A Fuzzy Approach of the Correspondence of Microcalcifications

Maria Rifqi Jean-Yves Delort
Bernadette Bouchon-Meunier

LIP6 - CNRS
Université Pierre et Marie Curie
case 169 – 4, Place Jussieu, 75252 Paris Cedex 05
{Maria.Rifqi, Jean-Yves.Delort, Bernadette.Bouchon-Meunier}@lip6.fr

Abstract

This paper addresses the correspondence problem in fuzzy image processing. This problem is a generic one in computer vision, arising whenever a match must be found between the features of two related patterns.

The final goal of our problem is to propose an aid when a biopsy using stereotactic guidance is used to confirm the diagnostic of a breast cancer. This implies a matching process of microcalcifications of two views. The matching process of microcalcifications is divided into four steps: the detection of all microcalcifications on the two views (left and right); the construction of a correspondence matrix which indicates all the possible combinations of microcalcifications; the construction of a resemblance matrix, and the exploitation of this matrix. This paper focuses on the two last steps of the process. We propose an approach that uses the general framework of comparison of fuzzy values [3] and the general assignment problem.

KEYWORDS: matching, correlation, correspondence, resemblance

1 Introduction

The general topic of our work is detection of cancer on mammograms. The signs of cancer visible on mammograms can be separated in two groups: microcalcifications and fibrous lesions. The microcalcifications are small depositions of radiologically very opaque materials which can be seen on mammography exams as small bright spots.

In this paper, we focus on the problem of the correspondence between two views of microcalcifications: we have to decide for each microcalcification detected in view 1 where it is located in view 2 (with the possibility that it is not visible in the second view). This task is needed when breast lesions containing ambiguous microcalcifications are detected: a biopsy using stereotactic guidance can be used to confirm the diagnostic. The stereotactic exam consists in taking two views of the breast from two different directions. If the object to be punctured – typically a microcalcification – can be detected on both views and the geometry of the acquisition is known, the 3D position of the object can be calculated. A positioning device can then be used to introduce a needle at the exact position of the object, to obtain a tissue sample.

We present an original solution based on the resemblance measure [3] and the generalized assignment problem and we compare it with some others existing methods.

2 Microcalcifications Matching

The problem can be divided into four steps :

1. the detection of all microcalcifications on the two views (left and right)

2. the construction of a correspondence matrix which indicates all the possible combinations of microcacifications
3. the construction of a resemblance matrix
4. the exploitation of this matrix in order to associate each microcalcification of one view to a microcalcification of the other view.

An overview of the matching process we propose is shown in figure 1.

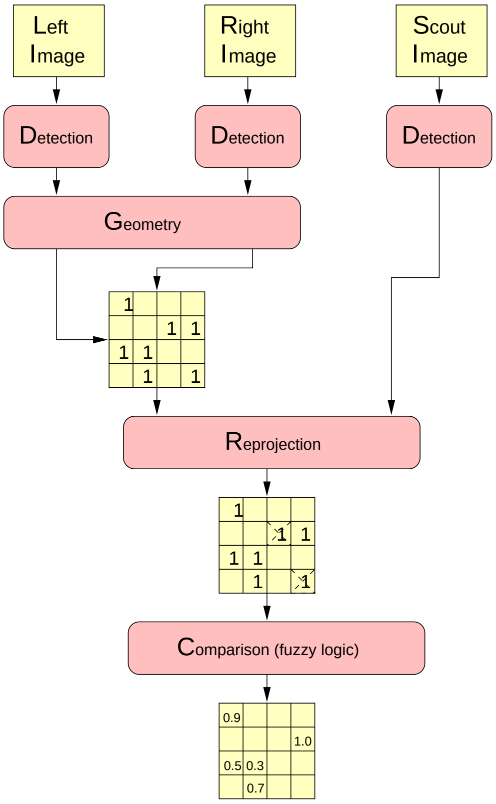

Figure 1: Matching process for microcalcifications

The first step of this process will not be detailed in this paper. This was described in [2] or [11]. What the reader has to know is the result of this step: a set of microcalcifications described by means of fuzzy attributes like surface or compacity (see an example in figure 2). The given membership function of surface means that they are two possible contours for the considered microcalcification: one has a surface of approximately 18 pixels and the other approximately 67 pixels. Analogously, the uncertainty of contours appears in membership functions of the attributes concerning the compacity, with values approximately equal to 122 or approximately equal to 201. We can see that the obtained membership functions are not standard (*i.e.* triangular or trapezoidal). It implies a delicate management of membership functions.

2.1 Construction of a correspondence matrix

After the detection of all microcalcifications on all views, we construct a correspondence matrix which has one row for every candidate in the left image and one column for every candidate in the right image. This correspondence matrix is initialized to 1.0 for each element to indicate that all matches between any pair of candidates are possible. The geometrical constraints from the known acquisition geometry (epipolar line) are used to eliminate all candidate combinations which are not possible by setting to zero the corresponding

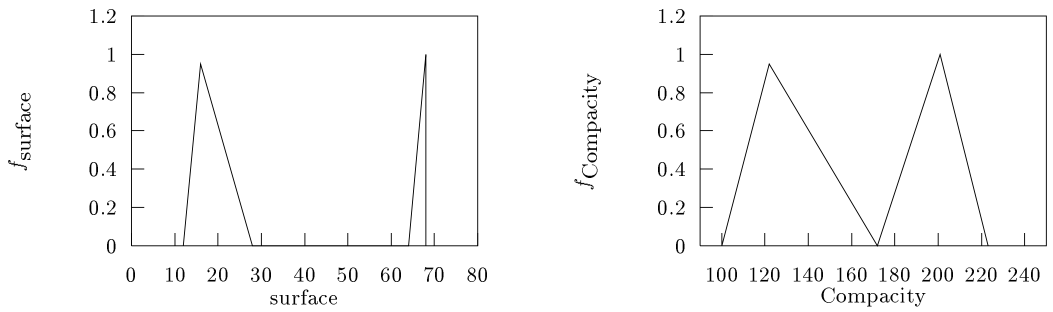

Figure 2: Description of a `small` microcalcification by means of fuzzy values.

matrix component. If a scout view was taken (a third view with a centered tube angle) it can be used to eliminate further candidate combination which do not project at the right place on the third view.

2.2 Construction of a resemblance matrix

The remaining candidates are compared by using fuzzy attributes calculated on the fuzzy contours of the calcifications. All candidates from one view are compared to all remaining candidates in the second view by calculating similarity measures based on the fuzzy attribute.

2.2.1 Notations

We denote the value of attribute j of the microcalcification k for the view i by $m^i_{k,j}$. A microcalcification on the first view is described by $m^1_{k,1\cdots J}$ or in short m^1_k. In the second view, a microcalcification is described by $m^2_{l,1\cdots J}$ or in short m^2_l. The number of microcalcifications in the first view is N whereas it is P in the second view. We suppose that $N \leq P$.

2.2.2 Measure of resemblance

We have proposed [3] a general form for measures of comparison of objects in order to provide a means to choose a measure of comparison relevant and appropriate for a given problem. The framework we have introduced includes the majority of all the measures, indices and relations of comparison used in the literature and helps the user for the creation of new quantities. This framework is compatible with the concepts intoduced in psychology by Tversky [12].

Formally, for any set Ω of elements, let $F(\Omega)$ denote the set of fuzzy subsets of Ω, f_A the membership function of any description A in $F(\Omega)$ and M a fuzzy set measure (like the maximum of the height, the sigma count, the area, ...).

A *measure of comparison* on Ω is a mapping $S(A, B)$ that depends on three arguments: the common features measured by $M(A \cap B)$, and the distinctive features measured by $M(A - B)$ and $M(B - A)$. A measure of comparison is a general family that includes measures of satisfiability, inclusion, resemblance or dissimilarity.

We focus on the measure of resemblance.

A *measure of resemblance* r is used for a comparison between the descriptions of two objects, of the same level of generality, to decide if they have many common characteristics. It is defined as a measure of comparison non decreasing in $M(A \cap B)$, non increasing in $M(A - B)$ and $M(B - A)$, reflexive and symmetrical.

Examples of measures of resemblance are the following:

- $r(A, B) = exp(-\beta|d_r(A, B)|)$ [10] where $\beta > 0$ and $d_r(A, B) = (\sum |f_A - f_B|^r)^{1/r}$, for $r \geq 1$, the generalized geometric distance for fuzzy sets.

- $r(A, B) = M(A \cap B)/M(A \cup B)$ [4] for M such that: $M(A \cup B) = M(A \cap B) + M(A - B) + M(B - A)$.

– $r(A,B) = 1 - \frac{1}{|\Omega|}\sum_x |f_A(x) - f_B(x)| = 1 - \frac{1}{|\Omega|}(M(A-B) + M(B-A))$ [4], $M(A) = \sum_x f_A(x)$.

2.2.3 Method

For a given microcalcification on the first view m_k^1 and for each non-zero cell in the correspondent column (i.e. a matched microcalcification m_l^2 in the second view), the resemblance $r(m_{k,j}^1, m_{l,j}^2)$ is computed for all j. The resemblance index $R(m_k^1, m_l^2)$ is obtained by aggregating the degrees $r(m_{k,j}^1, m_{l,j}^2)$. For the sake of simplicity, we will denote now $R(m_k^1, m_l^2)$ by $R_{k,l}$.

2.3 Exploitation of the resemblance matrix

The proposed solution to the problem of correspondence microcalcifications of two views is inspired by the Generalized Assignment Problem. Intuitively, we define the best solution as the solution that maximises the sum of the values of resemblance of all the possible pairs.

We introduce the variables X_{kl} defined as:

$$X_{kl} = \begin{cases} 1 & \text{if the microcalcification } k \text{ of the view 1 is assigned to the microcalcification } l \text{ of the view 2} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

The hypotheses of the problem are:

1. Each object of the first view has to be assigned to a unique object of the second view.
2. Each object of the second view can not be assigned to more than one object of the first view.

The problem can then be formulated as:

$$\begin{cases} \text{Max} \sum_{k=1,\ldots,N; l=1,\ldots,P} R_{kl} X_{kl} \\ \sum_{k=1,\ldots,N} X_{kl} \leq 1 \quad \forall l = 1,\ldots,P \\ \sum_{l=1,\ldots,P} X_{kl} = 1 \quad \forall k = 1,\ldots,N \\ X_{kl} \in \{0,1\} \end{cases} \tag{2}$$

This classical problem can be solved by the Lagrange relaxation and the sub-gradient method [8].

3 Comparisons with other methods

The trivial approach consists in taking the maximum value in the resemblance matrix and displaying it, removing the object from both views (taking out the row and column of the matrix) and continuing with the next maximum. Except this simple solution, our problem of correspondence of microcalcifications is common to two domains: image processing and sensor fusion. In this section, we briefly present the chosen methods for a comparison process. The reader can see it in details in the cited references.

3.1 Matching by Matching Likelihood Coefficient

Methods used for the correspondence problem in image processsing use in general geometrical knowledge like the relative position of detected objects, line segments, etc. This kind of knowledge is also present in our problem (such as epipolar line constraint) but it is not central: it was used to avoid impossible combinations of microcalcifications. Other methods are intensity-based where the correlation is done pixel by pixel. These methods are not suitable for us. One method is suitable for our problem: matching by Matching Likelihood Coefficient [6].

In this approach, with $R = [R_{kl}]$ the resemblance matrix, for all $k = 1,\ldots,N$ the aim is to:

- find the indice $l_k = \arg\max_l R_{kl}$

- for all $l = 1, \ldots, P$, find the index q_k such that $R_{kq_k} = \max_{l \neq l_k} R_{kl}$
- define $d_k = R_{kl_k}(R_{kl_k} - R_{kq_k})$ as the matching likelihood coefficient.

For the d_{k_0} maximum, the pair that has the maximum resemblance $R_{k_0 l}$ is linked. The process is repeated for all $k \neq k_0$.

3.2 Hierarchical clustering

Methods used in sensor fusion [1] are more adaptable than those seen in the previous section. In particular, we focus on the hierarchical clustering by complete linkage approach [5], [9] used by [7]. The result is a correspondence tree or a dendrogram which shows how objects are grouped together. The first level shows all objects as singleton clusters. As we increase levels, more and more objects are clustered together in a hierarchical manner. The complete linkage method (also called farthest neighbor or maximum distance method) is relatively sensitive to outliers and dependent on the metric. The crucial point for the decision of the good clusters lies on fixing beforehand a level of resemblance. This is a parameter to control together with the metric. If the level is low, then the partitions are large, that means that there are multiple associations (if the level is equal to 0, all combinations are possible). That is, for the problem we are concerned, the very weak point of the hierarchical approach.

3.3 Tests

The validation was performed on 13 matrix of size between ... and ... The following table 1 resumes the results obtained from different approaches: MAX refers to the maximum method, MLC to the Matching Likelihood Coefficient, HC for Hierarchical Clustering ang GAP for Generalized Assignment Problem. For the HC method, several results are given because of the impact of the level of partitionning (α). With the percentage of good matching, the percentage of multi-assignment (MA) is given. This percentage with $\alpha = 0$ (94.84) is the best result but the percentage of multi-assignment is 44.84 that means that 44.84 of microcalcifications of the first view are assigned more than once to the objects of the second view.

Methods	Percentage of success
MAX	62.70
MLC	59.33
HC	38.69 ($\alpha = 1$; MA=0); 76.19 ($\alpha = 0.5$; MA=25.60); 94.84 ($\alpha = 0$; MA=44.84)
GAP	63.89

Table 1: Results from different approaches

4 Conclusion

This paper proposes a new method based on the Generalized Assignment Problem for the correlation problem (in sensor fusion terms) or in correspondance problem (in image processing terms). The validation step shows a good behaviour compared with other methods.

Aknowledgment

We are grateful to Andreas Rick and Serge Muller from General Electric Medical Systems Europe for providing us with the correspondence matrix.

References

[1] A. Ayoun and P. Smets. Data association in multi-target detection using the transferable belief model. *International Journal of Intelligent Systems*, to appear.

[2] S. Bothorel, B. Bouchon, and S. Muller. A fuzzy logic-based approach for semiological analysis of microcalcification in mammographic images. *International Journal of Intelligent Systems*, 12(11–12):819–848, 1997.

[3] B. Bouchon-Meunier, M. Rifqi, and S. Bothorel. Towards general measures of comparison of objects. *Fuzzy Sets and Systems*, 84(2):143–153, 1996.

[4] D. Dubois and H. Prade. *Fuzzy Sets and Systems, Theory and Applications.* Academic Press, New- York, 1980.

[5] R. Duda and P. Hart. *Pattern Classification and Scene Analysis.* Wiley, 1973.

[6] J. Flusser. Object matching by means of matching likelihood coefficients. *Pattern Recognition Letters*, 16:893–900, 1995.

[7] M. Grabisch and H.Prade. The correlation problem in sensor fusion in a possibilistic framework. *International Journal of Intelligent Systems*, to appear.

[8] M. Minoux. *Mathematical Programming: Theory and Algorithms.* John Wiley and Sons, 1986.

[9] S. Miyamoto. *Fuzzy Sets in Information Retrieval and Cluster Analysis.* Kluwer Academic, 1990.

[10] S. V. Ovchinnikov. Representations of transitive fuzzy relations. In H. J. Skala, S. Termini, and E. Trillas, editors, *Aspects of Vagueness*, pages 105–118. D. Reidel Publishing Company, 1984.

[11] A. Rick, S. Bothorel, B. Bouchon-Meunier, S. Muller, and M. Rifqi. Fuzzy techniques in mammographic image processing. In Etienne Kerre and Mike Nachtegael, editors, *Fuzzy Techniques in Image Processing*, Studies in Fuzziness and Soft Computing. Springer Verlag, 2000.

[12] A. Tversky. Features of similarity. *Psychological Review*, 84:327–352, 1977.