[22] R. Kimmel, A. Amir, and A. M. Bruckstein, "Finding shortest paths on surfaces using level sets propagation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 635–640, 1995.

[23] M. P. Do Carmo, *Differential Geometry of Curves and Surfaces*. Englewood Cliffs, NJ: Prentice-Hall, 1976.

[24] C. L. Epstein and M. Gage, *The Curve Shortening Flow, Wave Motion: Theory, Modeling, and Computation*, A. Chorin and A. Majda, Eds. New York: Springer, 1987.

[25] D. Chopp, "Computing minimal surfaces via level set curvature flow," *J. Comput. Phys.*, vol. 106, pp. 77–91, 1993.

[26] D. Adalsteinsson and J. A. Sethian, "A fast level set method for propagating interfaces," *J. Comput. Phys.*, vol. 118, pp. 269–277, 1995.

[27] G. Borgefors, "Distance transformations in arbitrary dimensions," *Comput. Vis., Graph., Image Process.*, vol. 27, pp. 321–345, 1984.

[28] J. Neider, T. Davis, and M. Woo, *OpenGL Programming Guide*. Reading, MA: Addison-Wesley, 1993.

[29] L. D. Cohen, "On active contour models and balloons," *Comput., Vis., Graph., Image Process.: Image Understand.*, vol. 53, pp. 211–218, 1991.

Estimation of Generalized Mixture in the Case of Correlated Sensors

Wojciech Pieczynski, Julien Bouvrais, and Christophe Michel

***Abstract*—This paper deals with unsupervised Bayesian classification of multidimensional data. We propose an extension of a recent method of generalized mixture estimation to correlated sensors case. The method proposed is valid in the independent data case, as well as in the hidden Markov chain or field model case, with known applications in signal processing, particularly speech or image processing. The efficiency of the method proposed is shown via some simulations concerning hidden Markov fields, with application to unsupervised image segmentation.**

***Index Terms*—Bayesian classification, image segmentation, Markov fields, mixture estimation, multisensor data.**

I. Introduction

The aim of this paper is to deal with the following problem. We are faced with m series of real data produced by m sensors. For each sensor $1 \leq j \leq m$ the data are denoted by $y_1^j, \cdots, y_n^j$. We assume that for each point $1 \leq s \leq n$ the data $y_s^1, \cdots, y_s^m$ correspond to a certain class, among k classes $\omega_1, \cdots, \omega_k$, and the problem is to find which class it is. In other words, the problem is to classify each point $1 \leq s \leq n$ from the data available. The probabilistic approach, which will be our approach in this paper, consists in assuming that the class of the point $1 \leq s \leq n$ is a realization of a random variable X_s, and the data $y_s^1, \cdots, y_s^m$ produced by the m sensors are a realization of a random vector $Y_s = (Y_s^1, \cdots, Y_s^m)$. Thus the problem is to estimate the unobserved realizations of a random process $X = (X_1, \cdots, X_n)$ from the observed realization of a random process $Y = (Y_1, \cdots, Y_n)$. Different methods of such a statistical classification exist once the distribution $P_{(X,Y)}$ of (X, Y) is known. When $P_{(X,Y)}$ is not known, one has to identify it from $Y = y$, the only data available. The aim of our paper is to generalize to correlated sensors the method proposed in [8].

Manuscript received June 23, 1998; revised February 16, 1999. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Robert J. Schalkoff.

The authors are with the Département Signal et Image, Institut National des Télécommunications, 91000 Evry, France (e-mail: wojciech.pieczynski@int-evry.fr).

Publisher Item Identifier S 1057-7149(00)01257-4.

Let us first consider the case of one sensor. When $P_{(X,Y)}$ depends on an unknown parameter θ, the problem is to estimate θ from Y. This problem, which is known as the mixture estimation problem, is a very general and important one [17]. The pioneering method of mixture estimation is the expectation-maximization (EM) algorithm [6], [16], which admits theoretical justifications and gives very good results in classical cases, such as independent Gaussian mixtures. Other methods like stochastic gradient [18] or iterative conditional estimation (ICE, [12]) can also be used. In particular, their use in the Markov random field model (MFR) context leads to the unsupervised image segmentation [2], [9], [18], among others. In fact, once the parameters are estimated, the segmentation can be performed with simulated annealing [7], maximum posterior mode (MPM [11]), or iterated conditional mode (ICM [1]).

All these methods are easily generalizable to the multi-sensor case when the noise is Gaussian. When the noise is not Gaussian and the sensors are independent, one may use the ICE-general mixture (ICE-GEMI) algorithm, valid in the following context [8]. We have k classes, and so we have to find the k probability densities $f_1, \cdots, f_k$ on R^m. Because of the independence, each of these densities f_i is written

$$f_i(y_1^1, y_1^2, \cdots, y_1^m) = f_i^1(y_1^1) f_i^2(y_1^2) \cdots f_i^m(y_1^m). \tag{1.1}$$

ICE-GEMI allows one to find the form of the km functions f_i^j, and estimate their parameters, once we know that each f_i^j belongs to a given set of forms. For instance, in the case of three classes and two sensors, in which each component can be exponential or Gaussian, there are sixty-four possibilities and ICE-GEMI makes possible to search what case the data lie in. In this paper, we propose the following generalization of ICE-GEMI: each f_i of the densities $f_1, \cdots, f_k$ is searched in the set of possible densities of the distribution of a random vector $Y_i = A_i Z_i$, where A_i is $m \times m$ matrix and Z_i a random vector with independent components. Roughly speaking, we add matrices $A_1, \cdots, A_k$ making one possible to deal with correlated sensors. The distribution of Z_i is thus given by

$$g_i(z_1^1, z_1^2, \cdots, z_1^m) = g_i^1(z_1^1) g_i^2(z_1^2) \cdots g_i^m(z_1^m) \tag{1.2}$$

where each g_i^j belongs to a given set of forms. The density of the distribution of Y_i is then written

$$\begin{aligned} f_i(y_1^1, y_1^2, \cdots, y_1^m) = (\det A_i) g_i^1(a_{11}^i y_1^1 + \cdots + a_{1m}^i y_1^m) \\ \cdots g_i^m(a_{m1}^i y_1^1 + \cdots + a_{mm}^i y_1^m). \end{aligned} \tag{1.3}$$

We can see how (1.3) generalizes (1.1). Finally, the method we propose allows one to find the form of each g_i^j, estimate its parameters, and estimate the k matrices $A_1, \cdots, A_k$. It is valid in the independent data case, as well as in the hidden Markov chain or field model case. The efficiency of the method proposed is shown via some simulations concerning hidden Markov fields, with application to unsupervised image segmentation.

The organization of the paper is as follows. In the next section we specify the assumptions needed and describe the run of the proposed method. Section III is devoted to simulation results. The final section contains some concluding remarks and perspectives.

II. Generalized Correlated Sensors Mixture Estimation

Let $X = (X_s)_{s \in S}$, $Y = (Y_s)_{s \in S}$, be two random processes, where each X_s takes its values in the set of classes $\Omega = \{\omega_1, \cdots, \omega_k\}$ and each Y_s takes its values in the set of observations R^m. The distribution of (X, Y), denoted by $P_{(X,Y)}$, is defined by P_X, the distribution of X, and the set of distributions of Y conditional on X. We assume that P_X, which is a probability on the finite space Ω^n, with $n = Card(S)$, depends on a parameter α. The set of random variables X can be a Markov field, a Markov chain, a set of independent variables, or still have any other structure, once the assumptions below are verified. The random variables $(Y_s)_{s\in S}$ are assumed to be independent conditionally to X, and the distribution of each Y_s conditional on X is equal to its distribution conditional on X_s. The distributions of Y_s conditional on $X_s = x_s \in \{\omega_1, \cdots, \omega_k\}$ are given by the densities $f_1, \cdots, f_k$ with respect to Lebesgue measure, respectively. Furthermore, we assume that there exist k triangular matrices $A_1, \cdots, A_k$, with

$$A_i = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{21}^i & 1 & 0 & \cdots & 0 \\ a_{31}^i & a_{32}^i & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & 1 & 0 \\ a_{m1}^i & a_{m2}^i & \cdots & a_{mm-1}^i & 1 \end{bmatrix} \quad (2.1)$$

such that for $1 \le i \le k$, the components of $Z_s = A_i Y_s$ are independent conditionally to $X_s = \omega_i$. Denoting by $g_i^1, g_i^2, \cdots g_i^m$ the densities of the distributions of these components we assume that each of them is of a form being in the finite set of forms $\Psi = \{F_1, \cdots, F_M\}$, each form F_j being a parametrized family of densities on R. Thus the problem is to find the km densities (g_i^l), the k matrices $A_1, \cdots, A_k$, and the parameter α. We assume the following:

(A_1) An estimator $\hat{\alpha} = \hat{\alpha}(X)$ of α from X is available; (A_2) One may simulate realizations of X according to its distribution conditional to Y; (A_3) Each family F_j of $\Psi = \{F_1, \cdots, F_M\}$ is characterized by a parameter β^j, i.e., $F_j = \{g_{\beta^j}\}_{\beta^j \in B^j}$. In practice B^j is a subset of R^{n_j} with n_j depending on F_j; for instance $n_j = 2$ if F_j are Gaussian; (A_4) M estimators $\hat{\beta}^1, \cdots, \hat{\beta}^M$ are available such that if a sample $z = (z_1, \cdots, z_r)$ is generated by a distribution g_{β^j} in F_j, then $\hat{\beta}^j = \hat{\beta}^j(z)$ estimates β^j; (A_5) A decision rule D is available, such that for any sample $z = (z_1, \cdots, z_r)$ and any $(g_1, \cdots, g_M) \in F_1 \times \cdots \times F_M$, the rule D associates to z the "best suited" density among $g_1, \cdots, g_M$, according to some criterion.

Roughly speaking, the method we propose resembles ICE-GEMI, except that we use, at each iteration, some estimates of the matrices $A_1, \cdots, A_k$ in order to "decorrelate" the sensors. Thus the method proposed here, which we will call ICE-COR (COR for "correlate"), is an iterative method and runs as follows. At each iteration:

1) Simulate x^q, a realization of X, according to its α^q and $f_1^q, \cdots, f_k^q$ based distribution conditional to $Y = y$.
2) Calculate $\alpha^{q+1} = E_q[\hat{\alpha}(X)|Y = y]$, where $E_q[\cdot|Y = y]$ denotes the conditional expectation given $\alpha = \alpha^q$ and $(f_1, \cdots, f_k) = (f_1^q, \cdots, f_k^q)$. If this calculation is not possible, calculate $\alpha^{q+1} = \hat{\alpha}(x^q)$.
3) For $i = 1, \cdots, k$, consider $S_i^q = \{s \in S / x_s^q = \omega_i\}$. Let $y_i^q = (y_s)_{s\in S_i^q} = (y_s^1, \cdots, y_s^m)_{s\in S_i^q}$ and $y_i^{q,r} = (y_s^r)_{s\in S_i^q}$. For each $i = 1, \cdots, k$ calculate, from $y_i^q = (y_s)_{s\in S_i^q}$, the empirical covariance matrix $\hat{\Gamma}_i^q$ and consider A_i^q, of the form (2.1) and such that $A_i^q \hat{\Gamma}_i^q (A_i^q)^T$ is diagonal. For each $s \in S_i^q$, put $z_s = A_i^q y_s$ and consider $z_i^q = (z_s)_{s\in S_i^q}$.
4) For each $r = 1, \ldots, m$ and each class $i = 1, \ldots, k$, calculate M parameters $\beta_i^{1,r} = \hat{\beta}^1(z_i^r), \ldots, \beta_i^{M,r} = \hat{\beta}^M(z_i^r)$, which give the densities $g_i^{1,r}, \ldots, g_i^{M,r}$. Put, for each $r = 1, \ldots, m g_i^{r,q+1} = D(g_i^{1,r}, \ldots, g_i^{M,r})$ which give $g_i^{1,q+1}, \ldots, g_i^{m,q+1}$.
5) Put

$$f_i^{q+1}(y_s^1, \cdots, y_s^m) = g_i^{1,q+1}(z_s^1) \cdots g_i^{m,q+1}(z_s^m), \quad \text{with}$$
$$z_s = \begin{bmatrix} z_s^1 \\ \cdots \\ z_s^m \end{bmatrix} = A_i^q \begin{bmatrix} y_s^1 \\ \cdots \\ y_s^m \end{bmatrix}.$$

Concerning the points 3) and 5), let us make the following remarks. The solution of 3), which consists in finding A_i^q of the form (2.1) and such that $A_i^q \hat{\Gamma}_i^q (A_i^q)^T$ is diagonal. Given Γ, A such that $A\Gamma A^T$ is diagonal can be obtained from the LDU decomposition of Γ. So, this procedure is applied to each $1 \le i \le k$ by the use of $\hat{\Gamma}_i^q$. We will assume the needed hypothesis according to which all $\hat{\Gamma}_i^q$ are positive definite, which means, in the context considered, that there is no deterministic link among the sensors considered. The equality in 5) is a particular case of the following property. If Y and Z are two random vectors taking their values in R^m, if f_Y and f_Z are densities of their distributions, and if $Z = AY$ with A a matrix, then $f_Y(y) = \det(A) f_Z(Ay)$. Given that $\det(A_i) = 1$ for any i, we get 5).

Finally, the generalization of ICE-COR with respect to ICE-GEMI is following. In real situations we have, for k classes and m sensors, k unknown densities $f_1, \cdots, f_k$ on R^m. In the previous modeling, each of them is of the form $f_i(y_1^1, y_1^2, \cdots, y_1^m) = f_i^1(y_1^1) f_i^2(y_1^2) \cdots f_i^m(y_1^m)$, where each f_i^j belongs to one among the families $F_1, \cdots, F_M$. In the present modeling we have

$$f_i(y_s^1, \ldots, y_s^m) = g_i^1(y_s^1) g_i^2(a_{21}^i y_s^1 + y_s^2) \ldots g_i^m \cdot (a_{m1}^i y_s^1 + \ldots + a_{mm-1}^i y_s^{m-1} + y_s^m) \quad (2.2)$$

where each g_i^j belongs to one among the families $F_1, \cdots, F_M$. Thus, in order to better approximate the real unknown densities $f_1, \cdots, f_k$, we introduce $km(m-1)/2$ supplementary parameters a_{jl}^i, and, when they are all equal to zero, we find again the previous model. Furthermore, when they are all equal to zero ICE-COR becomes ICE-GEMI.

III. Simulation Results

Let $X = (X_s)_{s\in S}$ be a Markov field and let us consider the case of two classes $(k = 2)$ and two sensors $(m = 2)$. Thus each X_s takes its values in $\Omega = \{\omega_1, \omega_2\}$ and an observation is a realization of a random field $Y = (Y_s)_{s\in S}$, with

$$Y_s = \begin{bmatrix} Y_s^1 \\ Y_s^2 \end{bmatrix}.$$

According to the general conditions of the previous section, we consider two matrices

$$A_1 = \begin{bmatrix} 1 & 0 \\ a_1 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 0 \\ a_2 & 1 \end{bmatrix}$$

such that the components Z_s^1, Z_s^2 of the vector

$$Z_s = \begin{bmatrix} Z_s^1 \\ Z_s^2 \end{bmatrix} = A_1 \begin{bmatrix} Y_s^1 \\ Y_s^2 \end{bmatrix}$$

are independent conditionally to $X_s = \omega_1$, and the components U_s^1, U_s^2 of the vector

$$U_s = \begin{bmatrix} U_s^1 \\ U_s^2 \end{bmatrix} = A_2 \begin{bmatrix} Y_s^1 \\ Y_s^2 \end{bmatrix}$$

TABLE I

GEEG: g_1^1 AND g_2^2 GAUSSIAN, AND EXPONENTIAL. ρ_1, ρ_2: CORRELATION'S IN Γ_1, Γ_2. m_1^1, m_1^2, m_2^1, AND m_2^2: MEANS OF g_1^1, g_1^2, AND g_2^2, RESPECTIVELY. VARIANCE OF THE FOUR DISTRIBUTION EQUAL TO 1. τ_1 AND τ_2: ERROR RATES OF BAYESIAN CLASSIFICATION WITHOUT (τ_1) AND WITH (τ_2) TAKING THE CORRELATION INTO ACCOUNT

Case	Laws	ρ_1	ρ_2	m_1^1	m_2^1	m_1^2	m_2^2	τ_1	τ_2
1	GEEG	0.0	0.0	-1.0	1.0	1.0	-1.0	1.12%	1.18%
2	GEEG	0.8	0.0	-1.0	1.0	1.0	-1.0	2.5%	1.04%
3	GEEG	0.4	0.4	-1.0	1.0	1.0	-1.0	1.73%	1.12%
4	GEEG	0.8	0.8	-1.0	1.0	1.0	-1.0	3.12%	1.12%
5	GEEG	0.8	0.0	-0.5	0.5	0.7	-0.7	38.00%	25.25%
6	GEGG	0.8	0.8	-0.5	0.5	0.7	-0.7	46.27%	3.03%

are independent conditionally to $X_s = \omega_2$. Let us denote by $g_1(z) = g_1(z^1, z^2) = g_1^1(z^1)g_1^2(z^2)$ the density of the distribution of $Z_s = A_1 Y_s$ conditional to $X_s = \omega_1$, and $g_2(u^1, u^2) = g_2^1(u^1)g_2^2(u^2)$ the densities of the distribution of $U_s = A_2 Y_s$ conditional to $X_s = \omega_2$. The densities of the distribution of Y conditional to the classes ω_1, ω_2, respectively, are then $g_1(A_1 y)$, $g_2(A_2 y)$. Furthermore, we consider $\Psi = \{F_1, F_2\}$, where F_1 are Gaussian laws and F_2 are exponential laws. Thus each of the densities g_1^1, g_1^2, g_2^1, g_2^2 can be exponential or Gaussian. Likely to Gaussian densities which are easily determined from the mean and the variance, an exponential density is of the form $f(z) = be^{-b(z-a)}1_{[a,+\infty[}(z)$, and thus depends on two parameters, which can easily be determined from the mean and the variance.

Concerning the structure of the Markov field X, we adopt the simple Potts model, which is Markovian with respect to four nearest neighbors and whose distribution depends on just one real parameter α. Although there exist numerous estimators of α from X, we will fix α in the simulations below and focus on the noise densities recognition.

Finally, n being the number of pixels, we have available a sample

$$(y_1, y_2, \cdots, y_n) = \left(\begin{pmatrix} y_1^1 \\ y_1^2 \end{pmatrix}, \begin{pmatrix} y_2^1 \\ y_2^2 \end{pmatrix}, \cdots \begin{pmatrix} y_n^1 \\ y_n^2 \end{pmatrix} \right)$$

and we must identify the forms of g_1^1, g_1^2, g_2^1, g_2^2 and estimate their parameters and estimate a_1, a_2. The algorithm is as follows:

1) Initialization:

Assume the sensors independent (matrices A_1 and A_2 are identity) and all densities Gaussian. Calculate, from y_1^1, $y_2^1, \cdots, y_n^1$, the empirical mean and variance M_0^1, Σ_0^1 of first sensor, and calculate, from y_1^2, $y_2^2, \cdots, y_n^2$, the empirical mean and variance M_0^2, Σ_0^2 of the second one. Put $m_1^1 = M_0^1 - (\Sigma_0^1/2)$ and $m_2^1 = M_0^1 + (\Sigma_0^1/2)$ for means of f_1^1, f_2^1, and $(\sigma_1^1)^2 = (\sigma_2^1)^2 = \Sigma_0^1$ for their variances. Proceed in the same way to calculate m_1^2, m_2^2, $(\sigma_1^2)^2$, $(\sigma_2^2)^2$, the means and variances of f_1^2, f_2^2, from M_0^2, Σ_0^2. Of course, numerous other initializations could be used but, on the contrary to EM technics, the initialization is not essential because of the stochastic aspect of iterations.

2) At each iteration

i) Using the Gibbs sampler, simulate a realization $X = x^q$ in $\Omega^n = \{\omega_1, \omega_2\}^n$ according to the posterior distribution.

ii) Calculate, from $S_1^q = (s \in S/x_s^q = \omega_1\}$ and $S_2^q = \{s \in S/x_s^q = \omega_2\}$, the empirical covariance matrices $\hat{\Gamma}_1^q, \hat{\Gamma}_2^q$. Calculate A_1^q, A_2^q $\left(\text{take } A = \begin{bmatrix} 1 & 0 \\ -\dfrac{\rho}{\sigma_1^2} & 1 \end{bmatrix} \text{ for } \Gamma = \begin{bmatrix} \sigma_1^2 & \rho \\ \rho & \sigma_2^2 \end{bmatrix}\right)$.

iii) Consider

$$z_1 = \begin{pmatrix} z_s^2 \\ z_s^1 \end{pmatrix}_{s \in S_1^q} = A_1^q \begin{pmatrix} y_s^2 \\ y_s^1 \end{pmatrix}_{s \in S_1^q}$$

$$z_2 = \begin{pmatrix} z_s^2 \\ z_s^1 \end{pmatrix}_{s \in S_2^q} = A_2^q \begin{pmatrix} y_s^2 \\ y_s^1 \end{pmatrix}_{s \in S_2^q}$$

and use the samples $(z_s^1)_{s \in S_1^q}$, $(z_s^2)_{s \in S_1^q}$, $(z_s^1)_{s \in S_2^q}$, and $(z_s^2)_{s \in S_2^q}$ to identify the forms of g_1^1, g_1^2, g_2^1, g_2^2 and estimate their parameters. The latter is done as follows: to identify the form and the parameters of g_1^1, estimate the mean and the variance from $(z_s^1)_{s \in S_1^q}$, which gives a Gaussian density h^1 on the one hand, and an exponential density h^2 on the other hand. Calculate the histogram $\hat{h}$ from $(z_s^1)_{s \in S_1^q}$ and consider $d_i = \int_R [h^i(z) - \hat{h}(z)]^2 \, dz$ for $i = 1, 2$. Put $g_1^1 = h^1$ if $d_1 \leq d_2$ and $g_1^1 = h^2$ if $d_1 \geq d_2$. Proceed in the same way for g_1^2, g_2^1, and g_2^2. Of course, numerous other rules D could be used.

iv) Determine the densities f_1, f_2 with $f_1(y) = g_1(A_1 y)$, $f_2(y) = g_2(A_2 y)$. Calculate the posterior distribution.

The results of some simulations are as follows. In order to clarify the presentation in Table I, let us specify, as an example, the case 2. Laws GEEG means that g_1^1 is Gaussian, g_2^1 and g_1^2 are Exponential, and g_2^2 is Gaussian. In all simulations the variance of the four densities g_1^1, g_2^1, g_1^2, and g_2^2 is 1, thus they are defined by their means m_1^1, m_2^1, m_1^2, and m_2^2. Recall that f_1 is defined by g_1^1, g_1^2, and $A = \begin{bmatrix} 1 & 0 \\ a_1 & 1 \end{bmatrix}$. On the other hand,

$$A = \begin{bmatrix} 1 & 0 \\ -\dfrac{\rho}{\sigma_1^2} & 1 \end{bmatrix} \quad \text{for} \quad \Gamma = \begin{bmatrix} \sigma_1^2 & \rho \\ \rho & \sigma_2^2 \end{bmatrix}$$

$$\text{As } \Gamma_1 = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix} \quad \text{and} \quad \Gamma_2 = \begin{bmatrix} 1 & \rho_2 \\ \rho_2 & 1 \end{bmatrix}$$

we have $a_1 = -\rho_1$ and $a_2 = -\rho_2$. Finally, f_1 is defined by the type of g_1^1, g_1^2, their means m_1^1, m_1^2, and ρ_1. Recalling that $f_1(y) = g_1(A_1 y)$, the exact form of f_1 is (the index s is omitted)

$$\begin{aligned} f_1(y) &= g_1^1(y^1) g_1^2(-\rho_1 y^1 - y^2) \\ &= \frac{1}{\sqrt{2\pi}} e^{-(y^1+1)^2/2} e^{-(-0.8y^1+y^2-1)} 1_{[0,+\infty[} \\ &\quad \cdot (-0.8y^1 + y^2). \end{aligned} \tag{3.1}$$

In the same way, f_2 is defined by the type of g_2^1, g_2^2, their means m_2^1, m_2^2, and ρ_2. According to the data on the line 2 of the Table I, its exact form is

$$\begin{aligned} f_2(y) &= g_2^1(y^1) g_2^2(y^2) \\ &= e^{-(y^1-1)} 1_{[1,+\infty[}(y^1) \frac{1}{\sqrt{2\pi}} e^{-(y^2+1)^2/2} \end{aligned} \tag{3.2}$$

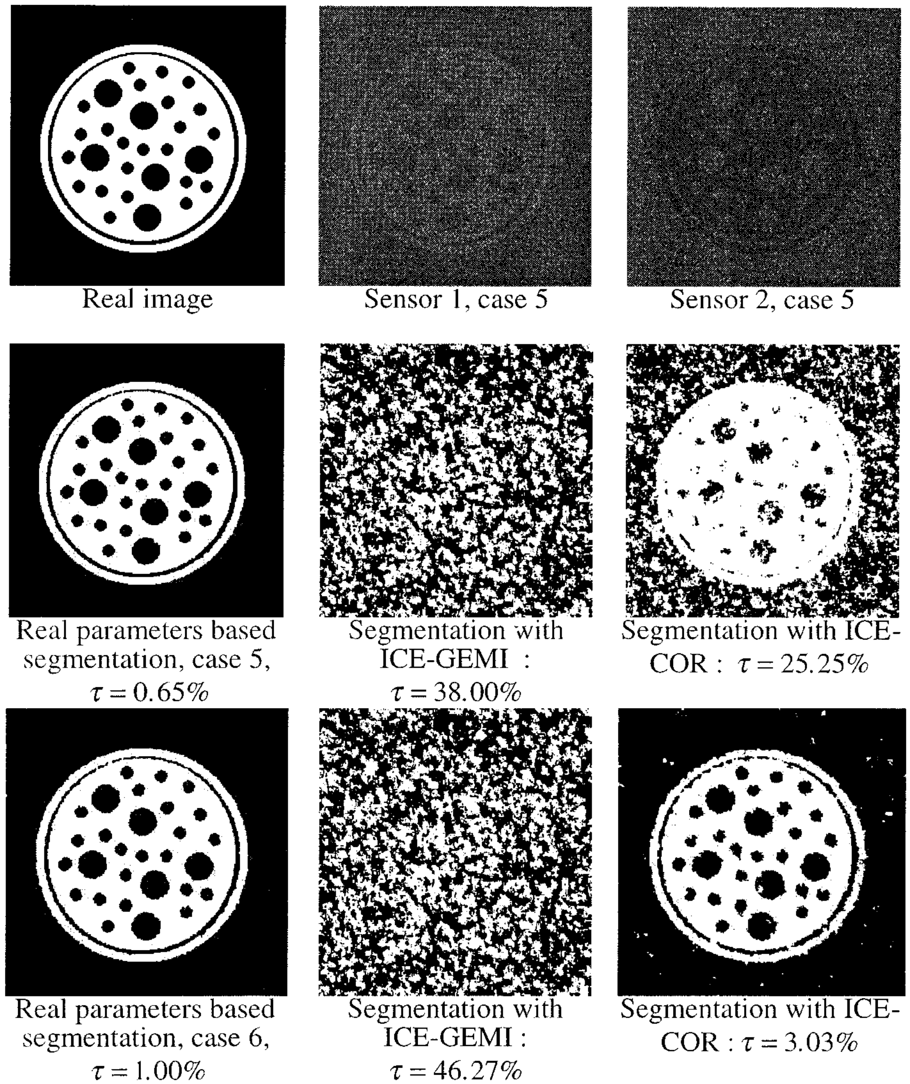

Fig. 1. Hand written image, its noisy versions in the two sensors, and Markov Fields based segmentations: Cases 5 and 6, Table I. τ is real parameter based Bayesian classification error rate.

The results presented in Table I show that taking into account the correlation of the sensors may be advantageous when considering unsupervised Bayesian multisensor data classification. In some situations, as the case 3, its advantage seems limited and in some others, as the case 6, its advantage is determining. Of course, the number of unsupervised classification of multivariate data problems is extremely wide and, for a given problem, we have numerous possibilities of considering an ICE-COR method. So we have to be careful in claming anything about the general quality of ICE-COR; however, this short simulation study shows that ICE-COR may be of interest.

Some visual effects are presented in Fig. 1.

Let us notice that the computer time cannot be indicated precisely because it depends on different subjective parameters, like the rule D or number of different iterations. However, in the hidden Markov fields context the ICE-COR methods have to be seen as rather time consuming.

IV. Conclusions

In this paper we presented a new method for estimating multivariate mixtures, with some applications to unsupervised Bayesian classification of multisensor data. The novelty is that the sensors are possibly non-Gaussian and can be correlated. The method proposed is a generalization of the recently published method ICE-GEMI [8]. In the latter the sensors were assumed independent; however, the densities attached with each class were allowed to be, for each sensor, of different form, and they were allowed to vary, for a given class, with sensors. All we required was that each density be of a form belonging to a given set of forms. Such mixtures are called "generalized mixtures" and thus ICE-GEMI allows one to estimate such mixtures in a quite general context, which includes Markov hidden chains or fields. The method proposed in this paper, which we call ICE-COR (classical ICE is found in the classical mixture case and COR for correlated sensors mixture), generalizes ICE-GEMI in that the sensors can be correlated, and, in the particular case when they are independent, ICE-COR gives ICE-GEMI.

In the same way as ICE-GEMI, the method proposed is valid in a quite general setting; in particular, hidden Markov fields or hidden Markov chains can be treated, with known applications to image or signal restoration. As hidden Markov chains have been studied in [8], we proposed in this paper some simulation results concerning the hidden Markov field model based multisensor image segmentation.

Different simulations allow us to asses the existence of some situations in which the use of ICE-COR is of interest.

Unfortunately, we have no theoretical result to present concerning the asymptotic behavior of ICE-COR. We hope to devote to this important subject some efforts in our future works.

REFERENCES

[1] J. Besag, "On the statistical analysis of dirty pictures," *J. R. Statist. Soc. B*, vol. 48, pp. 259–302, 1986.

[2] J. M. Boucher and P. Lena, "Unsupervised Bayesian classification, application to the forest of Paimpont (Brittany)," *Photo Interp.*, vol. 32, no. 1994/4, 1995/1-2, pp. 79–81, 1995.

[3] H. Caillol, W. Pieczynski, and A. Hillon, "Estimation of fuzzy Gaussian mixture and unsupervised statistical image segmentation," *IEEE Trans. Image Processing*, vol. 6, no. 3, pp. 425–440, 1997.

[4] R. Chellapa and A. Jain, *Markov Random Fields, Theory and Application*, A. Jain, Ed. San Diego, CA: Academic, 1993.

[5] Y. Delignon, A. Marzouki, and W. Pieczynski, "Estimation of generalized mixture and its application in image segmentation," *IEEE Trans. Image Processing*, vol. 6, pp. 1364–1375, Oct. 1997.

[6] M. M. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. B*, vol. 39, pp. 1–38, 1977.

[7] S. Geman and G. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 721–741, June 1984.

[8] N. Giordana and W. Pieczynski, "Estimation of generalized multisensor hidden Markov chains and unsupervised image segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, no. 5, pp. 465–475, 1997.

[9] Z. Kato, J. Zerubia, and M. Berthod, "Unsupervised parallel image classification using Markovian models," *Pattern Recognit.*, vol. 32, pp. 591–604, 1999.

[10] S. Lakshmanan and H. Derin, "Simultaneous parameter estimation and segmentation of Gibbs random fields," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 799–813, 1989.

[11] J. Marroquin, S. Mitter, and T. Poggio, "Probabilistic solution of ill-posed problems in computational vision," *J. Amer. Statist. Assoc.*, vol. 82, pp. 76–89, 1987.

[12] W. Pieczynski, "Statistical image segmentation," *Mach. Graph. Vis.*, vol. 1, no. 1/2, pp. 261–268, 1992.

[13] W. Pieczynski, J. Bouvrais, and C. Michel, "Unsupervised Bayesian fusion of correlated sensors," in *1st Int. Conf. Multisource–Multisensor Information Fusion*, Las Vegas, NV, July 6–9, 1998.

[14] W. Qian and D. M. Titterington, "Stochastic relaxations and EM algorithms for Markov random fields," *J. Statist. Comput. Simulat.*, vol. 40, pp. 55–69, 1992.

[15] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[16] R. A. Redner and H. F. Walker, "Mixture densities, maximum likelihood and the EM algorithm," *SIAM Rev.*, vol. 26, pp. 195–239, 1984.

[17] D. M. Titterington, A. F. M. Smith, and U. E. Makov, *Statistical Analysis of Finite Mixture Distribution*. New York: Wiley, 1985.

[18] L. Younes, "Parametric inference for imperfectly observed Gibbsian fields," *Prob. Theor. Relat. Fields*, vol. 82, pp. 625–645, 1989.

Separating Touching Objects in Remote Sensing Imagery: The Restricted Growing Concept and Implementations

Leen-Kiat Soh and Costas Tsatsoulis

***Abstract*—This paper defines the restricted growing concept (RGC) for object separation and provides an algorithmic analysis of its implementations. Our concept decomposes the problem of object separation into two stages. First, separation is achieved by shrinking the objects to their cores while keeping track of their originals as masks. Then the core is grown within the masks obeying the guidelines of a restricted growing algorithm. In this paper, we apply RGC to the remote sensing domain, particularly the synthetic aperture radar (SAR) sea ice images.**

***Index Terms*—Morphology, object separation, remote sensing imagery, restricted growing.**

I. INTRODUCTION

When two gray level objects touch with shared boundaries, it makes shape analysis and recognition difficult in areas such as industrial vision applications [3], in aerial image and terrain analysis [7] or in shape analysis [5]. The objectives of our work are to achieve object separation, and to preserve (or approximate as closely as possible) the object's original shape and size. The tradeoff between separation and preservation of size and shape is inherent in all object separation algorithms. To address this problem, we have designed a technique based on the restricted growing concept (RGC), that achieves separation and, then, reestablishes the sizes and shapes of the objects lost or distorted during the separation process by performing restricted growing.

In this paper, we present the restricted growing concept and address the issues of preserving details through different designs of masks, investigate the use of morphological reconstruction h-domes in extracting cores, compare the differences between the performance of the morphological operators in synthetic and remotely sensed images, and describe a reverse skeletonization algorithm to guide the growth of object pixels in the image. We finally present twelve algorithms of RGC and examine their weaknesses and strengths when applied to synthetic aperture radar (SAR) sea ice images.

II. RESTRICTED GROWING CONCEPT

The main idea behind the RGC is to decompose the object separation problem into two steps: The first step achieves separation, accomplished by *shrinking* objects such that each object is separated from its touching neighbors. The second step preserves size and shapes by growing the shrunk objects to restore them. To ensure that the separation established after the first stage is not disturbed, our growing process is *restricted*.

In our approach, a *mask object* is a version of the original object such that the original size of the object is preserved. Mask objects are usually interconnected and can encompass one or more core objects. An image with mask objects is a *mask image*. A *core object* is a version of the original object such that its linkages to neighboring objects are disconnected, satisfying object separation. Such an object is reduced in

Manuscript received September 10, 1998; revised June 16, 1999. This work was supported in part by the Naval Research Laboratory under Contract N00014-95-C-6038. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Robert J. Schalkoff.

The authors are with the Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, KS 66044 USA (e-mail: lksoh@ittc.ukans.edu).

Publisher Item Identifier S 1057-7149(00)01258-6.