

Multimodal medical case retrieval using the Dezert-Smarandache Theory

Gwénolé Quéllec, Mathieu Lamard, Guy Cazuguel, Christian Roux, Béatrice Cochener

Abstract—Most medical images are now digitized and stored with semantic information, leading to medical case databases. They may be used for aid to diagnosis, by retrieving similar cases to those in examination. But the information are often incomplete, uncertain and sometimes conflicting, so difficult to use. In this paper, we present a Case Based Reasoning (CBR) system for medical case retrieval, derived from the Dezert-Smarandache theory, which is well suited to handle those problems. We introduce a case retrieval specific frame of discernment θ , which associates each element of θ with a case in the database; we take advantage of the flexibility offered by the DSMT's hybrid models to finely model the database. The system is designed so that heterogeneous sources of information can be integrated in the system: in particular images, indexed by their digital content, and symbolic information. The method is evaluated on two classified databases: one for diabetic retinopathy follow-up (DRD) and one for screening mammography (DDSM). On these databases, results are promising: the retrieval precision at five reaches 81.8% on DRD and 84.8% on DDSM.

Index Terms—Case based reasoning, Image indexing, Dezert-Smarandache theory, Diabetic Retinopathy, Mammography

I. INTRODUCTION

In medicine, the knowledge of experts is a mixture of textbook knowledge and experience through real life clinical cases. Consequently, there is a growing interest in case-based reasoning (CBR), introduced in the early 1980s, for the development of medical decision support systems [1]. The underlying idea of CBR is the assumption that analogous problems have similar solutions, an idea backed up by physicians' experience. In CBR, the basic process of interpreting a new situation revolves around the retrieval of relevant cases in a case database. The retrieved cases are then used to help interpreting the new one.

We propose in this article a CBR system for the retrieval of medical cases made up of a series of images with contextual information: a class of CBR problems which has hardly been treated. The proposed system is applied to a Diabetic Retinopathy (DR) multimedia database built up in our laboratory; to diagnose DR, physicians analyze series of multimodal photographs together with contextual information such as the patient age, sex and medical history. To

show that the method is generic, we also applied it to DDSM, a public access database for screening mammography; to screen mammography, physicians analyze two views of each breast, with associated contextual information.

When designing a CBR system to retrieve such cases, several problems arise. We have to aggregate heterogeneous sources of evidence (images, contextual information) and to manage missing information. These sources may be uncertain and conflicting. As a consequence, we applied the Dezert-Smarandache Theory (DSMT) of plausible and paradoxical reasoning, proposed in recent years [2], which is well suited to fuse uncertain, highly conflicting and imprecise sources of evidence.

II. MATERIAL AND METHODS

A. Databases

1) *Diabetic retinopathy database*: the diabetic retinopathy (DR) database contains retinal images of diabetic patients, with associated anonymous information on the pathology. Diabetes is a metabolic disorder characterized by sustained inappropriate high blood sugar levels. This progressively affects blood vessels in many organs, which may lead to serious renal, cardiovascular, cerebral and also retinal complications. Different lesions appear on the damaged vessels, which may lead to blindness. The database is made up of 63 patient files containing 1045 photographs altogether. Images have a definition of 1280 pixels/line for 1008 lines/image. They are lossless compressed images. Patients have been recruited at Brest University Hospital since June 2003 and images were acquired by experts using a Topcon Retinal Digital Camera (TRC-50IA) connected to a computer. An example of an image series is given in figure 1.

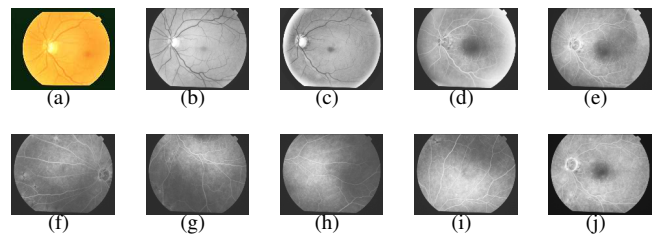


Fig. 1. Photograph series of a patient eye

Images (a), (b) and (c) are photographs obtained by applying different color filters. Images (d) to (j) form a temporal angiographic series: a contrast product is injected and photographs are taken at different stages (early (d), intermediate (e)-(i) and late (j)).

Gwénolé Quéllec, Guy Cazuguel and Christian Roux are with INSTITUT TELECOM; TELECOM Bretagne; UEB; Dpt ITI, Brest, F-29200 France; gweno.le.quelec@telecom-bretagne.eu

Mathieu Lamard and Béatrice Cochener are with Univ Bretagne Occidentale, Brest, F-29200 France;

Gwénolé Quéllec, Mathieu Lamard, Guy Cazuguel, Christian Roux and Béatrice Cochener are with Inserm, U650, IFR 148 ScInBioS - Science et Ingénierie en Biologie-Santé, Brest, F-29200 France;

Béatrice Cochener is with CHU Brest, Service d'Ophthalmologie, Brest, F-29200 France;

The contextual information available is the patients' age and sex and structured medical information (about the general clinical context, the diabetes context, eye symptoms and maculopathy). Thus, at most, patients records are made up of 10 images per eye (see figure 1) and of 13 contextual attributes; 12.1% of these images and 40.5% of these contextual attribute values are missing. The disease severity level, according to ICDRS classification [3], was determined by experts for each patient. The distribution of the disease severity among the above-mentioned 63 patients is given in table I.

2) *Digital Database for Screening Mammography (DDSM)*: the DDSM project [4] has built a mammographic image database for research on breast cancer screening. It is made up of 2277 patient files. Each one includes two images of each breast, associated with some patient information (age at time of study, rating for abnormalities, American College of Radiology breast density rating and keyword description of abnormalities) and imaging information. The following contextual attributes are taken into account in the system:

- the age at time of study
- breast density rating

Images have a varying definition, of about 2000 pixels/line for 5000 lines/image. An example of image sequence is given in figure 2. Each patient file has been graded by a physician. Patients are then classified in three groups: normal, benign and cancer. The distribution of grades among the patients is given in table I.

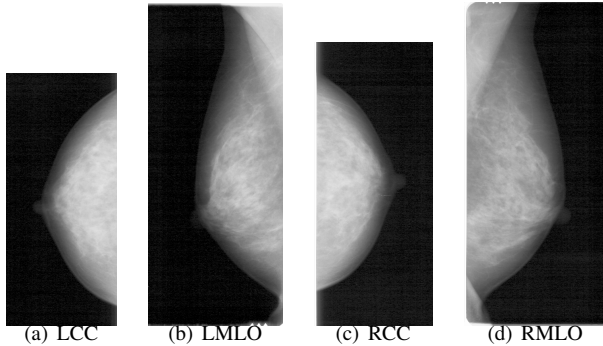


Fig. 2. Mammographic image sequence of the same patient. (a) and (b) are two views of the left breast, (c) and (d) are two views of the right one.

B. Including images in the retrieval system

To include images in the proposed retrieval system, we have to define a distance measure between images and to cluster images acquired at a given imaging modality into a finite number of groups. For this purpose, we follow the usual steps of Content-Based Image Retrieval (CBIR) [5]: 1) building a signature for each image (i.e. extracting a feature vector summarizing their numerical content), and 2) defining a distance measure between two signatures. Thus, measuring the distance between two images comes down to measuring the distance between two signatures. We can then cluster

TABLE I
PATIENT DISEASE SEVERITY DISTRIBUTION

database	disease severity	number of patients
DRD	no apparent diabetic retinopathy	7
	mild non-proliferative	9
	moderate non-proliferative	22
	severe non-proliferative	9
	proliferative	9
	treated/non active diabetic retinopathy	11
DDSM	normal	695
	benign	669
	cancer	913

similar image signatures according to the defined distance measure.

In previous studies, we proposed to compute a signature for images from their wavelet transform (WT) [6]. These signatures model the distribution of the WT coefficients in each subband of the decomposition. The associated distance measure d [6] computes the divergence between these distributions. We used these signature and distance measure to cluster similar images.

Any clustering algorithm can be used, provided that the distance measure between feature vectors can be specified. We used FCM (*Fuzzy C-Means*) [7], one of the most common algorithms, and replaced the Euclidian distance by d .

C. Dezert-Smarandache Theory

The Dezert-Smarandache Theory (DSmT) allows to combine any type of independent sources of information represented in term of belief functions. It is more general than probabilistic fusion or Dempster-Shafer theory (DST). It is particularly well suited to fuse uncertain, highly conflicting and imprecise sources of evidence [2].

1) *The fusion model*: let $\theta = \{\theta_1, \theta_2, \dots\}$ be a set of hypotheses under consideration for a fusion problem; θ is called the frame of discernment. In DST, these hypotheses are assumed incompatibles (constrained model $\mathcal{M}^0(\theta)$), while in DSmT they are not (free model $\mathcal{M}^f(\theta)$): if $\theta = \{\text{"blue"}, \text{"red"}\}$, we may model objects that are both blue and red (magenta). Thus, in DSmT, a belief mass $m(A)$ is assigned to each element A of the hyper-power set $D(\theta)$, i.e. the set of all composite propositions built from elements of θ with \cap and \cup operators, such that $m(\emptyset) = 0$ and $\sum_{A \in D(\theta)} m(A) = 1$; m is called a generalized basic belief assignment (gbba). Nevertheless, it is possible to introduce constraints in the model [2] (hybrid model $\mathcal{M}(\theta)$): we can specify pairs of incompatible hypotheses (θ_a, θ_b) , i.e. each subset A of $\theta_a \cap \theta_b$ must have a null mass, noted $A \in C(\theta)$.

2) *The fusion operators and the decision functions*: to fuse information in the DSmT framework, the user specifies a gbba m_j for each source of evidence S_j , $j = 1..N$. Then, these gbba are fused into a global gbba m_f , according to a given rule of combination. Several rules have been proposed to combine mass functions, including the hybrid rule of combination or the PCR (Proportional Conflict Redistribution) rules [2].

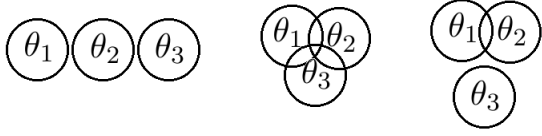


Fig. 3. Considering the frame of discernment $\theta = \{\theta_1, \theta_2, \theta_3\}$, the figure represents, from left to right, the Venn diagram of the constrained, the free and a hybrid model (in which θ_3 is incompatible with the other two hypotheses). A non-empty intersection between circles means the corresponding hypotheses are compatible.

3) *The decision functions*: once the fused gbba m_f has been computed, a decision function (such as the credibility, the plausibility or the pignistic probability) is used to evaluate the probability of each hypothesis. The pignistic probability $BetP$, a compromise between the other two functions, is used since it provides the best system performance:

$$BetP(A) = \sum_{B_i \in D(\theta), B_i \notin V(\theta)} \frac{\mathcal{C}_{\mathcal{M}}(B_i \cap A)}{\mathcal{C}_{\mathcal{M}}(B_i)} m(B_i) \quad (1)$$

where $\mathcal{C}_{\mathcal{M}}(B_i)$ is the cardinality of B_i in the Venn diagram of $\mathcal{M}(\theta)$: in the examples of figure 3, $\mathcal{C}_{\mathcal{M}^0}(\theta_1) = 1$, $\mathcal{C}_{\mathcal{M}^f}(\theta_1) = 4$ and $\mathcal{C}_{\mathcal{M}}(\theta_1) = 2$.

D. Dezert-Smarandache Theory Based Retrieval

1) *Outline of the method*: let c_q be a case placed as a query by the user. We want to rank the cases in the database by decreasing order of relevance for c_q . In that purpose, for each case c_i in the database, $i = 1..M$, we estimate the degree of relevance of c_i for c_q , noted $DR(c_i, c_q)$; then we rank the cases c_i by decreasing order of $DR(c_i, c_q)$, and the top five results are returned to the user. To compute $DR(c_i, c_q)$, we see each case descriptor as a source of evidence: the degrees of relevance are first estimated for each case feature F_j , $j = 1..N$, noted $DR_j(c_i, c_q)$ (as described in section II-D.3), they are then translated into gbba (see section II-D.4) which are fused in the DSMT framework. Finally, $DR(c_i, c_q)$ is estimated from the fused gbba by the pignistic probability (see equation 1). The procedure is summed up in figure 5.

2) *The model*: the following frame of discernment is used for the fusion problem: $\theta = \{C_1, C_2, \dots, C_M\}$, where C_i is the hypothesis “ c_i is relevant for c_q ”, $i = 1..M$. The cardinal of $D(\theta)$ is hyper-exponential in M (it is majored by 2^{2^M}). As a consequence, from computational considerations, it is necessary to include constraints in the model. These constraints are also justified by logical considerations. Indeed, it is not possible for the case c_q to be similar to both cases c_a and c_b if

- c_a and c_b are dissimilar, or
- c_a and c_b have different severity levels

In those cases, we set $A \in C(\theta) \forall A \subset C_a \cap C_b$. To build the model $\mathcal{M}(\theta)$, we first define a non-oriented graph $G_c = (V, E)$, that we call compatibility graph; each vertex $v \in V$ represents an hypothesis and each edge $e \in E$ a couple of compatible hypothesis. To build G_c , we link the

hypothesis associated with each case c_i to the hypothesis associated with its l nearest neighbors at the same severity level. The distance measure used to find the nearest neighbors is simply a linear combination of heterogeneous distance functions (one for each case feature F_j - see section II-B for images), managing missing values [8]. We noticed that the complexity of the fusion operators mainly depends on the cardinality of the largest clique in G_c (a clique is a set of vertices V such that for every two vertices in V , there exists an edge connecting the two). The number l is closely related to the cardinality of the largest clique in G_c and consequently to the complexity of the fusion operation. Finally, the hybrid model $\mathcal{M}(\theta)$ is built from G_c by identifying all the cliques in G_c (see figure 4).

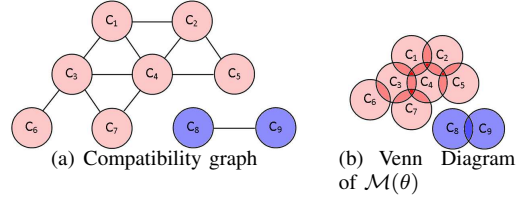


Fig. 4. Building the model $\mathcal{M}(\theta)$ from the compatibility graph. An example of compatibility graph is shown on figure (a). Hypotheses associated with cases at different severity levels are represented with different colors.

3) *The degree of relevance*: to compute the degree of relevance $DR_j(c_i, c_q)$ that a case c_i is relevant for c_q , given the feature F_j , we first define a finite number of states f_{jk} for F_j . Then we compare the membership degree of c_i ($\alpha_{jk}(c_i)$) and of c_q ($\alpha_{jk}(c_q)$) to each state f_{jk} . If F_j is a discrete variable, we associate a state with each possible value for F_j and α_{jk} is a Boolean function. If F_j is an image attribute, $\alpha_{jk}(c)$ is the membership degree of a case c to the k^{th} image cluster (see section II-B - the same clustering algorithm is used for simple continuous variables such as age). We assume that the state of the cases in the same class are predominantly in a subset of states for F_j . So, in order to estimate the conditional probabilities, we use a correlation measure $S_{jk_1 k_2}$ between two feature states f_{jk_1} and f_{jk_2} , regarding the class of the cases at these states. To compute $S_{jk_1 k_2}$, we first compute the mean membership $D_{jk_1 c}$ (resp. $D_{jk_2 c}$) of cases in a given class c to the state f_{jk_1} (resp. f_{jk_2}) (equation 2):

$$\begin{cases} D_{jk c} = \beta \frac{\sum_x \delta(x, c) \alpha_{jk}(x)}{\sum_x \delta(x, c)} \\ \sum_{c=1}^C (D_{jk c})^2 = 1, \forall (j, k) \end{cases} \quad (2)$$

where $\delta(x, c) = 1$ if x is in class c , $\delta(x, c) = 0$ otherwise, and β is a normalizing factor. $S_{jk_1 k_2}$ is given by equation 3:

$$S_{jk_1 k_2} = \sum_{c=1}^C D_{jk_1 c} D_{jk_2 c} \quad (3)$$

the expression of $DR_j(c_i, c_q)$ is given by the following equation:

$$DR_j(c_i, c_q) = \sum_{k_1} \sum_{k_2} \alpha_{jk_1}(c_i) S_{jk_1 k_2} \alpha_{jk_2}(c_q) \quad (4)$$

4) *The generalized basic belief assignments (gbba)*: The gbba m_j is then defined as follows:

- We identify the cases $(c_i)_{i=1..M' \leq M}$ such that $DR_j(c_i, c_q) > \tau_j$,
- We set $m_j \left(\bigcup_{i=1}^{M' \leq M} C_i \right) = m_{j0}$,
- We set $m_j \left(\bigcup_{i=1}^M C_i \right) = 1 - m_{j0}$.

To find τ_j , we define the following test T_j = “if $DR_j(c_i, c_q) > \tau_j$ then C_i is true, otherwise C_i is false”. The sensitivity (resp. the specificity) of T_j represents the degree of confidence in a positive (resp. negative) answer to T_j . T_j is relevant if it is both sensitive and specific. As τ_j increases, sensitivity increases and specificity decreases. So, we set τ_j as the intersection of the two curves “sensitivity according to τ_j ” and “specificity according to τ_j ”, using a dichotomic search. We set m_{j0} as the sensitivity of T_j . $\bigcup_{i=1}^M C_i$ is assigned the degree of uncertainty of T_j . Note that a single setup (τ_j, m_{j0}) is used for each feature, whatever c_i and c_q .

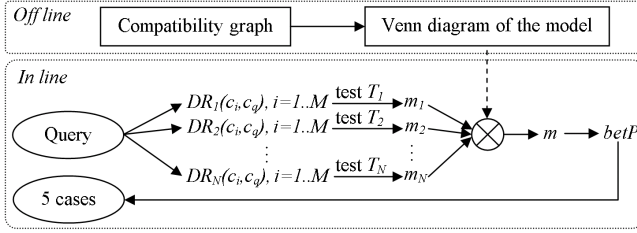


Fig. 5. Outline of the method

III. RESULTS

The mean precision at five (mp5) of the system, i.e. the mean number of relevant cases among the top five results, is 78.6% for DRD and 82.1% for DDSM. As a comparison, the mp5 obtained by CBIR (when cases are made up of a single image), with the same image signatures, is 46.1% for DRD and 70.0% for DDSM [6]. To evaluate the contribution of the proposed system for the retrieval of heterogeneous and incomplete cases, it is compared to the linear combination of heterogeneous distance functions [8] that was used to build the model. A mp5 of 52.3% was achieved by this method for DRD and of 71.4% for DDSM. To assess the robustness of the method regarding missing information, 1) we generated artificial cases from each case in the database by removing attributes, 2) we placed sequentially each artificial case as a query to the system and 3) we plotted on figure 2 the precision at five of these queries according to the number of available attributes.

IV. DISCUSSION AND CONCLUSION

In this article, we introduce a method to include image series, with contextual information, in CBR systems. DSMT is used to fuse the output of several sensors (cases features). We introduced a case retrieval specific frame of discernment θ , which associates each element of θ with a case in the database; we use the flexibility offered by the DSMT's hybrid models $\mathcal{M}(\theta)$ to finely model the database. On DRD, the

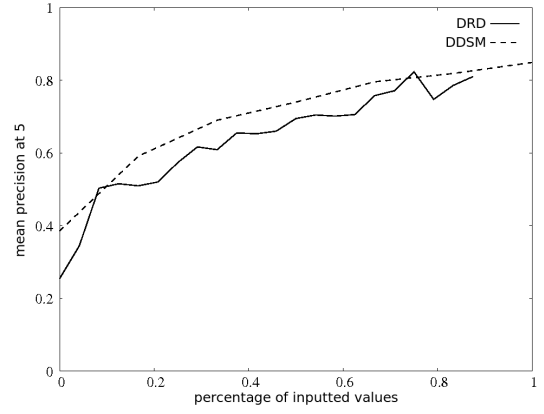


Fig. 6. Robustness with respect to missing values. Note that cases are returned at random when no attributes are inputted (0 on the X axis).

method largely outperforms our first CBIR algorithm (78.6% / 46.1%). This stands to reason because experts generally need more than one image to correctly diagnose the patient's disease severity level. The improvement is also interesting on DDSM (82.1% / 70.0%). This non-linear retrieval method is significantly more precise than a simple linear combination of heterogeneous distances on both databases (78.6% / 52.3% on DRD, 82.1% / 71.4% on DDSM). Moreover, if we use a Bayesian network to infer the missing values, prior to estimating the degree of relevance (described in section II-D.3), the mp5 becomes 81.8% on DRD and 84.8% on DDSM. It is a possible alternative to the decision tree based retrieval system we proposed previously [9] (giving a performance of 79.5% in mp5 on DRD). Finally, the proposed framework is also interesting for being generic: any multimedia database may be processed, as far as a procedure to cluster cases is provided for each new modality (sound, video, etc).

REFERENCES

- [1] I. Bichindaritz and C. Marling, “Case-based reasoning in the health sciences: What's next?” *Artificial Intelligence in Medicine*, vol. 36, no. 2, pp. 127–135, january 2006.
- [2] F. Smarandache and J. Dezert, *Advances and Applications of DSMT for Information Fusion II*. American Research Press Rehoboth, 2006.
- [3] C. Wilkinson, F. Ferris, and R. K. et al., “Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales,” *Ophthalmology*, vol. 110, no. 9, pp. 1677–1682, 2003.
- [4] M. Heath, K. Bowyer, and D. K. et al., “Current status of the digital database for screening mammography,” *Digital Mammography, Kluwer Academic Publishers*, pp. 457–460, 1998.
- [5] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, “Content-based image retrieval at the end of the early years,” *IEEE Transactions on PAMI*, vol. 22, no. 12, pp. 1349–1380, December 2000.
- [6] M. Lamard, G. Quéllec, L. Bekri, G. Cazuguel, B. Cochener, and C. Roux, “Content based image retrieval based on wavelet transform coefficients distribution,” in *Proceedings of the 29th annual international conference of the IEEE EMBS*, august 2007.
- [7] J. Bezdek, “Fuzzy mathematics in pattern classification,” Ph.D. dissertation, Applied Math. Center, Cornell University, Ithaca, 1973.
- [8] D. Wilson and T. Martinez, “Improved heterogeneous distance functions,” *Journal of Artificial Intelligence Research*, vol. 6, pp. 1–34, 1997. [Online]. Available: citeseer.ist.psu.edu/wilson97improved.html
- [9] G. Quéllec, M. Lamard, L. Bekri, G. Cazuguel, B. Cochener, and C. Roux, “Multimedia medical case retrieval using decision trees,” in *29th annual international conference of the IEEE EMBS*, august 2007, pp. 4536–4539.