UNSUPERVISED MULTISENSOR DATA FUSION APPROACH

Fabien Salzenstein
Laboratoire PHASE, ULP Strasbourg I
23 rue du Loess
67037 Strasbourg, France
salzenst@phase.c-strasbourg.fr

Abdel—Ouahab Boudraa
IRENav, Ecole Navale
Lanveoc Poulmic BP 600
29240 Brest-Naval, France
boudra@ecole-navaie.fr

ABSTRACT

A new iterative approach of multisensor data fusion based on the Dempster-Shafer formalism is presented. Mass functions, formalized by a Gaussian model, are estimated at each iteration using the output fused image and the source images. The effectiveness of the method is demonstrated on synthetic images.

1. INTRODUCTION

Multisensor data fusion is an evolving technology that is analogous to the ongoing cognitive process used by human to integrate data from their senses (sights, sounds, smells,...) continually and make inferences about the external world [1]. The information provided by one sensor is usually limited and sometimes of low accuracy. The use of multiple sensors is an alternative to improve accuracy and provide the user with additional information of increased reliability about the environment in which the sensors operates. Data fusion is an important process in the areas of environmental systems, surveillance, automation, medical imaging and robotics. Different strategies such as Bayesian inference, fuzzy logic or Dempster-Shafer (DS) theory [2] have been developed for data management. DS theory is flexible mathematical tool for handling uncertain and incomplete information through the definition of two dual non additive measures: Plausibility and Belief. These measures are derived from a mass function which is attributed to each subset of frame of discernment. Mass function (MF) definition remains a difficult problem to apply DS theory to practical applications such as image processing. The aim of this work is to estimate automatically MF for unsupervised multisensor classification. MF are derived, at pixel level, from the probabilities. We make assumption that pixels in each region in the image to be segmented have Gaussian distributed gray-levels and that the number of the scene image is known.

2. DEMPSTER-SHAFER THEORY

DS theory is developed as an attempt to generalize probability theory. This theory is suitable to reason with uncertainty and has been developed to overcome the limitations of conventional probability theory by distinguishing between uncertainty and ignorance. DS theory is suited for the combination of information from different sensors.

In DS theory, there is a fixed set of $\mathcal{G} = \{\emptyset, \mathcal{C}_1, \ldots, \mathcal{C}_r\}$ mutually exclusive and exhaustive elements, called the frame of discernment, which is symbolized by:

$$\Theta = \{\mathcal{C}_0, \mathcal{C}_1, \ldots, \mathcal{C}_r\}$$

The representation scheme $\Theta$ defines the working space for the desired application since it consists of all propositions for which the information sources can provide evidence. Information sources can distribute mass values on subsets of the frame of discernment $\mathcal{A}_c \in \mathcal{P}(\Theta)$ (1). An information source assigns mass values only to those hypotheses, for which it has direct evidence. If an information source cannot distinguish between two propositions, it assigns a mass value to the set including both propositions $\mathcal{C}_0 \leftarrow \mathcal{C}_1 \cup \mathcal{C}_2$. The derivation of the mass distribution is the most crucial step since it represents the knowledge about the actual application as well as the uncertainty incorporated in the selected information source.

The mass distribution has to fulfill the following conditions:

- $m_s(\emptyset) = 0$
- $\sum_{A_c \in \Theta} m_s(A_c) = 1$

Mass distribution from different information sources, $m_{s_i} = \{\mathcal{C}_1, \ldots, \mathcal{C}_r\}$, are combined with Dempster’s orthogonal rule (2). The result is a new distribution $m_{\text{new}}(A_c) = \bigoplus_{s_i} m_{s_i}(A_c)$, which incorporates the joint information.
provided by the sources.

\[ m(A_k) = (1 - K)^{-1} \times \sum_{A_1 \cap A_2 \cap \ldots \cap A_n = A_k} \left( \prod_{s \in S} m_s(A_s) \right) \]  

(2)

\[ K = \sum_{A_1 \cap A_2 \cap \ldots \cap A_n = \emptyset} \left( \prod_{s \in S} m_s(A_s) \right) \]  

(3)

\( K \) is often interpreted as a measure of conflict between the different sources (3) and is introduced as a normalization factor. From a mass distribution, numerical values can be calculated that characterize the uncertainty and the support of certain hypotheses. Belief (4) measures the minimum or necessary support whereas plausibility (5) reflects the maximum or potential support for that hypothesis. These two measures, derived from mass values, are respectively defined from \( 2^\mathcal{P} \) to \([0, 1]\):

\[ \text{Bel}(A_k) = \sum_{A_j \subset A_k} m(A_j) \]  

(4)

\[ \text{Pla}(A_k) = \sum_{A_j \supset A_k} m(A_j) \]  

(5)

3. ITERATIVE DATA FUSION METHOD

Suppose that a set of images of the same scene have to be fused in order to classify the scene into \( \mathcal{C} \) classes of interest. Let \( \mathcal{I} \) be the set of images to be fused and defined as random vectors \( X^j = [y_{ij}]^T, j = 1, 2, \ldots, \mathcal{I}; \mathcal{A} = 1, 2, \ldots, \mathcal{N}. y_{ij} \) is a gray level value at site \( j \) of \( i \)th image and \( \mathcal{N} \) is the number of sites of each image. For image segmentation \( \Theta \) denotes the set of hypotheses about pixel class. Let \( \mathcal{L} \) be the set of \( \mathcal{C} \) labels: \( \mathcal{L} = \{ \omega_1, \omega_2, \ldots, \omega_\mathcal{C} \} \). The fusion result is a segmented image \( X = [x_{ij}] \in \mathcal{L}^\mathcal{N} \) into \( \mathcal{C} \) classes. An advantage to apply DS theory to classification problems is that not only single classes (simple hypotheses) but also union of classes (compound hypotheses) can be represented. If there is no ambiguity between classes, a null mass is affected to their union. This means that the different classes are well discriminated in the image and each pixel is affected to only one class. Conversely, when two or more classes are not distinguishable by a sensor, a non null mass is affected to their union. We make assumption that the distribution of \( y_{ij}^k \) conditionally to a class \( \omega_k \) is a Gaussian distribution:

\[ P(y_{ij}^k | x_{ij} \in \omega_k) = \frac{1}{\alpha_{ij}^k \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{(y_{ij}^k - \mu_{ij}^k)^2}{\alpha_{ij}^k} \right)} \]  

(6)

where \( \mu_{ij}^k \) and \( \sigma_{ij}^k \) represent respectively the mean and the standard deviation of the \( \omega_k \) class corresponding to the \( i \)th sensor. MF are estimated for each \( y_{ij}^k \) and for the \( \mathcal{I} \) images. For the choice of MF of two classes \( H_p \) and \( H_p' \), the following strategy is used:

\[ m(y_{ij}^k, H_p \cup H_p') = m(y_{ij}^k | x_{ij} \in H_p) / M \]  

(8)

\[ M = \sum_{x_{ij} \in H_p \cup H_p'} m(y_{ij}^k | x_{ij} \in H_p) \]  

(9)

\[ \psi = \arg \{ \max_{x_{ij} \in H_p} m(y_{ij}^k | x_{ij} \in H_p) \} \]  

(10)

\[ \psi' = \arg \{ \max_{x_{ij} \in H_p'} m(y_{ij}^k | x_{ij} \in H_p) \} \]  

(11)

where \( M \) is a normalization factor. Our method operates on the image histogram and thus the calculated MF (7)-(8) are sensitive to noise. To reduce the noise influence, spatial information is taken into account for each pixel. For each pixel \( m(y_{ij}^k) \) values are re-estimated using the mass values of its neighbourhood \( V_r \) centered at \( x_{ij} \) and including the site \( x_{ij} \) as follows:

\[ \tilde{m}(y_{ij}^k, H_p) = \frac{\sum_{x_{ik} \in V_r} m(y_{ik}^k, H_p)}{|V_r|} \]  

(12)

\[ \tilde{m}(y_{ij}^k, H_p \cup H_p') = \frac{\sum_{x_{ik} \in V_r} m(y_{ik}^k, H_p \cup H_p')}{|V_r|} \]  

(13)

\( \forall \mathcal{K}, h \in \{1, \ldots, \mathcal{C}\}, |V_r| \) is the cardinal of \( V_r \). Before the estimated MF are combined using the orthogonal sum (2) \( K \) is calculated according to (14). To correctly fuse the \( \mathcal{I} \) images, a precise geometrical correspondence between the \( \mathcal{I} \) images is required in order to ensure that each \( \mathcal{I} \)-uple of pixels in the \( \mathcal{I} \) images does represent the same physical point in the object. This correspondence is based on approach using labels configuration of a given reference image \( \mathcal{L}_{\mathcal{K}} \). The labels of each image are permuted and the corresponding MF estimated so that the average conflict is minimum:

\[ K'(x_{ij}) = \sum_{A_1 \cap A_2 \cap \ldots \cap A_n = \emptyset} \left( \prod_{s \in S} \tilde{m}_s(A_s) \right) \]  

(14)

\[ K'(\mathcal{K}) = \sum_{x_{ij} \in V_r} K'(x_{ij}) / N \]  

(15)

The best image reference is given by:

\[ \mathcal{L}_{\mathcal{K}} = \arg \{ \max_{x_{ij} \in V_r} K'(x_{ij}) \} \]  

(16)

Once the conflict is calculated the masses are combined using relation (2). Since the MF are derived from Gaussian probabilities, simple hypotheses are never null [3]. Thus, a decision rule such as maximum of belief over all hypotheses will always favor compound hypotheses. In this case, a decision rule taken over simple hypotheses is preferred to compound ones:

\[ x_{ij} \in \omega_p \iff \text{Bel}(x_{ij}, H_p) = \max_{x_{ij} \in \mathcal{L}} \text{Bel}(x_{ij}, H_p) \]  

(17)
The fusion resulting image is \( \mathbf{X}^n = [\mathbf{a}, \mathbf{b}] \) where \( \mathbf{a} \) denotes the \( \mathbf{a}^{th} \) iteration. Thus, at the \( (\mathbf{a}+1) \) iteration the couple of images \( \{\mathbf{X}^n, \mathbf{Y}^n\} \) is used to re-estimate the statistical parameters of the Gaussian probabilities:

\[
\mu^n_{(\mathbf{a}+1)} = \frac{\sum_{i=1}^{N} \mu^n_i \cdot 1_{[z \in \mathcal{V}_i]}}{\sum_{i=1}^{N} 1_{[z \in \mathcal{V}_i]}} \tag{18}
\]

\[
(\sigma^n_{(\mathbf{a}+1)})^2 = \frac{\sum_{i=1}^{N} (\mu^n_{(\mathbf{a}+1)})^2 \cdot 1_{[z \in \mathcal{V}_i]}}{\sum_{i=1}^{N} 1_{[z \in \mathcal{V}_i]}} \tag{19}
\]

4. RESULTS

Synthetic images with different contrast are used to illustrate our method. Two images, corrupted by Gaussian noise, simulating weak and strong X-rays acquisitions are shown in Fig. 1a and b respectively. Each image contains four regions \( (C = 4) \). In the first image (Fig. 1a), one region (smallest thickness) is confused with the background and in the second one (Fig. 1b) the greatest thickness is underestimated and the thicker regions are not well distinguished. These images are combined in order to provide a classification of the image into four classes. The fused output image is shown in Fig. 1e where the \( C \) regions are well identified and discriminated. Note the very low noise effect on the segmentation result (Fig. 1e). This result shows that the informations complementarity provided by the sensors are well exploited by our method and the adequate mass modelling of the information associated to the different hypotheses. The result is obtained with \( |\mathcal{V}_i| \geq 9 \) and with 10 iterations. Our method has been compared with a classical segmentation method where each image is segmented separately using Fuzzy c-Means (FCM) algorithm. Fuzzy segmented images, shown in Fig. 1c and d, demonstrate that none of the images provides complete and reliable information compared to DS fusion approach (Fig. 1e). Fig. 2 shows two synthetic images (Fig. 2a-b) where the contrast has been reversed and the "Ground-truth" image is known (Fig. 2c). Each sensor identifies only one class. Fig. 2d and e show that the three regions are well brought out. The fusion result shown in Fig. 2d and e are performed with 5 and 20 iterations corresponding to 4.15% and 3.77% of pixels mis-classification respectively. To show the effect of the spatial information on fusion result, different \( |\mathcal{V}_i| \) values are tested. Fig. 3a shows output fused image obtained without taking the spatial information. One may notice the high number of mis-classified pixels. This result demonstrates the interest of the neighbourhood information in estimating MF. The effect of three neighbourhood of 9, 25 and 49 pixels is illustrated in Fig. 3b, c and d respectively. Note the significant segmentation improvement. In particular, as the \( |\mathcal{V}_i| \) increases the homogeneity of the segmented regions increases too. These results are obtained with 5 iterations.

5. CONCLUSIONS

In this paper an iterative unsupervised multisensor classification method using DS theory is described. MF, formalized by a Gaussian model, are estimated at each iteration using the output fused image and the source images. The obtained results on synthetic images are encouraging. Extensive tests on real data are necessary in order to evaluate the method.

6. ACKNOWLEDGMENTS

We would like to thank Pr. Bloch from ENST for her helpful comments and suggestions.

7. REFERENCES


Figure 1: Two images simulating (a) Strong and (b) Weak X-rays acquisitions respectively. (c) Fuzzy segmentation of weak X-rays image. (d) Fuzzy segmentation of strong X-rays image. (e) Output fused image.

Figure 2: Second synthetic images. (a) Image of the first sensor. (b) Image of the second sensor. (c) "Ground-truth" image. (d) and (e) Output fused images obtained with 5 and 20 iterations respectively.

Figure 3: Effect of neighborhood size. (a) Output fused image without pixel spatial information. (b), (c) and (d) Output fused images obtained with neighbourhood of 9, 25 and 49 pixels respectively.