IEEE INT. CONF. ON IMAGE PROCESSING, LAUSANNE, SWITZERLAND, SEPTEMBER 1996

# STATISTICAL MODEL-BASED SEGMENTATION OF DEFORMABLE MOTION

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

*Fabrice HEITZ*
ENSPS/LSIIT-URA CNRS 1871
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 here on medical X-ray image sequences.

## 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 objects 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. Deformable model-based segmentation schemes usually use spatial gradient information, related to intensity edges, to extract and track object boundaries in image sequences [3, 5, 6, 10]. 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. 1). Motion-based segmentation schemes, relying on spatio-temporal image gradients and motion models, enable to extract such homogeneous regions. Motion-based segmentation is however known to be ill-conditioned and requires some regularization. A Markov Random Field (MRF) model is for instance used in [4]. It is unfortunately difficult to incorporate shape information in standard MRFs models [4, 7].

In this paper, the regularization in the motion-based segmentation scheme, is conducted by incorporating statistical constraints about shape, yielding robust image partitions. The shape representation relies on a statistical description of deformations applied to a prototype shape ("template") [8, 10]. Deformations are modeled using a Karhunen-Loeve (KL) expansion of the distorsions observed on a representative population [5]. This model is used to constrain the *a priori* structure and variability of the shapes to be extracted. Besides, within local regions in space and time, the optical flow field is described by linear models providing a trade-off between accuracy and concision of the representation [2, 4, 12]. These models are used to partition the image into two regions: the inside of the deformable template and the outside of the template corresponding to the background. A Maximum Likelihood (ML) estimate of shape and linear motion parameters is derived. An *Adaptive Segmentation Algorithm* (ASA) using a simulated annealing procedure as a major component is implemented to perform this ML estimation. The 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 using this method on X-ray medical image sequences, as shown in the last section of this paper.

## 2. DEFORMABLE SHAPE MODEL

The statistical deformable model under concern here has been introduced by the authors in [8]. The object of interest is represented by a "deformable template" which incorporates *a priori* knowledge on the structure of the object and its variability. A particular shape $\mathbf{x}$ is represented by a set of $n$ labeled points (landmarks) which approximate its outline. A cubic B-spline shape representation, with $n$ control points corresponding to the "landmarks", is defined ; it describes most real-world shapes rather realistically [3, 1].

To represent the deformations of the shape, a modal analysis technique described by Cootes *el al.* [5] is adopted. The deformations of shape $\mathbf{x}$ are characterized by a displacement vector $\mathbf{dx} = \mathbf{x} - \bar{\mathbf{x}}$ with respect to a pre-computed mean shape ("template") $\bar{\mathbf{x}}$. A KL expansion of the displacement vectors observed on a representative population allows to obtain a good representation of deformations on a low dimension eigenspace [5, 8]. Five to ten parameters are usually adequate to obtain an accurate description of deformations. If $\mathbf{\Phi}$ designates the matrix of the $m$ unit eigenvectors corresponding to the $m$ largest eigenvalues, and if $\mathbf{b}$ denotes the modal amplitudes vector ($m \times 1$) corresponding to the deformation parameters associated to the $m$ most significant deformation modes, the deformable template is represented by [8]:

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

Global transformations from the similarity group (rotation of angle $\theta$, scale change by a factor $k$ and translation $\mathbf{T}$) are taken into account in this model. Matrix $\mathbf{\Phi}$ and mean vector $\bar{\mathbf{x}}$ are estimated *off-line* using the KL expansion. In our approach, the deformable B-spline curve is thus parametrized by the $(4+m)$ parameter vector $\mathbf{\Theta} = [\,\mathbf{M}(k,\theta),\;\mathbf{T},\;\mathbf{b}\,]$ and not by its control points as done usually [1].

## 3. DEFORMABLE MOTION MODEL

Parametrized 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, 2, 4]. Within regions showing homogeneous movements, the following affine model is often a good approximation:

$$\mathrm{v(s)} \;=\; \begin{bmatrix} \mathsf{a}_1 & \mathsf{a}_2 \\ \mathsf{a}_4 & \mathsf{a}_5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \;+\; \begin{bmatrix} \mathsf{a}_3 \\ \mathsf{a}_6 \end{bmatrix}, \qquad (2)$$

where $\mathsf{a}_i, i = 1,\ldots,6$ designate the parameters of the affine model and $\mathrm{v(s)}$ represents the optical flow vector at spatial location $s = [\,x,\;y\,]$.

The standard gradient-based formulation of the optical flow measurement problem is based on the assumption that the brightness $I$ of a moving point is constant through time, i.e. $\frac{dI}{dt} = 0$, leading to the following classical image flow constraint equation:

$$\nabla I\,(s) \cdot \mathrm{v}(s) \;+\; I_t\,(s) \;=\; 0, \qquad (3)$$

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 $n_{\mathrm{v}}(s)$ and $n_{I_t}(s)$ (with variance $\sigma^2_{I_t}$), as described in [13]:

$$\nabla I\,(s) \cdot (\mathrm{v}(s) \;+\; n_{\mathrm{v}}(s)) \;\;=\;\; -\,I_t\,(s) \;+\; n_{I_t}(s). \qquad (4)$$

The following conditional probability is easily derived at site $s$ and time $t$:

$$\mathsf{P}\,(I_t(s)\,|\,\mathrm{v}(s), \nabla I(s)) \;\;\propto\;\; \exp - \frac{(\,\nabla I(s)\cdot \mathrm{v}(s) + I_t(s))^2}{\sigma^2_{\mathrm{v}}\,\|\nabla I(s)\|^2 \;+\; \sigma^2_{I_t}}, \qquad (5)$$

where we choose the covariance matrix of $n_{\mathrm{v}}(s)$ to be diagonal with diagonal entry $\sigma^2_{\mathrm{v}}$ [13].

## 4. DEFORMABLE MOTION-BASED SEGMENTATION

The image is partitioned into two distinct regions: the inside of the deformable B-spline model $\Gamma^I_{\mathbf{\Theta}}$ and the background $\Gamma^O_{\mathbf{\Theta}}$. Let $\mathrm{v}^I_{\mathcal{A}}(s)$ and $\mathrm{v}^O_{\mathcal{A}}(s)$ denote the optical flow vector at point $s$ belonging to regions $\Gamma^I_{\mathbf{\Theta}}$ and $\Gamma^O_{\mathbf{\Theta}}$ respectively:

$$\mathrm{v}^I_{\mathcal{A}}(s) \;=\; \mathbf{S}(s)\cdot \mathcal{A}^I, \qquad (6)$$
$$\mathrm{v}^O_{\mathcal{A}}(s) \;=\; \mathbf{S}(s)\cdot \mathcal{A}^O, \qquad (7)$$

where:
$$\mathbf{S}(s) \;=\; \begin{bmatrix} x & y & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x & y & 1 \end{bmatrix} \qquad (8)$$

and $\mathcal{A}^I$ and $\mathcal{A}^O$ are the affine parameters corresponding to regions $\Gamma^I_{\mathbf{\Theta}}$ and $\Gamma^O_{\mathbf{\Theta}}$:

$$\mathcal{A}^I \;=\; [\,\mathsf{a}^I_1, \mathsf{a}^I_2, \mathsf{a}^I_3, \mathsf{a}^I_4, \mathsf{a}^I_5, \mathsf{a}^I_6\,]^T, \qquad (9)$$
$$\mathcal{A}^O \;=\; [\,\mathsf{a}^O_1, \mathsf{a}^O_2, \mathsf{a}^O_3, \mathsf{a}^O_4, \mathsf{a}^O_5, \mathsf{a}^O_6\,]^T. \qquad (10)$$

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

$$(\mathbf{\Theta}^*, \mathcal{A}^{I^*}, \mathcal{A}^{O^*}) = \arg \max_{\mathbf{\Theta},\,\mathcal{A}^I,\,\mathcal{A}^O} \mathsf{P}\,(\mathbf{I}_t\,|\,\mathbf{\Theta}, \mathcal{A}^I, \mathcal{A}^O, \nabla \mathbf{I}), \qquad (11)$$

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

$$\mathsf{P}\ (\mathbf{I}_t \,|\, \boldsymbol{\Theta}, \mathcal{A}^I, \mathcal{A}^O, \nabla\mathbf{I}) \ \propto \ \exp - \sum_{s\in\Gamma^I_{\boldsymbol{\Theta}}} \frac{(\nabla I(s)\cdot \mathrm{v}^I_{\mathcal{A}}(s) + I_t(s))^2}{\sigma_{\mathrm{v}}^2\,\|\nabla I(s)\|^2 \ + \ \sigma_{I_t}^2} \cdot \exp - \sum_{s\in\Gamma^O_{\boldsymbol{\Theta}}} \frac{(\nabla I(s)\cdot \mathrm{v}^O_{\mathcal{A}}(s) + I_t(s))^2}{\sigma_{\mathrm{v}}^2\,\|\nabla I(s)\|^2 \ + \ \sigma_{I_t}^2}. \quad (12)$$

where the partition function does not depend on parameters $\boldsymbol{\Theta}, \mathcal{A}^I$ and $\mathcal{A}^O$.

The computation of the exact ML estimate (11) is usually untractable. In practice we estimate alternately the shape of the deformable structure controlled by $\boldsymbol{\Theta}$ 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 [11]:

$$\begin{cases} (\boldsymbol{\Theta}^*) & = \ \arg\max_{\boldsymbol{\Theta}} \mathsf{P}\,(\mathbf{I}_t \,|\, \boldsymbol{\Theta}, \mathcal{A}^{I^*}, \mathcal{A}^{O^*}, \nabla\mathbf{I}), \\ (\mathcal{A}^{I^*}, \mathcal{A}^{O^*}) & = \ \arg\max_{\mathcal{A}^I,\, \mathcal{A}^O} \mathsf{P}\,(\mathbf{I}_t \,|\, \boldsymbol{\Theta}^*, \mathcal{A}^I, \mathcal{A}^O, \nabla\mathbf{I}). \end{cases} \quad (13)$$

The *Adaptive Segmentation Algorithm* (ASA) described in [11] is used to compute such an estimate. The ASA procedure is basically a simulated annealing procedure over $\boldsymbol{\Theta}$, which is interrupted at regular intervals to get an ML estimate of $\mathcal{A}^I$ and $\mathcal{A}^O$.

Let $\hat{\boldsymbol{\Theta}}_T$ designate the current estimate of $\boldsymbol{\Theta}$ obtained by running a Gibbs Sampler [7] at temperature $T$. The stochastic Gibbs sampling procedure is interrupted at regular intervals before decreasing the temperature parameter $T$. $\hat{\boldsymbol{\Theta}}_T$ is used to estimate the affine parameters $\mathcal{A}^I$ and $\mathcal{A}^O$ according to:

$$\hat{\mathcal{A}}^I \ = \ \arg\min_{\mathcal{A}^I} \sum_{s\in\Gamma^I_{\hat{\Theta}_T}} \frac{(\nabla I(s)\cdot \mathrm{v}^I_{\mathcal{A}}(s) + I_t(s))^2}{\sigma_v^2\,\|\nabla I(s)\|^2 \ + \ \sigma_{I_t}^2},$$

$$\hat{\mathcal{A}}^O \ = \ \arg\min_{\mathcal{A}^O} \sum_{s\in\Gamma^O_{\hat{\Theta}_T}} \frac{(\nabla I(s)\cdot \mathrm{v}^O_{\mathcal{A}}(s) + I_t(s))^2}{\sigma_v^2\,\|\nabla I(s)\|^2 \ + \ \sigma_{I_t}^2}. \quad (14)$$

For this model, $\hat{\mathcal{A}}^I$ and $\hat{\mathcal{A}}^O$ simply correspond to standard Linear Least-Squares estimates. The theoretical convergence of this algorithm to a partial optimal solution (Eq. 13) has been established in [11]. For theoretical convergence, the ASA algorithm must satisfy some constraints related to the temperature schedule [7, 11]. 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.

## 5. EXPERIMENTAL RESULTS

The performance of the motion segmentation algorithm is evaluated here on a medical X-ray image sequence showing the movement of the left ventricle of a heart. X-Ray images generally show poor intensity edges (Fig. 1) and consequently are difficult to partition using standard intensity-based segmentation algorithms [3, 5, 6, 10]. When available, motion information may thus provide an useful alternative 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 shapes extracted from a cardiac cycle sequence and $m = 7$ deformation modes have been considered. This statistical knowledge has been integrated in our model as *a priori* knowledge on the shape and deformations of the beating heart.

The motion-based segmentation algorithm was able to extract and track the left ventricle reliably over the whole 40 image sequence, by initializing the deformable model in the current frame by the final configuration of $\boldsymbol{\Theta}$ estimated on the previous frame (Fig. 1a-c). Let us notice that the initial configuration of the deformable template in the first image of the sequence was defined at random.

Manual segmentations (Fig. 1d-f) performed by an expert clinician enable to evaluate the ability of the algorithm to provide suitable results, as far as this application is concerned. The optimization procedure based on the Gibbs sampler dynamics leads to CPU times of less than 10 mn for one frame, on a Sun Sparc10 workstation. Simulated annealing may be used here with an acceptable cost, thanks to the reduced number of parameters $(4 + m)$ that have to be estimated. This performance may be compared to the computationaly demanding scheme associated to the statistical approach described in [6].

## 6. CONCLUSION

In this paper, we have presented a deformable template and motion-based segmentation method for extracting and tracking non-rigid structures in long image sequences. Our approach combines statistical constraints on the shape, deformations and velocity fields characterizing the movement of the deformable structure. The proposed modeling and algorithmic framework yields promising future prospects for the characterization of the dynamic behavior of a large class of deformable objects.

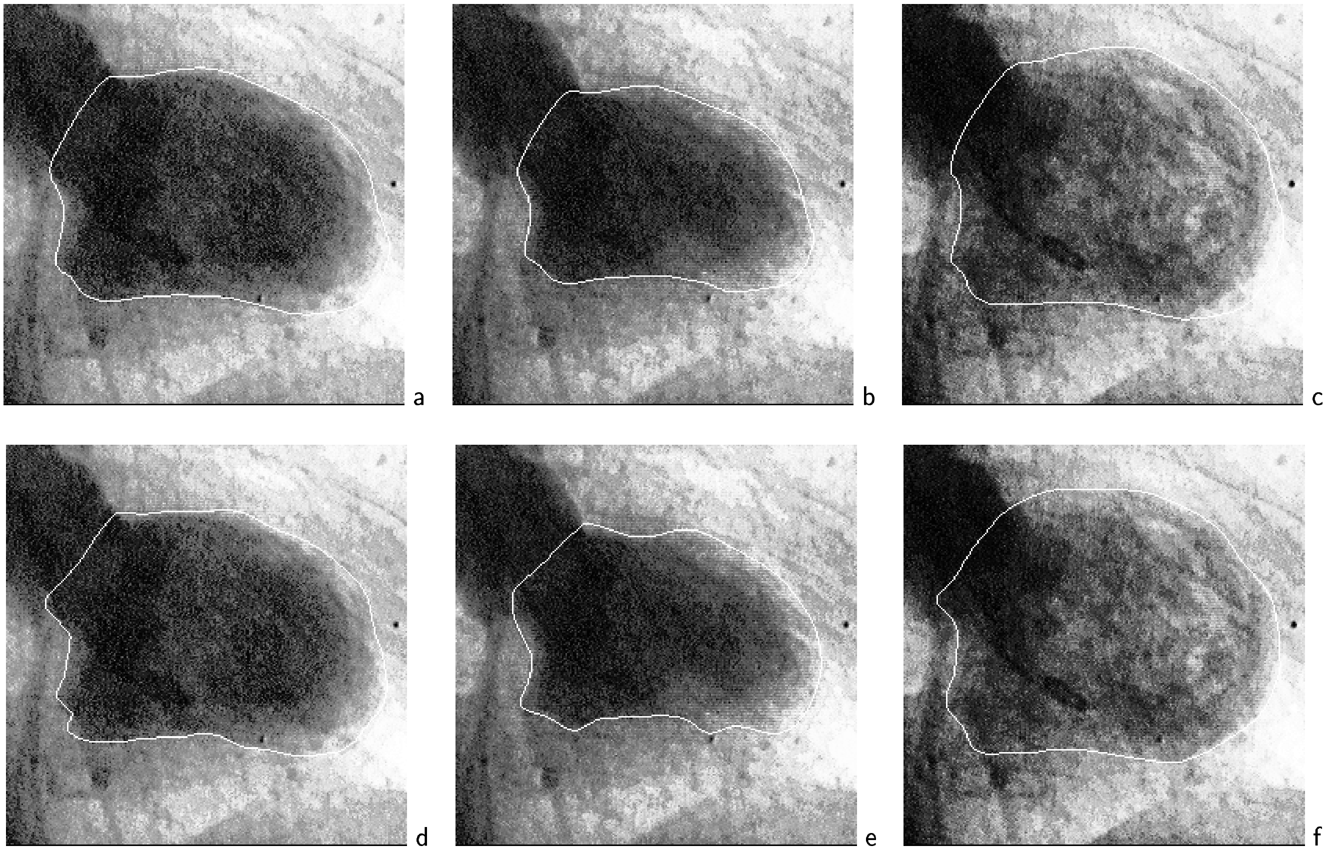


Figure 1: Motion-based segmentation of the left ventricle (256×256 images) a-c) results of the ASA algorithm ; d-f) results provided by an expert clinician (by courtesy of *LTSI- Université Rennes I*).

## 7. REFERENCES


[1] B. BASCLE, P. BOUTHEMY, N. DERICHE and F. MEYER. – Tracking complex primitives in an image sequence. – In *Proc. Int. Conf. on Pattern Recognition*, pages 426–431, Jerusalem, Israel, Oct. 1994. –

[2] 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. Comp. Vis.*, pages 374-381, Boston, USA, June 1995. –

[3] 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, Oct. 1993. –

[4] P. BOUTHEMY and E. FRANCOIS. – Motion segmentation and qualitative dynamic scene analysis from an image sequence. – *Int. J. Comput. Vision*, Vol. 10, No 2: pp. 157–182, 1993. –

[5] 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. –

[6] U. GRENANDER, Y. CHOW and D.M. KEENAN. – Hands. A Pattern Theoretic Study of Biological Shapes. – Springer, New-York, 1991. –

[7] 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, nov. 1984. –

[8] C. KERVRANN and F. HEITZ. – A Hierarchical Statistical Framework for the Segmentation of Deformable Objects in Image Sequences. – In *Proc. IEEE Conf. Comput. Vis. Pat. Rec.*, pages 724–728 , Seattle, USA, June 1994. –

[9] C. KERVRANN and F. HEITZ. – Robust tracking of stochastic deformable models in image sequences. – In *Proc. Int. Conf. Image Processing*, Vol. 3, pages 88–92 , Austin, USA, Nov. 1994. –

[10] C. KERVRANN and F. HEITZ. – Statistical model-based estimation and tracking of non-rigid motion. – In *Proc. Int. Conf. on Pattern Recognition*, Vienna, Austria, Aug. 1996.

[11] S. LAKSHAMNAN 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, Aug. 1989. –

[12] F. MEYER, R.T. CONSTABLE, A.J. SINUSAS and J.S. DUNCAN. – Tracking myocardial deformation using spatially-constrained velocities. – In *Proc. Int. Conf. on Inf. Proc. in Med. Im.*, pages 177–188 , Brest, France, June 1995. and shape optimization. –

[13] E. SIMONCELLI, E. ADELSON and D.J. HEEGER. – Probability distributions of optical flow – In *Proc. IEEE Conf. Comput. Vis. Pat. Rec.*, pages 310–315, Maui, Hawai, USA, may 1991. –