Dempster-Shafer's Theory as an aid to Color Information Processing Application to Melanoma Detection in Dermatology

Patrick VANNOORENBERGHE, Olivier COLOT and Denis DE BRUCQ
Laboratoire **P**erception **S**ystèmes **I**nformation, UPRES EA 2120
Université/INSA de ROUEN
Place Emile Blondel, BP 08
76131 Mont Saint-Aignan Cedex, France.
Patrick.Vannoorenberghe@univ-rouen.fr, Olivier.Colot@insa-rouen.fr

Abstract

In this paper, we first propose a color image segmentation method based on the Dempster-Shafer's theory. The tristimuli R, G and B are considered as three independent information sources which can be very limited or weak. The basic idea consists in modeling the color information in order to have the features of each region in the image. This model, obtained on training sets extracted from the intensity, allows to reduce the classification errors concerning each pixel of the image. The proposed segmentation algorithm has been applied to biomedical images in order to detect a kind of skin cancer (melanoma). In a second step, features concerning the lesion are extracted using color information. These features are used in order to classify the begnin lesions (naevus) from the other. Results, including the management of false alarms and no detections, allow to demonstrate the effectiveness of the proposed methodology.

1 Introduction

In color image segmentation, color of a pixel is given as three values corresponding to the well known tristimuli R (Red), G (Green) and B (Blue). Different kinds of colors spaces have been developped by several authors [11], [14], [16], [22]. They are derived from this representation of the color using linear and nonlinear transformations. In the framework of segmentation, each color model is more or less convenient, efficient or reliable [17]. The major problem consists in choosing the adapted color model for a specific application. In our study, we choose to work only with the tristimuli (R, G and B) given by the sensor. Each color plane is considered as an information source which can be imprecise or uncertainty. The basic idea of our purpose consists in combining these three information sources using the Dempster-Shafer's theory of evidence [18]. Traditionally, probability theory, which is inadequate in some cases as well known [1], is used for dealing with uncertain data. In the recent past, other models have been developed for handling imprecise knowledge (theory of fuzzy sets [23], possibility theory [7, 8]) or uncertain information (probability theory, theory of belief functions [18]). The use of belief functions as an alternative to subjective probabilities for representing uncertainty was later justified axiomatically by Smets [19, 20] who introduced the Transferable Belief Model, providing a clear and coherent interpretation of the various concept underlying the theory. This well known tool in classification problems [6] provides a convenient framework which allows modeling uncertainty in situations where the available evidence is limited or weak. Areas of application of the Dempster-Shafer's theory are numerous. Some works related to image processing propose to use this approach derived from the confidence measure theory [2]. The paper is organized as follows. Section 2 introduces the problem we want to solve in the framework of Dermatology. We present respectively in sections 3 and 4 the proposed segmentation scheme and the classification procedure. Finally, some experimental results are proposed in the section 5.

2 Color Image Processing in Dermatology

In Dermatology Science, melanoma is an increasing form of cancer. It has increased twice times for 15 years in Canada and it is now 3% of cancers in the USA. The rates of clinical diagnostic accuracy are about 65% at the very best. In particular, it is very difficult to distinguish some atypical lesions - which are benign - from melanoma because they have the same properties according to the well known ABCDE rules used by dermatologists [12]. There is a vi-

sual inspection problem for the atypical lesions class. Unnecessary excisions are often practise for these lesions. The variability of colours and shapes (see Figure 1) can lead to several interpretation by different dermatologists. However,

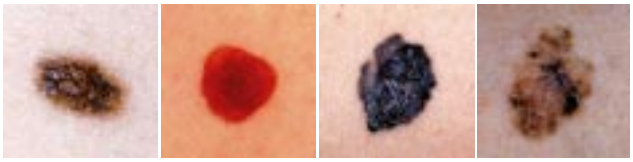

Figure 1. Original images of lesions

melanoma is well suited for color image processing because it is on the skin. Some researches [10] have shown the advantages to use image processing in dermatology. Furthermore, the image processing by computer ensures the reproductibility of the analysis. However, the essential difficulty is to design robust and relevant parameters to ensure the separation between melanoma and benign lesions, in particular the atypical lesions (benign), called naevus, which can be clinically mistaken for melanoma. At first, the lesion border is the feature to identify. It is the first step of the processing to engage in order to extract information about the lesion. So, the border extraction or identification is a critical step in computerized vision analysis in skin cancer as pointed out in [10]. Then, the segmentation step which takes place in any classification processing has to be really accurated. In the framework of our application, only two regions are considered. So, the problem is to separate lesions from the surrounding safe skin. So as to obtain geometric and colorimetric information on a lesion, it is necessary to run a segmentation process which will allow to extract the pixels belonging to the lesion from the image. Dempster-Shafer's theory of evidence [18] is also used in two different steps of the detection system. It is first used in the segmentation scheme but managing uncertainty in the classification procedure is very important. Figure 2 illustrates this both utilisation. This paper is organized as follows. The

Figure 2. Dempster-Shafer's Theory

proposed methodology can be decomposed in two different steps. First one corresponds to the segmentation scheme and second one is the problem of classification of features extracted from the images.

3 The segmentation scheme

A segmentation of an image I is a partition of I into disjoint nonempty subsets $\mathcal{R}_u$ for $u = 1, 2, ..., U$ such as :

$$I = \bigcup_{u=1}^{U} \mathcal{R}_u \tag{1}$$

Under the assumption that images containing only two regions, we can compute a single threshold on the gray level image obtained by means of the Maximum Entropy Principle (MEP) [15]. This coarse segmentation gives two training sets containing pixels which belong surely to one of the considered regions. This first segmentation, based only on the use of gray level image, induces some classification errors. The proposed method is based on the color information contained in the image. It is decomposed in three steps :

- Modeling the belief on the training sets,
- Combining the Q information sources with the Dempster's rule,
- Taking a decision to classify each pixel to a region $\mathcal{R}_u$.

For our segmentation scheme, we choose to work with $Q = 3$ where the different information sources are the tristimuli R, G and B.

3.1 Modeling the belief on the training sets

Let Θ represents the finite set of regions $\mathcal{R}_u$ such as :

$$\Theta = \{\mathcal{R}_u\} \text{ for } u = 1, 2, ..., U \tag{2}$$

Each color plane is assimilated to an information source S_q for $q \in 1, ..., Q$. Let us consider a basic belief assignment m^{S_q} defined as :

$$m^{S_q} : 2^{\Theta} \longmapsto [0, 1] \tag{3}$$

with $m^{S_q}(\emptyset) = 0$ and $\sum_{\mathcal{R}_u \subseteq \Theta} m^{S_q}(\mathcal{R}_u) = 1$. Under the assumption of Gaussian distributions, basic belief functions can be written :

$$m^{S_q}(\mathcal{R}_u) = \frac{1}{\sigma_u \sqrt{2\pi}} \exp^{-\frac{(x^{S_q} - \mu_u)^2}{2\sigma_u^2}} \tag{4}$$

where x^{S_q} is a realization of a Q-dimensional random variable X. In our case, x^{S_q} is the value of a pixel $P_{(i,j)}$ for one of the three color planes. The values $\mu_u = E(X)$ and $\sigma_u^2 = E(X - E(X))^2$ are respectively the mean and the variance on the region $\mathcal{R}_u$. These values are replaced by their statistical approximations computed on the training sets.

Subsets $\mathcal{R}_u$ of Θ such that $m^{S_q}(\mathcal{R}_u) > 0$ are called **focal elements** of m^{S_q}. The union of all the focal elements of a mass function is called the **core** of the mass function (equation 5).

$$\mathcal{F}^q = \{\mathcal{R}_u \subseteq \Omega \mid m^{S_q}(\mathcal{R}_u) > 0\} \quad (5)$$

The advantage of Dempster-Shafer theory lies in representing uncertainty by means of a belief on the whole frame of discernement. This basic belief assignment allows to define $m^{S_q}(\Theta)$ with the following equation :

$$m^{S_q}(\Theta) = \frac{1}{\sigma_\Theta \sqrt{2\pi}} \exp^{-\frac{(x^{S_q} - \mu_\Theta)^2}{2\sigma_\Theta^2}} \quad (6)$$

with $\mu_\Theta = (\mu_1 + \mu_2)/2$ and $\sigma_\Theta = \max(\sigma_1, \sigma_2)$.

3.2 Belief function attenuation

An additionnal aspect of the Dempster-Shafer's theory concerns the attenuation of the basic belief assignment m by a coefficient α_q. The attenuated belief function can be written as :

$$m^{S_q}_{\alpha_q}(\mathcal{R}_u) = \alpha_q . m^{S_q}(\mathcal{R}_u) \quad \forall\, \mathcal{R}_u \in 2^\Theta \quad (7)$$

$$m^{S_q}_{\alpha_q}(\Theta) = 1 - \alpha_q + \alpha_q . m^{S_q}(\Theta). \quad (8)$$

The problem consists in evaluating for each source, the coefficient α_q in order to have the more certain information to aggregate. After the learning step, the main idea is to resume the information contained in each source S_q by means of an optimum histogram computed on the set $\bigcup_{u \in \mathcal{R}_u} \mathcal{X}_{(u;j)}$ in the sense of the maximum likelihood and of a mean square cost. This histogram will be used in order to establish the relevance of a source of information. First, we have to build an approximation of the unknown probability distribution with only the samples given in each source. That is done by means of a histogram building which is led by the use of an information criterion. We will see that different information criteria initially designed for model selection can be used [3, 4]. Once this histogram is obtained, we use the Hellinger's distance between the approximated distribution computed on the set $\mathcal{X}_{(u;j)}$ and the approximated distribution computed on the set $\mathcal{X}_{(u';j)}$. This distance gives a dissimilarity between the two probability densities that is to say the ability of the source to distinguish the two regions $\mathcal{R}_u$ and $\mathcal{R}_{u'}$.

3.3 Fusion of several sources

The Dempster-Shafer's theory allows the fusion of several sources using the Dempster's combination operator. It is defined like the orthogonal sum (commutative and associative) following the equation :

$$m(\mathcal{R}_u) = m^{S_1}(\mathcal{R}_u) \oplus m^{S_2}(\mathcal{R}_u) \oplus .... \oplus m^{S_Q}(\mathcal{R}_u) \quad (9)$$

For two sources S_q and $S_{q'}$, the data fusion can be written as :

$$m(\mathcal{R}_u) = \frac{1}{\mathcal{K}} \sum_{\mathcal{R}_v \cap \mathcal{R}_w = \mathcal{R}_u} m^{S_q}(\mathcal{R}_v) . m^{S_{q'}}(\mathcal{R}_w) \quad (10)$$

where $\mathcal{K}$ is defined by :

$$\mathcal{K} = 1 - \sum_{\mathcal{R}_v \cap \mathcal{R}_w = \emptyset} m^{S_q}(\mathcal{R}_v) . m^{S_{q'}}(\mathcal{R}_w). \quad (11)$$

The normalization coefficient $\mathcal{K}$ evaluates the conflict between two sources. $\mathcal{K} = 0$ corresponds to the case where the sources are totally in conflict.

3.4 Decision rule

The credibility Bel and the plausibility Pl can be computed from the basic belief assignment using following equations :

$$Bel(\mathcal{R}_u) = \sum_{\mathcal{R}_v \subseteq \mathcal{R}_u} m^{S_q}(\mathcal{R}_v) \quad (12)$$

$$Pl(\mathcal{R}_u) = \sum_{\mathcal{R}_u \cap \mathcal{R}_v \neq \emptyset} m^{S_q}(\mathcal{R}_v). \quad (13)$$

A belief function is then transformed into a pignistic probability function defined for all $\mathcal{R}_u \subseteq \Omega$ as :

$$P_p(\mathcal{R}_u) = \sum_{\mathcal{R}_v \subseteq \Omega, \mathcal{R}_v \neq \emptyset} m^{S_q}(\mathcal{R}_v) \frac{\mid \mathcal{R}_u \cup \mathcal{R}_v \mid}{\mid \mathcal{R}_v \mid} \quad (14)$$

In this operation, the mass $m(\mathcal{R}_v)$ is distribued equally among the elements of $\mathcal{R}_u$. Finally, the decision is made by assigning a pixel P to a region $\mathcal{R}_u$ with the maximum pignistic probability. Some results are presented in the section 5.

4 Classification scheme

The classification scheme, used to discriminate the begnin lesions from melanoma, can be decomposed in two different steps. The extraction of features representing the shape and the color of lesion corresponds to the first step. Second step is the classification procedure.

4.1 Extraction of features

For the classification procedure, we propose p primitives - geometric as well as photometric - which reveals to be robust and relevant [5, 9]. The different primitives extracted from the images segmented are derived from :

- Geometric features :

1. The compactness of the shape,
2. The expanse,
3. The mean deviation from the gravity center,
4. The regularity of the contour,
5. The lengthening;

- Photometric features :
  1. The homogeneity of color planes,
  2. The symetry of color planes,
  3. The deviation between the mean of color plane on the safe skin and the mean of color plane on the lesion,

For example, the homogeneity is defined as the sum of the transitions clear zone/dark zone and dark zone/clear zone of a color plane when the lesion is described horizontally and vertically. The dark zone is defined as the third darker area in the lesion and the clear zone is defined as the third clearer area in the lesion. Examples of these kinds of images are presented in the figure 3. The photometric features,

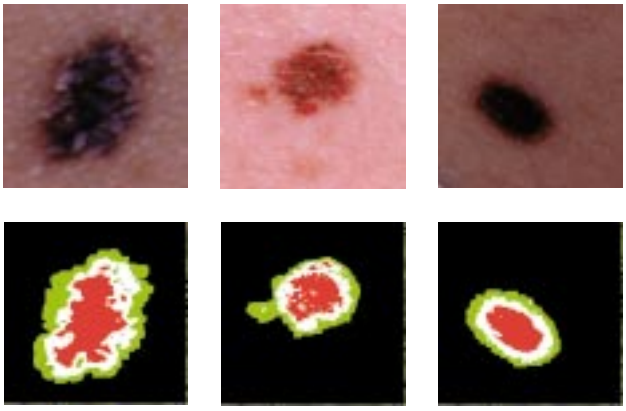

Figure 3. Images illustrating the homogeneity

homogeneity and symetry, are indexes of polychromy.

These different features derives from clinical features or are original features. It is possible to increase the number of primitives but only robust and efficient parameters are needed in a classification step [5, 9]. So, we use 16 parameters in our study taking into account geometric and photometric features. Each primitive gives less or more information about the lesion nature. Then, to build a classification rule, we have to consider all these information sources the primitives are. So, we have to make information fusion in order to decide if a lesion is or not malignant. The classification rule is based on the fact that the malignant lesions have some primitive values which are significantly different from those obtained on the benign lesions. However, the primitives which have such a behaviour can be different from a malignant lesion to another. So, we have to make fusion of the information (fusion of the primitives) to bring out a particular behaviour for the malignant melanoma.

4.2 Classification procedure

After extraction of these different features concerning the lesion, a classification procedure is computed. It is based on the same principle and used the Dempster-Shafer's Theory in order to manage uncertainty on data. Figure 4 gives an example of this histogram building step. The proposed methodology consists in intializing the belief functions with probability densities obtained by learning. By means of information criteria (Akaike's criterion), we determine the attenuation of the belief assignment based on the dissimilarity between probability distributions. Finally, the decision is made by assigning a lesion to the hypothesis melanoma or begnin lesion with the maximum credibility or with the maximum plausibility. Some results on the classification procedure are presented in the section 5.

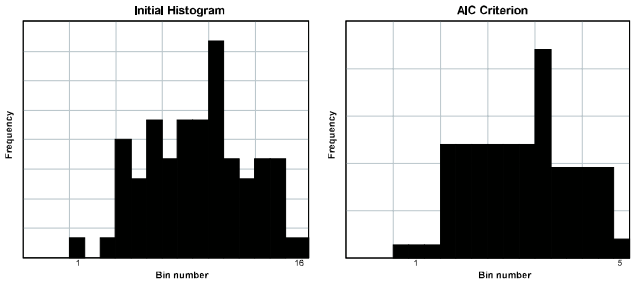

Figure 4. Initial and optimum histograms

5 Experimental Results

This section is devoted to present some results concerning the segmentation scheme and the classification procedure in order to evaluate the methodology.

5.1 Segmentation process

The proposed segmentation algorithm has been applied to biomedical images in order to illustrate the methodology. Some images, in the context of dermatology, are presented in the figure 5. First row corresponds to the original color images. The respective two other rows represent the segmentation scheme results with the decision concerning the maximum of credibility (second row) and the maximum of plausibility (third row). We can note that the lesion (red color) is correctly extracted from the safe skin (white color). The blue color corresponds to pixels which cannot be classify either to the safe skin or to the lesion.

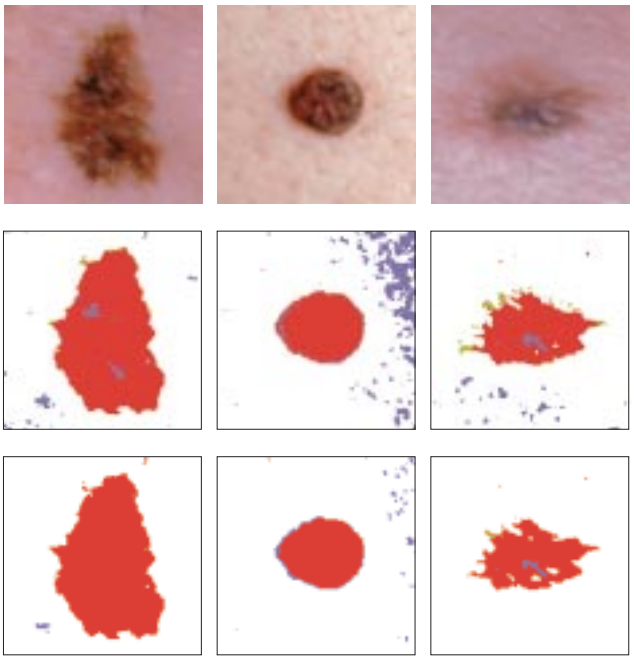

Figure 5. Results images 1

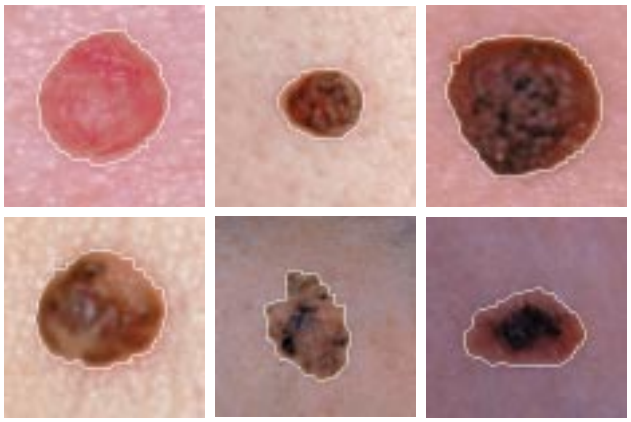

Figure 6. Results images 2

	Naevus	Melanoma
Decision Naevus	80	6 (ND)
Decision Melanoma	14 (FA)	20
Total	94	26

Figure 7. Decision results of the classification rule

5.2 Classification procedure

We applied the proposed classification process on a set of 120 lesions: 94 benign lesions (naevi) and 26 malignant lesions (melanoma). On the training database of 120 lesions (94 benign lesions distributed on the 5 subclasses; 26 malignant melanoma), we obtained 23% of non detection (none malignant melanoma was undetected) and 15% of false alarm which agrees with the theorical rate of false alarm (see table on figure 7). These first results have to be reinforced by a test engaged on a larger database.

6 Conclusion

In this paper, we have presented an original color image segmentation procedure using both information criteria and Dempster-Shafer's theory. The proposed methodology consists in intializing the belief functions with probability densities obtained by learning. By means of information criteria, we determine the attenuation of the belief assignment based on the dissimilarity between probability distributions. After this segmentation step, color information is modelized in order to extract features. The same methodology is applied to the classification procedure. This framework allows to use the whole information contained in the image as better as possible. Future work is concerned with analysis of several decision rules using uncertainty measures proposed by Klir [13, 21].

7 Acknowledgements

The authors would like to acknowledge the support of the Normandy regional council. The co-operation of P. Joly of the clinic of dermatology of Charles Nicolle's hospital of Rouen in supplying diagnosis on melanoma and acquisition of image lesions, is also gratefully acknowledged.

References

[1] J. Bezdek. Fuzziness vs. probability - the n-th round. *IEEE Trans. on Fuzzy Systems*, 2(1):1–42, 1994.

[2] I. Bloch and H. Maître. Fusion de données en traitement d'images: modèles d'informations et décisions. *Traitement du Signal*, Numéro Spécial: Fusion de données(11):435–446, 1994.

[3] O. Colot. *Apprentissage et détection automatique de changements de modèles - Application aux signaux électro-encéphalographiques*. PhD thesis, Université de Rouen, 1993.

[4] O. Colot and al. Information criteria and abrupt changes in probability laws. In M. Holt, C. Cowan, P. Grant, and W. Sandham, editors, *Signal Processing VII: Theories and Applications*, pages 1855–1858. EUSIPCO'94, September 1994.

[5] O. Colot and Al. *A Colour Image Processing Method for Melanoma Detection*. Lecture Notes in Computer Science, 1496. Springer-Verlag, October 1998.

[6] T. Denoeux. Analysis of evidence-theory decision rules for pattern classification. *Pattern Recognition*, 30(7), 1998.

[7] D. Dubois. Belief structures, possibility theory and decomposable confidence measures on finite sets. *Comput. Artif. Intell.*, 5(5):403–416, 1986.

[8] D. Dubois, J. Lang, and H. Prade. Automated reasoning using possibilistic logic: Semantics, belief revision, and varaible certainty weights. *IEEE Trans. on Knowledge and Data Engineering*, 6:64–71, 1994.

[9] C. O. et Al. Une méthode de traitement et d'analyse d'image pour la classification de lésions dermatologiques. In *Proc. Of the Conf.*, pages 54–57, Marrakech, September 1998. MCEA'98.

[10] G. J. E. et Al. Boundary detection in skin tumor images: An overall approach and radial search algorithm. *Pattern Recognition*, 23(11):1235–1247, 1990.

[11] R. W. G. Hunt. *Measuring Color*. Applied science and industrial technology. Ellis Horwood, 2 edition, 1991.

[12] F. R. J. Early detection of melanoma: The role of physician examination and self examination of the skin. *CA*, 35:130–151, 1985.

[13] G. Klir and T. Folger. *Fuzzy Sets, Uncertainty, and Information*. Prentice Hall P T R, Englewood Cliffs, New Jersey 07632, 1988.

[14] M. Miyahara and Y. Yoshida. Mathematical transform of (r,g,b) color data to munsell (h,s,v) color data. In *SPIE Proceedings : Visual Communications and Image Processing*, volume 1001, pages 650–657, San-Jose, 1988. SPIE.

[15] K. J. N. A new method for grey-level picture thresholding using entropy of the histogram. *CVGIP*, 29:273–285, 1985.

[16] N. Ohta. Correspondance between cielab and cieluv color differences. 2(4):178–182, 1977.

[17] Y. Ohta, T. Kanade, and T. Sakai. Color information for region segmentation. *Comput. Graph. and Imag. Proc.*, 13:222–241., 1980.

[18] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[19] P. Smets. Belief functions: The disjonctive rule of combination and the generalized bayesian theorem. *Int. J. of Approximate Reasoning*, 9:1–35, 1993.

[20] P. Smets and R. Kennes. The transferable belief model. *Artificial Intelligence*, 66:191–234, 1994.

[21] Z. Wang and G. Klir. *Fuzzy Measure Theory*. Plenum Publishing Corporation, 233, Spring Street New York, 1992.

[22] G. Wyszecki and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. John Wiley & sons, second edition, 1982.

[23] L. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.