Localization of Shapes Using Statistical Models and Stochastic Optimization

François Destrempes, Max Mignotte, and Jean-François Angers

Abstract—In this paper, we present a new model for deformations of shapes. A pseudolikelihood is based on the statistical distribution of the gradient vector field of the gray level. The prior distribution is based on the Probabilistic Principal Component Analysis (PPCA). We also propose a new model based on mixtures of PPCA that is useful in the case of greater variability in the shape. A criterion of global or local object specificity based on a preliminary color segmentation of the image is included into the model. The localization of a shape in an image is then viewed as minimizing the corresponding Gibbs field. We use the Exploration/Selection (E/S) stochastic algorithm in order to find the optimal deformation. This yields a new unsupervised statistical method for localization of shapes. In order to estimate the statistical parameters for the gradient vector field of the gray level, we use an Iterative Conditional Estimation (ICE) procedure. The color segmentation of the image can be computed with an Exploration/Selection/Estimation (ESE) procedure.

Index Terms—Shape localization, statistical model, stochastic optimization, Exploration/Selection (E/S) algorithm, Probabilistic Principal Component Analysis (PPCA).

1 INTRODUCTION

L OCALIZATION of shapes is an important problem in Image Processing with applications to segmentation and classification [1], [2], pattern recognition [3], motion tracking [4], and 3D reconstruction [5], [6].

Localization of a shape consists [7] of matching a deformable template to the data in the image. In the free-form formulation, the shape of the template is constrained only by local geometric criteria, such as smoothness and continuity. The elastic deformable model [8], [9], the active contour model [10], the models presented in [11] and [12], and the Mumford-Shah functional [13] are examples of that approach. In the parametric formulation, the prior global information of the shape is used to model the deformations. In that approach, one can represent the template either by 1) a family of parametrized curves or by 2) one or several prototype templates together with their parametric spaces of deformations, as formulated in the pattern theory of [14], [15]. The models of [16], [17] follow the first scheme, whereas [1], [2], [3], [7], [15], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32] are examples of the second scheme.

An important idea exploited in [7], [11], [12], [15], [20], [32] is the formulation of the localization problem in a Bayesian framework. Namely, one searches for the optimal deformation of the shape in the sense of the Maximum A Posteriori (MAP). Nevertheless, in the probabilistic model presented in [7], the likelihood is defined in an ad hoc manner: It depends

Recommended for acceptance by R. Basri.

on a smoothing parameter ρ , which is not estimated from the image. Moreover, it is based on a binary detection of contours so that the localization step might be penalized by flaws occurring at the edge-detection step. In contrast, the likelihood functions presented in [11], [12], [20] are based on the distribution of the gray level. However, in those models, the prior distribution is not learned from a sample of shapes. In the active shape model [19], [25], the prior distribution of the deformations of the shape is learned from a sample of curves representing the same shape. The classical Principal Component Analysis (PCA) is used to reduce the dimensionality of the deformation space and the distribution of the reduced parameters is modeled by a mixture of Gaussian distributions. A statistical shape prior can also be incorporated into a Mumford-Shah segmentation process [21], [22], [26], [27], a geodesic active contours method [28], or level sets representations [31]. Models based on textures have been introduced in [33], [34], [35].

In [36], we presented a model for localization of shapes based solely on contours. A shape is viewed as a deformable template and the distribution of the deformation parameters is modeled with the Probabilistic Principal Component Analysis (PPCA) [37] obtained from a training set aligned as in [19]. This gives the prior distribution of deformations of a shape. A pseudolikelihood of deformations of a given shape is based on the statistical distribution of the gradient vector field of the gray level [38] (similar to [39], [40]). We use an Iterative Conditional Estimation (ICE) procedure [41], [42] for the estimation of the parameters, as explained in [38]. Localization of a shape can then be viewed as minimizing the Gibbs field corresponding to the pseudoposterior distribution. The E/S optimization algorithm [43] is then used to find the optimal solution. From [43], the internal parameters of the E/S algorithm that ensure asymptotic convergence to a global optimal solution are known and are practical. This is not the case, as far as we know, for genetic algorithms or the simulated annealing. Furthermore, as mentioned in [44], the known convergence results [45] for the particle filtering algorithms require that

[•] F. Destrempes and M. Mignotte are with the Département d'Informatique et de Recherche Opérationnelle (DIRO), Université de Montréal, Pavillon André-Aisenstadt, C.P. 6128 Succursale centre ville, H3C 3J7 Montreal (Quebec), Canada. E-mail: {destremp, mignotte}@iro.umontreal.ca.

J.-F. Angers is with the Département de Mathématiques et de Statistique, Université de Montréal, C.P. 6128 Succursale centre ville, H3C 3J7 Montreal (Quebec), Canada. E-mail: angers@dms.umontreal.ca.

Manuscript received 9 Mar. 2006; revised 6 Oct. 2006; accepted 27 Nov. 2006; published online 17 Jan. 2007.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-0217-0306. Digital Object Identifier no. 10.1109/TPAMI.2007.1157.

the number of particles tend to infinity, although working with a finite number of particles might be useful in practice [24] in the context of localization of shapes. Furthermore, deterministic algorithms such as the iterative model refinement method [25] work well in some contexts, but we do not know of any theorems of convergence in a general setting (for instance, with complex backgrounds, as considered in some examples reported here). Also, the graph cuts algorithms yield an optimal solution and are useful [46], but, generally speaking, they apply only to the optimization of special types of functions [47]. Although elegant and useful, the formulation of the problem in level sets representations [31], active contours [28], or variational functionals [27] deserve further study, in particular with respect to the convergence of algorithms to an optimal solution of the model. Finally, the jump-diffusion random sampling algorithm presented in [15] and the variant [32] of the reversible jump Monte Carlo Markov Chain (RJMCMC) [48] are two algorithms that converge to a sampling of the desired posterior distribution. However, in this paper, we do not wish to sample from the posterior distribution, but, rather, to find its mode. Thus, we may rely on the (simpler) E/S optimization algorithm.

Yet, due to its complexity, the Gibbs field of the posterior distribution is rather hard to optimize directly, even if the optimization algorithm converges asymptotically. Thus, one has to use an initialization procedure to explore plausible regions of the deformation space in order to speed up convergence to a solution. In [36], we optimized a heuristic potential function similar to [2] for the initialization procedure. In [49], we used a clustering of contour segments in order to find initial positions of the shape. In [40], we used for the initialization procedure a clustering of the shapes. The main problem with these approaches lies in fitting a one-dimensional object (the shape) with a one-dimensional subset of the image, i.e., the contours.

In this paper, we present an extension (that appears in an earlier version of Destrempes' PhD thesis [50, Chapter 8]) of the above model that takes into account a segmentation of the image into regions. Namely, we require that the region labels inside and outside of the shape be distinct. We call this property the *object specificity*. This simple idea turns out to make the model a lot easier to estimate, as the tests reported in this paper indicate. Indeed, the task now consists of fitting a two-dimensional object (the interior of the shape) with a twodimensional subset of the image, i.e., the color regions forming the object. In particular, we no longer need the initialization procedures of [36], [40], [49]. We base the image segmentation on the model presented in [51]. A procedure called the ESE procedure [44] is used to compute the segmentation, as explained with full details in [51]. We also make comparisons with a few other segmentation models. We also propose a *local object specificity* property that is useful in the case of multiple occurrences of the object in the image. Furthermore, we include a generalization [37] of the PPCA that uses mixtures of Gaussian kernels. This model for reduction of dimensionality offers more flexibility in the case of a shape with greater variability. Note, however, that, in [37], the reconstruction operator depends on the Gaussian kernel, whereas here it is fixed.

Altogether, the method proposed here for localizing shapes is organized in four steps:

- 1. The first step consists of learning offline the deformations of the shape from a sample of curves. This means learning the mean shape, the nonlinear deformations of the shape, and the prior distribution of deformations of the shape. This training step is based on the PPCA [37] in its simplest form or on the new model for reduction of dimensionality presented in Section 5.
- 2. The second step consists of estimating online the statistical distribution of the gradient vector field of the gray level from the observed data in a given image, using an ICE procedure.
- 3. The third step consists of segmenting online the image based on colors using the ESE procedure.
- 4. The fourth step consists of localizing the shape in the image, by minimizing with the E/S algorithm [43] the Gibbs field of the pseudoposterior distribution of deformations of the shape, with the global constraint of (local) object specificity taken into account.

In our opinion, the main contribution of this paper is to bring together interesting ideas of previous work with the novelty of a new model for deformations of shapes and new global constraints for the localization.

The remainder of this paper is organized as follows: In Section 2, we recall the statistical model [38] for the gradient vector field of the gray level, as well as the statistical model of the colors [51]. Section 3 presents the statistical model for deformations of a shape, with a brief review of the PPCA. The stochastic method for localization of a shape is explained in Section 4. The new model for reduction of dimensionality is presented in Section 5. The local object specificity property is presented in Section 6. In Section 7, we discuss experimental results. We conclude in Section 8.

2 STATISTICAL MODELS FOR THE IMAGE

2.1 Statistical Model for the Gradient Vector Field

In this section, we recall the model introduced in [38] (similar to what we have proposed in [39], [40]) for the distribution of the gradient of the gray level in an image. It gives much more than a binary detection of contours (i.e., a classification of pixels as being on or off contours): It also gives, for each pixel, its likelihood of being on contours and its likelihood of being off contours. The likelihood of deformations of a shape presented in Section 3.4 is based on this model.

Given an image of size $M \times N$, $G = (V_G, E_G)$ will denote the nonoriented graph consisting of the $M \times N$ pixels of the image together with the edges given by the usual 8-neighbors. If s and t are adjacent sites of G, we denote the edge joining s and t by (s,t) (so, $(s,t) \in E_G$).

For each $s \in V_G$, the random variable Y_s represents the norm of the gradient of the gray level at the corresponding pixel. If $(s,t) \in E_G$, the random variable Y_{st} represents the absolute value of the angle between the mean of the gradient at s and t and the normal to the vector from s to t. This angle is normalized between $-\pi/2$ and $\pi/2$ before taking its absolute value. As explained in [38], the variable Y_{st} can be viewed as the angle in absolute value between a level curve of the gray level function and the curve going through the segment (s,t). Finally, for each $s \in V_G$ and $(s,t) \in E_G$, the random variables Z_s and Z_{st} take their values in the set

 $\{e_1, e_2\}$, where e_1 denotes the class "off contours" and e_2 the class "on contours."

As in [38], we adopt the following distributions for the marginals of Y_s and Y_{st} on the classes e_1 and e_2 :

- $P(y_s|z_s = e_1, C, \alpha)$ is a shifted Weibull law [52] $\mathcal{W}(y_s; \min, C, \alpha)$ defined by $C(y_s \min)^{C-1} / \alpha^C e^{-((y_s \min)/\alpha)^C}$, where $y_s > \min = -0.001$, and $C, \alpha > 0$.
- $P(y_s|z_s = e_2, w_j, \mu_j, \sigma_j)$ is a mixture $\mathcal{M}(y_s; w_j, \mu_j, \sigma_j) = \sum_{j=1}^{3} w_j \mathcal{N}(y_s; \mu_j, \sigma_j)$ of three Gaussian kernels truncated on the interval $(0, \infty)$.
- $P(y_{st}|z_{st} = e_1)$ is a uniform distribution $\mathcal{U}(y_{st}; 0, \frac{\pi}{2})$ on $[0, \frac{\pi}{2}]$.
- $P(y_{st}|z_{st} = e_2, \alpha_0)$ is a truncated exponential law $k_0 \mathcal{E}(y_{st}; \alpha_0)$ on the interval $[0, \frac{\pi}{2}]$, with

$$k_0 = \left\{ 1 - \exp\left(-\frac{\pi}{2\alpha_0}\right) \right\}^{-1}.$$

As explained in [38], the distribution on Y_s reflects the hypothesis that the norm of the gradient of the gray level tends to be larger for points on edges than for points off edges. Also, the distribution on Y_{st} models the hypothesis that there is no privileged angle between a level curve of the gray level function and a segment off edges, whereas this angle is expected to be near 0 whenever the segment is located on edges.

See [38] for an example of estimated distributions and empirical distributions. We use an ICE procedure [41], [42], as explained in [38], in order to estimate the hyperparameters C, α , w_j , μ_j , σ_j , and α_0 for a given image. The choice of three Gaussian kernels is arbitrary; it could be estimated using the Bayesian Information Criterion (BIC) [53] for instance.

2.2 Statistical Model for the Field of Colors

In this section, we recall the model introduced in [51] for the distribution of the colors in an image.

For each $s \in V_G$, the random variable W_s represents the Luv components at the corresponding pixel. For each $s \in V_G$, the random variable X_s takes its values in a finite set of K region labels $\{f_1, f_2, \ldots, f_K\}$.

We adopt the following distributions for the marginals of W_s on a region class f_k :

• $P(w_s|x_s = f_k, \pi_{(i,k)}, \tilde{\mu}_i, \tilde{\Sigma}_i)$ is a mixture of K_1 Gaussian kernels $\sum_{i=1}^{K_1} \pi_{(i,k)} \mathcal{N}(w_s; \tilde{\mu}_i, \tilde{\Sigma}_i)$, where $\sum_i^{K_1} \pi_{(i,k)} = 1$ and $\pi_{(i,k)} \ge 0$.

We use the ESE procedure [44], as explained in [51], in order to estimate the parameters and obtain a segmentation xfor a given image. Here, we take K = 30 region labels and $K_1 = 30$ Gaussian kernels. The initial number of labels is Kand the ESE procedure allows this number to be decreased. See Fig. 1 for an example of segmentation. Note that this is an unsupervised method. Also, in practice, each region class f_k is defined by the proportions $\pi_{(1,k)}, \ldots, \pi_{(K_1,k)}$, some of which are near 0. Thus, we can fix the upper bound K_1 and let the procedure estimate the actual contribution of each Gaussian kernel. The choice of the upper bounds K and K_1 is arbitrary.



Fig. 1. Segmentations of image (25) of Fig. 8 of Appendix I (which can be found at http://computer.org/tpami/archives.htm), based on color models. (a) The color model of Section 2.2. (b) The Mean Shift algorithm. (c) The color model of [44]. (d) The K-means algorithm (with K = 30).

3 STATISTICAL MODEL FOR DEFORMATIONS OF A SHAPE

In this section, we present a statistical model for deformations of a shape. The training phase and the prior distribution are based on the PPCA [37]. The likelihood is based on the statistical model of the gradient vector field of the gray level presented in Section 2.1. Thus, the prior distribution can be learned from a sample of curves representing the shape, and the likelihood can be estimated from the observable data given by the image. The global constraint is based on the image segmentation method presented in Section 2.2. The material presented in Sections 3.1, 3.2, 3.3, and 3.4 first appeared in [36]. Sections 3.1, 3.2, 3.3 were generalized to the three-dimensional case int the context of object reconstruction in [6]. Sections 3.1 and 3.3 are extended to a more flexible model in Section 5.

3.1 Training Phase

A curve is represented as a template, that is, a sequence of points

$$\gamma = (a_0, b_0, a_1, b_1, \dots, a_m, b_m), \tag{1}$$

where the m + 1 points $(a_0, b_0), (a_1, b_1), \ldots, (a_m, b_m)$ are equally spaced on the curve between certain key-points (see Fig. 6 of Appendix I for an example, which can be found at http://computer.org/tpami/archives.htm). Those keypoints are only used at the training phase. Although not essential, they are useful in the semiautomatic editing of the templates corresponding to a database of the shape. All steps of the localization procedure itself are automatic and do not use those key-points. In our experiments, we segment the contour of the shape with the "intelligent scissors" algorithm [54]. Once this is done, an automatic procedure chooses m + 1 points on the curve (equally spaced between the keypoints). In our experiments, we chose m rather arbitrarily so that the approximated curves are "satisfying visually." Instead, a method could be designed that determines m as a function of a given error between the original curve and its approximated curve of m + 1 points.

Given a sample $\gamma_1, \ldots, \gamma_N$ of curves with the same number of points and representing the deformations of a same shape, we resort to the procedure [19] to align this training set on a mean shape. Viewing $T = (A_0, B_0, \ldots, A_m, B_m)$ as a random vector and setting d = 2m + 2, we consider the Probabilistic Principal Component Analysis (PPCA) model [37]: T = $W \Xi + \nu + \varepsilon$, with W a $d \times q$ matrix, ν a vector of dimension d, $\Xi \sim \mathcal{N}(0, I_q)$, and $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_d)$. If t_1, \ldots, t_N is an i.i.d. sample of the variable T, we obtain an optimal model [37] in the sense of the ML upon taking

$$\sigma^{2} = \frac{1}{2d - q} \sum_{i=q+1}^{d} \lambda_{i}, \quad \nu = \bar{t} = \frac{1}{N} \sum_{i=1}^{N} t_{i},$$

$$W = U_{q} (\Lambda_{q} - \sigma^{2} I_{q})^{1/2},$$
(2)

where $\lambda_1, \ldots, \lambda_d$ are the eigenvalues of the sample covariance matrix in decreasing order, Λ_q is the diagonal matrix with entries $\lambda_1, \ldots, \lambda_q$, and U_q is the $d \times q$ matrix with columns equal to the corresponding eigenvectors, normalized so that they have euclidean norm equal to 1 (i.e., the columns of U_q span the principal subspace of the sample covariance matrix). See [26], [30], [35] for a similar use of the SVD decomposition. Note that *T* has distribution function

$$p_T(t) = \mathcal{N}(t; \nu, \sigma^2 I_d + WW^t). \tag{3}$$

If N < d, the sample covariance matrix of the full data might not be positive-definite. However, the matrix $\sigma^2 I_d + WW^t$ is positive-definite as long as $\sigma^2 > 0$. This case occurs provided $\lambda_{q+1} > 0$, even if the sample covariance matrix is not full-rank [26] (i.e., $\lambda_d = 0$). Thus, we can actually consider the case where N < d, as in our tests. The corresponding reconstruction operator is

$$\gamma_{ppca}(\xi) = U_q (\Lambda_q - \sigma^2 I_d)^{1/2} \xi + \nu.$$
(4)

We view the vector ξ as a vector of shape deformation.

3.2 Deformation Parameters

From the training phase, we obtain the mean shape $\nu = \overline{t}$ and its statistical deformations $\gamma_{ppca}(\xi)$, where ξ is a vector in the parameter space Ξ of dimension q. In addition, we consider projective deformations of the form $(x', y') = (\cos(\psi_y)x + \sin(\psi_y)\sin(\psi_x)y,\cos(\psi_x)y)$ applied pointwise to a curve. Such a transformation is obtained by applying to a point (x, y, 0) a rotation by an angle ψ_x around the *x*-axis, then a rotation by an angle ψ_y around the *y*-axis, and, finally, projecting the resulting point onto the *xy*-plane. Namely, we have

$$\begin{bmatrix} \cos(\psi_y) & 0 & \sin(\psi_y) \\ 0 & 1 & 0 \\ -\sin(\psi_y) & 0 & \cos(\psi_y) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\psi_x) & -\sin(\psi_x) \\ 0 & \sin(\psi_x) & \cos(\psi_x) \end{bmatrix}$$
$$\times \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} \cos(\psi_y)x + \sin(\psi_y)\sin(\psi_x)y \\ \cos(\psi_x)y \\ -\sin(\psi_x)x + \cos(\psi_y)\sin(\psi_x)y \end{bmatrix}.$$

We also consider rigid deformations given by scaling *s*, rotation ψ , and translation (τ_x, τ_y) applied pointwise to a curve lying in the *xy*-plane. This yields a vector of deformation

$$\theta = (\tau_x, \tau_y, s, \psi, \psi_y, \psi_x, \xi_1, \dots, \xi_q) \tag{5}$$

of dimension q + 6. The transformation θ is computed in the following order: First, calculate the statistical deformation $\gamma_{ppca}(\xi)$; next, apply the projective transformation (ψ_y, ψ_x) ; then, apply *s* and ψ ; finally, translate by τ . The resulting template is denoted γ_{θ} . Note that, although (4) defines a linear operator (from \mathbb{R}^q to \mathbb{R}^d), it does not describe in general an affine operator (on \mathbb{R}^2) applied pointwise to a curve. In this sense, it yields a nonaffine deformation of the curve.

3.3 Prior Distribution

Let Θ be the random variable corresponding to the vector of deformations. We model the distribution of Θ by

$$P(\theta) = \mathcal{U}(\tau_x, \tau_y, s, \psi, \psi_y, \psi_x) p_T(\gamma_{ppca}(\xi)), \tag{6}$$

where \mathcal{U} denotes the uniform distribution and T represents the full data. From the lemma of [36], we obtain

$$p_T(\gamma_{ppca}(\xi)) \propto \exp(-\frac{1}{2}\xi^t(I_q - \sigma^2 \Lambda_q^{-1})\xi).$$
(7)

3.4 Pseudolikelihood

Let *Y* be the random field $\{Y_s, Y_{st}\}$, where the random variables Y_s, Y_{st} are as in Section 2.1. Given an image *I*, *y* is the observed realization of Y = Y(I). Given a path $c = (s_0, \ldots, s_m)$ in the graph *G* of Section 2.1, let E(c) denote the set of edges $\{(s_{i-1}, s_i) : i = 1, 2, \ldots, m\}$. We consider, as we have proposed in [40], the pseudolikelihood P(y|c) given by

$$\prod_{s \notin c} P(y_s | z_s = e_1) \prod_{s \in c} P(y_s | z_s = e_2) \\ \times \prod_{(s,t) \notin E(c)} P(y_{st} | z_{st} = e_1) \prod_{(s,t) \in E(c)} P(y_{st} | z_{st} = e_2) \\ = k(y) \prod_{s \in c} \frac{P(y_s | z_s = e_2)}{P(y_s | z_s = e_1)} \prod_{(s,t) \in E(c)} \frac{P(y_{st} | z_{st} = e_2)}{P(y_{st} | z_{st} = e_1)} \\ = k(y) \prod_{i=0}^{m} \frac{\mathcal{M}(y_{s_i}; w_j, \mu_j, \sigma_j)}{\mathcal{W}(y_{s_i}; \min, C, \alpha)} \prod_{i=1}^{m} \frac{k_0 \mathcal{E}(y_{s_{i-1}, s_i}; \alpha_0)}{\mathcal{U}(y_{s_{i-1}, s_i}; 0, \frac{\pi}{2})},$$
(8)

where

$$k(y) = \prod_{s \in V_G} P(y_s \mid z_s = e_1) \prod_{(s,t) \in E_G} P(y_{st} \mid z_{st} = e_1).$$
(9)

The distributions $\mathcal{W}(y_s; \min, C, \alpha)$, $\mathcal{M}(y_s; w_j, \mu_j, \sigma_j)$, $\mathcal{U}(y_{st}; 0, \frac{\pi}{2})$, and $k_0 \mathcal{E}(y_{st}; \alpha_0)$ are as in Section 2.1.

We call P(y|c) a pseudolikelihood since, in general, $\int P(y(I)|c)dI \neq 1$ because the random variables of the random field Y(I) are dependent. Note, however, that $\int_y P(y|c)dy = 1$. Also, we have dropped the dependence of P(y|c) on the vector of statistical parameters of the gradient model for simplicity of notation.

The map of $\log(P(y_s|z_s = e_1)/P(y_s|z_s = e_2))$ is presented in Fig. 5 of Appendix I, which can be found at http:// computer.org/tpami/archives.htm.

Note that the scaling factor k(y) depends only on the observed data (and not on the curve) and that the principal factor

DESTREMPES ET AL.: LOCALIZATION OF SHAPES USING STATISTICAL MODELS AND STOCHASTIC OPTIMIZATION

$$\prod_{i=0}^{m} \frac{\mathcal{M}(y_{s_i}; w_j, \mu_j, \sigma_j)}{\mathcal{W}(y_{s_i}; \min, C, \alpha)} \prod_{i=1}^{m} \frac{k_0 \mathcal{E}(y_{s_{i-1}, s_i}; \alpha_0)}{\mathcal{U}(y_{s_{i-1}, s_i}; 0, \frac{\pi}{2})}$$
(10)

is invariant under affine transformations of the gray levels in the following sense: Let $I = (I_s)$ be the realization of the gray levels and consider the image $I' = (I'_s)$ obtained by a pointwise affine transformation $I'_s = aI_s + b$. Then, $y'_s = ay_s$ and $y'_{s,t} = y_{s,t}$. Under the change of variable $y'_s = ay_s$, the distributions for y'_s correspond to $\mu'_j = a\mu_j$, $\sigma'_j = a\sigma_j$, $\alpha' = a\alpha$, min' = $a \min$, whereas ω_j and C will remain unchanged. One can then check directly that $\frac{\mathcal{M}(ay_s;w_j,a\mu_j,a\sigma_j)}{\mathcal{W}(ay_s;a\min,C,a\alpha)} = \frac{\mathcal{M}(y_s;w_j,\mu_j,\sigma_j)}{\mathcal{W}(y_s;\min,C,\alpha)}$. Finally, the second factor remains identical. Note, however, that, in our tests, we fix min = -0.001 so that the invariance property holds only approximately.

We can then define the pseudolikelihood of a deformation θ by

$$P(y|\theta) = P(y|c_{\theta}), \tag{11}$$

where c_{θ} is obtained by interpolation of the polygonal curve with vertices given by the points of γ_{θ} .

3.5 Pseudolikelihood of the Augmented Data

Let $x = (x_s)$ be a classification of the pixels in the image into K equivalence classes according to the colors as in Section 2.2, with $x_s \in \Lambda = \{f_1, \ldots, f_K\}$. We consider the global constraint

$$V(x,\theta) = \sum_{k=1}^{K} p_k(\theta) \left| \left\{ s : x_s = f_k, s \notin c_{\theta}^{int} \right\} \right|, \qquad (12)$$

where $p_k(\theta)$ is the proportion of the points in the interior c_{θ}^{int} of the curve having label f_k . Note that the second factor is equal to $|\{s : x_s = f_k\}| - Np_k(\theta)$, where $N = |\{s : x_s = f_k, s \in c_{\theta}^{int}\}|$. So, only $p_k(\theta)$ needs to be computed dynamically, whereas $|\{s : x_s = f_k\}|$ can be precomputed. Thus, $V(x, \theta)$ is minimal whenever the region labels inside an object are specific to that object, i.e., the region labels inside an object do not occur outside that object and vice versa. Thus, we are working under the following hypothesis:

Object specificity: the color labels inside and outside the object are distinct. (13)

In Section 6, a variant of this global constraint is presented in the context of multiple occurrence of the object. In Section 7.2, we suggest a variant to handle the case of strong occlusions of the object. In [19], [25], [29], [35], the distribution of the gray levels of the object is also learned from a training set. In this paper, we consider the problem where the object has a graylevel appearance that might vary according to the image. On the other hand, we assume that the color distribution of the object differs sufficiently from the background, as expressed by the object specificity constraint.

We consider the couple of random fields (Y, X). The hidden discrete random field X is not observable, but it can nevertheless be deduced from the observable random field of colors W, for instance as in Section 2.2. We define the pseudolikelihood of the augmented data by

$$P(y, x|\theta) \propto P(y|\theta)e^{-V(x,\theta)}$$
 (14)

Again, for an image *I*, there is a statistical dependence between the random field Y = Y(I) and the hidden discrete field X = X(I) that is estimated from the random field of colors jointly with its statistical parameters (Section 2.2). In order to have a likelihood distribution from (14), we would need a normalizing factor

$$Z(\theta) = \int P(y(I)|\theta) e^{-V(x(I),\theta)} dI.$$
(15)

The likelihood would then be equal to

$$P(I|\theta) = \frac{1}{Z(\theta)} P(y(I), x(I)|\theta).$$
(16)

Note also that, since the contours and the colors are correlated, one should estimate jointly the statistical parameters of the gradient and the color models, as well as the segmentation x. However, we preferred to estimate them independently in this paper for the sake of simplicity, thus obtaining a hierarchical approximation of the full estimate.

3.6 Pseudoposterior Distribution

We consider the pseudoposterior distribution of a deformation conditional to (y, x),

$$P(\theta|y,x) \propto P(y|\theta)e^{-V(x,\theta)}P(\theta).$$
(17)

Writing $c_{\theta} = (s_0, s_1, \dots, s_m)$, let $V_1(\theta, y, x | \chi_g)$ be the Gibbs field defined by

$$\sum_{i=0}^{m} \log \left(\mathcal{W}(y_{s_{i}}; \min, C, \alpha) / \mathcal{M}(y_{s_{i}}; w_{j}, \mu_{j}, \sigma_{j}) \right) \\ + \sum_{i=1}^{m} \log \left(\mathcal{U}(y_{s_{i-1}, s_{i}}; 0, \frac{\pi}{2}) / k_{0} \, \mathcal{E}(y_{s_{i-1}, s_{i}}; \alpha_{0}) \right) \\ + \sum_{k=1}^{K} p_{k}(\theta) \left| \left\{ s : x_{s} = f_{k}, s \notin c_{\theta}^{int} \right\} \right| + \frac{1}{2} \xi^{t} (I_{q} - \sigma^{2} \Lambda_{q}^{-1}) \xi.$$
(18)

Then, the pseudoposterior distribution is given by

$$P(\theta|y,x) \propto \exp(-V_1(\theta,y,x)), \tag{19}$$

where the factor depends only on the image and not on the deformation of the shape.

4 STOCHASTIC LOCALIZATION OF A SHAPE

We view, similarly to [11], [12], [7], and [20], the localization of a shape c_0 in an image as finding its deformation θ that maximizes the pseudoposterior distribution $P(\theta|y, x)$. So, in order to localize a shape in an image, we want to minimize the Gibbs field $V_1(\theta, y, x)$ of Section 3.6, as a function of θ . Our assumption is that an optimal solution for the function V_1 is the desired deformation of the shape. This will be the case, as shown in our tests, under the hypothesis of object specificity (see (13) of Section 3.5), provided there is not a strong occlusion of the object.

In order to solve the optimization problem, we resort to the stochastic algorithm [43] for which the asymptotic convergence has been proven. In this paper, we use the version presented in Table 1. Its asymptotic convergence follows from [44]. We had already used the ES algorithm in the context of localization of shapes in [36], [40], [49] and in the context of 3D-reconstruction in [6]. The main point here is that the use of the global constraint of Section 3.5 makes the algorithm converge a lot more rapidly than before. In

TABLE 1 Version of the E/S Algorithm Used in Section 4

$\mid E$	a finite discrete subset (using the ε -machine)					
12	of the Cartesian product of Section IV:					
	of the Catesian product of Section 1^{v} .					
	$[0, M-1] \times [0, N-1] \times [\rho_1 a, \rho_2 a]$					
	$\times [0, 2\pi] \times [-\psi_0, \psi_0]^2 \times [-1, 1]^q;$					
l	number of coordinates (here $l = q + 6$);					
V_1	the Gibbs field of equation (18) defined on E :					
r	a real number in the interval $(0, 1)$ called the					
	radius of exploration with r greater than the					
	a mashing (here $n = 1$):					
D 1	ε -machine (nere, $r = \frac{1}{8}$),					
$D = \frac{1}{2r}$	the diameter of exploration;					
au	the initial temperature (here, $\tau = 15$);					
n	an integer greater than 1 (here, $n = 30$);					
$\vec{\theta}$	an element $(\theta_1, \dots, \theta_m)$ of E^n , called a popu-					
-	lation:					
la la	the current iteration:					
ĸ						
p_k	the probability of exploration.					
Goal: to m	inimize V. on F.					
Guai. to m	$\lim_{n \to \infty} V_1 \cup V_1 \cup V_2 \cup$					
Initialization	<i>i</i> : Initialize randomly $\theta_i^{(i)}$, for $i = 1,, n$. $k \leftarrow$					
1.	196000 M					
while $k \leq 4$	36 do					
Update k	$\leftarrow k+1.$					
Determine	the best current solution $\alpha(\overline{\theta}^{[k]}) = \theta_i^{[k]}$, where					
i						
is defined	by $V_1(\theta_{i}^{[k]}) > V_1(\theta_{i}^{[k]})$ for $i < i$, and					
$V(o^{[k]}_{j}) > V(o^{[k]}_{j}) f = i > i$						
$v_1(\sigma_j) \ge v_1(\sigma_i)$ for $j > i$.						
for $i = 1, 2,, n$ do						
Let u be a random number between 0 and 1.						
if $u \leq p_k = k^{-\frac{1}{ au}}$ then						
<i>Exploration:</i> draw m according to the binomial						
distribution $b(D-1, \frac{1}{2})$. Set $\theta = \theta^{[k+1]} = \theta^{[k]}$						
for $m \perp 1$ times do						
Se	$t \theta - \theta^{[k+1]}$					
Bouless with probability 1 each of the $1 1$						
K	proce with probability $\frac{1}{l}$ each of the <i>l</i> coordi-					
nates of θ by a random number within a distance						
of	r to the coordinate.					
W	ith probability $\frac{1}{2}$, add to the fourth coordinate					
(re	presenting the angle of rotation) a random					
m	ılti-					
nle	e of $\frac{\pi}{2}$. Let $\theta^{[k+1]}$ be the resulting solution					
end	for					
plep						
Solor	etion: Set $\theta^{[k+1]} = \alpha(\vec{\theta}^{[k]})$					
selection. Set $v_i = \alpha(v^{i,j})$.						
enu n						
end uchil-						
end while						

particular, we do not need anymore the (complicated) initialization procedures proposed in [36], [40], [49].

Given an image of dimension $M \times N$, we consider the function V_1 of (18) on a domain of the form

$$(\tau_x, \tau_y, s, \psi, \psi_x, \psi_y) \in [0, M - 1] \times [0, N - 1] \times [\rho_1 d, \rho_2 d] \times [0, 2\pi] \times [-\psi_0, \psi_0]^2$$
(20)
$$\xi = (\xi_1, \dots, \xi_q) \in [-1, 1]^q,$$

where $d = \sqrt{M^2 + N^2}$ is the diameter of the image, $0 < \rho_1 < \rho_2 \le 0.5$, and $0 \le \psi_0 < \frac{\pi}{2}$. The choice for the interval associated to the scalar factor *s* implicitly assumes an a priori on the dimension of the shape. This is reasonable since, in applications of localization of shapes, one usually knows the relative size of the shape (for instance, an anatomic object in a medical image). We experimented with $\rho_1 = \rho_2/2$ (see Fig. 7 in the Appendix, which can be found at http://computer.org/ tpami/archives.htm). Also, the choice for the interval of ψ_x and ψ_y is guided by the restriction on projective deformations that is allowed for the shape. In our tests, $\psi_0 = \frac{\pi}{8}$. In our implementation in C++, the domain of (20) is discretized naturally by the precision of the representation of the floats.

Note that it seems preferable, in practice, to do a few trials with different seeds for the generator of pseudorandom numbers, rather than increase the number of iterations with the same seed. In our tests, 20 stochastic optimizations are performed with different seeds. Each search is limited to 436 iterations. The parameters of the E/S algorithm are set equal to $r = \frac{1}{8}$, $\tau = 15$, and n = 30 in Table 1. For the two hard examples of Sections 7.1 and 7.2, after the 20 stochastic optimizations on the full image, we perform 20 more stochastic optimizations with τ_x and τ_y restricted to a rectangular neighborhood of the position coordinates of the best solution previously found (the radius is M/8 for τ_x and N/8 for τ_y).

5 GENERAL MODEL FOR DEFORMATIONS OF A SHAPE

In this section, we present a statistical model for deformations of a shape that extends the PPCA of Section 3. The main point is to replace a single Gaussian distribution by a more flexible mixture of Gaussian kernels.

5.1 Modified Training Phase

Let $T = (T_1, \ldots, T_d) = (A_0, B_0, \ldots, A_m, B_m)$ be the random vector of a template as in (1). We consider the model $T = W\Xi + \nu + \varepsilon$, with W a $d \times q$ orthonormal matrix, ν a vector of dimension $d, \Xi \sim \sum_{i=1}^{\ell} \pi_i \mathcal{N}(\zeta_i, \Sigma_i)$, and $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_d)$. Note that, unlike the PPCA, we require that W be orthonormal (i.e., $W^tW = I_q$). But then, the distribution of Ξ is allowed to be any mixture of Gaussian kernels. One can show that T has distribution

$$p_T(t) = \sum_{i=1}^{\ell} \pi_i \mathcal{N}(t; W\zeta_i + \nu, \sigma^2 I_d + W\Sigma_i W^t).$$
(21)

Now, let t_1, \ldots, t_N be an i.i.d. sample of the variable *T*. We consider the Principal Component Analysis (PCA) reconstruction operator

$$\gamma_{pca}(\xi) = U_q \xi + \nu, \qquad (22)$$

where U_q and $\nu = \bar{t}$ are as in (2). Thus, we now take $W = U_q$ instead of $U_q(\Lambda_q - \sigma^2 I_d)^{1/2}$. That choice minimizes the reconstruction error. The reduced dimension q is chosen so that the minimal reconstruction error is smaller than a certain value.

Our goal is then to estimate the vector of parameters $\chi = (\pi_i, \zeta_i, \Sigma_i, \sigma^2)$. For instance, in the case of $\ell = 1$, the estimation of the parameters in the sense of the ML is given by $\zeta_1 = 0$, $\Sigma_1 = \Lambda_q - \sigma^2 I_q$, and $\sigma^2 = \frac{1}{(d-q)} \sum_{i=q+1}^d \lambda_i$, as follows from (2). Thus, in that case, we recover the PPCA.

TABLE 2 Gibbs Sampler Used in Section 5.1



Note that the proposed mixture model differs from the mixture of PPCA [37] used in [6]. In our case, there is only one reconstruction operator (W, ν) , whereas, in [37], the reconstruction operator varies with the kernel.

Under the Bayesian paradigm, we set the following usual priors [55] on the parameters π_i , ζ_i , Σ_i , and σ^2 ,

1. A Dirichlet prior on the mixture parameters π_i

$$(\pi_1,\ldots,\pi_\ell) \sim \mathcal{D}(A_0;\alpha_1,\ldots,\alpha_\ell),$$
 (23)

where $A_0 = \ell$, and $\alpha_1 = \ldots = \alpha_\ell = \frac{1}{\ell}$. Recall that the Dirichlet distribution is defined by $\mathcal{D}(\pi_1, \ldots, \pi_\ell | A_0; \alpha_1, \ldots, \alpha_\ell) = \frac{\Gamma(A_0)}{\prod_{i=1}^{\ell} \Gamma(A_0 \alpha_i)} \prod_{i=1}^{\ell} \pi_i^{A_0 \alpha_i - 1}$.

2. An inverted Wishart prior on the covariance matrix Σ_i

$$\Sigma_i \sim \mathcal{IW}(\Lambda_0, d_0),$$
 (24)

where Λ_0 and d_0 are as in Table 2. Recall that the inverted Wishart distribution of dimension q is defined by

$$\mathcal{IW}(\Sigma|\Lambda_0, d_0) = |\Lambda_0|^{\frac{1}{2}d_0}|\Sigma|^{-\frac{1}{2}(d_0+q+1)}e^{-\frac{1}{2}tr\Lambda_0\Sigma^{-1}}2^{\frac{1}{2}d_0q}\Gamma_q\left(\frac{1}{2}d_0\right)^{-1},$$

where $\Gamma_q(\frac{1}{2}d_0) = \prod_{j=1}^q \Gamma(\frac{1}{2}d_0 - \frac{1}{2}(j-1)).$ 3. A Gaussian conditional prior on the mean ζ_i

$$\zeta_i | \Sigma_i \sim \mathcal{N}\left(\zeta_0, \frac{1}{k_0} \Sigma_i\right),$$
(25)

where ζ_0 and $k_0 = 0.01$ are as in Table 2. 4. A noninformative improper prior on σ^{-2}

$$\sigma^{-2} \sim 1. \tag{26}$$

We are then interested in the mean estimator of χ

$$\hat{\chi} = \int_{\chi} \chi P(\chi | t_1, \dots, t_N) d\chi.$$
(27)

We compute an approximate value of the estimator using a Monte Carlo Markov Chain (MCMC) algorithm. In order to do so, we consider a latent discrete variable *C* taking its values in the set of hidden labels $\{g_1, \ldots, g_\ell\}$ indicating the Gaussian kernel. Furthermore, we consider Ξ as a latent continuous variable. This is equivalent to the following distributions:

$$t \mid \xi, \sigma^2 \sim \mathcal{N}(W\xi + \nu; \sigma^2 I_d)$$

$$\xi \mid c = g_i \sim \mathcal{N}(\zeta_i, \Sigma_i) \qquad (28)$$

$$c \sim \mathcal{M}(\pi_1, \dots, \pi_\ell),$$

where $\ensuremath{\mathcal{M}}$ denotes the multinomial distribution. We deduce the distributions

$$\begin{aligned} \xi \mid t, c &= g_i, \sigma^2 \sim \mathcal{N}(\tilde{\zeta}, \tilde{\Sigma}), \\ \tilde{\zeta} &= (\sigma^{-2}I_q + \Sigma_i^{-1})^{-1} \big(\Sigma_i^{-1}\zeta_i + \sigma^{-2}W^t(t-\nu) \big), \\ \tilde{\Sigma} &= (\sigma^{-2}I_q + \Sigma_i^{-1})^{-1}; \\ \sigma^{-2} \mid t, \xi \sim \mathcal{G} \bigg(d/2 + 1, \frac{1}{2} \| t - W\xi - \nu \|^2 \bigg), \end{aligned}$$
(29)

where $\mathcal{G}(x; a, b)$ denotes the Gamma distribution $\frac{b^a}{\Gamma(a)}x^{a-1}e^{-bx}$. We can now simulate from the *augmented* posterior distribution of $(\xi_j, c_j, \pi_i, \zeta_i, \Sigma_i, \sigma^2)$ conditional to t_1, \ldots, t_N by the straightforward Gibbs sampler of Table 2. The model parameters are learned from a sample set upon iterating the Gibbs sampler for a few hundred iterations and then taking the average value of each parameter. The initialization of the simulation is presented in Table 2.

5.2 Modified Gibbs Energy

Let Θ be the random variable corresponding to the vector of deformations of Section 3.2. We model the distribution of Θ by

$$P(\theta) = \mathcal{U}(\tau_x, \tau_y, s, \psi, \psi_y, \psi_x) p_T(\gamma_{pca}(\xi)), \qquad (30)$$

where the first factor is as in (6), whereas p_T is now given by (21).

We have the following lemma.

Lemma 1. The distribution $p_T(\gamma_{pca}(\xi))$ is equal to

$$\sum_{i=1}^{\ell} \pi_i \frac{1}{(2\pi)^d |\sigma^2 I_d + U_q \Sigma_i U_q^t|^{1/2}} \times \exp\left(-\frac{1}{2\sigma^2} (\xi - \zeta_i)^t \left(I_q - \left(I_q + \sigma^2 \Sigma_i^{-1}\right)^{-1}\right) (\xi - \zeta_i)\right).$$

Proof. See [50].

Accordingly, the Gibbs energy of (18) is replaced by $V_2(\theta, y, x)$ equal to

$$\sum_{i=0}^{m} \log(\mathcal{W}(y_{s_{i}};\min,C,\alpha)/\mathcal{M}(y_{s_{i}};w_{j},\mu_{j},\sigma_{j})) + \sum_{i=1}^{m} \log\left(\mathcal{U}\left(y_{s_{i-1},s_{i}};0,\frac{\pi}{2}\right)/k_{0} \mathcal{E}(y_{s_{i-1},s_{i}};\alpha_{0})\right) + \sum_{k=1}^{K} p_{k}(\theta) \left| \left\{ s : x_{s} = f_{k}, s \notin c_{\theta}^{int} \right\} \right|$$

$$- \log\left\{ \sum_{i=1}^{\ell} \pi_{i} \frac{1}{(2\pi)^{d} |\sigma^{2}I_{d} + U_{q}\Sigma_{i}U_{q}^{t}|^{1/2}} \times \exp\left(-\frac{1}{2\sigma^{2}} (\xi - \zeta_{i})^{t} \left(I_{q} - (I_{q} + \sigma^{2}\Sigma_{i}^{-1})^{-1}\right) (\xi - \zeta_{i})\right) \right\}.$$
(31)

Now, we consider ℓ auxiliary change of variables, one for each Gaussian kernel

$$\xi = \varphi_i(\tilde{\xi}) = A_i \tilde{\xi} + \zeta_i, \tag{32}$$

for $i = 1, ..., \ell$, where $\Sigma_i = A_i A_i^t$. Instead of (20), we now consider the domain

$$\begin{aligned} (\tau_x, \tau_y, s, \psi, \psi_x, \psi_y) &\in [0, M-1] \times [0, N-1] \\ &\times [\rho_1 d, \rho_2 d] \times [0, 2\pi] \times [-\psi_0, \psi_0]^2 \\ &\tilde{\xi} \in [-1, 1]^q. \end{aligned}$$
(33)

This is equivalent to a parameterization of the domain for ξ . In particular, it would not make sense to require that $\xi \in [-1, 1]^q$. In the E/S algorithm of Table 1, the choice of the auxiliary change of variables is modified randomly with probability 1/2.

6 GLOBAL CONSTRAINT IN THE CASE OF MULTIPLE OCCURRENCE OF THE OBJECT

When an object is expected to appear in multiple instances in the image, the specificity hypothesis (13) is replaced by the following hypothesis:

Local object specificity: the color labels inside and outside the object are distinct within a neighborhood (34) of the object.

Formally, let $x = (x_s)$ be a classification of the pixels in the image into K equivalence classes according to the colors as in Section 2.2, with $x_s \in \Lambda = \{f_1, \ldots, f_K\}$. Consider the global constraint

$$V_{loc}(x,\theta) = \sum_{k=1}^{K} p_k(\theta) \big| \big\{ s : x_s = f_k, s \in c_{\theta}^{nbhd} \setminus c_{\theta}^{int} \big\} \big|, \quad (35)$$

where $p_k(\theta_k)$ is the proportion of the points inside the curve γ_{θ} having label f_k , c_{θ}^{int} is the interior of the curve, and c_{θ}^{nbhd} is a neighborhood of the curve. Thus, $V_{loc}(x,\theta)$ is minimal whenever the region labels inside an object are *locally* specific to that object. In our tests, we took a neighborhood of radius of 25 pixels. Accordingly, (31) is replaced by $V_3(\theta, y, x)$:

$$\sum_{i=0}^{m} \log(\mathcal{W}(y_{s_{i}};\min,C,\alpha)/\mathcal{M}(y_{s_{i}};w_{j},\mu_{j},\sigma_{j})) + \sum_{i=1}^{m} \log(\mathcal{U}(y_{s_{i-1},s_{i}};0,\frac{\pi}{2})/k_{0} \mathcal{E}(y_{s_{i-1},s_{i}};\alpha_{0})) + \sum_{k=1}^{K} p_{k}(\theta)|\{s:x_{s}=f_{k},s\in c_{\theta}^{nbhd}\setminus c_{\theta}^{int}\}|$$

$$\log\left\{\sum_{i=1}^{\ell} \pi_{i}\frac{1}{(2\pi)^{d}|\sigma^{2}I_{d}+U_{q}\Sigma_{i}U_{q}^{t}|^{1/2}}\right\}$$
(36)

× exp
$$\left(-\frac{1}{2\sigma^2}(\xi-\zeta_i)^t \left(I_q - (I_q + \sigma^2 \Sigma_i^{-1})^{-1}\right)(\xi-\zeta_i)\right)\right\}.$$

See Fig. 4 for an example of multiple occurence of an object.

7 EXPERIMENTAL RESULTS

Since Destrempes' PhD thesis [50], some of the tests had to be rerun with an improved implementation. Furthermore,



Fig. 2. Top row: The four mean shapes for the MPCA prior in the example of two gymnastic movements (flying bird and weight-lifting); the rightmost image represents the mean shape of the single PPCA model. Second and third rows: Examples of localization of a shape obtained by stochastic optimization of the Gibbs field based on the contour parameters estimated by the ICE procedure and the segmentation computed with the ESE procedure using the MPCA model of four kernels. Fourth row: Five of the six solutions that were counted as failures over 50 tests. Bottom row: The 30 solutions that were chosen randomly to initialize the E/S algorithm with one of the seeds for the generator of pseudorandom numbers.

we now include a quantitative evaluation of the method. Thus, some of the results reported here differ from [50].

The localization procedure crucially depends on the color segmentation of the image. We have experimented the procedure with the following four segmentation models:

- 1. the color model of Section 2.2 (with $K = K_1 = 30$),
- 2. the Mean Shift algorithm [56], with a spatial bandwidth of 7 and a range bandwidth of 6.5,
- 3. the color model of [44] (with K = 30), and
- 4. the K-means algorithm (with K = 30).

See Fig. 1 for examples of segmentation according to each of the four models. We have also experimented with the following schemes:

- 5. the color model of Section 2.2 without the contour model of Section 2.1, and
- 6. the contour model of Section 2.1 and no segmentation model.

When comparing a solution c_{θ} with the ground-truth *C* (obtained manually), we consider the following error measure

$$\delta(c_{\theta}, C) = \frac{1}{diam(C)|C|} \sum_{p \in C} d(p, c_{\theta}), \qquad (37)$$

where |C| is the number of points on C, $d(p, c_{\theta})$ is the euclidean distance between $p \in C$ to the closest point of c_{θ} ,

and diam(C) is the diameter of C (i.e., the greatest euclidean distance between two points of C). We considered that a failure corresponds to an error δ that is greater than 0.08. This criterion is of course arbitrary, but we actually checked that the resulting classification corresponded to our own (personal) judgment of an "acceptable" localization.

We now present two challenging examples. Three additional examples can be found in Appendix I, which can be found at http://computer.org/tpami/archives.htm.

7.1 Example 1: Gymnastic Movements

A first experiment uses a rather variable shape. Namely, we took two video sequences of two gymnastic movements of the arms: A flying bird movement and a weight-lifting movement. The database consists of 180 frames of the first execution of the two movements in the video sequences. Each curve is represented by a template of 200 points. The training phase of Section 5.1 yields a reduced dimension of q = 6 for the statistical deformations with less than 0.3 percent for the relative reconstruction error. We fixed the number ℓ of Gaussian kernels to 4.

We have tested the localization procedure with the deformation model of (31) on 10 images with five different initial seeds for the generator of pseudorandom numbers. The scaling factors were $\rho_1 = 0.15$ and $\rho_2 = 0.30$. The 10 images were taken in the second execution of the two movements in the video sequences. See Fig. 2 for examples of localization.



Fig. 3. First and second rows: Examples of localization of a shape obtained by stochastic optimization of the Gibbs field based on the contour parameters estimated by the ICE procedure and the segmentation computed with the ESE procedure, using the MPCA model of four kernels. Bottom row: The five solutions that were counted as failures over 50 tests.



Fig. 4. Examples of localization for the saxophone shape of Appendix I-C (which can be found at http://computer.org/tpami/archives.htm) with multiple occurrence of the object.

For image (1) of Fig. 2, the whole procedure takes about 63 minutes. The segmentation method takes 90 percent of that time. When using the K-means algorithm, the whole procedure takes only 9 minutes and 9 seconds.

The error rate was 12 percent with the segmentation method of Section 2.2. Image (4) presents three failures, image (5) presents two failures, and image (10) presents one failure. The error rate was 8 percent with the same color model but with the contour energy term dropped. However, the error rate was as much as 75 percent with that strategy in the example of Appendix I-B, which can be found at http:// computer.org/tpami/archives.htm.

In the case of the segmentation model of Section 2.2, out of the five initial seeds, the lowest value of the Gibbs energy corresponded to a good localization. Thus, the error rate for the five initial seeds is actually 0 percent. In the case of the same color model but with the contour energy term dropped, the error rate for the five initial seeds is also 0 percent on this data set.

We compared the MPCA shape prior with four kernels with a MPCA consisting of only one kernel. The error rate was 38 percent. Thus, we are inclined to think that, for very variable shapes, a simple PPCA (or PCA) is not adequate to capture the statistics of the deformations of the shape.

Method	Constraint	Shape prior	(1)	(2)	(3)	(4)	(5)	(6)
gym	global	MPPCA ($\ell = 4$)	12%	84%	32%	44%	8%	60%
gym	global	MPPCA $(\ell = 1)$	38%	Τ	Ι	Ι	Ι	-
gym+occlusion	global	MPPCA ($\ell = 4$)	10%	96%	24%	22%	86%	94%
gym+occlusion	global	MPPCA ($\ell = 1$)	36%	_	_	_	-	-
guitar	global	PCA $(\ell = 1)$	4.7%	38.7%	16.7%	7.3%	13.3%	45.3%
van	global	MPPCA $(\ell = 2)$	5.8%	15.8%	4.2%	8.3%	75%	19.2%
saxophone	local	MPPCA $(\ell = 1)$	0.0%	16.7%	6.7%	0%	63.3%	13.3%
saxophone	global	MPPCA ($\ell = 1$)	70%	_	-	_	_	-

TABLE 3 Proportions of Wrong Solutions

(i.e., with a relative error between the answer and the solution found by the algorithm that is greater than 0.08) for the examples of Sections 7.1 and 7.2 and Appendix I-A, I-B, and I-C (which can be found at http://computer.org/tpami/archives.htm). Segmentation models (1) to (6) as in Section 7.

7.2 Example 2: Gymnastic Movements with Occlusions

A second experiment uses the same shape model as in the previous example. This time, the localization procedure is tested on 10 images where there are partial occlusions of the shape, with five different initial seeds for the generator of pseudorandom numbers. The scaling factors were $\rho_1 = 0.15$ and $\rho_2 = 0.30$. The 10 images were taken in the execution of the two gymnastic movements in two video sequences. See Fig. 3 for examples of localization. The error rate was 10 percent with the segmentation method of Section 2.2. Image (1) presents one failure, image (2) presents one failure, image (5) presents two failures, and image (9) presents one failure. The error rate was 22 percent with the K-means algorithm.

In the case of the segmentation model of Section 2.2, out of the five initial seeds, the lowest value of the Gibbs energy corresponded to a good localization except for two images. Thus, the error rate for five initial seeds is actually 20 percent in this case. So, we suggest the following modification of the object specificity constraint in order to handle strong occlusions:

$$V(x,\theta) = \sum_{k=1}^{K} B(p_k(\theta)) p_k(\theta) \big| \big\{ s : x_s = f_k, s \notin c_{\theta}^{int} \big\} \big|, \quad (38)$$

where B(x) could be of the form $1 - \frac{1}{1+(x/x_0)^{2n}}$ (see the Butterworth filter) in order to attenuate the influence of smaller proportions $p_k(\theta)$ due to occlusions. We have not tested this constraint as of now.

We compared the MPCA shape prior with four kernels with a MPCA consisting of only one kernel. The error rate was 36 percent. Thus, we are inclined to think that, for very variable shapes, a simple PPCA (or PCA) is not adequate to capture the statistics of the deformations of the shape.

As can be seen from Table 3, the proposed color model outperforms the other models, except for time considerations. In our opinion, this is due to the fact that, for models (2)-(4), the distribution of each color region is unimodal, whereas, in the case of the proposed model, the region distributions are multimodal. It follows that methods (2)-(4) yield to an oversegmentation of the image, which in turn makes the optimization task too difficult. Also, contours or regions alone are not sufficient to efficiently localize the shape. In [36], [40], [49], the success rates were much higher for most images because initialization procedures were used. In this paper, we

have dropped those initialization procedures in order to show that a localization procedure based solely on contours is a much harder optimization problem. Furthermore, despite various initialization procedures, there were still a few images that presented a 0 percent success rate, whereas now each image has at least a 20 percent success rate. So, we are inclined to think that the proposed model is more adequate than one that is based solely on contours.

8 CONCLUSION

In this paper, we have presented a coherent statistical model for deformations of shapes. We have brought together the following ideas: The prior distribution of deformations can be learned using the PPCA or the proposed mixture of PPCA. The pseudolikelihood distribution of deformations is based on a statistical model for the gradient vector field of the gray level in the image and can be estimated using an ICE procedure. A criterion of global or local object specificity makes the localization of the shape a lot easier. This criterion is based on a color segmentation of the image that can be computed with an ESE procedure. The optimization E/S algorithm converges asymptotically to an optimal solution, in the sense of the MAP in our context. The error rates with our method were 12 percent, 10 percent, 4.7 percent, 5.8 percent, and 0 percent for five sets of experiments. In future work, we intend to develop a (much) more efficient version of the segmentation method used in this paper in terms of the computational time.

ACKNOWLEDGMENTS

The authors thank FCAR (Fonds formation chercheurs & aide recherche, Quebec, Canada) for financial support of this work. They thank the music stores Archambault and Steve's Music in Montreal for kindly allowing them to take pictures of classical guitars. The authors are grateful to the three anonymous reviewers for their numerous comments and suggestions that helped improve both the scientific content and the presentation of this paper.

REFERENCES

 M.-P.D. Jolly, S. Lakshmanan, and A.K. Jain, "Vehicle Segmentation and Classification Using Deformable Templates," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, no. 3, pp. 293-308, Mar. 1996.

- M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Hybrid [2] Genetic Optimization and Statistical Model-Based Approach for the Classification of Shadow Shapes in Sonar Imagery," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 2, pp. 129-141, Feb. 2000.
- A.K. Jain and D. Zongker, "Representation and Recognition of [3] Handwritten Digits Using Deformable Templates," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 12, pp. 1386-1391, Dec. 1997.
- [4] C. Kervrann and F. Heitz, "A Hierarchical Markov Modeling Approach for the Segmentation and Tracking of Deformable Shapes," Graphical Models and Image Processing, vol. 60, no. 3, pp. 173-195, 1998.
- S. Benameur, M. Mignotte, H. Labelle, and J.A. de Guise, "A [5] Hierarchical Statistical Modeling Approach for the Unsupervised 3D Biplanar Reconstruction of the Scoliotic Spine," IEEE Trans. Biomedical Eng., vol. 52, no. 12, pp. 2041-2057, 2005.
- S. Benameur, M. Mignotte, F. Destrempes, and J.A. de Guise, "3D [6] Biplanar Reconstruction of Scoliotic Rib Cage Using the Estimation of a Mixture of Probabilistic Prior Models," IEEE Trans. Biomedical Eng., vol. 52, no. 10, pp. 1713-1728, 2005.
- A.K. Jain, Y. Zhong, and S. Lakshmanan, "Object Matching Using [7] Deformable Templates," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18, no. 3, pp. 267-278, Mar. 1996.
- [8] D.J. Burr, "Elastic Matching of Line Drawings," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 3, no. 6, pp. 708-713, 1981.
- M. Moshfeghi, S. Ranganath, and K. Nawyn, "Three-Dimensional [9] Elastic Matching of Volumes," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 16, no. 2, pp. 128-138, Feb. 1994.
- M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active Contour [10] Models," Int'l J. Computer Vision, vol. 1, no. 4, pp. 321-331, 1988.
- [11] M. Figueiredo and J. Leitao, "Bayesian Estimation of Ventricular Contours in Angiographic Images," IEEE Trans. Medical Imaging, vol. 11, no. 3, pp. 416-429, 1992.
- [12] G. Storvik, "A Bayesian Approach to Dynamic Contours through Stochastic Sampling and Simulated Annealing," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 16, no. 10, pp. 976-986, Oct. 1994.
- [13] D. Mumford and J. Shah, "Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems," Comm. Pure Applied Math., vol. 42, pp. 577-685, 1989.
- [14] U. Grenander, Pattern Synthesis: Lectures in Pattern Theory. Springer, 1976.
- U. Grenander and M.I. Miller, "Representations of Knowledge in [15] Complex Systems," J. Royal Statistical Soc. (series B), vol. 56, no. 4, pp. 549-603, 1994.
- [16] L.H. Staib and J.S. Duncan, "Boundary Finding with Parametric Deformable Models," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 14, no. 11, pp. 1061-1075, Nov. 1992.
- [17] A. Chakraborty, L.H. Staib, and J.S. Duncan, "Deformable Boundary Finding Influenced by Region Homogeneity," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 624-627, June 1994.
- [18] T.F. Cootes, C.J. Taylor, and J. Haslam, "The Use of Active Shape Models for Locating Structures in Medical Images," Image and Vision Computing, vol. 12, no. 6, pp. 355-366, 1994.
- [19] 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, 1995.
- [20] M. Mignotte, J. Meunier, and J.-C. Tardif, "Endocardial Boundary Estimation and Tracking in Echocardiographic Images Using Deformable Templates and Markov Random Fields," Pattern Analysis and Applications, vol. 4, no. 4, pp. 256-271, 2001.
- [21] D. Cremers, T. Kohlberger, and C. Schnorr, "Shape Statistics in Kernel Space for Variational Image Segmentation," Pattern Recognition, vol. 36, no. 9, pp. 1929-1943, 2003.
- D. Cremers and C. Schnorr, "Statistical Shape Knowledge in [22] Variational Motion Segmentation," Image and Vision Computing, vol. 21, no. 1, pp. 77-86, 2003.
- [23] M. Bergtholdt, D. Cremers, and C. Schnorr, "Variational Segmentation with Shape Priors," Math. Models in Computer Vision: The Handbook, N. Paragios, Y. Chen, and O. Faugeras, eds., Springer, 2005.
- [24] M. de Bruijne and M. Nielsen, "Shape Particle Filtering for Image Segmentation," Proc. Medical Image Computing and Computer-Assisted Intervention, vol. 1, pp. 168-175, 2004.

- [25] T.F. Cootes, G.J. Edwards, and C.J. Taylor, "Active Appearance Models," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 6, pp. 681-685, June 2001.
- [26] D. Cremers, F. Tischhauser, J. Weickert, and C. Schnorr, "Diffusion Snakes: Introducing Statistical Shape Knowledge into the Mumford-Shah Functional," Int'l J. Computer Vision, vol. 50, no. 3, pp. 295-313, 2002.
- D. Cremers, N. Sochen, and C. Schnorr, "A Multiphase Dynamic Labeling Model for Variational Recognition-Driven Image Segmentation," Int'l J. Computer Vision, vol. 66, no. 1, pp. 67-81, 2006.
- M.E. Leventon, W.E.L. Grimson, and O.D. Faugeras, "Statistical [28] Shape Influence in Geodesic Active Contours," Proc. Int'l Computer Vision Pattern Recognition, pp. 1316-1323, 2000.
- J. Luettin and N.A. Thacker, "Speechreading Using Probabilistic Models," Computer Vision and Image Understanding, vol. 65, no. 2, pp. 163-178, 1997.
- B. Moghaddam and A. Pentland, "Probabilistic Visual Learning for [30] Object Representation," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 696-710, July 1997.
- N. Paragios and M. Rousson, "Shape Priors for Level Set Representations," Proc. European Conf. Computer Vision, pp. 78-92, [31] 2002.
- [32] H. Rue and M.A. Hurn, "Bayesian Object Identification," *Biometrika,* vol. 86, no. 3, pp. 649-660, 1999. Y. Zhong and A.K. Jain, "Object Localization Using Color, Texture
- [33] and Shape," Pattern Recognition, vol. 33, pp. 671-684, 2000.
- L. Liu and S. Sclaroff, "Deformable Shape Detection and [34] Description via Model-Based Region Grouping," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 5, pp. 475-489, May 2001.
- [35] M.J. Black and A. Jepson, "EigenTracking: Robust Matching and Tracking of Articulated Objects Using a View-Based Representation," Int'l J. Computer Vision, vol. 26, no. 1, pp. 63-84, 1998.
- [36] F. Destrempes and M. Mignotte, "Unsupervised Localization of Shapes Using Statistical Models," Proc. Fourth IASTED Int'l Conf. Signal and Image Processing, pp. 66-71, Aug. 2002.
- M.E. Tipping and C.M. Bishop, "Mixtures of Probabilistic Principal Component Analyzers," Neural Computation, vol. 11, no. 2, pp. 443-482, 1999.
- F. Destrempes and M. Mignotte, "A Statistical Model for Contours [38] in Images," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, no. 5, pp. 626-638, May 2004.
- [39] F. Destrempes and M. Mignotte, "Unsupervised Detection and Semi-Automatic Extraction of Contours Using a Statistical Model and Dynamic Programming," Proc. Fourth IASTED Int'l Conf. Signal and Image Processing, pp. 60-65, Aug. 2002.
- [40] F. Destrempes, "Détection Non-Supervisée de Contours et Localisation de Formes à l'Aide de Modèles Statistiques," master's thesis, Université de Montréal, Apr. 2002.
- [41] W. Pieczynski, "Champs de Markov Cachés et Estimation Conditionnelle Itérative," Revue Traitement Du Signal, vol. 11, no. 2, pp. 141-153, 1994.
- F. Salzenstein and W. Pieczynski, "Unsupervised Bayesian Seg-[42] mentation Using Hidden Markovian Fields," Proc. Int'l Conf. Acoustics, Speech, and Signal Processing, vol. 4, pp. 2411-2414, May 1995.
- [43] O. François, "Global Optimization with Exploration/Selection Algorithms and Simulated Annealing," Annals of Applied Probability, vol. 12, no. 1, pp. 248-271, 2002.
- F. Destrempes, M. Mignotte, and J.-F. Angers, "A Stochastic [44] Method for Bayesian Estimation of Hidden Markov Random Field Models with Application to a Color Model," IEEE Trans. Image Processing, vol. 14, no. 8, pp. 1096-1124, 2005.
- [45] D. Crisan and A. Doucet, "A Survey of Convergence Results on Particle Filtering Methods for Practitioners," IEEE Trans. Signal Processing, vol. 50, no. 3, pp. 736-746, 2002.
- Y.Y. Boykov and M.-P. Jolly, "Interactive Graph Cuts for Optimal [46] Boundary & Region Segmentation of Objects in N-D Images," Proc. Int'l Conf. Computer Vision, vol. I, pp. 105-112, July 2001.
- K. Kolmogorov and R. Zabih, "What Energy Functions Can Be [47] Minimized via Graph Cuts?" Proc. European Conf. Computer Vision, vol. III, pp. 65-81, 2002.
- [48] P.J. Green, "Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination," Biometrika, vol. 82, no. 4, pp. 711-732, 1995.

- [49] F. Destrempes and M. Mignotte, "Unsupervised Statistical Method for Edgel Clustering with Application to Shape Localization," Proc. Third Indian Conf. Computer Vision, Graphics, and Image Processing, pp. 411-416, Dec. 2002.
- [50] F. Destrempes, "Estimation de Paramètres de Champs Markoviens Cachés avec Applications à la Segmentation d'Images et la Localisation de Formes," PhD thesis, Université de Montréal, Feb. 2006.
- [51] F. Destrempes, J.-F. Angers, and M. Mignotte, "Fusion of Hidden Markov Random Field Models and Its Bayesian Estimation," *IEEE Trans. Image Processing*, vol. 15, no. 10, pp. 2920-2935, 2006.
- [52] N.S. Matloff, Probability Modeling and Computer Simulation. PWS-KENT Publishing, 1988.
- [53] G. Schwartz, "Estimating the Dimension of a Model," Annals of Statistics, vol. 6, pp. 461-464, 1978.
- [54] E.N. Mortensen and W.A. Barrett, "Interactive Segmentation with Intelligent Scissors," *Proc. Graphical Models and Image Processing*, vol. 60, no. 5, pp. 349-384, Sept. 1998.
- [55] T.W. Anderson, An Introduction to Multivariate Statistical Analysis, second ed. John Wiley & Sons, 1971.
- [56] D. Comaniciu and P. Meer, "A Robust Approach Toward Feature Space Analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603-619, May 2002.
- [57] J.O. Berger, Statistical Decision Theory and Bayesian Analysis, second ed. Springer-Verlag, 1985.



François Destrempes received the BSc degree in mathematics from the Université de Montréal (1985) and the MSc (1987) and PhD (1990) degrees in mathematics from Cornell University. He was a postdoctoral fellow at the Centre de Recherche Mathématiques (CRM) of the Université de Montréal from 1990-1992. He has taught mathematics at Concordia University, the University of Ottawa, the University of Toronto, and the University of Alberta. He also received a

postgraduate degree (2000) in applied computer science and the MSc (2002), and PhD (2006) degrees in computer science from the Université de Montréal. He is presently a postdoctoral fellow at the Laboratoire de Biorhéologie et d'Ultrasonographie Médicale, Centre de Recherche du CHUM, Hôpital Notre-Dame, Université de Montréal. His current research interests include statistical methods for image segmentation, parameters estimation, detection of contours, localization of shapes, applications of stochastic optimization to computer vision, and ultrasonic image segmentation.



Max Mignotte received the DEA (postgraduate degree) in digital signal, image, and speech processing from the INPG University, France (Grenoble), in 1993 and the PhD degree in electronics and computer engineering from the University of Bretagne Occidentale (UBO) and the digital signal laboratory (GTS) of the French Naval academy, France, in 1998. He was an INRIA postdoctoral fellow at University of Montréal (DIRO), Canada (Québec), from 1998 to

1999. He is currently with DIRO at the Computer Vision & Geometric Modeling Lab as an assistant professor (professeur adjoint) at the University of Montreal. He is also a member of LIO (Laboratoire de Recherche en Imagerie et Orthopédie, Centre de Recherche du CHUM, Hôpital Notre-Dame) and a researcher at CHUM. His current research interests include statistical methods and Bayesian inference for image segmentation (with hierarchical Markovian, statistical templates, or active contour models), hierarchical models for high-dimensional inverse problems from early vision, parameters estimation, tracking, classification, shape recognition, deconvolution, 3D reconstruction, and restoration problems.



Jean-François Angers received the BSc (1981) and MSc (1984) degrees in applied mathematics from the Université de Sherbrooke, and the PhD degree (1987) in statistics from Purdue University. He was an assistant professor of statistics at the Université de Sherbrooke from 1987 to 1990 and he is now a full professor at the Université de Montréal. He is also a member of the Centre de Recherche Mathématique and the Centre de Recherche sur les Transports de

l'Université de Montréal. His current research interests include nonparametric Bayesian functional estimation, hierarchical Bayesian model, and Bayesian robust estimation.

▷ For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.