rely on spatial gradient information, related to intensity edges [2, 4, 5, 11, 12, 13, 17] or texture [7, 18, 19, 22, 6], to extract and track object boundaries in image sequences. In several important application fields, however, the structures of interest are not delineated by sharp intensity edges, but are rather characterized by homogeneous motion fields (see for instance Fig. 6 herein). A deformable body may in this case be described by its boundary and the movement of its surface [16]. Motion-based segmentation schemes, relying on spatio-temporal image gradients and motion models, enable to extract such homogeneous moving regions. Motion-based segmentation is however known to be ill-conditioned and requires some regularization. Markov Random Field (MRF) models have for instance been used in [3, 10], but it is generally difficult to incorporate specific shape information in standard MRFs models. In [20], an elegant method for reconstructing the boundary of a moving object using piecewise constant motion field models and local shape optimization techniques has been described. However in the local shape optimization procedure, the deformable curve delimiting the object is not constrained, yielding noisy segmentations in real-world situations.

The major contribution of the present work is the definition of a unified framework that uses both motion field information and shape information to provide a motion-based segmentation of scenes containing non-rigid objects [14]. The regularization in the motion-based segmentation scheme is conducted by incorporating statistical constraints about the shape of interest. This enables robust image partitions. The shape representation relies on a statistical description of deformations applied to a prototype

shape (“template”) [13]. Following [4], deformations are modeled using a Karhunen-Loeve expansion of the distortions observed on a representative population. This model is used to constrain the *a priori* structure and variability of the shapes to be extracted. The model parameterization guarantees regularity of the represented shape and the introduction of an energy term which penalizes deformations is thus not necessary. Besides, within local regions in space and time, the optical flow field is described by parametric models [1, 3, 16]. We have adopted affine velocity field models, that provide a good trade-off between accuracy and complexity and also enable the representation of a large class of deformable motions, as demonstrated in [1]. In the proposed approach the affine model is used to partition the image into two regions: the inside of the deformable template and the outside of the template corresponding to the background. The segmentation problem is formulated in a statistical estimation framework where both shape and motion parameters are considered unknown. A Maximum Likelihood (ML) estimate of shape and linear motion parameters is derived. An *Adaptive Segmentation Algorithm* (ASA) [15] using a fast simulated annealing procedure [9] as a major component is implemented to perform this ML estimation. Our method combines the advantages of global optimization techniques with a compact description of deformations ; in particular, no human interaction is required to initialize the model. Robust and accurate motion-based segmentations have been obtained on synthetic data and on medical X-ray image sequences, as shown in the last section of this correspondence.

2 Deformable shape model

The statistical deformable model under concern here has been introduced by Cootes *et al.* in [4]. It is known, in the literature as the “Point Distribution Model” (PDM). The object of interest is represented by a “deformable template” which incorporates *a priori* knowledge on the structure of the object and on its variability. A particular shape \mathbf{x} is represented by a set of n labeled points (landmarks) which approximate its outline:

$$\mathbf{x}_k = [x_{k1}, y_{k1}, x_{k2}, y_{k2}, \dots, x_{kn}, y_{kn}]^T.$$

Following [4], the shape belonging to the learning population are normalized in scale, and aligned with respect to a common reference frame. To represent the deformations of the shape, a modal analysis technique described in [4, 13] is adopted. The deformations of shape \mathbf{x} are characterized by a displacement vector $d\mathbf{x} = \mathbf{x} - \bar{\mathbf{x}}$ with respect to a pre-computed mean shape (“template”) $\bar{\mathbf{x}}$. The main deformation modes of the template model \mathbf{X} , are then described by the eigenvectors Φ of the covariance matrix \mathbf{C} , with the largest eigenvalues [4]. A configuration of the deformable template is described by the following model [13]:

$$\mathbf{X} = \mathbf{M}(k, \theta) [\bar{\mathbf{x}} + \Phi \mathbf{b}] + \mathbf{T}. \tag{1}$$

- \mathbf{T} and $\mathbf{M}(k, \theta)$ account for rigid transformations of the template in the image plane (\mathbf{T} is a global translation vector, and $\mathbf{M}(k, \theta)$ performs a rotation by θ and a scaling by k),

- $\Phi = [\phi_1, \phi_2, \dots, \phi_m]$ is the matrix of the first m ($m < 2n$) eigenvectors associated to the m largest eigenvalues λ_i , $i = 1, \dots, m$ and $\mathbf{b} = [b_1, \dots, b_m]^T$ is a vector of weights controlling these m deformation modes.

A global configuration of the deformable template is thus described by $4 + m$ parameters corresponding to rigid transformations (4 parameters) and m modal amplitudes b_j , $j = 1, \dots, m$. Only 5 to 7 modes are generally necessary to stand for more than 90% of the variability observed on a typical population [4]. Matrix Φ and the mean vector $\bar{\mathbf{x}}$ are estimated *off-line* on a training population using standard statistical estimators [4].

Finally, a cubic B-spline shape $\mathbf{X}(u)$ (with $u \in [1, n]$), with n control points, is created from the PDM model. The B-spline curve is piecewise polynomial and is written as a linear combination of B-spline basis functions $\mathcal{B}_i(\cdot)$ with control points $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$:

$$\mathbf{X}(u) = \sum_{i=1}^n \mathcal{B}_i(u) X_i. \quad (2)$$

This representation enables to describe smooth real-world shapes rather realistically. Let us notice that in this approach, the deformable B-spline curve is parameterized by the $(4 + m)$ parameter vector $\Theta = [\mathbf{M}(k, \theta), \mathbf{T}, \mathbf{b}]$ rather than by its control points, as usually [2, 6].

Fig. 1 shows an example of reconstructed shapes corresponding to a $n = 30$ points PDM model of a heart left ventricle. The heart model, that will be used in section 5 for the segmentation of angiographic image sequences, has been trained on a set of 10 shapes extracted from a typical cardiac cycle (see section 5). The template deformations shown Fig. 1 are obtained by varying the first modal amplitude b_1 within suitable limits (i.e. in the range $[-3\sqrt{\lambda_1}, 3\sqrt{\lambda_1}]$, where λ_1 is the eigenvalue associated to the first mode of variation: most of the population is here assumed to lie within three standard deviations of the mean shape $\bar{\mathbf{x}}$).

3 Deformable motion model

Parameterized models of image motion make simplifying assumptions about the spatio-temporal variations of the optical flow field: the motion field is for instance represented by a low-order polynomial [1, 3, 16]. Within regions showing homogeneous movements, the following affine model is often a good approximation:

$$\mathbf{v}(s) = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \\ \mathbf{a}_4 & \mathbf{a}_5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \mathbf{a}_3 \\ \mathbf{a}_6 \end{bmatrix}, \quad (3)$$

where $\mathbf{a}_i, i = 1, \dots, 6$ designate the parameters of the affine model and $\mathbf{v}(s)$ represents the optical flow vector at spatial location $s = [x, y] \in S$. While affine models are popular for estimating motion of rigid objects, their application to non-rigid motion may appear unconventional. In fact, the parameters $\mathbf{a}_i, i = 1, \dots, 6$ have simple interpretations in terms of rigid and non-rigid image motion. \mathbf{a}_3 and \mathbf{a}_6

represent horizontal and vertical translations respectively. In addition, we can express the isotropic expansion (divergence), the rotation about the viewing direction and deformations including squashing and stretching [1, 3] as combinations of the \mathbf{a}_i :

$$\textit{divergence} = \mathbf{a}_1 + \mathbf{a}_5, \quad \textit{rotation} = \mathbf{a}_4 - \mathbf{a}_2, \quad \textit{deformation} = \mathbf{a}_1 - \mathbf{a}_5.$$

Assuming that the brightness $I(s)$ of a moving point is constant through time, i.e. $\frac{dI(s)}{dt} = 0$, the following classical image flow constraint equation is derived:

$$\nabla I(s) \cdot \mathbf{v}(s) + I_t(s) = 0, \quad (4)$$

where $\nabla I(s)$ and $I_t(s)$ are the spatial and temporal derivatives of the image $I(s)$ at site s . This relation does not strictly hold if noise, changes in lightning or surface orientation occur. Small deviations with respect to the constant brightness assumption may be described by introducing in the image flow constraint equation independent gaussian noise terms $n_v(s)$ and $n_{I_t}(s)$, as described in [21]:

$$\nabla I(s) \cdot [\mathbf{v}(s) - n_v(s)] = -I_t(s) + n_{I_t}(s). \quad (5)$$

The noise term $n_{I_t}(s)$ (with variance $\sigma_{I_t}^2$) describes errors in the temporal measurement. The second noise term $n_v(s)$ describes other error sources, resulting for instance from an inaccurate motion model. The following conditional probability is then easily derived at site s and time t [21]:

$$\mathbb{P}(I_t(s) | \mathbf{v}(s), \nabla I(s)) = \frac{1}{[2\pi(\sigma_v^2 \|\nabla I(s)\|^2 + \sigma_{I_t}^2)]^{\frac{1}{2}}} \exp - \frac{[\nabla I(s) \cdot \mathbf{v}(s) + I_t(s)]^2}{2[\sigma_v^2 \|\nabla I(s)\|^2 + \sigma_{I_t}^2]} \quad (6)$$

where we choose the covariance matrix of $n_v(s)$ to be diagonal with diagonal entry σ_v^2 [21]. We have adopted the following values, for the parameters : $\sigma_{I_t}^2 = 1$. and $\sigma_v^2 = 0.1$ (the final segmentation showed relatively insensitive to variations in these parameters).

4 Deformable motion-based segmentation

The image lattice S is partitioned into two distinct regions: the *inside* of the deformable B-spline model Γ_{Θ}^I and the *outside* Γ_{Θ}^O (background). Let $\mathbf{v}_{\mathcal{A}^I}(s) = \mathbf{S}(s) \cdot \mathcal{A}^I$ and $\mathbf{v}_{\mathcal{A}^O}(s) = \mathbf{S}(s) \cdot \mathcal{A}^O$ denote the optical flow vector at point s belonging to regions Γ_{Θ}^I and Γ_{Θ}^O respectively, where:

$$\mathbf{S}(s) = \begin{bmatrix} x & y & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x & y & 1 \end{bmatrix} \quad (7)$$

and \mathcal{A}^I and \mathcal{A}^O are the affine parameters corresponding to regions Γ_{Θ}^I and Γ_{Θ}^O :

$$\mathcal{A}^I = [\mathbf{a}_1^I, \mathbf{a}_2^I, \mathbf{a}_3^I, \mathbf{a}_4^I, \mathbf{a}_5^I, \mathbf{a}_6^I]^T \quad \text{and} \quad \mathcal{A}^O = [\mathbf{a}_1^O, \mathbf{a}_2^O, \mathbf{a}_3^O, \mathbf{a}_4^O, \mathbf{a}_5^O, \mathbf{a}_6^O]^T. \quad (8)$$

The ML estimate of the deformable model parameters Θ and of the affine motion parameters $(\mathcal{A}^I, \mathcal{A}^O)$ is defined by:

$$(\Theta^*, \mathcal{A}^{I*}, \mathcal{A}^{O*}) = \arg \max_{\Theta, \mathcal{A}^I, \mathcal{A}^O} \mathbb{P}(\mathbf{I}_t | \Theta, \mathcal{A}^I, \mathcal{A}^O, \nabla \mathbf{I}), \quad (9)$$

where $\nabla\mathbf{I}$ and \mathbf{I}_t designate spatial and temporal image derivative vectors. Under standard independence assumptions, the conditional probability of \mathbf{I}_t may be expressed as:

$$P(\mathbf{I}_t | \Theta, \mathcal{A}^I, \mathcal{A}^O, \nabla\mathbf{I}) = \frac{1}{Z} \exp - \left[\sum_{s \in \Gamma_{\Theta}^I} \frac{[\nabla I(s) \cdot v_{\mathcal{A}^I}(s) + I_t(s)]^2}{2[\sigma_v^2 \|\nabla I(s)\|^2 + \sigma_{I_t}^2]} + \sum_{s \in \Gamma_{\Theta}^O} \frac{[\nabla I(s) \cdot v_{\mathcal{A}^O}(s) + I_t(s)]^2}{2[\sigma_v^2 \|\nabla I(s)\|^2 + \sigma_{I_t}^2]} \right] \quad (10)$$

where Z is a normalization constant. Let us notice that, contrary to [11], the normalization constant Z does not depend on the parameters Θ, \mathcal{A}^I and \mathcal{A}^O of the model:

$$Z = \prod_{s \in S} [2\pi(\sigma_v^2 \|\nabla I(s)\|^2 + \sigma_{I_t}^2)]^{\frac{1}{2}}. \quad (11)$$

The computation of the exact ML estimate (9) is generally untractable. In practice, we estimate alternately the shape of the deformable structure controlled by Θ and the affine motion parameters $\mathcal{A}^I, \mathcal{A}^O$, yielding a *partial optimal solution*, as proposed by Lakshmanan *et al.* for unsupervised image segmentation [15]:

$$\begin{cases} \hat{\Theta} & = \arg \max_{\Theta} P(\mathbf{I}_t | \Theta, \hat{\mathcal{A}}^I, \hat{\mathcal{A}}^O, \nabla\mathbf{I}), \\ (\hat{\mathcal{A}}^I, \hat{\mathcal{A}}^O) & = \arg \max_{\mathcal{A}^I, \mathcal{A}^O} P(\mathbf{I}_t | \hat{\Theta}, \mathcal{A}^I, \mathcal{A}^O, \nabla\mathbf{I}). \end{cases} \quad (12)$$

An *Adaptive Segmentation Algorithm* (ASA) inspired by the parameter estimation algorithm described in [15] is used to compute such an estimate. The ASA procedure is basically a simulated annealing procedure over Θ , which is interrupted at regular intervals to get an ML estimate of \mathcal{A}^I and \mathcal{A}^O . For model (10), $\hat{\mathcal{A}}^I$ and $\hat{\mathcal{A}}^O$ simply correspond to standard *Linear Least-Squares estimates*. A detailed synopsis of the proposed segmentation algorithm, with the estimation formulae, are presented in Fig. 2.

The theoretical convergence of this algorithm to a *partial optimal solution* (Eq. 12) has been established in [15]. For theoretical convergence, the ASA algorithm must satisfy some constraints related to the temperature schedule [9, 15]. In our experiments, we have adopted a fast suboptimal exponential decreasing temperature schedule. Although this does not ensure convergence to the optimal estimate, it yields satisfactory final segmentations in practice.

In the next section, we present two experiments conducted on synthetic data and on a real medical image sequence.

5 Experimental results

Synthetic image sequence In order to illustrate the features of the deformable model and motion-based algorithm, we first present the segmentation of a synthetic sequence composed of 20 (256×256) frames showing a parameterized curve, corresponding to a superquadric, undergoing a non-rigid motion in a virtual room. In this sequence, the camera is translating to the left and to the bottom of the room, inducing a translation on the background (Fig. 4). The superquadric in the foreground translates to the right and moves parallel to the axis of view (hence a dilatation is observed). In addition, the

parameterized curve changes into a sphere over the sequence. The superquadric animated by such a non-rigid motion can be considered as a deformable object to be extracted from a moving background.

In our experiments, the superquadric has been modeled using a 12 point PDM model describing its outline (see Fig. 3). This model has been computed from a training set of 6 configurations selected in the sequence. $m = 2$ deformation modes enabled here to take into account more than 95 % of the observed variability. Fig. 3 shows the first deformation mode captured by the trained deformable model. The contours estimated by the motion-based segmentation algorithm are superimposed on six typical images of the sequence in Fig. 4. The initial configuration of the deformable template is defined at random on the first image of the sequence. The motion-based segmentation algorithm was able to extract and track the synthetic object reliably over the whole 20 image sequence, by initializing the deformable model configuration in the current frame by the final estimate of Θ obtained on the previous frame (Fig. 4). The solution provides a good fit to the visible contours of the deformable object, as can be seen in Fig. 4. Fig. 5 depicts the piecewise affine optical flow fields that were obtained for the two frames corresponding to Fig. 4c and Fig. 4f. These results demonstrate the ability of the algorithm to provide an accurate partition of the image, using only motion information (let us recall that intensity edge information is not used here). CPU time is about 2 mn on a SUNSPARC20 workstation to process a 256×256 frame.

Real medical image sequence The performance of the motion segmentation algorithm is also assessed on a medical X-ray image sequence showing the movement of the left ventricle of a heart. Left ventricular angiography is an important procedure in the evaluation of many cardiac diseases, providing valuable information concerning anatomic and functional aspects. X-Ray images generally show poor intensity edges (Fig. 6) and consequently are difficult to partition using standard intensity-based segmentation algorithms [2, 4, 5, 7]. When available, motion information may thus provide an useful alternative (or additional) clue for the segmentation task.

In our experiments, $n = 30$ landmarks have been selected manually by an expert clinician on a representative population of 10 left ventricles extracted from a cardiac cycle sequence and $m = 5$ deformation modes that take into account 95 % of the observed variability, have been considered in this case. This statistical knowledge has been integrated in the model as *a priori* knowledge on the shape and deformations of the beating heart.

The motion-based segmentation algorithm was able to track the left ventricle over the whole 40 image sequence, by initializing, as previously, the deformable model in the current frame by the shape parameters Θ estimated on the previous frame (Fig. 6a-c). The deformable model was also initialized *at random* on the first frame, yielding a completely data-driven segmentation algorithm.

Manual segmentations of the left ventricle (Fig. 6d-f), performed by an expert physician were used to evaluate the ability of the algorithm to provide suitable results, as far as this application is concerned. As can be seen on Fig. 6, the segmentations obtained by the algorithm are close to the partitions

defined by the human expert: motion is a relevant clue for the segmentation in this case. Fig. 7a and Fig. 7b show the piecewise affine optical flow fields associated to the segmentation maps appearing Fig. 6a and Fig. 6b. In Fig.7, a motion activity is clearly observed in the cardiac cavity whereas the background tissues are approximately static. The optimization procedure based on the Gibbs sampler dynamics requires CPU times of less than 5 mn for one frame, on a SUNSPARC20 workstation. Simulated annealing is used here with an acceptable cost, thanks to the reduced number of parameters ($4+m$) to be estimated. This performance may be compared to the computationally demanding schemes associated to the stochastic approaches described in [7, 8, 22].

6 Conclusion

In this correspondence, we have presented a deformable template and motion-based segmentation method for extracting and tracking non-rigid structures in image sequences. Our approach combines statistical constraints on the shape, deformations and velocity fields models characterizing the movement of the deformable structure. The segmentation problem is formulated as a statistical parameter estimation problem. The maximization of the likelihood function is implemented by a data-driven algorithm that performs simultaneous shape and motion parameter estimation. The proposed modeling and algorithmic framework yields promising future prospects for the characterization of the dynamic behavior of deformable objects. The main limitation of the approach is that it enables to extract only one deformable object from the sequence. For the future, we plan to address the segmentation problem of multiple objects [8]. A multiple object extension of the previously described scheme is mathematically straightforward, when there are no interactions (e.g. occlusions) between multiple objects and when the number of objects and the object classes are known *a priori*. Coping with situations where this knowledge is not available is an intricate issue, which has been addressed recently by Grenander *et al.* in [8].

References

- [1] M.J. Black, and Y. Yacoob. – Tracking and recognizing rigid and non-rigid facial motions using local parametric models of image motion. – In *Proc. Int. Conf. Computer Vision*, pages 374-381, Boston, USA, June 1995. –
- [2] A. Blake, R. Curwen and A. Zisserman. – A framework for spatiotemporal control in the tracking of visual contours. – *Int. J. Computer Vision*, Vol. 11, No 2: pp. 127–145, 1993. –
- [3] P. Bouthemy and E. Francois. – Motion segmentation and qualitative dynamic scene analysis from an image sequence. – *Int. J. Computer Vision*, Vol. 10, No 2: pp. 157–182, 1993. –
- [4] T. Cootes, C.J. Taylor, D.H. Cooper and J. Graham. – Active shape models - Their training and application. – *CVGIP: Image Understanding*, Vol. 1, No 1: pp. 38–59, 1994. –
- [5] M.T. Figueiredo and J.M.N. Leitaó. – Bayesian estimation of ventricular contours in angiographic images. – *IEEE trans. on Medical Imaging*, Vol. 11, No 3: pp. 416–429, 1992. –

- [6] M.A.T. Figueiredo, J.M.N. Leitao and A.K. Jain. – Adaptive B-splines and boundary estimation. – In *Proc. Conf. Computer Vision Pattern Recognition*, pages 724–730, San Juan, Puerto Rico, USA, June 1997. –
- [7] U. Grenander, Y. Chow and D.M. Keenan. – Hands. A Pattern Theoretic Study of Biological Shapes. – Springer, New-York, 1991. –
- [8] U. Grenander and M.I. Miller. – Representations of knowledge in complex systems. – *J. R. Statist. Soc. B*, Vol. 56, No 3: pp. 549–603, 1994. –
- [9] S. Geman and D. Geman. – Stochastic relaxation, Gibbs distributions and the bayesian restoration of images. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 6, No 6: pp. 721–741, 1984. –
- [10] F. Heitz and P. Bouthemy, “Multimodal estimation of discontinuous optical flow using Markov Random Fields,” *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 15, No. 12, pp. 1217–1232, 1993.
- [11] A.K. Jain, Y. Zhong and S. Lakshmanan. – Object matching using deformable templates. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 18, No 3: pp. 267–278, 1996. –
- [12] M. Kass, A. Witkin and D. Terzopoulos. – Snakes: active contour models. – *Int. J. Computer Vision*, Vol. 1, pp. 259–268, 1987. –
- [13] C. Kervrann and F. Heitz. – A Hierarchical Statistical Framework for the Segmentation of Deformable Objects in Image Sequences. – In *Proc. Conf. Computer Vision Pattern Recognition*, pages 724–728, Seattle, USA, June 1994. –
- [14] C. Kervrann and F. Heitz. – Statistical model-based segmentation of deformable motion. – In *Proc. Int. Conf. Image Processing*, pages 937–940, Lausanne, Switzerland, September 1996. –
- [15] S. Lakshmanan and H. Derin. – Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 11, No 8: pp. 799–813, 1989. –
- [16] F. Meyer, R.T. Constable, A.J. Sinusas and J.S. Duncan. – Tracking myocardial deformation using spatially-constrained velocities. – In *Proc. Int. Conf. on Information Processing in Medical Imaging*, pages 177–188, Brest, France, June 1995.
- [17] C. Nastar and N. Ayache. – Frequency-based nonrigid motion analysis: application to four dimensional medical images. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 18, No 11: pp. 1067–1079, 1996. –
- [18] Ph. Refregier, O. Germain and T. Gaidon. – Optimal snake segmentation of target and background with independent Gamma density probabilities, application to speckled and preprocessed images. – *Opt. Com.*, Vol. 137, pp. 382-388, 1997. –
- [19] R. Ronfard. – Region-based strategies for active contour models. – *Int. J. Computer Vision*, Vol. 13, pp. 229–251, 1994. –
- [20] C. Schnorr. – Computation of discontinuous optical flow by domain decomposition and shape optimization. – *Int. J. Computer Vision*, Vol. 8, No 2, pages 153–165, 1992. –
- [21] E. Simoncelli, E. Adelson and D.J. Heeger. – Probability distributions of optical flow – In *Proc. Conf. Computer Vision Pattern Recognition*, pages 310–315, Maui, Hawaii, USA, May 1991. –
- [22] G. Storvik. – A bayesian approach to dynamic contours through stochastic sampling and simulated annealing. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 16, No 10: pp. 976–986, 1994. –

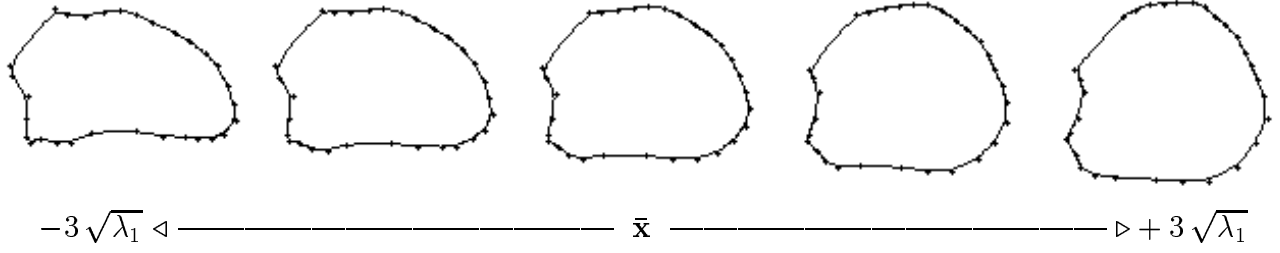


Figure 1: Deformations captured by the first mode of a heart ventricle model.

| | |
|--|---|
| Algorithm MOTION-BASED SEGMENTATION $(T_0, c, \Theta^0, P(\cdot), \nabla I, \mathbf{I}_t)$ | |
| T_0 | <i>initial temperature</i> |
| c | <i>annealing parameter ($0 < c < 1$)</i> |
| Θ^0 | <i>initial (random) configuration</i> |
| $P(\mathbf{I}_t \Theta, \mathcal{A}^I, \mathcal{A}^O, \nabla I)$ | <i>likelihood function</i> |
| $(\nabla I, \mathbf{I}_t)$ | <i>measurements</i> |
| <ul style="list-style-type: none"> • $T \leftarrow T_0$; • <i>counter</i> $\leftarrow 0$; • $\hat{\Theta} \leftarrow \Theta^0$; • initialization of $\hat{\mathcal{A}}^I$ and $\hat{\mathcal{A}}^O$ using Eq. (13) and (14) ; • while <i>counter</i> $\geq \tau$ repeat : <ul style="list-style-type: none"> – Run a Gibbs sampler at temperature T to estimate $\hat{\Theta}$ given $P(\mathbf{I}_t \Theta, \hat{\mathcal{A}}^I, \hat{\mathcal{A}}^O, \nabla I)$ – Update $(\hat{\mathcal{A}}^I, \hat{\mathcal{A}}^O)$ given $\hat{\Theta}$ according to: <div style="text-align: center; margin-top: 10px;"> $\hat{\mathcal{A}}^I = \arg \min_{\mathcal{A}^I} \sum_{s \in \Gamma_{\hat{\Theta}}^I} \frac{(\nabla I(s) \cdot v_{\mathcal{A}^I}(s) + I_t(s))^2}{\sigma_v^2 \ \nabla I(s)\ ^2 + \sigma_{I_t}^2}, \quad (13)$ </div> <div style="text-align: center; margin-top: 10px;"> $\hat{\mathcal{A}}^O = \arg \min_{\mathcal{A}^O} \sum_{s \in \Gamma_{\hat{\Theta}}^O} \frac{(\nabla I(s) \cdot v_{\mathcal{A}^O}(s) + I_t(s))^2}{\sigma_v^2 \ \nabla I(s)\ ^2 + \sigma_{I_t}^2}. \quad (14)$ </div> – Decrease temperature: $T \leftarrow c.T$ and increment <i>counter</i> ; • end while ; | |
| end of algorithm. | |

Figure 2: Motion-based segmentation algorithm.

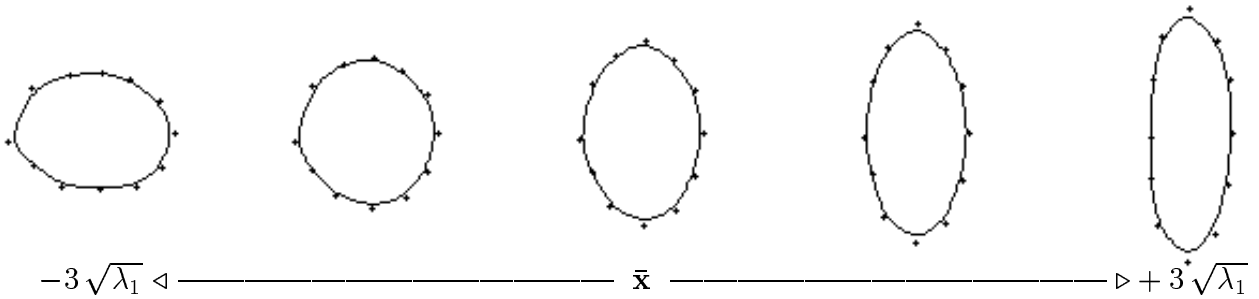


Figure 3: Deformation captured by the first deformation mode (superquadric model).

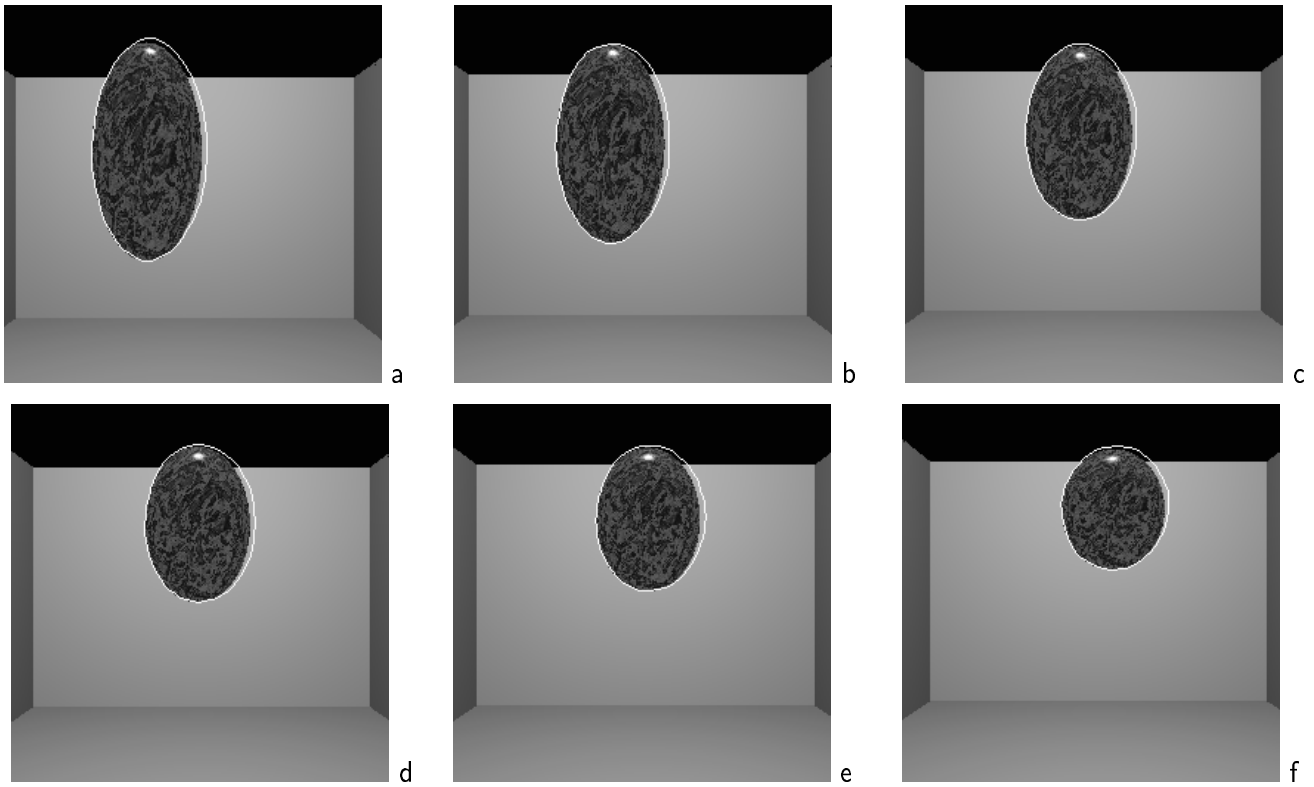


Figure 4: Motion-based segmentation of the moving superquadric object (256×256 images).

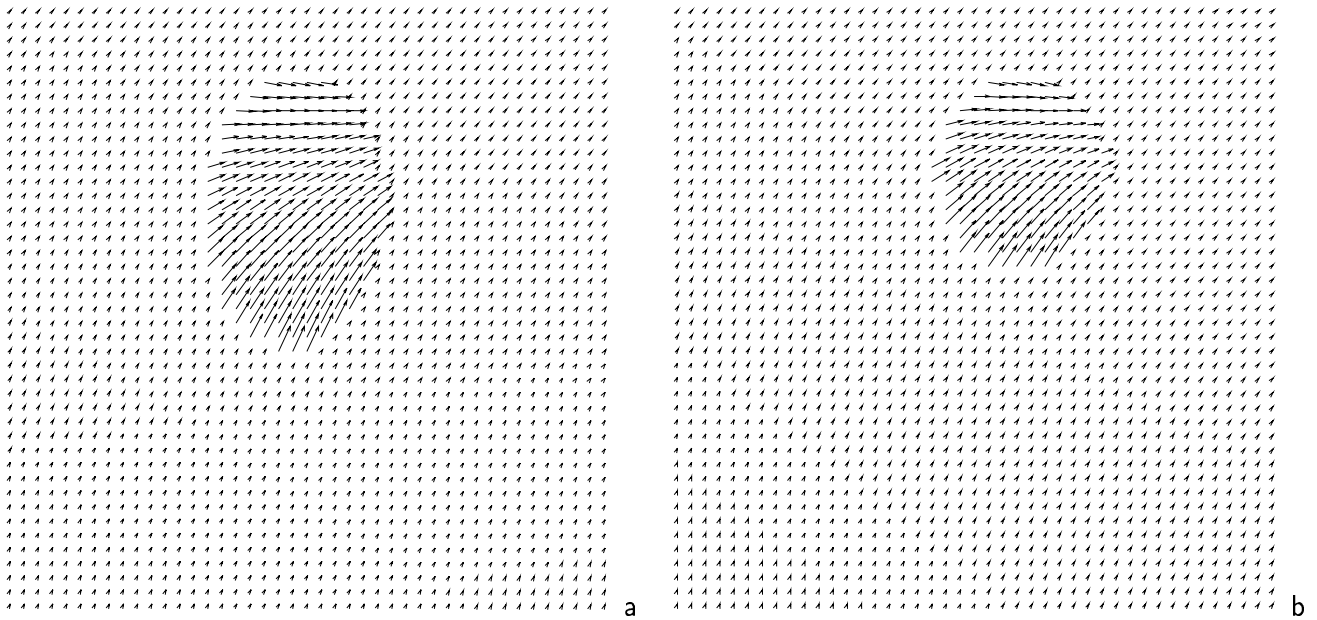


Figure 5: Piecewise affine optical flows fields associated to Fig. 4c and Fig. 4f (horizontally and vertically subsampled by 6 – flow vectors are magnified by a factor 7).

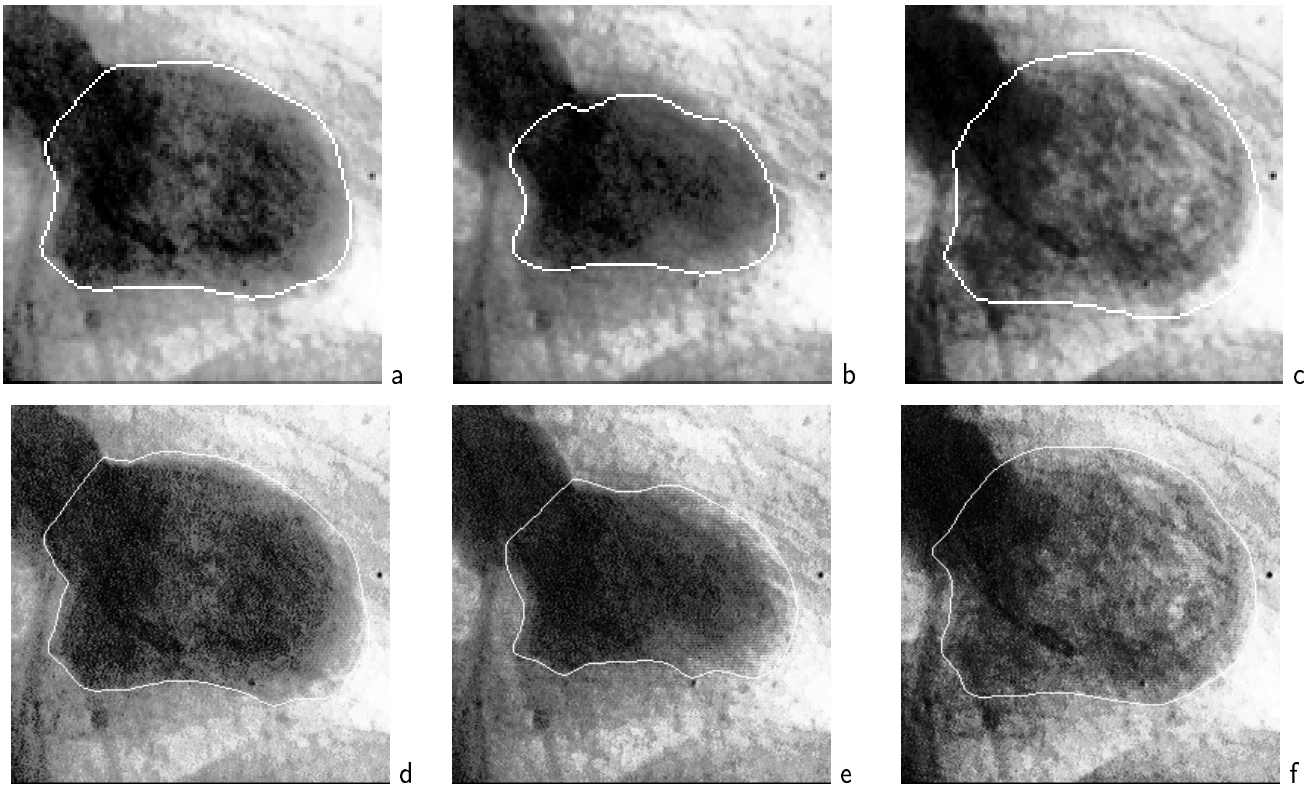


Figure 6: Deformable model-based segmentation of the left ventricle movement (X-ray image sequence, 256×256 images) ; a-c) results of the data-driven segmentation algorithm ; d-f) results provided by an expert physician (by courtesy of *LTSI - Rennes I University, France*).

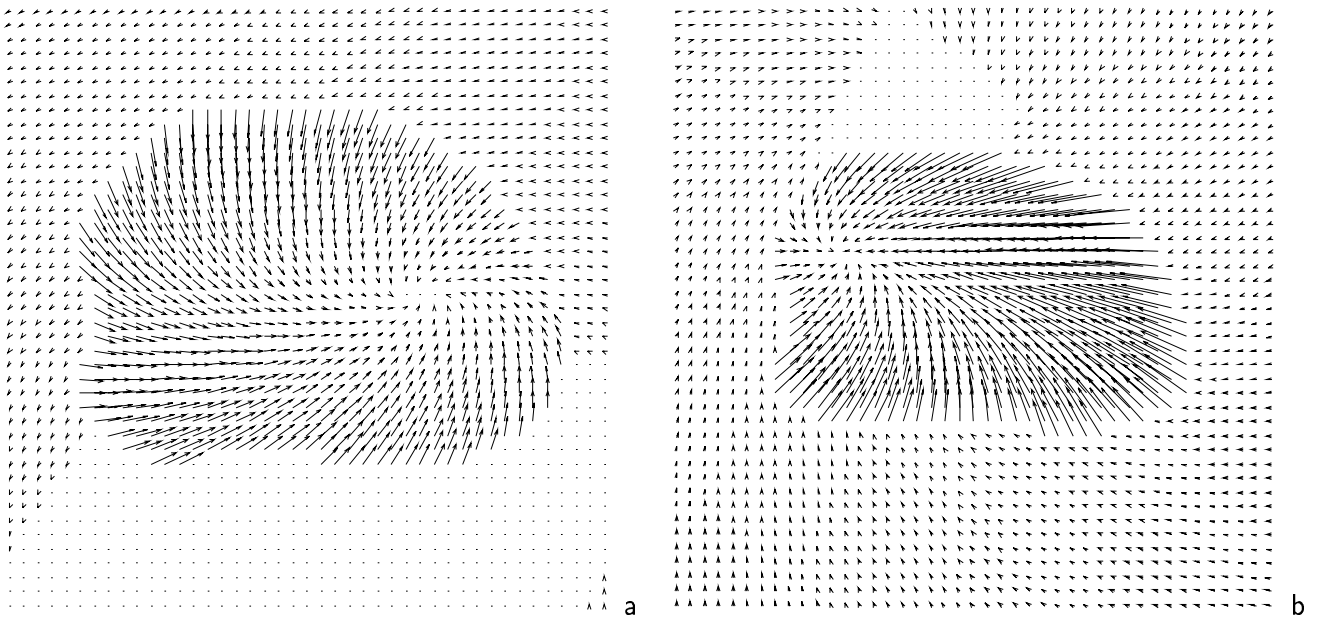


Figure 7: Piecewise affine optical flows fields associated to Fig. 6a and Fig. 6b (horizontally and vertically subsampled by 6 – flow vectors are magnified by a factor 4).