Unsupervised Dempster-Shafer Fusion of Dependent Sensors

Wojciech Pieczynski

Département Signal et Image
Institut National des Télécommunications,
9, rue Charles Fourier, 91000 Evry, France

E-mail Wojciech.Pieczynski@int-evry.fr

Abstract

This paper deals with the problem of statistical unsupervised fusion of dependent sensors with its potential applications to multisensor image segmentation. On the one hand, Bayesian fusions can be of great efficiency, particularly when using hidden Markov models. On the other hand, we give some examples showing that there are situations in which the Dempster-Shafer fusion can be usefully integrated in the classical Bayesian models. The contribution of this paper is then to show how a recent parameter estimation of probabilistic models, valid in the case dependent and possible non Gaussian sensors case, can be extended to the situations in which some of sensors can be evidential. The proposed method allows one to imagine different unsupervised segmentation methods, valid in the Dempster-Shafer framework for dependent and possibly non Gaussian sensors.

1. Introduction

In this paper we consider the problem of unsupervised Dempster-Shafer fusion of correlated sensors with applications to unsupervised multisensor statistical image segmentation. When the sensors are independent and the noise distribution known, it is possible to use different Bayesian methods like Maximum Posterior Mode (MPM) or Maximum a Posteriori (MAP), once we have adopted some model for the class process like, for instance, Markov field [6, 11]. These methods can be made unsupervised by estimating the model parameters by some general methods like Stochastic Gradient (SG, [16]), or Iterative Conditional Estimation (ICE, [8, 9]). When the nature of the noise is not known exactly, it is still possible to estimate the model using some recent extension of ICE [8]. Finally, considering the multisensor case with sensors neither independent nor Gaussian, one can still estimate the model parameters by a further recent extension ICE [12].

On the other hand, the Dempster-Shafer theory of evidence can provide some interesting extension of the classical probabilistic models [1, 3, 5, 7, 15].

The aim of this paper is to propose a family of parameter estimation methods valid in some Dempster-Shafer extensions of the classical probabilistic models. Extending the method proposed in [12], this family is valid in the case of non necessarily independent or Gaussian sensors case.

2. An example

Let us first consider an example of the problem of satellite image segmentation in the case of two classes (forest and water) and two sensors (optical and radar). Let S be the set of pixels, $X = (X_s)_{s \in S}$ the random field of classes (each X_s takes its values in the set of classes $\Omega = \{\omega_1, \omega_2\}$), and $Y = (Y_s)_{s \in S}$ the random field of observations (each $Y_s = (Y_s^1, Y_s^2)$ takes its values in R^2). A Markovian structure of $X = (X_s)_{s \in S}$ will be considered in the following; however, for the moment we assume the random variables (X_s) independent.

Let us temporarily assume that Y_s^1, Y_s^2 are independent conditionally on X, and that their distribution conditional to X is their distribution conditional to X_s. Furthermore, their distributions conditional to $X_s = \omega_1, \omega_2$ are given by densities f_1^1, f_1^2, f_2^1, and f_2^2, respectively.

When priors $\pi_1 = P[X_s = \omega_1]$, $\pi_2 = P[X_s = \omega_2]$ are known, the posterior probability distribution of X_s is :

$$
\begin{aligned}
P[X_s = \omega_1 \big| (Y_s^1, Y_s^2) = (y_s^1, y_s^2)] &\propto \pi_1 f_1^1(y_s^1) f_1^2(y_s^2) \\
P[X_s = \omega_2 \big| (Y_s^1, Y_s^2) = (y_s^1, y_s^2)] &\propto \pi_2 f_2^1(y_s^1) f_2^2(y_s^2)
\end{aligned}
\qquad (1)
$$

Now, let us assume that there are clouds and, when a pixel is concealed with a cloud, the first sensor, which is an optical one, cannot give any information about the class lying at the ground. It just gives a grey level of clouds, whose probability distribution is given with a density f_c^1 (c for "clouds"). How does one integrate the presence of clouds in the model to obtain posterior distribution analogous to (1)? The Dempster -Shafer theory of evidence allows one to treat this problem as follows. Following one of the possible models [1], which we adopt in this paper, the observed $Y_s^1 = y_s^1$ defines a probability q on $\Omega^* = \{\{\omega_1\},\{\omega_2\},\{\omega_1,\omega_2\}\}$ with:

$$q[\{\omega_1\}] = \frac{f_1^1(y_s^1)}{f_1^1(y_s^1) + f_2^1(y_s^1) + f_c^1(y_s^1)}$$

$$q[\{\omega_2\}] = \frac{f_2^1(y_s^1)}{f_1^1(y_s^1) + f_2^1(y_s^1) + f_c^1(y_s^1)} \qquad (2)$$

$$q[\{\omega_1,\omega_2\}] = \frac{f_c^1(y_s^1)}{f_1^1(y_s^1) + f_2^1(y_s^1) + f_c^1(y_s^1)}$$

The probability q, which is called the "mass function" in the theory of evidence language, may then be fused, using the Dempster-Shafer combination rule (or "fusion"), with the posterior probability

$P[X_s = \omega_1 | Y_s^2 = y_s^2] \propto \pi_1 f_1^2(y_s^2)$,

$P[X_s = \omega_2 | Y_s^2 = y_s^2] \propto \pi_2 f_2^2(y_s^2)$,

provided by the second sensor. We obtain:

$$P[X_s = \omega_1 | (Y_s^1, Y_s^2) = (y_s^1, y_s^2)] \propto \pi_1 [f_1^1(y_s^1) + f_c^1(y_s^1)] f_1^2(y_s^2)$$

$$(3)$$

$$P[X_s = \omega_2 | (Y_s^1, Y_s^2) = (y_s^1, y_s^2)] \propto \pi_2 [f_2^1(y_s^1) + f_c^1(y_s^1)] f_2^2(y_s^2)$$

The aim of this paper is to handle the above problem once the sensor independence hypothesis is relaxed. In the absence of clouds, $f_1^1(y_s^1) f_1^2(y_s^2)$ in (1) is replaced with $f_1(y_s^1, y_s^2)$, and $f_2^1(y_s^1) f_2^2(y_s^2)$ is replaced with $f_2(y_s^1, y_s^2)$. In the presence of clouds, (3) becomes :

$$P[X_s = \omega_1 | (Y_s^1, Y_s^2) = (y_s^1, y_s^2)] \propto \pi_1 [f_1(y_s^1, y_s^2) + f_{c,1}(y_s^1, y_s^2)]$$

$$(4)$$

$$P[X_s = \omega_2 | (Y_s^1, Y_s^2) = (y_s^1, y_s^2)] \propto \pi_2 [f_2(y_s^1, y_s^2) + f_{c,2}(y_s^1, y_s^2)]$$

These posterior probabilities can then be used in different processing, like, for example, Bayesian segmentation. In the unsupervised processing case that interests us in this paper, the main problem is then to find the four functions $f_1, f_2, f_{c,1}, f_{c,2}$, which are probability densities on R^2, and the priors π_1, π_2. In other words, we have to estimate a mixture of four components $f_1, f_2, f_{c,1}, f_{c,2}$ on R^2 and the priors π_1, π_2.

We specify in this paper how the "correlated ICE-GEMI" method proposed in [12] (which is itself a generalization to the dependent sensor case the ICE-GEMI algorithm proposed in [8]) can be employed in a theory of evidencec ontext.

The correlated ICE-GEMI principle is as follows. We assume that $f_1, f_2, f_{c,1}, f_{c,2}$ are probability laws of random vectors U_1, U_2, U_3, U_4, with, for each $i \in \{1,2,3,4\}$,

$$U_i = \begin{pmatrix} 1 & \theta_i \\ 0 & 1 \end{pmatrix} \begin{pmatrix} V_i^1 \\ V_i^2 \end{pmatrix} \qquad (5)$$

and V_i^1, V_i^2 uncorrelated (the parameters θ_1, ..., θ_4 are computed from the correlations of the components of U_1, ..., U_4). The variables V_i^1, V_i^2 are then assumed independent (this is an approximation). Furthermore, the distributions of the random variables (V_i^j) are given with densities (g_i^j), and each density g_i^j belongs to one among families $F_1, \ldots, F_M$. For instance, if we have $F_1 = \{\text{Gaussian Laws}\}$, $F_2 = \{\text{Gamma Laws}\}$, $F_3 = \{\text{Beta Laws}\}$, each of the eight densities g_i^j can be Gaussian, Gamma, or Beta (giving 3^8 possibilities). The correlated ICE-GEMI can then be used to determinate the nature of each density g_i^j, estimate the parameters for each of them (for instance we need the mean and the variance if g_i^j is Gaussian), and estimate the four parameters θ_i. So, given (5), the correlated ICE-GEMI method, whose general walk will be detailed in the next section, allows one to estimate the four densities $f_1, f_2, f_{c,1}, f_{c,2}$.

3. General case

In a general way, let us consider a set of classes $\Omega = \{\omega_1, \ldots, \omega_k\}$, the power set $\Omega^* = \{\Omega_1, \ldots, \Omega_{2^k}\}$ of Ω, and $m+1$ mass functions $M_0, M_1, \ldots, M_m$, which are probabilities on Ω^*. Recall that if a mass function only charges the singletons, it can be assimilated to a classical probability on Ω : such a mass function will be called "Bayesian" or "probabilistic". Roughly speaking, M_0, which will be assumed probabilistic in this paper, will model the prior information and $M_1, \ldots, M_m$ will model the information contained in the observation of m sensors. In the case of independent pieces of information, the Dempster-Shafer combination rule, which enables one to aggregate theses different pieces of information, is as follows:

$$M(A) \propto \sum_{A_0 \cap \ldots \cap A_m = A \neq \emptyset} [\prod_{j=0}^{m} M_j(A_j)] \qquad (6)$$

The combination rule (6) corresponds to independent sensors (fusion (3) in the previous section). When sensors are dependent, (6) becomes (see (4) above):

$$M^{sen}(A) \propto \sum_{A_1 \cap \ldots \cap A_m = A \neq \emptyset} M'(A_0 \times \ldots \times A_m) \qquad (7)$$

where M' is a mass function on the Cartesian product $(\Omega^*)^{m+1}$. This defines a mass function on Ω^* which will be denoted by $M = M_0 \otimes M_1 \otimes \ldots \otimes M_m$. We then have the following well known property:

Proposition 3.1

If at least one mass function among $M_0, M_1, \ldots, M_m$ is probabilistic, $M = M_0 \otimes M_1 \otimes \ldots \otimes M_m$ is probabilistic.

In the case we are interested in, the mass function M_0 is the prior distribution π_1, ..., π_m, and the mass function M' is provided by m sensors Y_s^1, ... , Y_s^m. Each sensor Y_s^j can distinguish Ω_1^j, ..., $\Omega_{r(j)}^j$, which are $r(j)$ subsets of Ω. The distribution of Y_s^j conditional to Ω_l^j will be denoted by f_l^j. In other words, the law distribution of Y_s^j is some mixture of f_1^j, ..., $f_{r(j)}^j$. For each $t = (t_1, \ldots t_m)$, with $1 \le t_1 \le r(1)$, ..., $1 \le t_m \le r(m)$, let $\Omega^t *$ be the set of elements $\Omega^t = (\Omega_{t_1}^1, \ldots, \Omega_{t_m}^m)$ and, for each $\omega_i \in \Omega$, let $\Omega_i^t *$ be the subset of $\Omega^t *$ of elements $\Omega_i^t = (\Omega_{i,t_1}^1, \ldots, \Omega_{i,t_m}^m)$ such that $\omega_i \in \Omega_{i,t_1}^1$, ..., $\omega_i \in \Omega_{i,t_m}^m$. So for each $\Omega_i^t = (\Omega_{i,t_1}^1, \ldots, \Omega_{i,t_m}^m)$ there exists a probability density $f_{\Omega_i^t}$ on R^m. Let us recall that this density is not necessarily a product of marginal densities. For an observed $y_s \in R^m$, we have to fuse the priors with the mass function M^{sen} resulting from the Dempster-Shafer fusion of the m sensors. The result, which is a generalization of (4), is then a probability dinstribution defined on Ω by:

$$\pi_i^{y_s} \propto \sum_{\omega_i \in A} \pi_i M^{sen}(A) \propto \pi_i \sum_{\omega_i \in \Omega_{t_1}^1, \ldots, \omega_i \in \Omega_{t_m}^m} f_{\Omega^t}(y_s) \qquad (8)$$

(the sums above are made over A such that ω_i, which is fixed, is included in A).

So the problem is to estimate the priors $\pi = (\pi_1, \ldots, \pi_k)$ and $K = r(1)r(2)\ldots r(m)$ densities (f_{Ω^t}) on R^m.
Following [12], we assume that there exist K triangular matrices A^{Ω^t},with

$$A^{\Omega^t} = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ a_{21}^{\Omega^t} & 1 & 0 & \ldots & 0 \\ a_{31}^{\Omega^t} & a_{32}^{\Omega^t} & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & 1 & 0 \\ a_{m1}^{\Omega^t} & a_{m2}^{\Omega^t} & \ldots & a_{mm-1}^{\Omega^t} & 1 \end{bmatrix}, \qquad (9)$$

such that for each Ω^t, when (f_{Ω^t}) is the density distribution of Y_s, then the components of $Z_s = A^{\Omega^t} Y_s$ are uncorrelated. Let us denote by $g_{\Omega^t}^1, g_{\Omega^t}^2, \ldots g_{\Omega^t}^m$ the densities of the distributions of the components. We assume that for each i, the form of $g_{\Omega^t}^i$ belongs to the finite set of forms $\Psi = \{F_1, \ldots, F_M\}$, 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_{Ω^t}), and the priors $\pi = (\pi_1, \ldots, \pi_k)$.
We assume that the following general points hold:
(A_1) An estimator of $\pi = (\pi_1, \ldots, \pi_k)$ 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 $\Psi = \{F_1, \ldots, F_M\}$ is characterized by a parameter β^j, i.e., $F_j = \{g_{\beta^j}\}_{\beta^j \in B^j}$. In practice β^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 $\hat{\beta}^1, \ldots, \hat{\beta}^M$ are available such that if a sample $z = (z_1, \ldots, z_r)$ 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 $z = (z_1, \ldots, z_r)$ and any $(g_1, \ldots, g_M) \in F_1 \times \ldots \times F_M$, the rule D associates to z the "best suited" density among $g_1, \ldots, g_M$, according to some criterion.
A_1 is obvious but it will be useful in the next section.
We propose the following iterative generalization of the ICE-COR method:
(i) initialize π^0, $(f_{\Omega^t}^0)$ in some way;
(ii) determine π^{q+1} and $(f_{\Omega^t}^{q+1})$ from π^q, $(f_{\Omega^t}^q)$, and the observed $Y = y$ according to:
1. For each $s \in S$, calculate $\pi^{y_s,q}$ and put $\pi^{q+1} = (\sum_{s \in S} \pi^{y_s,q}) / Card(S)$;

2. Simulate a realization x^q of X according to its current posterior distribution $\pi^{y,q}$ (given by π^q, $(f^q_{\Omega^t})$, and y). For each $\omega_i \in \Omega$, use the subset $y^{i,q} = \{y_s / x^q_s = \omega_i\}$ to estimate, by using the classical ICE-COR specified in [12], the densities $(f_A)_{A \in \Omega^t_i *}$, obtaining $(\hat{f}_A)_{A \in \Omega^t_i *}$. Put $f^{q+1}_A = \hat{f}_A$ for each $A \in \Omega^t_i *$.

We notice that the generalization proposed consists on adding a novel level : at each iteration, a classical ICE-COR iteration is used k times on sub-samples.

Example

Let us consider three classes : grass on dry ground (ω_1), rice in water (ω_2), and water (ω_3), with possible presence of clouds. The priors are π_1, π_2, π_3. Considering an optical sensor and a radar one, we can imagine that the optical sensor distinguishes $\{\omega_1, \omega_2\}$, $\{\omega_3\}$, and $\{\omega_1, \omega_2, \omega_3\}$, because it cannot make any difference between grass on dry ground and rice on water; moreover, when there is a cloud, it does not distinguish anything. On the other hand, radar sensor distinguishes $\{\omega_1\}$ and $\{\omega_2, \omega_3\}$, because it cannot make any difference between rice in water and water.

So, for the first sensor we have $r(1) = 3$, and for the second one $r(2) = 2$. We thus have six possible couples $t = (t_1, t_2)$, which implies that a mixture of six components on R^2 is to be estimated: $f_{(\{\omega_1,\omega_2\},\{\omega_1\})}$, $f_{(\{\omega_1,\omega_2\},\{\omega_2,\omega_3\})}$, $f_{(\{\omega_3\},\{\omega_2,\omega_3\})}$, $f_{(\{\omega_1,\omega_2,\omega_3\},\{\omega_1\})}$, $f_{(\{\omega_1,\omega_2,\omega_3\},\{\omega_1\})}$, and $f_{(\{\omega_1,\omega_2,\omega_3\},\{\omega_2,\omega_3\})}$. Once these components are estimated by ICE-COR as above, the posterior probabilities are:

$$\pi_1^{y_s} \propto \pi_1[f_{(\{\omega_1,\omega_2\},\{\omega_1\})}(y_s) + f_{(\{\omega_1,\omega_2,\omega_3\},\{\omega_1\})}(y_s)]$$
$$\pi_2^{y_s} \propto \pi_2[f_{(\{\omega_1,\omega_2\},\{\omega_2,\omega_3\})}(y_s) + f_{(\{\omega_1,\omega_2,\omega_3\},\{\omega_2,\omega_3\})}(y_s)]$$
$$\pi_3^{y_s} \propto \pi_3[f_{(\{\omega_3\},\{\omega_2,\omega_3\})}(y_s) + f_{(\{\omega_1,\omega_2,\omega_3\},\{\omega_2,\omega_3\})}(y_s)]$$

4. Other distributions for X.

We have assumed above that the random variables (X_s, Y_s) were independent (and thus the random variables (X_s) were independent too), but more complex structures for the distribution of the class random field X can be considered. Roughly speaking, every structure is suitable under the following conditions : (a) the distribution of X depends on a parameter α which can be estimated from X; (b) it is possible to simulate realizations of X according to its distribution conditional to observations.

Let $\hat{\alpha}$ be the estimator of α from X. Then the point (ii) 1. in ICE-COR above is replaced by $\alpha^{q+1} = E_q[\hat{\alpha}(X) | Y = y]$, if this conditional expectation is calculable, and by $\alpha^{q+1} = \hat{\alpha}(x^q)$ if it is not.

Several classical probabilistic models for the X can be used because of the following result:

Proposition 4.1

Let Π be the prior probability distribution of X and M^{sen} the mass function provided by the Dempster-Shafer fusion of possibly dependent sensors. Let us consider the following assumptions : (i) the random variables (Y_s) are independent conditionally to X; (ii) the distribution of (Y_s) conditional to X is equal to its distribution conditional to X_s.

The probability distribution $\Pi \otimes M^{sen}$ on Ω^N is the same as the *a posteriori* probability distribution of the random field X classically noised by the independent noise given by

$$\varphi_{x_s}(y_s) \propto \sum_{x_s \in A} M_s^{sen}(A) \tag{10}$$

Proof

This result is a generalisation of the particular case of Markov fields showed in [2]; the proof is quite analogous.

So, roughly speaking, for observed $Y = y$, the observations are fused "pixel by pixel", and the result of the fusion is then converted by using (10) into a "classical" noise. This is very useful because in numerous models it is then possible to simulate realizations of X according to its distribution conditional to observations. Let us view some of them :

- Hidden Markov fields model : X is a Markov field. So the fused probability distribution $\Pi \oplus M^{sen}$ is also a Markov field, which which can thus be simulated by the Metropolis algorithm or the Gibbs sampler. This model has ben introduced and discussed in [2].
- Hidden Markov chain model. It is well known that the posterior distribution of a Markov chain noised with an independent noise is a Markov chain distribution. Realizations of X can then be performed. These models are widely used in different signal processing problems and they can also be used in the statistical unsupervised images segmentation [8].
- Hidden Markov trees. More recent models introduce networks in form of "trees", in which the set of pixels S forms the "leaves" [10, 14]. Such models are pyramids in which S forms the basis. In these models X is not a Markov field but realizations of X according to its distribution conditional to observations can be performed. Let us notice that Markov trees can also be used to model evidential priors [13].

- Truncated Markov trees. This recent model, which seems to be very promising, is a truncated pyramid, in which the top is a Markov field and conditionally on the top, the basis is formed by trees whose roots are in the top. Realizations of X according to its distribution conditional to observations can still be simulated in such models [4].

5. Conclusion

We addressed in this paper the problem of unsupervised Dempster-Shafer fusion of possibly correlated and non Gaussian sensors, with application to Bayesian methods of multisensor image segmentation. Such methods, based on the probabilistic fusion of sensors, can be of great efficiency in numerous situations. Furthermore, such methods are based on some probabilistic model like hidden Markov field, and they can be made unsupervised by estimating, in a previous step, all model parameters. On the other hand, Dempster-Shafer fusion can be useful, and we have given some examples of such situations.
The contribution of this paper was to propose, extending the methods valid in a probabilistic frame [12], some methods of the model parameters estimation valid in the Dempster-Shafer framework. The proposed methods are valid in different structures for the class process, such as Markov field, Markov chain, Markov tree, or even Markov truncated tree. Furthermore, the sensors can be dependent and the method determine the very nature of their noise, which can be other than Gaussian.

6. References

[1] A. Appriou, Probabilités et incertitude en fusion de données multisenseurs, *Revue Scientifique et Technique de la Défense*, No. 11, pp. 27-40, 1991.

[2] A. Bendjebbour and W. Pieczynski, Multisensor Evidential Hidden Markov Fields and Image Segmentation, 2nd IEEE International Conference on Intelligent Processing Systems (ICIPS'98), Gold Coast, Australia, August 4-7, 1998.

[3] D. Dubois, S. Moral and H. Prade. Belief change rules in ordinal and numerical uncertainty theories. In D.M. Gabbay and Ph Smets (eds), Handbook of defeasible reasoning and uncertainty management systems, vol. 3, 311-392, Kluwer.

[4] A. Chardin and P. Pérez, Semi-Iterative Inference With Hierarchical Energy-Based Models for Image Analysis, International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition, York, United Kingdom, July 1999, pp. 83-98, Springer LNCS 1654.

[5] T. Denoeu, Analysis of evidence-theoretic decision rule for pattern classification, *Pattern Recognition*, Vol. 30, No. 7, pp. 1095-1107, 1997.

[6] S. Geman, G. Geman, Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *IEEE Trans. on PAMI*, Vol. 6, No. 6, pp. 721-741, 1984.

[7] J. Guan and D.A. Bell, Evidence Theory and Its Applications, North-Holland, Amsterdam, 1991.

[8] N. Giordana and W. Pieczynski, Estimation of Generalized Multisensor Hidden Markov Chains and Unsupervised Image Segmentation, *IEEE Transactions on PAMI*, Vol. 19, No. 5, pp. 465-475, 1997.

[9] Z. Kato, J. Zerubia, and M. Berthod, "Unsupervised parallel image classification using Markovian models, *Pattern Recognition*, Vol.32, n.4, pp.591-604, Avril 1999.

[10] J.-M. Laferté, P. Pérez, and F. Heitz, Discrete Markov Image Modeling and Inference on Quadtree, *IEEE Transactions on IP*, , 2000.

[11] J. Marroquin, S. Mitter, T. Poggio, Probabilistic solution of ill-posed problems in computational vision, *Journal of the American Statistical Association*, 82, pp. 76-89, 1987.

[12] W. Pieczynski, J. Bouvrais, and C. Michel, Estimation of Generalized Mixture in the Case of Correlated Sensors, *IEEE Transactions on IP*, Vol. 09, No. 2, February 2000.

[13] W. Pieczynski, Hidden Evidential Markov Trees and Image Segmentation, IEEE International Conference on Image Processing, "ICIP'99", 24-28 October 1999, Kobe, Japan.

[14] P. Rostaing, J.-N. Provost, and C. Collet, Unsupervised Multispectral Image Segmentation Using Generalized Gaussian Noise Model, International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition, York, United Kingdom, July 1999, pp. 142-156, Springer LNCS 1654.

[15] G. Shafer, A mathematical theory of evidence, Prinston University Press, Prinston, 1976.

[16] L. Younes, Estimation and annealing for Gibbsian fields, *Annales de l'Institut Henri Poincaré*, Vol. 24, No. 2, 1988, pp. 269-294.