Learning structure and deformation modes of nonrigid objects in long image sequences

Charles Kervrann and Fabrice Heitz

IRISA/INRIA, Campus Universitaire de Beaulieu,
35042 Rennes Cedex, France
E-mail: kervrann@irisa.fr, heitz@irisa.fr

International Workshop
on Automatic Face- and Gesture-Recognition
IWAFGR95
June 26-28, 1995, Zurich, Switzerland

Learning structure and deformation modes of nonrigid objects in long image sequences

Charles Kervrann
IRISA / INRIA
Campus Universitaire de Beaulieu
35042 Rennes Cedex, France
E-mail: kervrann@irisa.fr

Fabrice Heitz
ENSPS / LSIIT
Boulevard Sébastien Brant
67400 Illkirch, France
E-mail: heitz@irisa.fr

Abstract

In this paper, we present an original approach for an unsupervised learning of the structure and deformation modes of 2D moving objects in long image sequences. The object representation relies on a statistical description of the deformations applied to a prototype shape. The optimal bayesian estimate of the deformation process is obtained by maximizing a non-linear joint probability distribution using stochastic and deterministic optimization techniques. The estimates obtained at time t are integrated in the deformation model as a priori knowledge for the segmentation at time $t+1$. Deformation modes are updated on line using a Principal Component Analysis of the distorsions computed from the shapes estimated previously in the image sequence. The approach yields robust segmentations and is demonstrated on real-world image sequences showing the tracking of hands and lips undergoing complex movements.

1 Introduction

Up to the middle of the eighties, the models developed in computer vision have essentially been specified for the description and the analysis of rigid objects undergoing rigid movements. On the other hand, the representation and processing of deformations has recently gained considerable popularity in many application fields [13, 10, 11]. The introduction of deformable models to represent shapes and dynamic phenomena yields promising future prospects as far as the characterization and the interpretation of complex object movements is concerned. General-purpose active contours [12] or *snakes* enforce constraints controlled by elastic forces based on local structure, inflating forces and image based potentials but are not adapted to constrain the deformations for a particular object class [1] or application. On the other hand, application-specific models [16] have been proposed, but the specification of the structure and the deformation modes of the model require an off-line supervised training step [4].

Deformation modes may be identified by matching sets of feature points over time [2, 14]. Cootes *et al.* [3] consider a training set of fixed feature points specified manually to capture the variations of the shapes of interest. The Principal Component Analysis (PCA) allows to determine the main variation modes superimposed on a pre-computed mean shape. This approach has been extended recently [5]. In this method, a large training set yields a robust and compact representation of the object class. Unfortunately, the method proposed in [5] is not robust to occlusions and cannot handle large rotations. In [14] Pentland *et al.* describe an alternate method for establishing point correspondences between two shapes and for providing a robust canonical description of shapes using physically-based modeling techniques [14]. This method may however be subject to mismatches in presence of partial occlusions. In this paper, we introduce a modeling framework to learn automatically deformation modes associated to the object of interest from real image data without any human interaction (such as manual point correspondence). The method requires that the object be described as the deformations of a single prototype object [2, 6, 5].

The model relies on a statistical description of shape deformations, in which two deformation processes are considered [8]. The global description of deformations relies on the modal decomposition introduced recently by Cootes *et al.* in [2]. Local deformations are modeled as stochastic perturbations and are assumed to follow a first order Markov process [6]. At the beginning of the image sequence, no training has been performed and the local deformation process only is identified. It is *local* in the sense that it models deformations involving a point (and its neighbors). After a few frames, the Principal Component analysis of the shapes associated to the local deformation process allows to identify and update global deformation modes as well as the shape structure over time. For each frame a Maximum A Posteriori (MAP) estimate of both local deformations and global deformation modes is obtained by maximizing a highly non-linear joint probability distribution describing the interactions between observations (spatial or temporal gradients extracted from the image) and the deformation process. Global optimization techniques are necessary to obtain optimal solutions for the local deformation process (unless a good initial guess is available) [6]. Due to the large size of the space of configuration, the computation of the MAP estimate of these local deformations is generally computationally demanding [6] but is reasonable since it only concerns the first frames of the sequence (where no *a priori* knowledge on global deformations is available). On the other hand, global deformation modes can be estimated efficiently on the subsequent frames using a deterministic algorithm thanks to the reduced number of parameters of this representation and thanks to the good initialization provided by a temporal prediction scheme. This procedure dynamically updates the shape structure as well as the deformation modes and it is no longer necessary to resort to the local deformation process when the gathered knowledge is sufficient to characterize completely the global deformations associated to the object class.

In Section 2, we describe the statistical deformable model considered in this paper. The different optimization procedures used in the bayesian estimation of deformations are presented in Section 3. The unsupervised method for training the deformation modes over time is presented in Section 4. Segmentations and tracking on real image sequences showing moving hands with partial occlusions and mouth movements are reported in the Section 5.

2 A stochastic deformable model

The approach relies on the description of the object class of interest using a "deformable template" which incorporates *a priori* knowledge on the structure of the object and its variability [2]. A particular shape $\mathbf{X}_t$ is represented at time t by a set of n labeled points which approximate its outline. The variations of shape $\mathbf{X}_t$ are represented by a displacement vector $\mathbf{dX}_t$ with respect to a mean shape (the "template") $\mathbf{X}_t^\star$ [2]. A PCA on the displacement vectors computed from the shapes up to time t allows to determine the m_t deformation modes. If $\mathbf{P}_t$ designates the matrix of the m_t unit eigenvectors corresponding to the m_t largest eigenvalues, and if $\mathbf{b}_t$ denotes the vector $(m_t \times 1)$ corresponding to the m_t most significant deformation modes, the deformable template is represented by the following model [8]:

$$\mathbf{X}_t \;=\; \mathbf{M}(k_t,\theta_t)\,[\mathbf{X}_t^\star + \mathbf{P}_t\mathbf{b}_t] \;+\; \mathbf{T}_t. \qquad (1)$$

Global transformations from the similarity group (rotation of angle θ_t, scale change by a factor k_t and translation $\mathbf{T}_t$) are also taken into account in this model.

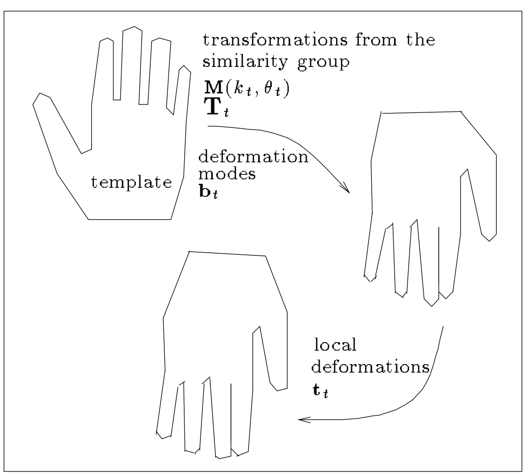

Figure 1: *Description of deformations.*

In addition to this global deformation process, a local deformation process is introduced to obtain an accurate description of the observed deformations. This local deformation process is prevailing at the beginning of the sequence. It becomes a simple refinement of the global deformation modes when the latter have been reliably identified. Local deformations are modeled as random perturbations on the location of the points belonging to the globally deformed pattern. The local deformation process $\mathbf{t}_t$ is modeled as a Markov random process, which takes into account interactions between neighboring points. The global statistical model becomes:

$$\mathbf{Y}_t \;=\; \mathbf{M}(k_t,\theta_t)\,[\mathbf{X}_t^\star + \mathbf{P}_t\mathbf{b}_t] \;+\; \mathbf{T}_t \;+\; \mathbf{t}_t, \qquad (2)$$

where $\mathbf{t}_t$ is the random local deformation process applied on the n labeled points. $\mathbf{t}_t$ is assumed to be a zero-mean first-order Gauss-Markov random process:

$$p(\mathbf{t}_t) \;=\; \frac{1}{C}\, exp\, -\frac{1}{2}\,\mathbf{t}_t{}^T\,\mathbf{R}^{-1}\,\mathbf{t}_t \qquad (3)$$

where $\mathbf{R}$ is the covariance matrix of $\mathbf{t}_t$ and C is the partition function.

3 Bayesian estimation of deformations

3.1 Segmentation task

We have considered here the particular problem of the extraction of moving objects from images sequences in the case where the camera is static. In this context, we rely on observations related to spatiotemporal gradients extracted from the image sequence.

Let $\mathbf{O} \;=\; (O_s,\; s \in S)$ designate an observation field defined on a rectangular lattice S related to the spatiotemporal variations of the intensity function at time t. The Maximum A Posteriori (MAP) estimate of the deformable template is defined by:

$$\widehat{\mathbf{Y}}_t \;=\; \arg\max_{\mathbf{Y}_t}\; p(\mathbf{O}|\mathbf{Y}_t)\; p(\mathbf{Y}_t). \qquad (4)$$

The distribution $p(\mathbf{O}|\mathbf{Y}_t)$ describes the interactions between the observations and the deformations to estimate.

According to the assumption on the statistics of $\mathbf{t}_t$ (see Equ. 3), $\mathbf{Y}_t$ follows a first order Gauss-Markov process:

$$p(\mathbf{Y}_t) = \frac{1}{C}\, exp\, -\frac{1}{2}\,(\mathbf{Y}_t - \mathbf{X}_t)^T\,\mathbf{R}^{-1}\,(\mathbf{Y}_t - \mathbf{X}_t). \quad (5)$$

This prior distribution controls the local and global deformations of the original template. The similarity transformations $\mathbf{M}(k_t,\theta_t)$, $\mathbf{T}_t$ and the global deformation modes $\mathbf{b}_t$ are considered as deterministic parameters of this probabilistic model.

For a given configuration of the template, the image can be partitioned into two regions: the inside of the template $R^-(\mathbf{Y}_t)$ and the outside of $R^+(\mathbf{Y}_t)$ corresponding to the background. A (Gibbs) distribution $p(\mathbf{O}|\mathbf{Y}_t)$ is specified to describe the interactions between local observations $\mathbf{O}(s)$ and the configuration of the deformable model. In [8] for instance, binary-valued observations corresponding to thresholded temporal gradients have been used for the extraction of moving objects (Fig. 2b):

$$p(\mathbf{O}|\mathbf{Y}_t) = \frac{1}{C'}\, exp\, -[\sum_{s \in R^-(Y_t)} |O(s) - 1| + \sum_{s \in R^+(Y_t)} |O(s) - 0|\;] \quad (6)$$

where C' is a normalization constant.

This distribution tends to enclose moving points inside the deformable model and to reject static points belonging to the background outside the outline of the model. The global distribution is a highly non linear function of the model parameters $\mathbf{M}(k_t, \theta_t)$, $\mathbf{T}_t$ and $\mathbf{b}_t$. When no initial guess is available for these parameters the computation of the MAP estimate requires global optimization techniques [8].

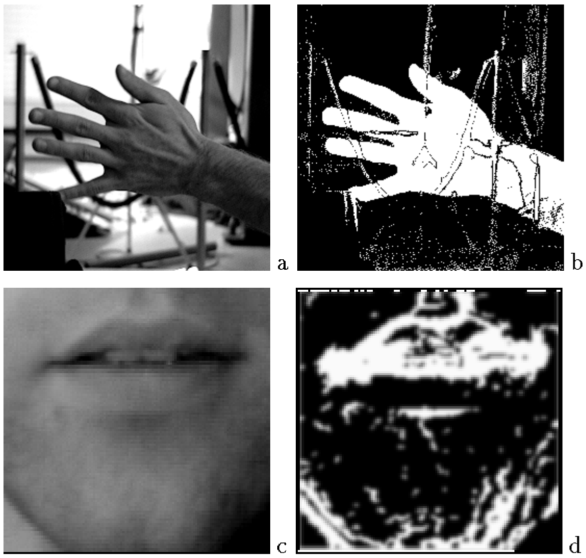

Figure 2: *Fig. 2a-b - observation map $O(s)$ used for the segmentation of a moving hand; Fig. 2c-d - observation map $O(s)$ used for the segmentation of a mouth.*

3.2 Global optimization

In order to simplify the notations, the stochastic deformable model will be expressed at a given time as:

$$\mathbf{Y} \;=\; \mathbf{X}(\Theta) \;+\; \mathbf{t} \tag{7}$$

where $\Theta = (\,\mathbf{M}(k,\theta),\,\mathbf{T},\,\mathbf{b}\,)$ is the *hyperparameter* vector of the model. Since Θ is unknown, the problem is to estimate an pair $(\,\widehat{\Theta},\,\widehat{\mathbf{Y}}\,)$ according to a criterion of optimality:

$$(\,\widehat{\Theta},\,\widehat{\mathbf{Y}}\,) \;=\; \arg\max_{\Theta,\mathbf{Y}}\; p(\mathbf{Y},\mathbf{O}\,/\,\Theta). \tag{8}$$

This criterion satisfies the global maximum of the joint distribution with respect to $\mathbf{Y}$ and Θ but remains extremely difficult to implement. We have adopted a suboptimal procedure recently described by Lakshmanan *et al.* [15] for estimating global model parameters in Markov Random Field-based unsupervised image segmentation. Model hyperparameters are estimated alternately with the configuration of the model, yielding a *partial optimal solution*:

$$\widehat{\mathbf{Y}}_k \;=\; \arg\max_{\mathbf{Y}}\; p(\mathbf{Y},\mathbf{O}\,/\,\widehat{\Theta}_k),$$

$$\widehat{\Theta}_{k+1} \;=\; \arg\max_{\Theta}\; p(\widehat{\mathbf{Y}}_k,\mathbf{O}\,/\,\Theta). \tag{9}$$

The log-likelihood of Θ with respect to $(\,\mathbf{Y},\,\mathbf{O}\,)$ may be defined from Equ.5 and Equ.6 as:

$$\log\, p(\mathbf{Y},\mathbf{O}\,/\,\Theta) \;\propto\; -U_{\Theta}(\mathbf{Y},\,\mathbf{O}) \tag{10}$$

where the energy function is specified as:

$$U_{\Theta}(\mathbf{Y},\mathbf{O}) = \sum_{s\,\in\, R^-(Y)} |O(s)-1\,| + \sum_{s\,\in\, R^+(Y)} |O(s)-0\,|$$

$$+\,\frac{1}{2}\,(\mathbf{Y}-\mathbf{X}(\Theta))^T\mathbf{R}^{-1}(\mathbf{Y}-\mathbf{X}(\Theta)) \tag{11}$$

with at the initial step, $\mathbf{Y} = \mathbf{X}(\Theta)$.

In practice, the estimate of Θ is performed using a fast deterministic relaxation scheme corresponding to a modified version of the ICM algorithm.

The estimation of $\widehat{\mathbf{Y}}$ at step $k+1$ is easily derived given Θ_{k+1}. The criterion of optimality becomes:

$$\mathbf{Y}_{k+1} \;=\; \arg\max_{\mathbf{Y}}\; p(\mathbf{Y},\mathbf{O}\,/\,\Theta_{k+1}). \tag{12}$$

This expression may be interpreted as a Maximum A Posteriori (MAP) of $\mathbf{Y}$ with respect to observations $\mathbf{O}$. The problem consists in minimizing a global energy function:

$$\mathbf{Y}_{k+1} = \arg\min_{\mathbf{Y}} \sum_{s\,\in\, R^-(Y)} |O(s)-1\,| + \sum_{s\,\in\, R^+(Y)} |O(s)-0\,|$$

$$+\,\frac{1}{2}(\mathbf{Y}-\mathbf{X}(\Theta_{k+1}))^T\mathbf{R}^{-1}(\mathbf{Y}-\mathbf{X}(\Theta_{k+1})). \tag{13}$$

In most methods involving deformable templates (apart from the work of Grenander [6]), deterministic minimization algorithms are used to this end. They are known to be very sensitive to local minima of the objective function to minimize. The estimation of $\widehat{\mathbf{Y}}$ requires stochastic algorithm to converge to a satisfying solution when no knowledge about deformation modes is available (*i.e.* at the beginning of the sequence, or when partial occlusions are present).

In practice only one iteration of the optimization loop (Equ. 9) is performed. This procedure has shown to provide good and stable results and additional iterations have only produced slight improvements on the final estimates.

This optimization scheme may be modified over time using the temporal coherence of the movement of the deformable structure and the acquired knowledge about deformation modes to provide good initial estimates from one frame to the next. The Gibbs sampler used to estimate $\widehat{\mathbf{Y}}$ on the first frames may then be substituted by a fast deterministic ICM algorithm. The computational cost for the process decreases from 12 mn cpu time on a workstation on the first frames to less than 1mn when deformation modes are reliably identified. The updating of deformation modes is described in the next section.

4 Unsupervised learning of deformation modes

Given the above optimization scheme, the goal is to update the object (*i.e.* its structure and deformation modes) in order to converge towards a more and more reliable and compact representation for the segmentation task. Let us recall the general form of the deformable model:

$$\mathbf{Y}_t \;=\; \mathbf{M}(k_t,\theta_t)\,[\mathbf{X}_t^\star + \mathbf{P}_t\mathbf{b}_t] \;+\; \mathbf{T}_t \;+\; \mathbf{t}_t. \quad (14)$$

Given the estimate of $\mathbf{t}_t$ and the hyperparameters of the model at time t, deformations are analyzed in a common reference coordinate system. In this reference coordinate system the new model $\mathbf{y}_t$ becomes:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{M}^{-1}(k_t,\theta_t)\;[\mathbf{Y}_t \;-\; \mathbf{T}_t], \\ \mathbf{y}_t &= \mathbf{X}_t^\star \;+\; \mathbf{P}_t\mathbf{b}_t \;+\; \mathbf{M}^{-1}(k_t,\theta_t)\,\mathbf{t}_t. \quad (15) \end{aligned}$$

The template is easily updated at time $t+1$ according to the following rule:

$$\mathbf{X}_{t+1}^\star \;=\; \frac{N_t}{N_t+1}\,\mathbf{X}_t^\star \;+\; \frac{1}{N_t+1}\,\mathbf{y}_t, \quad (16)$$

$$\mathbf{X}_{t+1}^\star \;=\; \mathbf{X}_t^\star + \frac{1}{N_t+1}\left[\,\mathbf{P}_t\mathbf{b}_t + \mathbf{M}^{-1}(k_t,\theta_t)\,\mathbf{t}_t\,\right] \quad (17)$$

where N_t is the current number of processed frames (corresponding to the number of shapes analyzed from the beginning of the sequence).

The deformation modes of the model are described by the unit eigenvectors of the covariance matrix $\mathbf{S}_{t+1}$, defined at time $t+1$ by:

$$\mathbf{S}_{t+1} \;=\; \frac{1}{N_t+1}\sum_{i=1}^{t+1}(\mathbf{y}_i - \mathbf{X}_{t+1}^\star)^T(\mathbf{y}_i - \mathbf{X}_{t+1}^\star). \quad (18)$$

An accurate description of the main variation modes is obtained by retaining in matrix $\mathbf{P}_{t+1}$ only the m_{t+1} eigenvectors associated to the m_{t+1} largest eigenvalues. The number of eigenvalues retained in this representation is adjusted through time in order to set the loss of the information to a constant (and small) value. Typically, between 99% and 99.5% of the total variability is preserved in the truncated representation.

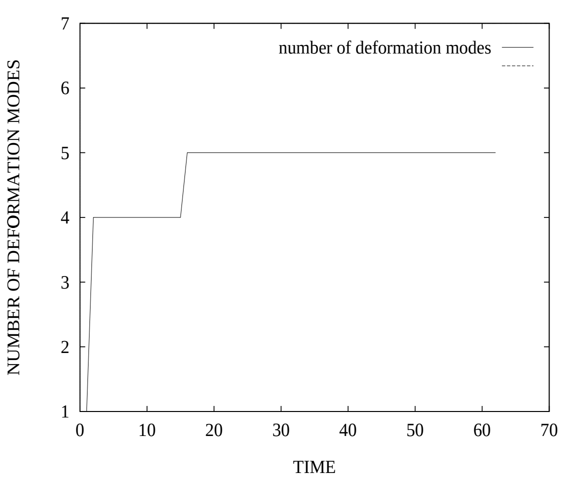

Figure 3: *Number of deformation modes over time in the unsupervised training procedure (moving hand sequence, see Fig. 5).*

The number of significant modes m_t is then observed to increase over time at the beginning of the sequence, before becoming constant when $t \to \infty$ (for instance the final number of deformation modes in Fig. 3 is five).

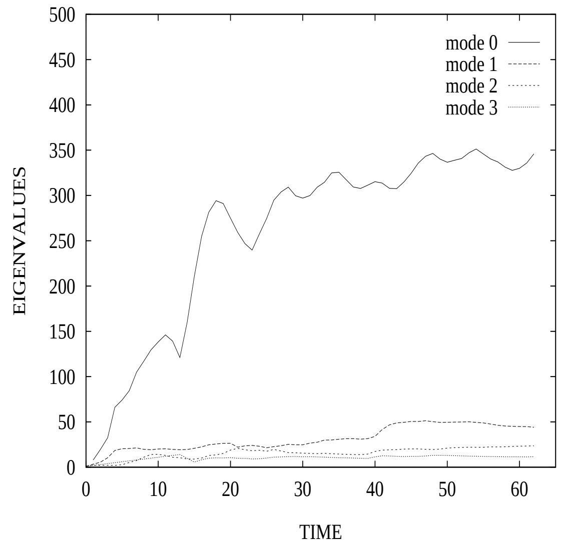

Figure 4: *Evolution over time of eigenvalues associated to the four first deformation modes (moving hand sequence, see Fig. 5).*

From Equ. 17 and Equ. 18, it is easily shown that $\mathbf{X}_{t+1}^\star$ and $\mathbf{S}_{t+1}$ converge to constant values when $t \to \infty$. In practice the infinite time corresponds to a sequence of more than one hundred frames where all representative deformations of the object class have occurred and thereby $\mathbf{P}_t$ and its associated eigenvalues are completely defined. Fig. 4 shows for instance the evolution of the eigenvalues, associated to the four first deformation modes, over a long image sequence composed of more than one hundred frames.

Besides, we take advantage of the temporal coherence of the object movements to predict the value $\widehat{\mathbf{b}}_{t+1}$ of the deformation modes at time $t+1$ from the estimate at time t and from the updated representation:

$$\widehat{\mathbf{b}}_{t+1} \;=\; \mathbf{P}_{t+1}^T\,(\mathbf{y}_t \;-\; \mathbf{X}_{t+1}^\star). \quad (19)$$

i.e.:

$$\widehat{\mathbf{b}}_{t+1} = \mathbf{P}_{t+1}^T\,(\mathbf{X}_t^\star + \mathbf{P}_t\mathbf{b}_t + \mathbf{M}^{-1}(k_t,\theta_t)\,\mathbf{t}_t - \mathbf{X}_{t+1}^\star). \quad (20)$$

When $t \to \infty$, $\mathbf{P}_{t+1}^T \simeq \mathbf{P}_t^T$ and $\mathbf{X}_{t+1}^\star \simeq \mathbf{X}_t^\star$, so $\widehat{\mathbf{b}}_{t+1}$ no longer depends on the template $\mathbf{X}_t^\star$:

$$\widehat{\mathbf{b}}_{t+1} \;=\; \mathbf{b}_t \;+\; \mathbf{P}_t^T\,\mathbf{M}^{-1}(k_t,\theta_t)\,\mathbf{t}_t. \quad (21)$$

$\widehat{\mathbf{b}}_{t+1}$ can be used as a good initial estimate for the deformation modes $\mathbf{b}_{t+1}$ in the next frame. By using the temporal coherence of the movement of the deformable structure, fast local optimization techniques can be used to obtain reliable MAP estimates. The experimental results show in this case that the optimal solution provided by global optimization techniques (stochastic algorithm) is indeed close to the initial estimate given by $\widehat{\mathbf{b}}_{t+1}$, $\mathbf{M}(k_t,\theta_t)$ and $\mathbf{T}_t$ at time $t+1$.

The method is able to provide an accurate and very compact representation of deformations (more compact than the method based on the alignment of extracted shapes according to a *Generalized Procrutes Analysis* [4, 5]). In Fig. 3 one can see that a very low number of deformation modes is required to integrate 99.5% of the total variability on a typical test sequence. Besides the MAP estimation technique is very robust to noise and to the presence of large occlusions [8, 9].

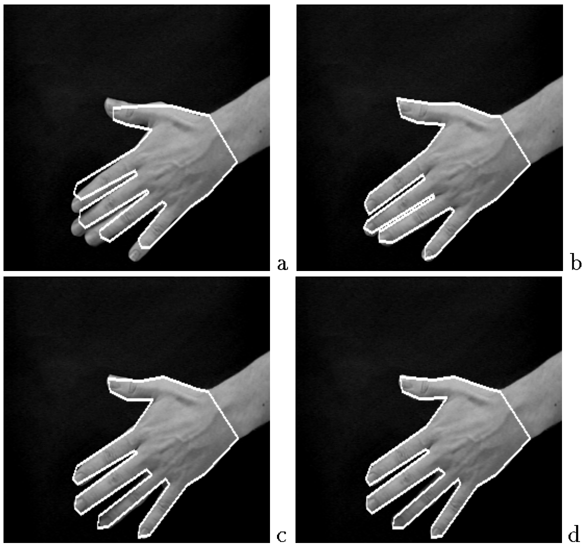

Figure 5: *Tracking of a moving hand against a uniform background (see text).*

Let us notice that the heuristic temporal prediction presented in Equ. 19 may be completed efficiently by a temporal tracking of the complete model which provides also good initial estimate for the location of the model in the next frame and enables to resort to fast local (deterministic) optimization procedures. In [9] we have developed a Kalman filter combined with a detection of abrupt changes [7] which performs prediction and filtering of the model *hyperparameters*. This tracking procedure ensures to converge faster to the optimal solution and is used here when the training of the deformation modes has converged (*i.e.* when $\mathbf{X}^{\star}_{t+1} \simeq \mathbf{X}^{\star}_{t}$ and $\mathbf{P}_{t+1} \simeq \mathbf{P}_{t}$).

5 Experimental results

In our experiments, we have considered the segmentation of deformable structures [1, 2, 4, 6] corresponding to hands (Fig. 5 and Fig. 6) and lips (Fig. 7).

Fig. 5a and Fig. 5b present respectively an intermediate step and the final result of the MAP segmentation on the second frame. Fig. 5a shows the estimation of the global transformations from the similarity group (see Equ. 2). Global deformation modes are not yet available on this early stage of the training process. Fig. 5b depicts the final segmentation including the local deformation process which contributes exclusively to the solution in this case.

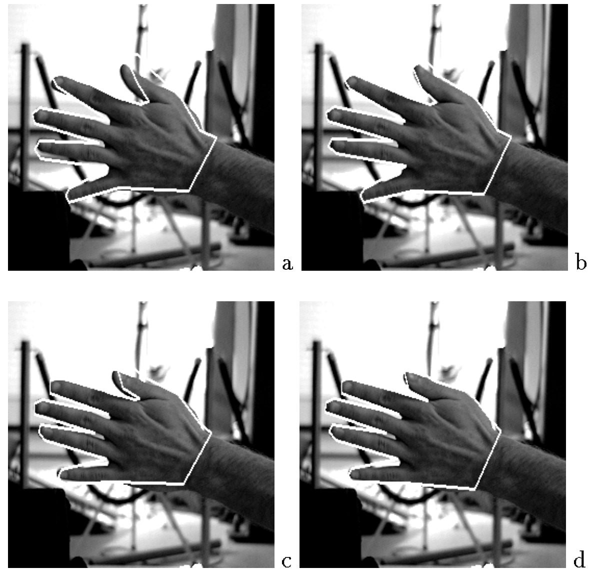

Figure 6: *Tracking of a moving hand against a textured background (see text).*

Fig. 5c and Fig. 5d show the similar result of the segmentation on the third frame in which the deformation modes have been updated for the first time with the estimation obtained from the second frame. As can be seen, the local deformation process (Fig. 5d) is nearly unutilized on this frame because the configuration of the shape is close to the previous one. Hence the deformation modes captured on the second frame provide an excellent training for the third frame.

Fig. 6 presents the same intermediate and final results for a second test sequence showing hand moving against a textured background. Fig. 6a-b and Fig. 6c-d correspond respectively to frame 6 and 7 in this sequence.

The last example (Fig. 7) presents the application of the complete segmentation and tracking procedure (including the Kalman filter) to a mouth structure. The distribution $p(\mathbf{O}|\mathbf{Y}_t)$ specified to the mouth tends to attract the deformable template toward salient features of the image corresponding to large spatial gradients [16]. The tracking over time provides valuable information about the global dynamic behavior of the deformable structure which might be used for interpretation purposes.

6 Conclusion

In this paper, we have presented a general framework for the modeling and unsupervised training of deformation modes of nonrigid objects. The technique relies on the definition of a prototype shape on which two deformation processes are applied. The deformations are described using statistical models and the optimal bayesian estimate of these deformations is computed using stochastic and deterministic optimization techniques.

The proposed modeling and algorithmic framework is comprehensive and suited to the representation of a large class of deformable objects. It may be adapted to segmentation problem based on other image attributes (luminance, color, texture, depth, etc.). The use of the learning procedure also yields promising future prospects as far as the characterization and the interpretation of the dynamic behavior of complex objects is concerned.

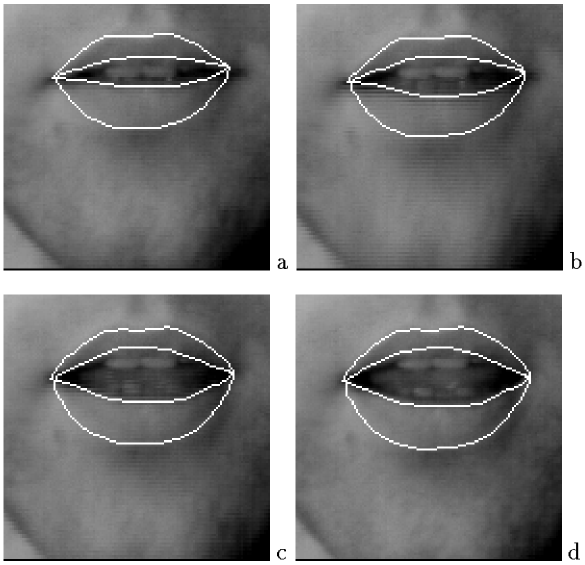

Figure 7: *Tracking of mouth movements (see text).*

References

[1] 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.

[2] T.F. COOTES, C.J. TAYLOR, D.H. COOPER and J. GRAHAM. – Training models of shape from sets of examples. – In *British, Machine Vision Conf.*, pages 9–18, Leeds, UK, Sept. 1992.

[3] T.F. COOTES and C.J. TAYLOR. – Active shape models : smart snakes. – In *British, Machine Vision Conf.*, pages 266–275, Leeds, UK, Sept. 1992.

[4] T.F. COOTES, C.J. TAYLOR, D.H. COOPER and J. GRAHAM. – Active shape models - Their training and application. – *Computer Vision and Image Understanding*, Vol. 61, No 1: pp. 38–59, January 1995.

[5] A. BAUMBERG and D. HOGG – Learning flexible models from image sequences. – In *3rd European Conference on Computer Vision*, pages 299–308, Stockolm, Sweden, May 1994.

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

[7] D. V. HINKLEY. – Inference about the change - point from cumulative sum - tests. – *Biometrika*, Vol. 58, No 3: pp. 509–523, 1971.

[8] C. KERVRANN and F. HEITZ. – A hierarchical Statistical Framework for the Segmentation of Deformable Objects in Image Sequences. – In *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 *IEEE First Int. Conf. Image Processing*, Vol. 3, pages 88–92 , Austin, USA, November 1994.

[10] D. METAXAS and D. TERZOPOULOS. – Shape and nonrigid motion estimation through physics-based synthesis. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 15, No 6: pp. 580–591, June 1993.

[11] C. NASTAR and N. AYACHE. – Fast segmentation, tracking and analysis of deformable objects. – *In Proc. 4th Int. Conf. Comp. Vis.*, pages 275-279, Berlin, Germany, May, 1993.

[12] M. KASS and A. WITKIN and D. TERZOPOULOS. – Snakes : Active Contour Models. – *In Proc. First Int. Conf. Comp. Vis.*, pages 259-268, London, UK, June, 1987.

[13] A. PENTLAND and B. HOROWITZ. – Recovery of nonrigid motion and structure. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 13, No 7: pp. 730–742, July 1991.

[14] A. PENTLAND and B. HOROWITZ. – A modal framework for correspondence and recognition. – *In Proc. 4th Int. Conf. Comp. Vis.*, pages 308-313, Berlin, Germany, May, 1993.

[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, August 1989.

[16] A. YUILLE, P. HALLINAN and D. COHEN. – Feature extraction from faces using deformable templates. – *Int. Journal of Comput. Vis.*, Vol. 8, No 2: pp. 99–111, 1991.