ESTIMATION OF CLASS MEMBERSHIP FUNCTIONS FOR GREY-LEVEL BASED IMAGE FUSION

Isabelle Bloch, Lars Aurdal, Domenico Bijno, Jens Müller

Ecole Nationale Supérieure des Télécommunications
Département Images - CNRS URA 820
46 rue Barrault, 75634 Paris Cedex 13, France
Tel: +33 1 45 81 75 85, Fax: +33 1 45 81 37 94, E-mail: bloch@ima.enst.fr

ABSTRACT

In this paper we propose a new unsupervised method for estimating class membership functions from statistical data. It combines in an original way information derived from the histogram as well as prior knowledge of the requirements that the functions must satisfy and that cannot be derived from the histogram. The method has been tested successfully on MR brain images, and applications to image fusion are illustrated.

1. INTRODUCTION

The aim of the work reported here is the estimation of membership functions for classes of interest in image fusion problems, more precisely low level fusion based on pixel characteristics. We consider mainly grey-level characteristics. Usually, the estimation aspect of the fusion problem is handled by supervised methods, in this paper we propose an unsupervised method. Our main idea is to introduce intuitive requirements on the membership function shapes. Typically, a pixel having a value which has a low occurrence frequency in the image may belong completely and without ambiguity to one class. This is generally not taken into account by methods that try to produce membership functions shaped as modes in the histogram.

This paper is organized as follows. In section 2 we present a short review of existing methods for transforming a histogram or a probability distribution into a possibility distribution or a membership function. In section 3 we describe the proposed method, based on criteria accounting for both the distance between distributions and for constraints on their shape. In section 4 we present a simple application of this method to an image fusion problem in MR brain imaging. We illustrate how the estimation can serve for several fusion methods.

D. Bijno and J. Müller worked on this project during their internship at ENST from Torino and Hannover respectively. The authors would like to thank Prof. C. Adamsbaum for providing the MR images.

2. THE TRANSFORMATION FROM HISTOGRAMS TO MEMBERSHIP OR POSSIBILITY FUNCTIONS

From a probability distribution to a possibility distribution: Several methods have been proposed in the literature for transforming a probability distribution p_k into a possibility distribution π_k (and conversely), both distributions being discrete ($1 \leq k \leq K$) [10], [9], [11]. The main constraints that are used for deriving such transformations are e.g.: preservation of ranking, normalization, probability-possibility consistency: $\pi_k \leq p_k$ [8] or $\sum_{k=1}^{K} p_k \pi_k = c$ where c is a constant in [0,1] [14] or uncertainty conservation using entropy [11]. More details on these methods and on their comparison can be found in [11].

From statistical data to membership functions: Defining membership functions from statistical data is typically done using the now classical C-means algorithms [2], although the shapes of the obtained membership functions are not always adequate for image fusion problems [3]. Another method consists in defining membership functions by optimizing certain criteria, e.g.: fuzzy entropy [6], minimum of specificity and consistency [7].

Using a posteriori criteria: Other approaches consist in minimizing classification errors. They are based on "a posteriori criteria", since they involve the whole classification process, including the decision rule. Such a method is proposed in [13], using fuzzy integrals for the combination step.

Limits of existing methods: The methods presented in the first two paragraphs are used to optimize a priori criteria in the sense that the estimation is performed without any knowledge of the subsequent steps. Therefore, the criteria are simply based on a "resemblance" between the histogram or the probability distribution and the desired function. On the contrary, the methods of the third paragraph include the whole fusion process in the estimation: a posteriori criteria are used in the form of a minimization of classification error. This assumes that the complete strategy of fusion, classification and decision is known before the estimation step is performed.

Since our aim has been to develop an estimation method for class membership functions that can be used with different fusion methods, we prefer methods of the first type. However, we found some drawbacks with existing methods (methods that are not dedicated to image fusion) when considering our application to image processing. Most criteria provide a function that depends on the shape of the histogram. Accounting for frequent situations where a pixel may belong completely and without any ambiguity to a class while having a grey-level with low occurrence frequency thus becomes difficult. Moreover, these methods usually compute one possibility distribution or membership function from a histogram, and do not allow for a direct computation of all class membership functions in a global way. In the next section, we propose criteria that deal with these problems.

3. FROM GREY-LEVELS TO CLASS MEMBERSHIP FUNCTIONS

In this section we propose a new method for the estimation of membership functions from grey-level histograms. This method is well adapted to the classification of homogeneous regions in images and can easily be extended to any other characteristic, including texture indices, e.g. for classification of textured regions. We assume that we have several images of the same scene that have to be fused in order to make a decision, the decision here consists in a classification of the scene into several classes of interest. The aim of the estimation is to define, for each considered image, a membership function for each class of interest represented in this image. We first define the criteria that the membership functions should verify, and then how to optimize them.

3.1. Criteria

We suggest to use two types of criteria simultaneously. The first type is based on a "resemblance" between the grey-level histogram and the membership function in the form of a distance between the two distributions. This type is very close to existing methods. The second type accounts for prior information on the expected shape of the membership function, in order to deal with problems mentioned above concerning low occurrence frequencies. This calls for a parametric representation of the functions. The combination of these two types of criteria leads to a simpler interpretation of the obtained functions that fits better with the intuitive notion of membership.

As membership functions we chose simple trapezoidal functions. They depend on a small number of parameters thus allowing for easy estimation while preserving the robustness expected from membership functions. We will denote by μ_i the membership function for class i ($1 \leq i \leq n$, where n is the number of classes). μ_i depends on four parameters a_i, b_i, c_i, d_i, as illustrated in figure 1. These parameters are subject to some ordering constraints: $\forall i, 1 \leq i < n, a_i \leq a_{i+1}$, and similar inequalities for the other parameters.

One of the prior information we want to introduce is the fact that pixels belonging to a class i without any ambiguity should have a membership value for that class equal to one, whatever the occurrence frequency of its associated characteristics. Therefore, we propose a two step approach: in the first step, we estimate functions μ'_i that minimize a distance criteria to the histogram; in the second step, we derive μ from μ' in order to include this prior information. The functions μ'_i are taken as trapezoidal functions having a maximum value h_i. We define the objective function by:

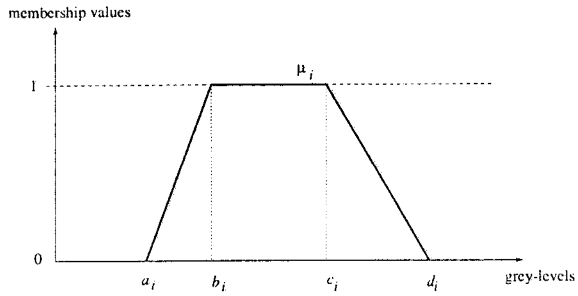

Figure 1: Shape of the membership functions: trapezoidal function depending on four parameters.

$$\sum_{i=1}^{n} \sum_{x \in D_i} [f(x) - \mu'_i(x)]^2, \qquad (1)$$

where x denotes grey levels, $f(x)$ its occurrence frequency in the image, and D_i the interval on which the squared differences are computed for each class i. Important features in this formulation are:

- All functions are computed simultaneously.
- For each class, only the points in the histogram that are concerned with this class are taken into account, through the interval D_i.

The computation of D_i is done according to the desired properties of the membership functions: (i) in the area of ambiguity between two classes (i.e. where two membership functions overlap), there should be no constraint on the fit between f and the μ'_i's; (ii) for the extreme classes (those with lowest and highest grey values respectively), the μ'_i's should not follow the shape of the histogram for the points showing no ambiguity (they usually have low frequency but should have high membership values). According to these constraints, each interval D_i is defined as follows:

1. For $1 < i < n$: we define $k_{i-1,i}$ and $k_{i,i+1}$ as the points at which the functions μ'_{i-1} and μ'_i, and μ'_i and μ'_{i+1} respectively intersect, i.e. such that:

   $$\mu'_{i-1}(k_{i-1,i}) = \mu'_i(k_{i-1,i}),$$

   $$\mu'_{i+1}(k_{i,i+1}) = \mu'_i(k_{i,i+1}).$$

   Thus D_i is defined by $[k_{i-1,i}, k_{i,i+1}]$.

2. For $i = 1$ (class with lowest membership values): we impose $a_1 = b_1$, and $D_1 = [c_i, k_{1,2}]$ where $k_{1,2}$ is the intersection point between μ'_1 and μ'_2.
3. For $i = n$ (class with highest membership values): we impose $c_n = d_n$, and $D_n = [k_{n-1,n}, b_n]$ where $k_{n-1,n}$ is the intersection point between μ'_{n-1} and μ'_n.

Once the functions μ'_i have been estimated in order to minimize the objective function, the membership functions are derived by:

$$\forall x, \mu_i(x) = \frac{\mu'_i(x)}{h_i}. \quad (2)$$

3.2. Optimization method

The optimization of the objective function is performed using simulated annealing in a way similar to the one suggested in [6]. Figure 2 illustrates the obtained results, for the estimation of 3 classes from the histogram of the image presented in the right part of figure 3.

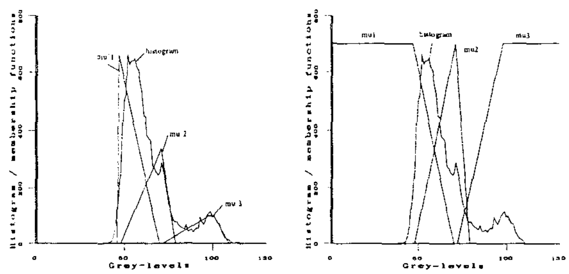

Figure 2: Result of the estimation of three classes (first step: estimation of μ'_i and second step: deriving μ_i) on the second image of figure 3.

As already stated, one of the advantages of the proposed method is that the optimization is performed globally for all classes. We only need to know the number of classes of interest, and the method is thus unsupervised.Moreover, it takes the histogram into account only in the areas where the histogram carries information about the membership fuctions. Elsewhere, the membership functions are defined by the parametric shape. We compared this method with a supervised method, where each membership function is estimated separately from a learning area for the corresponding class. The objective function is of the same type, but computed using the conditional histogram of the class only. We observed that this method is more time consuming, needs learning areas, and is highly sensitive to the choice of these areas.

4. APPLICATION TO BRAIN IMAGE FUSION

We applied the proposed estimation method to an image fusion problem in medical imaging, where we combine dual-echo brain MR images in order to provide a classification of the brain into 3 classes: brain, ventricles and CSF, and pathology. These images are shown in Figure 3. The membership functions for these 3 classes have been estimated in a completely unsupervised way on both images, as described before. We then use these membership functions in 3 different fusion schemes.

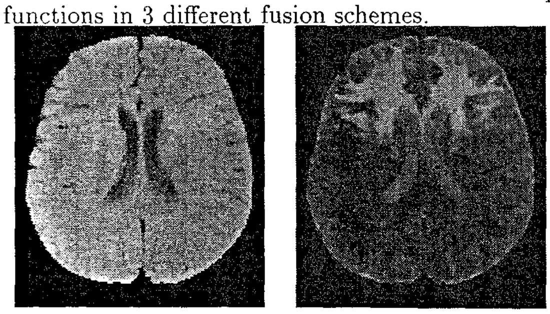

Figure 3: Dual echo MR image of the brain, showing three main classes: brain, ventricles and pathology (the white area on the right image).

Markovian fusion: A first fusion scheme has been developed in the framework of Markov random fields. It consists in combining in a conjunctive way information provided by both images and a regularization term (see [1] for more details on the fusion algorithm). Here, we define the potential functions used in the energy to be minimized as the fuzzy complementation of the estimated membership functions (i.e. a high membership value corresponds to a low energy and conversely). The results are shown in Figure 4.

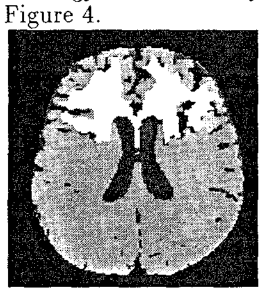

Figure 4: Classification result obtained using a Markovian method.

Fuzzy fusion: Next we use the results obtained with the automatic estimation of membership functions for combining these functions with fuzzy operators. One of the advantages of fuzzy fusion is the large set of operators to chose from depending on the type of information to be combined [4]. Here, since both images provide similar information about the ventricles, we use a mean operator to combine the membership functions obtained in both images for this class. Brain and pathology cannot be distinguished in the first echo and we obtain only one class for this image, denoted by μ_c^1. In the second image, we obtain two classes denoted by μ_c^2 and μ_{path}^2 respectively. We combine μ_c^1 and μ_c^2 using an arithmetical mean again. As for the pathology, we combine μ_c^1 and μ_{path}^2 using a symmetrical sum defined as: $\frac{ab}{1-a-b+2ab}$. This guarantees that no pathology is detected in the areas where $\mu^2_{path} = 0$, and this reinforces the membership to that class otherwise, in order to include the partial volume effect areas in the pathology (this corresponds to what radiologists do). After the combination, the decision is made according to the maximum of membership values. The result is shown in figure 5.

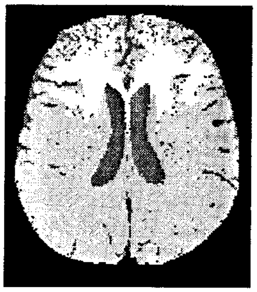

Figure 5: Final decision after fuzzy combination (note that the decision is taken at each pixel individually, without spatial regularization).

Dempster-Shafer fusion: Finally, we interpret the results of the automatic estimation as mass functions and combine them in the framework of Dempster-Shafer evidence theory [12]. We exploit an important feature of this theory that allows for a very flexible modeling of the situation at hand and does not force the introduction of information not contained in the images [5]. We do not assign any mass to the brain and to the pathology in the first image since it does not discriminate these classes, but we assign μ^1_c to the union of these two classes. The ambiguity will then be solved through the combination. The mass functions for the two images are combined using Dempster rule of combination, and decision is taken according to the maximum of belief. The result is shown in figure 6.

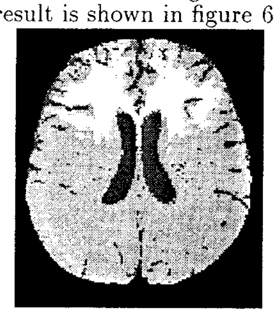

Figure 6: Dempster-Shafer fusion: result of decision after combination with Dempster rule (the results are quite similar to those obtained by fuzzy fusion, and better in the areas affected by partial volume effect around the pathology).

5. CONCLUSION

We have proposed in this paper an automatic method for estimating class membership functions from pixel characteristics. It takes into account histogram information while preserving characteristics of typical membership functions. We have shown that this method can be used for low-level image fusion, using different methods, and we obtained very promising results on dual-echo brain MR images. The method constitutes an improvement over existing approaches, where the estimation step is often performed in a supervised way. Moreover, applying the same estimation step for different fusion methods allows for a comparison of the other steps of these methods (combination and decision); the provided example already illustrates different combination behaviors depending on the chosen framwork. This will be the scope of future work.

6. REFERENCES

[1] L. Aurdal, X. Descombes, H. Maître, I. Bloch, C. Adamsbaum, and G. Kalifa. Fully Automated Analysis of Adrenoleukodystrophy from Dual Echo MR Images: Automatic Segmentation and Quantification. In *Computer Assisted Radiology CAR'95*, pages 35–40, Berlin, Germany, June 1995.

[2] J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms.* Plenum, New-York, 1981.

[3] I. Bloch. Fuzzy Sets in Image Processing. In *ACM Symposium on Applied Computing*, pages 175–179, Phoenix, Arizona, March 1994.

[4] I. Bloch. Information Combination Operators for Data Fusion: A Comparative Review with Classification. *IEEE Trans. on Systems, Man, and Cybernetics*, 26(1):52–67, 1996.

[5] I. Bloch. Some Aspects of Dempster-Shafer Evidence Theory for Classification of Multi-Modality Medical Images Taking Partial Volume Effect into Account. *Pattern Recognition Letters*, 17(8):905–919, 1996.

[6] H. D. Cheng and J. R. Chen. Automatically Determine the Membership Function based on the Maximum Entropy Principle. In *2nd Annual Joint Conf. on Information Sciences*, pages 127–130, Wrightsville Beach, NC, 1995.

[7] M. R. Civanlar and H. J. Trussel. Constructing Membership Functions using Statistical Data. *Fuzzy Sets and Systems*, 18:1–13, 1986.

[8] M. Delgado and S. Moral. On the Concept of Possibility-Probability Consistency. *Fuzzy Sets and Systems*, 21(3):311–318, 1987.

[9] B. Bharathi Devi and V. V. S. Sarma. Estimation of Fuzzy Memberships from Histograms. *Information Sciences*, 35:43–59, 1985.

[10] D. Dubois and H. Prade. Unfair Coins and Necessity Measures: Towards a Possibilistic Interpretation of Histograms. *Fuzzy Sets and Systems*, 10(1):15–20, 1983.

[11] G. J. Klir and B. Parviz. Probability-Possibility Transformations: A Comparison. *Int. J. General Systems*, 21:291–310, 1992.

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

[13] B. Yan and J. Keller. A Heuristic Simulated Annealing Algorithm of Learning Possibility Measures for Multisource Decision Making. *Fuzzy Sets and Systems*, 77:87–109, 1996.

[14] L. A. Zadeh. Fuzzy Sets as a Basis for a Theory of Possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.