International Journal of Sciences and Techniques of Automatic control & computer engineering
IJ-STA, Volume 2, N° 2, December 2008, pp. 818–829.

Color image segmentation using the Dempster-Shafer evidence theory for the fusion of uncertain information sources

Salim Ben Chaabane [(1)], Mounir Sayadi [(1) (2)], Farhat Fnaiech [(1) (2)]
and Eric Brassart [(2)]

[(1)] *SICISI Unit, ESSTT, 5 Av. Taha Hussein, 1008, Tunis, Tunisia*

[(2)] *Laboratory for Innovation Technologies (LTI-UPRES EA3899), Electrical Power Engineering Group (EESA), University of Picardie Jules Verne, 7, rue du Moulin Neuf - 80000 Amiens - France*

Emails: ben_chaabane_salim@yahoo.fr, mounir.sayadi@esstt.rnu.tn, fnaiech@ieee.org, eric.brassart@iut.u-picardie.fr

Abstract. The evidence theory aims to represent and handle uncertain information. An important property of this theory is its ability to merge different data sources in order to improve the quality of the information. In this paper, a color image segmentation approach based on the Dempster-Shafer's theory is presented. The three image components (Red, Green and Blue) are considered as uncertain information sources. An automatic thresholding approach is utilized for finding all major homogeneous regions in each images component at first stage. The evidence theory is then used for the fusion of information coming from the three information sources for the same image. The fusion process does not start from a single frame of discernment, as done in most previously reported works, but starts from first defining three independent frames of discernment associated with the three images to be fused, and then combining them for forming a new frame of discernment. The strategy used to define the mass distributions in the combined framework is discussed in detail. The proposed segmentation algorithm has been applied to textured and biomedical cell image in order to illustrate the methodology. The obtained results show the robustness of the method.

Keywords. *Segmentation, Dempster-Shafer evidence theory, data fusion.*

This paper was recommended for publication in revised form by the editor Staff.
Edition: CPU of Tunis, Tunisia, ISSN: 1737-7749

1. Introduction

Image segmentation is a critical and essential component of an image analysis and/or pattern recognition system, and is one of the most difficult tasks in image processing, which determines the quality of the final results of analysis. The goal of image segmentation is the partition of an image into a set of disjoint areas with uniform and homogeneous attributes such as intensity, color, tone or texture, etc. In most of the existing color image segmentation such as clustering, edge detection, fuzzy logic and neural networks [7] [10], the color image segmentation can be considered as a pixel labelling process in the sense that all pixels that belong to the same homogeneous region are assigned the same label while according to similar color.

In color image segmentation, color of a pixel is given as three values corresponding to the three component images R (Red), G (Green) and B (Blue). There are many papers dealing with segmentation using color and different kinds of colors spaces have been developed by several authors [11], [5] such as RGB space or their transformations (linear/non linear). The general segmentation problem consists in choosing the adapted color model for a specific application. In fact each color representation has its advantages and disadvantages [16]. Nonlinear color transformations such as HSI have essential singularities which are non-removable. The major problem of linear color spaces is the high correlation of the three components.

In this paper, a color image segmentation approach based on data fusion techniques is presented. The basic idea consists in combining these three information sources using the Dempster-Shafer evidence theory [6].

There are mainly three models of fusion operators cited in the science literature: probabilistic Bayesian models, fuzzy models and models resulting from the Dempster-Shafer evidence theory.

The probabilistic Bayesian models are the most cited models; the concept of fusion is deduced from the Bayes rule [12]. However, in the Bayesian models there is confusion between two antagonist concepts: the uncertainty and the inaccuracy. Moreover, we have to note that the performances of the Bayesian data fusion tend to be decrease when the number of information sources increases.

One of the most known non-probabilistic techniques is the fuzzy theory [9]. This technique introduced by L. Zadeh [17], represents information in the form of explicit functions of membership. The disadvantage of the fuzzy theory is that it characterizes the uncertainty in an implicit way; only the inaccurate property of information is presented [1]. In the same way the possibility theory [2], [3], derived from the fuzzy sets allows to process the inaccurate information. Although the Dempster-Shafer evidence theory allows to represent at the same time the inaccuracy and uncertainty using confidence, plausibility and credibility functions [6].

In the subsequent sections, section 2 is devoted to the background of the Dempster-Shafer's theory. The proposed method is described in section 3 and some results are given in section 4. Finally, conclusions are presented in section 5.

2. Background of the Dempster-Shafer evidence theory

The evidence theory, also called Dempster-Shafer theory, was first introduced by Dempster (*1967, 1968*) [4], and formalized by Shafer (*1976*) [13].
This theory is often described as a generalization of the Bayesian theory to represent at the same time the inaccuracy and uncertainty information. It defines a framework of understanding representing all the subsets of the classes spaces. The principal advantage of this theory is to affect a degree of confidence which is called mass function to all simple and composed classes, and to take into account the ignorance of the information.
The basic idea of this theory is to define a mass function on a hypotheses set Ω, called a frame of discernment. The mass function has to be set between values 0 and 1.
Let us note the hypotheses set Ω composed of n single mutually exclusive subset H_i, which is symbolized by:

$$\Omega = \{H_1, H_2, ...H_n\} \tag{1}$$

In order to express a degree of confidence for each proposition A of 2^Ω, it is possible to associate an elementary mass function *m(A)* which indicates all confidence that one can have in this proposition. The function m is defined from 2^Ω to [0, 1] by:

$$\begin{aligned} m : 2^\Omega &\rightarrow [0,1] \\ A &\rightarrow m(A) \end{aligned} \tag{2}$$

The mass distribution for all the hypotheses must fulfill the following conditions:

$$\begin{cases} m(\phi) = 0 \\ \sum_{A_n \subseteq \Omega} m(A) = 1 \end{cases} \tag{3}$$

The quantity m(A) is interpreted like the belief strictly placed on *A*. This quantity differs from a probability by the totality of the belief is distributed not only on the simple classes but also on the composed classes. This modelling shows the impossibility to dissociate several hypotheses. It is the principal advantage of this theory but it represents the principal difficulty of this method.
If $m(A) > 0$, A is called a focal elements.

The union of all the focal elements of a mass function is called the core N of the mass function (*equation 4*).

$$N = \{A \in 2^\Omega / m(A) \succ 0\} \tag{4}$$

From the basic beliefs assignment m, a credibility function $Cr(.)$ and plausibility function $Pl(.)$ can be computed using the equations:

$$Cr(H_n) = \sum_{A \subseteq H_n} m(A) \tag{5}$$

The value $Cr(H_n)$ quantifies the minimal degree of belief of the hypothesis H_n.

$$Pl(H_n) = \sum_{H_n \cap A \neq \phi} m(A) \tag{6}$$

The value $Pl(H_n)$ quantifies the maximal degree of belief of the hypothesis H_n.

The greatest advantage of *DS* theory is its robustness of combining information coming from various sources with the *DS* orthogonal rule.

Let us suppose in the presence of Q distinct and independent information sources. Each source is characterised by a mass function defined on the frame of discernment Ω. The Dempster's combination consists to determining the single masse $m(.)$ resulting from the fusion of these Q masses function $m_Q(.)$ by using the orthogonal rule. Then, the DS combination can be represented for *Q* information sources by the following orthogonal rule:

$$m(H_n) = m_1(H_n) \oplus m_2(H_n) \oplus ... \oplus m_Q(H_n) \tag{7}$$

In the case of two sources S_i and S_j, the *DS* combination can be represented as follows:

$$\forall H_n \subseteq \Omega,\ m(H_n) = \frac{1}{K} \sum_{A \cap B = H_n} m_i(A).m_j(B) \tag{8}$$

Where K is defined by:

$$K = 1 - \sum_{A \cap B = \phi} m_i(A) m_j(B) \tag{9}$$

Note that, this DS combination is commutative, associative, but not idempotent or continuous. In (9), the denominator in Dempster's rule, K, is a normalization factor. This has the effect of completely ignoring conflict and attributing any probability mass associated with conflict to the null set.

Some authors refer to this as a distinct rule, however, this is essentially the Dempster rule applied in Smets' Transferable Belief Model. Smet's model entails a slightly different conception and formulation of Dempster-Shafer theory, though it essentially distills down to the same ideas [14].

For instance, let us denote two mass distributions $m_i(.)$ and $m_j(.)$ which correspond to the two information sources S_i and S_j respectively. Smets proposes two types of combinations as conjunctive rule ($m = m_i \cap m_j$), where $\cap$ represents the conjunctive operator of Smets. Under these considerations, the total mass m assigned to a hypothesis H equals:

$$\forall H_n \subseteq \Omega,\ (m_i \cap m_j)(H_n) = \sum_{A\cap B=H_n} m_i(A).m_j(B) \qquad \textbf{(10)}$$

And disjunctive rule ($m = m_i \cup m_j$), where the total mass m assigned to a hypothesis H equals:

$$\forall H_n \subseteq \Omega,\ (m_i \cup m_j)(H_n) = \sum_{A\cup B=H_n} m_i(A).m_j(B) \qquad \textbf{(11)}$$

Therefore, the Dempster-Shafer evidence theory is a rich model to deal the uncertain information.

3. Proposed Method

The segmentation refers to the process of partitioning a digital image I into multiple regions (sets of pixels) R_i for $i = 1, 2, \ldots n$ (see equation 12). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images.

$$I = \bigcup_{i=1}^{n} R_i \qquad \textbf{(12)}$$

The result of image segmentation is a set of regions that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture. Indeed, Color image segmentation attracts more and more attention; hence, color image processing becomes increasingly prevalent nowadays. Compared to gray level scale, color provides additional information to intensity.

In this paper, a color image segmentation approach based on histogram thresholding and data fusion techniques is presented. Automatic histogram thresholding is utilised for finding all major homogeneous regions in each image component at first stage. Then the Dempster-Shafer theory of evidence is applied in order to fuse the information from these three images.

Under these considerations, we can compute a threshold of each image component by by means of the Maximum Interclass Variance Principle (MIVP). This technique is classified as automatic thresholding methods proposed in the literature [8]. Indeed, Histogram-based methods are very efficient when compared to other image segmentation methods because they typically require only one pass through the pixels. In this technique, a histogram is computed from all of the pixels in the image, and the peaks and valleys in the histogram are used to locate the clusters in the image. Color or intensity can be used as the measure. One disadvantage of the histogram-seeking method is that it may be difficult to identify significant peaks and valleys in the image [15]. In this case and to reduce the influence of the undesired factors on the first segmentation an averaging filter is applied to the histogram of each component image.

However, in this first segmentation, many unclassified pixels are present, reflecting the influence of lack information and high correlated of the three component images (R, G and B) of the final segmentation. In this context, data fusion techniques practiced by the Dempster-Shafer evidence theory is an appealing approach for the image segmentation. The mathematical theory is composed of three distinct parts: the definition of the mass functions, the combination process and the decision-making.

3.1. The definition of the mass functions

In the framework of our application, frame of discernment contains all the regions R_i covered each image component (R, G and B).

$$\Omega = \{R_i\} \qquad pour \quad i = 1,\, 2, \ldots n \tag{13}$$

Each image component is assimilated to an information source $S_q \quad for\ q \in 1, 2, \ldots, Q$, where the mass function m^{S_q} is defined as:

$$m^{S_q} : 2^{\Omega} \rightarrow [0,1] \tag{14}$$

With:

$$\begin{cases} m^{S_q}(\phi) = 0 \\ \sum_{A_n \subseteq \Omega} m^{S_q}(A) = 1 \end{cases} \tag{15}$$

The normal distribution, also called the Gaussian distribution, is an important family of continuous probability distributions, applicable in many fields. Each member of the family may be defined by two parameters: the mean ("average", μ µ) and variance (standard deviation squared) σ^2, respectively. Under the assumption of Gaussian distributions, the distribution of masses is determined as follows:

$$m^{S_q}(R_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp - \frac{(x^{S_q} - \mu_i)^2}{2\sigma_i^2} \tag{16}$$

Where, x^{S_q} represent the gray level of a pixel $P(i,j)$ covered each component image. These values $\mu_i = E(x)$ and $\sigma_i^2 = E(x - E(x))^2$ represent the mean and the variance on the region R_i.

In a finite discrete space, Dempster-Shafer theory can be interpreted as a generalization of probability theory where probabilities are assigned to sets as opposed to mutually exclusive singletons. In traditional probability theory, evidence is associated with only one possible event. In Dempster-Shafer theory, evidence can be associated with multiple possible events, e.g., sets of events.

The mass function assigned to the frame of discernment $m^{S_q}(\Omega)$ is defined by the following equation:

$$m^{S_q}(\Omega) = \frac{1}{\sigma_\Omega \sqrt{2\pi}} \exp - \frac{(x^{S_q} - \mu_\Omega)^2}{2{\sigma_\Omega}^2} \quad \textbf{(17)}$$

With:

$$\mu_\Omega = (\mu_1 + \mu_2)/2 \quad \textbf{(18)}$$

And

$$\sigma_\Omega = \max(\sigma_1, \sigma_2) \quad \textbf{(19)}$$

3.2. Evidence Combination

The greatest advantage of Dempster-Shafer theory is the robustness of its way of combining information coming from various sources with the Dempster-Shafer orthogonal rule, to extract a comprehensive knowledge and to apply a rule of decision. Indeed, the Dempster-Shafer combination can be represented by the following equation:

$$m(R_i) = m^{S_1}(R_i) \oplus m^{S_2}(R_i) \oplus \ldots \oplus m^{S_q}(R_i) \quad \textbf{(20)}$$

For two sources S_q and $S_{q'}$, the aggregation of evidence can be written :

$$m(R_i) = \frac{1}{K} \sum_{R_v \cap R_w = R_i} m^{S_q}(R_v).m^{S_{q'}}(R_w) \quad \textbf{(21)}$$

Where K i sis defined by:

$$K = 1 - \sum_{R_v \cap R_w = \phi} m^{S_q}(R_v).m^{S_{q'}}(R_w) \quad \textbf{(22)}$$

K is considered as a normalization factor and is interpreted as a measure of conflict between the various sources. In particular:

- If $K = 0$, then the sources are totally contradictory.
- If $K = 1$, then the sources are totally concordant.

3.3. The decision making

Unlike the Bayesian theory, where the decision criterion is often the maximum of likehood, the Dempster-Shafer theory gives many solutions (see equation 23 and 24) to take a decision.

Generally, the decision-making is carried out on simple hypotheses which represent the classes in the images. If we accept the composite hypotheses as final results in the decisional procedure, the obtained segmentation results would be more reliable but with a decreased precision.

- Maximum of plausibility

$$x \in C_i \quad si \ Pls(C_i)(x) = \max\{Pls(C_k)(x),\ 1 \le k \le n\} \tag{23}$$

- Maximum of credibility

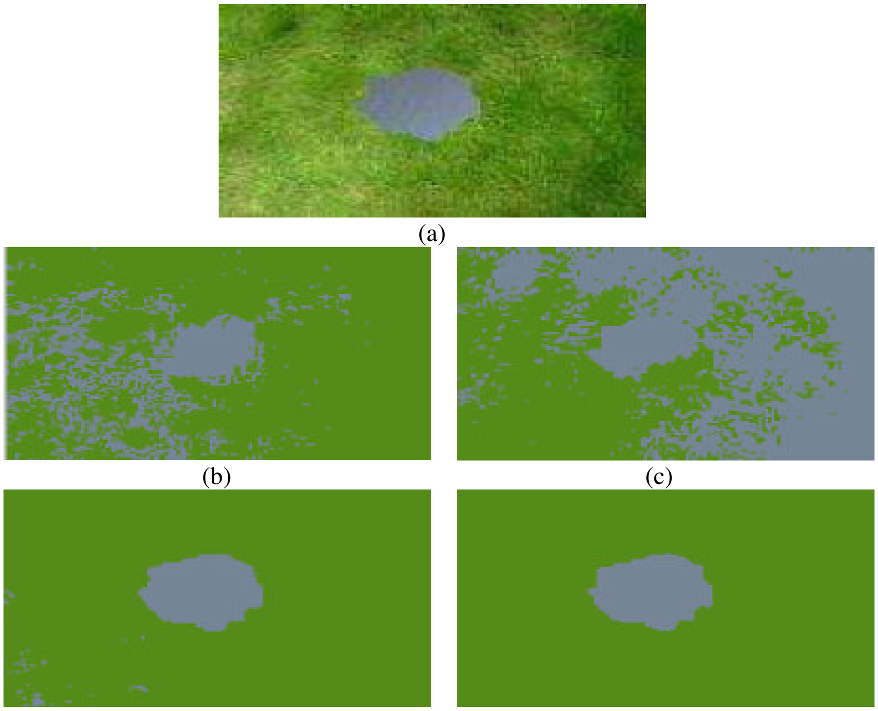
$$x \in C_i \quad si \ Cr(C_i)(x) = \max\{Pls(C_k)(x),\ 1 \le k \le n\} \tag{24}$$

4. Simulation Results

In order to evaluate the proposed approach, we have done experiments on a variety of color images. We have applied the proposed approach to color textured image and cells image. The results obtained for textured image are given in figure 1 and for cells image are presented in the figure 2. In all the following examples, the decision has been made using the criterion of plausibility.

4.1 Experiment 1 : Textured image segmentation

In Figure 1, the original image (top) is a textured image represented in the RGB color space, and contains two classes. Figure 1(e) shows the segmentation result obtained while combining the three information sources. It is observed that the two regions are well classified, showing that the complementary information provided by three images was well exploited by the fusion algorithm. This demonstrates that the calculated mass functions provide a good modelling of the available information associated to the different hypotheses.



(d) (e)

Figure 1. Result on color textured image

(a) Original image
(b) Results after first stage on red component;
(c) Results after first stage on green component;
(d) Results after first stage on blue component;
(e) Results after first stage and second stage

4.2 Experiment 2 : Cell image segmentation

The segmentation algorithm using the Dempster-Shafer evidence theory is applied to a color cell images (see Figure 2). In this application, we have to locate as precise as possible the cell present in the color image. These original images are shown in first line. The second, the third and the fourth line show the segmentation results obtained after the first stage for the component image R, G and B respectively.

We notice the presence of unclassified pixels resulting from the first segmentation. Indeed after the first stage and second stage, we can note that, the cell are properly extracted from the image (see the last line of figure 2). This big performance difference of segmentation results can also be easily assessed by visually comparing the segmentation results. Finally, figure 3 shows a binary mask which can be used to locate the edge of the cell.

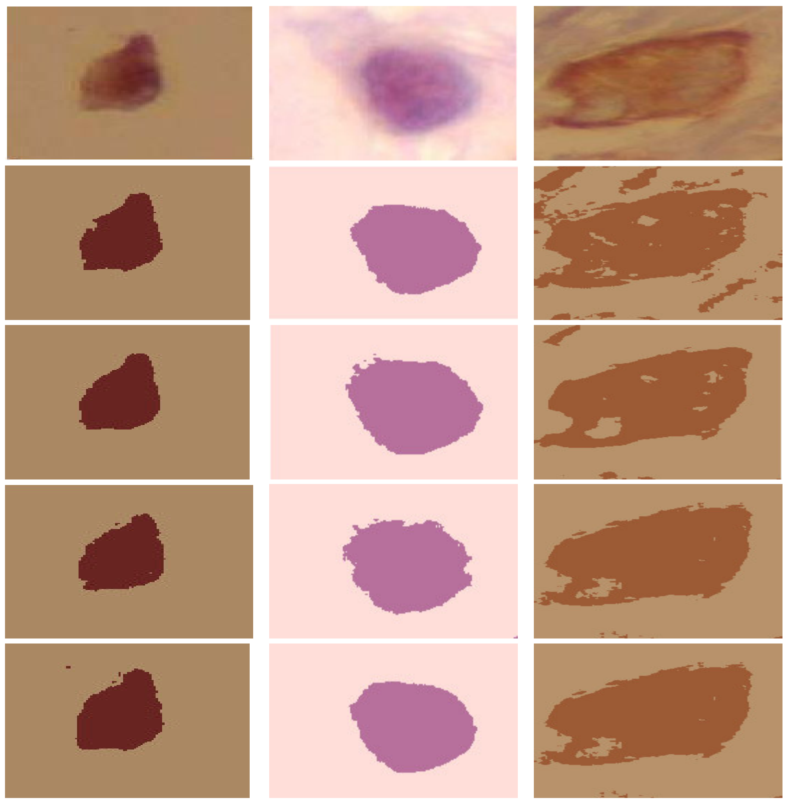


Figure 2. Result on cells image: (First line) Original image
(Second line) Results after first stage on red component;
(Third line) Results after first stage on green component;
(Fourth line) Results after first stage on blue component;
(Last line) Results after first stage and second stage

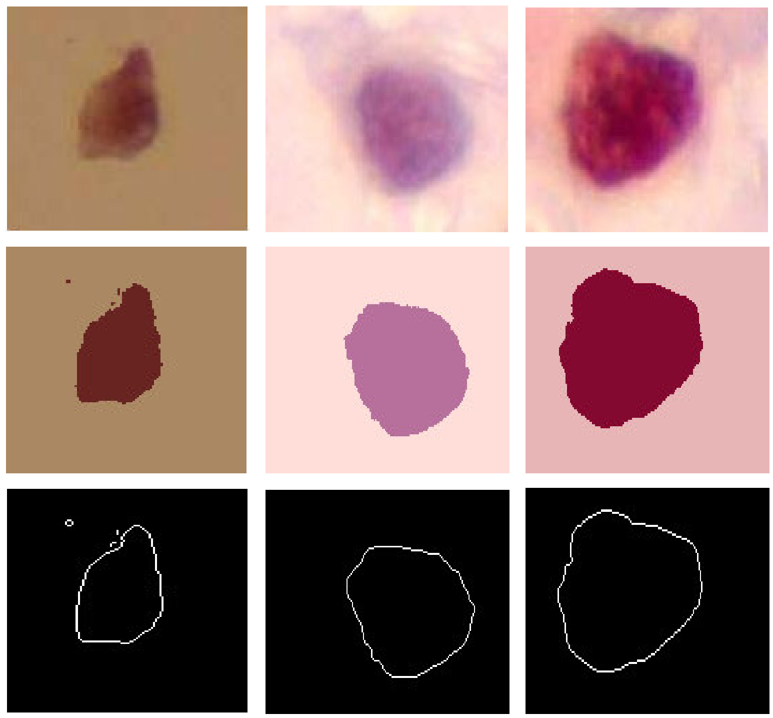


Figure 3. Edges Detection

5. Conclusion

In this paper, we have proposed a Dempster-Shafer's theory based fusion method for the image segmentation in the presence of multiple information. The average filter is used as a tool for analyzing the histograms of the three component images.

The methodology is based on the thresholding and the data fusion techniques. The first segmentation attempts to segment coarsely using the thresholding technique, while the second segmentation allows to reduce the classification errors concerning each pixel of the image based on the Dempster-Shafer evidence theory. The paradigm for deriving mass distributions associated with the images to be fused has been described in detail. In the future works, the proposed fusion method will be applied to a larger class of images.

References

1. I. Bloch, et Maître, H. Fusion de données en traitement d'image : Modèles d'informations et décisions, traitement du signal, vol. 6, p.1811-1823 (1994)
2. D. Dubois, J. Lang, H. Prade: Automated reasoning using possibilistic logic: Semantics, belief revision, and variable certainty weights. IEEE Trans. On Knowledge and Data Engineering, 6: 64-71 (1994)
3. D. Dubois: Belief structures, possibility theory and decomposable confidence measures on finite sets. Comput. Artif. Intell. 5(5): 403-416 (1986)
4. A. P. Dempster : A generalisation of Bayesian inference, Journal of the Royal Statistical Society, p. 205-247 (1968)
5. G. Wyszecki, W. S. Stiles, Color Science: Concepts and Methods, Quantitative Data and formulae. John Wiley and Sons, second edition (1982)
6. G. Shafer: A mathematical theory of evidence. Princeton University Press (1976)
7. J. Basak and D. Mahata: A Connectionist Model for Corner Detection in Binary and Gray Images. IEEE Transactions on neuronal networks, vol.11, no.5, September (2000)
8. K. Suzuki, I. Horiba, N. Sugie et M. Nanki: Extraction of Left ventricular Contours from Left Ventriculograms by Means of a Neural Edge Detector. IEEE transaction on Medical Imaging, Vol.23, NO 3, March (2004)
9. L. A. Zadeh: Fuzzy sets as a basis for a theory of possibility. Fuzzy sets ans systems, 1: 3-28 (1978)
10. M. C. Shin, D. B. Goldgof, K. W. Bowyer et S. Nikiforou. : Comparaison of Edge Detection Algorithms Using a Structure from Motion Task. IEEE Transactions on systems, Man, and cybernetics (SMC 01)-Part B: cybernetics, vol. 31, no. 4, August (2001)
11. N. Ohta: Correspondance between ceilab and cieluv color differences. 2(4): 178-182 (1977)
12. Peddle, D. R: An Empirical Comparaison of Evidential reasoning, Linear Discriminant Analysis and Maximum Likelihood Algorithms for Alpine Land Cover Classification, Canadian Journal of Remote Sensing, vol. 19, p. 31-44 (1993)
13. G. Shafer: A Mathematical Theory of Evidence, vol. Princeton, NJ (1976)
14. P. Smets : Data Fusion in the Transferable Belief Model. Third International Conference on Information Fusion, Paris, France. (2000)
15. V. Patrick, O. Colot and B. Denis: Color image segmentation using Dempster-Shafer's theory. Pro. ICIP'99, pp. 300-304, October (1999)
16. Y. I. Ohta, T. Kanade, T. Sakai: Color information for region segmentation. Comput. Graph. and Imag. Proc., 13:222-241 (1980)
17. L. A. Zadeh: Fuzzy sets Information and control, vol. 8, p. 338-352 (1965)