Unsupervised Localization of Shapes Using Statistical Models

François Destrempes and Max Mignotte
DIRO, Département d'Informatique et de Recherche Opérationnelle
Université de Montréal
C.P. 6128, Succ. Centre-Ville
Montréal, Canada (Quebec), H3C 3J7
email : destremp@iro.umontreal.ca

ABSTRACT

In this paper, we present a coherent Markov model for deformations of shapes based on the statistical distribution of the gradient vector field of the gray level and a probabilistic model for reduction of dimension. This allows us to infer the posterior distribution of deformations given the data. The localization of a shape in an image is then viewed as the minimizing of the corresponding Gibbs field. We use a stochastic optimization algorithm in order to find the optimal deformation, firstly with a simpler Gibbs field and then on the complete Gibbs field. This yields an unsupervised method for localization of shapes, provided that the image is not prohibitively too complex.

1 Introduction

Localization of shapes is an interesting problem in Image Processing with applications to detection of shapes and pattern recognition.

In [1], a first Markov model based on contours is presented. The same model has been used with variants in [2], [3] and [4]. Models based on textures have been introduced in [5] and [6]. None of those models takes into account the statistics of the image : the Gibbs fields considered are defined in an *ad hoc* way. However, in [7], a statistical model is presented for tracking deformable shapes.

In this paper, we present a Markov model based on the statistics of the image. We view a shape as a deformable template as in [1] but we model its density function using a probabilistic principal component analysis (PPCA) [8] obtained from a training set aligned as in [9]. This gives us the prior distribution of deformations of a shape. We base the likelihood of deformations on the statistical distribution of the gradient vector field of the gray levels introduced in [10]. We then obtain the posterior distribution of deformations conditional to the observable data. Localization of a shape can then be viewed as minimizing the posterior Gibbs field. We resort to a stochastic optimization algorithm [11] to find the optimal solution. In a first step, we use a simpler Gibbs field [2], more suitable for computational reasons. Then, we use the complete Gibbs field on a restricted domain given by the solution found in the first step.

The remaining part of this paper is organized as follows. In Section 2, we present in details a fundamental tool for our method : probabilistic principal component analysis [8]. In Section 3, we present the statistical model [10] for the gradient vector field of the gray level. Section 4 uses this model to define the likelihood for deformations of a shape, and uses the PPCA to define the prior distribution. The method for localization of shapes is explained in Section 5. In Section 6, we discuss briefly empirical results.

2 Probabilistic model for reduction of dimension

The following model has been introduced in [8]. We consider two random variables T and Ξ related by

$$T \;=\; W\Xi + \nu + \varepsilon$$

where $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_d)$, $\Xi \sim \mathcal{N}(0, I_q)$, and W is a $d \times q$ matrix (I_r denotes the $r \times r$ identity matrix). This is a special case of [12] as presented in [8]. The variable T represents the full data, whereas Ξ represents the reduced data. Usually, the dimension q of Ξ is significantly smaller than the dimension d of T.

Assuming that $\sigma^2 > 0$, one can show that T has distribution function

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

Now, let $t_1, ..., t_n$ be a sample of i.i.d. observations issued from T. We are interested in finding the optimal values of W, ν, σ^2 in the sense of the Maximum Likelihood (ML) ; i.e., which maximize the likelihood function

$$\mathcal{L}(W, \nu, \sigma^2) \;=\; \prod_{i=1}^{n} p_T(t_i).$$

Let $\lambda_1, ..., \lambda_d$ be the eigenvalues of the sample covariance matrix, in decreasing order, Λ_q be the diagonal matrix with entries $\lambda_1, ..., \lambda_q$, and U_q be 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). From [8], an optimal solution to the likelihood function is then given by

$$\begin{aligned}\sigma^2 &= \frac{1}{d-q}\sum_{i=q+1}^{d}\lambda_i\\ \nu &= \bar{t} = \frac{1}{n}\sum_{i=1}^{n} t_i\\ W &= U_q(\Lambda_q - \sigma^2 I_q)^{1/2}.\end{aligned}$$

The following useful fact follows from above by a simple calculation.

Lemma. *Let* $T = U_q(\Lambda_1 - \sigma^2 I_q)^{1/2}\Xi + \bar{t} + \varepsilon$ *with* $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_d)$ *and* $\Xi \sim \mathcal{N}(0, I_q)$. *Then,*

$$p_T(U_q(\Lambda_1 - \sigma^2 I_q)^{1/2}\xi + \bar{t}) \propto \exp\big(-\frac{1}{2}\xi^t(I_q - \sigma^2\Lambda_q^{-1})\xi\big).$$

Proof : First, observe that the inverse matrix of $\sigma^2 I_d + WW^t$ is

$$\sigma^{-2}\Big(I_d - W(W^tW + \sigma^2 I_q)^{-1}W^t\Big).$$

Now, substitute in the expression for p_T.

3 Statistical Model for the gradient vector field

The following model has been introduced in [10]. Given an image of size N, $G = (S, U)$ will denote the non-oriented graph consisting of the 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, $(s,t) \in U$). Each site s and edge (s,t) can be classified as "on" or "off" contour.

For each site $s \in S$, the random variable Y_s represents the norm of the gradient of the gray level at the corresponding pixel. Based on empirical results, we model $P_{\text{off}}(y_s)$ for the sites off contours by a Weibull law [13],

$$\mathcal{W}(y_s; \min, C, \alpha) = \frac{C}{\alpha^C}(y_s - \min)^{C-1}e^{-(\frac{y_s - \min}{\alpha})^C}$$

with $y_s > \min$. For points on contours, we model $P_{\text{on}}(y_s)$ by a mixture of two Gaussian laws,

$$\mathcal{M}(y_s; w_j, \mu_j, \sigma_j) = \sum_{j=1}^{2} w_j \mathcal{N}(y_s; \mu_j, \sigma_j).$$

If $(s,t) \in U$, we consider the distribution of the angle Y_{st} between the mean of the gradient at s and t with the normal to the vector from s to t. The angle is normalized between $-\pi/2$ and $\pi/2$. We have observed that $P_{\text{on}}(|Y_{st}|)$ follows a Weibull law up to a factor

$$k_0\mathcal{W}(y_{st}; \min_0, C_0, \alpha_0)$$

whenever the edge (s,t) belongs to a contour. We impose the condition that $C_0 \leq 1$ and the factor is taken so as to obtain a distribution in the interval $[0, \frac{\pi}{2}]$. In the case where (s,t) is off contours, we model $P_{\text{off}}(|Y_{st}|)$ by a uniform distribution on $[0, \frac{\pi}{2}]$,

$$\mathcal{U}(y_{st}; 0, \frac{\pi}{2}) = \frac{2}{\pi}.$$

An Iterated Conditional Estimation (ICE) procedure [14] is described in [10] for the estimation of the parameters.

4 Model for deformations

4.1 Deformable template representation

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

$$\gamma = (x_1, y_1, ..., x_d, y_d)$$

where the d points $(x_1, y_1), ..., (x_d, y_d)$ are equally spaced on the curve.

Given a sample $\gamma_1, ..., \gamma_n$ of curves with same number of points, we resort to the procedure [9] to align this training set. We then apply the PPCA of Section 2 to reduce the dimension to $q \ll 2d$. Viewing $T = (X_1, Y_1, ..., X_d, Y_d)$ as a random vector, we obtain an optimal model in the sense of ML of the form $T = W\Xi + \nu + \varepsilon$ with $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_d)$, W a $2d \times q$ matrix, and $\Xi \sim \mathcal{N}(0, I_q)$, by taking $W = U_q(\Lambda_q - \sigma^2 I_q)^{1/2}$, $\nu = \bar{t}$, and σ as in Section 2. This gives us non-linear deformations of the mean shape $\gamma_0 = \bar{t}$ and terminates the training phase.

We also consider rigid deformations given by translation (τ_x, τ_y), scaling s, and rotation ψ, applied point-wise to a curve. This yields a vector of deformation

$$\theta = (\tau_x, \tau_y, s, \psi, \xi)$$

of dimension $q + 4$.

The transformation θ is computed in the following order : first calculate $\gamma_\xi = U_q(\Lambda_q - \sigma^2 I_d)^{1/2}\xi + \bar{t}$; then, apply s and ψ ; finally, translate by τ. The resulting template is denoted γ_θ.

4.2 Prior distribution

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

$$P_\Theta(\theta) = \mathcal{U}(\tau_x, \tau_y, s, \psi)p_T(U_q(\Lambda_1 - \sigma^2 I_q)^{1/2}\xi + \bar{t})$$

where $\mathcal{U}$ denotes the uniform distribution and T represents the full data. From Section 2, this yields

$$P_\Theta(\theta) \propto \exp\big(-\frac{1}{2}\xi^t(I_q - \sigma^2\Lambda_q^{-1})\xi\big).$$

4.3 Likelihood

Given a path $c = (s_0, ..., s_m)$ in the graph G of Section 3, we consider as in [10] (and similarly to [15] and [16]) the likelihood $P_{Y/C}(y/c)$ given by

$$\begin{array}{c}\prod_{s\notin c} P_{\text{off}}(y_s) \prod_{s\in c} P_{\text{on}}(y_s) \\ \prod_{(s,t)\notin c} P_{\text{off}}(y_{st}) \prod_{(s,t)\in c} P_{\text{on}}(y_{st}) \\ = k(y) \prod_{s\in c} \frac{P_{\text{on}}(y_s)}{P_{\text{off}}(y_s)} \prod_{(s,t)\in c} \frac{P_{\text{on}}(y_{st})}{P_{\text{on}}(y_{st})} \\ \propto \prod_{i=0}^{m} \frac{\mathcal{M}(y_{s_i}; w_j, \mu_j, \sigma_j)}{\mathcal{W}(y_{s_i}; \min, C, \alpha)} \prod_{i=1}^{m} \frac{k_0 \mathcal{W}(y_{s_{i-1},s_i}; \min_0, C_0, \alpha_0)}{\mathcal{U}(y_{s_{i-1},s_i}; 0, \frac{\pi}{2})}\end{array}$$

where

$$k(y) \;=\; \prod_{s\in S} P_{\text{off}}(y_s) \prod_{(s,t)\in U} P_{\text{off}}(y_{st}).$$

We can then define the likelihood of a deformation θ by

$$P_{Y/\Theta}(y/\theta) \;=\; P_{Y/c_\theta}(y/c_\theta)$$

where c_θ is obtained by interpolation of the polygonal curve with vertices given by the points of γ_θ.

4.4 Posterior distribution

We deduce from above the posterior distribution of a deformation conditional to the observable data

$$P_{\Theta/Y}(\theta/y) \;\propto\; P_{Y/\Theta}(y/\theta) P_\Theta(\theta).$$

This model is entirely based on statistics.

5 Method for localization of shapes

We view similarly to [1] the localization of a shape γ_0 in an image as finding its deformation θ that maximizes the posterior distribution $P_{\Theta/Y}(\theta/y)$.

Equivalently, writing $c_\theta = (s_0, s_1, ..., s_m)$, we want to minimize the Gibbs field

$$\begin{aligned} U(\theta, y) \;=\; & \sum_{i=0}^{m} \big(\log(\mathcal{W}(y_{s_i}; \min, C, \alpha) \\ & - \log(\mathcal{M}(y_{s_i}; w_j, \mu_j, \sigma_j)\big) \\ + \; & \sum_{i=1}^{m} \big(\log(\mathcal{U}(y_{s_{i-1},s_i}; 0, \frac{\pi}{2}) \\ & - \log(k_0\, \mathcal{W}(y_{s_{i-1},s_i}; \min_0, C_0, \alpha_0)\big) \\ + \; & \frac{1}{2} \xi^t (I_q - \sigma^2 \Lambda_q^{-1}) \xi \end{aligned}$$

as a function of θ.

We proceed in two steps in order to find the optimal deformation. As a first step, we use the unsupervised method of [10] based on the distribution of the gradient of the gray levels to find the set C of contour points in the image. We then consider the simpler Gibbs field, similar to [2],

$$U'(\tau_x, \tau_y, s, \psi, y) \;=\; \sum_{i=0}^{m} -\max\Big(\frac{\exp(-\sigma d(s_i))}{d(s_i)}, 1\Big)$$

as a function of $(\tau_x, \tau_y, s, \psi)$ (i.e., $\xi = (0, ..., 0)$), where $d(s)$ is the distance from the point s to the set C. We use the stochastic optimization algorithm of [11] to find a global solution to this simple Gibbs field, with $\sigma = 10^{-4}$. If the image has dimension $M \times N$, we take

$$\begin{array}{rcl} 0 \le & \tau_x & \le M, \\ 0 \le & \tau_y & \le N, \\ 0.2\, d \le & s & \le 0.6\, d, \\ 0 \le & \psi & \le 2\pi \end{array}$$

for the range of U', where d is the diameter of the image. This choice of range involves an implicit *a priori* on the relative size of the shape in the image. A larger range will result in a more difficult optimization problem. It simply means that more iterations might be needed for convergence to a satisfactory initialization.

As a second step, we use the same optimization algorithm but on the complete Gibbs field restricted to a domain given by the best solution found in the first step. Namely, if $\tau_x(0), \tau_y(0), s(0), \psi(0)$ is the solution obtained from the first step, we take for the range of U

$$\begin{array}{rcl} \tau_x(0) - \frac{M}{20} \le & \tau_x & \le \tau_x(0) + \frac{M}{20}, \\ \tau_y(0) - \frac{N}{20} \le & \tau_y & \le \tau_y(0) + \frac{N}{20}, \\ 0.9\, s(0) \le & s & \le 1.1\, s(0), \\ \psi(0) - \frac{\pi}{16} \le & \psi & \le \psi(0) + \frac{\pi}{16}, \\ -3 \le & \xi_i & \le 3. \end{array}$$

We also add a term equal to 100 times the number of points of the shape lying outside the image to the functions U and U', in order to force the shape to be within the image range.

6 Experimental Results

The results obtained with a simple example are presented in this paper.

The shape is a 20th century classical guitar. The data base consists of 10 pictures. Each curve is represented by a template of 70 equally spaced points. The training phase yields a reduced dimension of 5 for the non-linear deformations with less than 5% error. Three different deformations are presented in Figure 2.

The parameters estimated by the ICE procedure for one of the images tested are presented in Figure 3, together with the simple Gibbs field as well as the map of $\log(p_{\text{off}}/p_{\text{on}})$.

We present in Figure 4 the values of the two Gibbs fields for three different localizations. It seems that the global minimum of either function is the desired deformation.

Finally, we present five examples of the entire procedure in Figure 5. The estimation step takes about 100s, the first localization step about 23s for 500 iterations, and the final localization step 41s for 500 iterations on a PC workstation 400MHz. We have tested the procedure on each image with 40 different seeds. The percentage of seeds yielding to a good result within a given number of iterations for the first step and a fixed number of 500 iterations for the second step is indicated in Figure 1. As one can see, if the image is relatively too complex, the number of iterations necessary to obtain convergence might be very large and is unpredictable. Therefore, the method cannot be used directly for detection of shapes. Future work will consist in improving the first step of the localization method.

7 Conclusion

In this paper, we presented a coherent Markov model for deformations of shapes based on the statistical distribution of the gradient vector field of the gray level and a probabilistic model for reduction of dimension. This allowed us to infer the posterior distribution of deformations given the data. The localization of a shape in an image is then viewed as the minimizing of the corresponding Gibbs field. We use a stochastic algorithm in order to find the optimal deformation. In a first step, we use a simpler Gibbs field, and then we use the complete Gibbs field on a restricted domain determined by the solution found in the first step. This yields an unsupervised method for localization of shapes, provided that the image is not prohibitively too complex.

Image	2500 iterations	5000 iterations	10000 iterations
1	90%	97.5%	100%
2	45%	52.5%	82.5%
3	7.5%	7.5%	7.5%
4	90%	100%	100%
5	22.5%	27.5%	32.5%

FIG. 1. *Proportion of the seeds for which convergence of the optimization algorithm occurred within the indicated number of iterations for each of the images of Fig. 5*

Références

[1] A. K. Jain, Y. Zhong, and S. Lakshmanan. Object matching using deformable templates. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pages 267–278, March 1996.

[2] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy. Hybrid genetic optimization and statistical model-based approach for the classification of shadow shapes in sonar imagery. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(2) :129–141, February 2000.

[3] A. K. Jain and D. Zongker. Representation and recognition of handwritten digits using deformable templates. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pages 1386–1391, December 1997.

[4] M.-P. Dubuisson Jolly, S. Lakshmanan, and A. K. Jain. Vehicle segmentation and classification using deformable templates. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pages 293–308, March 1996.

[5] Y. Zhong and A. K. Jain. Object localization using color, texture and shape. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 33 :671–684, 2000.

[6] L. Liu and S. Sclaroff. Deformable shape detection and description via model-based region grouping. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, June 1999.

[7] C. Kervrann and F. Heitz. A hierarchical markov modeling approach for the segmentation and tracking of deformable shapes. *Graphical Models and Image Processing*, 60(3) :173–195, 1998.

[8] M. E. Tipping and C. M. Bishop. Mixtures of probabilistic principal component analyzers. *Technical report NCRG/97/003*, July 1998.

[9] T. F. Cootes and C. J. Taylor. A mixture model for representing shape variation. *Image and Vision Computing*, 17 :567–573, 1999.

[10] F. Destrempes and M. Mignotte. Unsupervised detection and semi-automatic extraction of contours using a statistical model and dynamic programming. In *ICSIP' 2002*, accepted.

[11] O. François. An evolutionary strategy for global minimzation and its markov chain analysis. *IEEE Trans. on Evolutionary Computation*, 2 :77–90, 1998.

[12] D. J. Bartholomew. *Latent Variable Models and Factor Analysis*. Charles Griffin & Co. Ltd., London, 1987.

[13] N. S. Matloff. *Probability Modeling and Computer Simulation*. PWS-KENT Publishing Company, 1988.

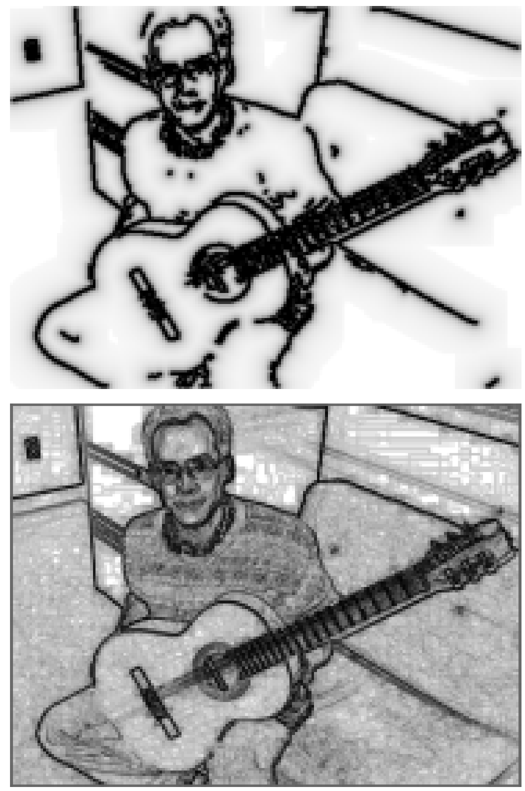
[14] F. Salzenstein and W. Pieczynski. Parameter estimation in hidden fuzzy markov random fields and image segmentation. *Graphical Models and Image Processing*, 59(4) :205–220, 1997.

[15] M. Pérez, A. Blake, and M. Gangnet. Jetstream : Probabilistic contour extraction with particles. In *Int. Conf. on Computer Vision, ICCV' 2001*, Vancouver, Canada, July 2001.

[16] D. Geman and B. Jedynak. An active testing model for tracking roads in satellite images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(1) :1–14, 1996.

Image	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5
top	0	0	0	0	0
center	-1.5	-1.5	-1.5	-1.5	-1.5
bottom	1.5	1.5	1.5	1.5	1.5

FIG. 2. *Example of non-linear deformations of a shape. Top : results obtained for three different sets of values, fixing the rigid transformation parameters. Bottom : Values of the non-linear parameters.*



min	C	α
-0.001	0.699618	3.59123
w_1	μ_1	σ_1^2
0.157434	52.8338	372.666
w_2	μ_2	σ_2^2
0.842566	29.7858	159.261
$\min_0$	C_0	α_0
-0.001	1	0.454012
k_0	ν	
1.03473	100.206	

FIG. 3. *From Top to Bottom : Simple Gibbs field obtained from the contours. Representation of the function* $\log(p_{off}(y_s)/p_{on}(y_s))$. *Values of the parameters estimated by the ICE procedure.*

Image	Simple Gibbs field	Complete Gibbs field
Top	-602.611	-1154.9
Center	-391.807	-289.084
Bottom	-297.302	1508.64

FIG. 4. *Values of the simple Gibbs field and the complete Gibbs field for three different localizations of the shape.*

FIG. 5. *Examples of localization of a shape in an image using stochastic optimization on the simpler Gibbs field for the initialization and then on the complete Gibbs field based on the parameters estimated by the ICE procedure.*