Integration of Strategies Based on Relevance Feedback into a Tool for the Retrieval of Mammographic Images

A. Fornells[1], E. Golobardes,[1] X. Vilasís[2], and J. Martí[3]

[1] Research Group in Intelligent Systems, Ramon Llull University
[2] Automatic and Control Section, Ramon Llull University
Enginyeria i Arquitectura La Salle, Quatre Camins 2-4, 08022 Barcelona (Spain)
{afornells, elisabet, xvilasis}@salle.url.edu
http://www.salle.url.edu/GRSI
[3] Computer Vision and Robotics Group, University of Girona,
Avda Lluís Santaló s/n, 17071 Girona (Spain)
joanm@eia.udg.es

Abstract. The incidence of breast cancer varies greatly among countries, but statistics show that every year 720,000 new cases will be diagnosed world-wide. However, a high percentage of these cases can be 100% healed if they are detected in early stages. Because symptoms are not visible as far as advanced stages, it makes the treatments more aggressive and also less efficient. Therefore, it is necessary to develop new strategies to detect the formation in early stages.

We have developed a tool based on a Case-Based Reasoning kernel for retrieving mammographic images by content analysis. One of the main difficulties is the introduction of knowledge and abstract concepts from domain into the retrieval process. For this reason, the article proposes integrate the human experts perceptions into it by means of an interaction between human and system using a Relevance Feedback strategy. Furthermore, the strategy uses a Self-Organization Map to cluster the memory and improve the time interaction.

Keywords: Breast Cancer, Bioinformatics Tools, Relevance Feedback, Knowledge Discovery & Retrieval Data, Case-Based Reasoning, Self-Organization Map.

1 Introduction

Breast cancer is the most common cancer among western women and is the leading cause of cancer-related death in women aged 15-54. Screening programs have proved to be good practical tools for prematurely detecting and removing breast cancer, and increasing the survival percentage in women [19]. In an attempt to improve early detection, a number of Computer Aided Diagnosis (CAD) techniques have been developed in order to help experts in the diagnosis.

We focus on a CAD tool for retrieving mammographic images by content analysis [12,5] called HRIMAC[1]. A mammographic image is like a breast

[1] HRIMAC is a project financed by Spanish government - TIC 2002-04160-C02-02.

E. Corchado et al. (Eds.): IDEAL 2006, LNCS 4224, pp. 116–124, 2006.

radiography, which allows the extraction information on the tissue composition. The main purpose is to allow human experts to access a certain type of digital mammographic images typology stored in several public databases. This way, the results returned by the system allow experts to enhance their interpretations, and consequently, to improve the reliability of their diagnosis.

The retrieval process is based on Case-Based Reasoning (CBR) [1] because it justifies results by means of similarity criterion. It is based on solving new problems using its experience, as humans do. The 'experience' is a case memory containing previously solved cases. The CBR cycle can be summarized in the next steps: (1) It retrieves the most similar cases, (2) It adapts them to propose a new solution to the new environment, (3) It revises whether the solution is valid, and finally, (4) It stores it following a learning policy. One of the main difficulties is the definition of the similarity criterion through a similarity function, which is the responsibility of comparing the cases and defining a value of similarity.

Nowadays, HRIMAC only uses the physical features extracted from *mammographic* images previously diagnosed, called microcalcifications (μCa). A sample can contain several μCa in a large variety of sizes and locations making it almost impossible to introduce these abstract concepts into the similarity function. We may improve the precision if we could model the problem domain. That is why we have previously studied a wide set of general purpose of similarity functions and strategies to define similarity functions [5]. However, the experts - using their experiences and their human abstraction abilities - can form concepts which cannot be detected by the system using only the physical information. For this reason, we want to make them participate in the retrieval process by means of a Relevance Feedback strategy [13]. Thus, they can lead the search depending on their points of view through an interactive iterative process. Also, with the aim of reducing the execution time of each interaction, we organize the CBR case memory using a Self-Organization Map.

The article is organized as follows. Section 2 surveys some related work about Relevance Feedback. Section 3 sets the background techniques needed to implement the strategy. Section 4 describes the experimentation. Finally, section 5 summarizes the conclusions and the further work.

2 Relevance Feedback

Relevance Feedback are strategies used to introduce human subjectivity in the retrieval process. This is done by an iterative process in which the expert and the system interact. First, the system shows a set of results parting from an initial question/query from the expert. Next, the expert marks the positives (relevant) and negatives (non-relevant) examples - according to his own perception - from the results. Finally, the system rebuilds the query and shows them again. The process ends when the expert finds what he is looking for. This way, the system auto-adjusts itself according to the perception of expert, in order to obtain more accurate results. Therefore, its application reduces the differences in the similarity concepts of the human expert and the system.

These techniques can be classified according to the next properties: (1) Low level features (using physical properties) versus High level concepts (using concepts or contexts, as humans do) [2,17]; (2) Learning from the feedback of the users or not. This is dangerous if feedback is not fully objective [3]; (3) Positives Feedback, negative Feedback, or both [13]; (4) Category (the user looks for concepts), Objective (the user searches for a specific set of examples), or Exploring (the user does not know what he is exactly looking for) search [3]; (5) Query by example (the query are images) or Query by word (the query are keywords) [2]. Some properties are more recommended than others depending on the domain and the experts. Even they were originally oriented to document retrieval, their usage has been extended to image retrieval (Content Base Image Retrieval).

3 Integration of Relevance Feedback into HRIMAC

3.1 Definition of the Strategy

The requirements needed in our Relevance Feedback strategy are the next: (1) Management of low level features because HRIMAC only uses physical data extracted from mammographic images. Thus, system uses low level features; (2) Positives and negatives Feedbacks; (3) Queries by example or by keyword; (4) Category search; (5) System does not learn from the interaction due to the complexity of the domain and the difference in perception and experiences of each expert when analysing mammographic images.

The time between interactions need to be short because they are done in real time. Retrieve phase is the main neck bottle because CBR compares with all cases from the case memory, which can be useful or not. For this reason, we need to minimize the number of comparison by means of a selective retrieval in which the system only compares with potentially useful cases. The best way is indexing or clustering the case memory depending on the properties of the cases. For this purpose we use a framework called SOMCBR [4] developed by us, which clusters the case memory using a Kohonen or Self-Organizing Map (SOM) [10].

3.2 Integration of SOMCBR into the Relevance Feedback Strategy

SOM [10] is one of the major unsupervised learning paradigms in the family of artificial neural networks. It has many important properties which make it useful for clustering [7]: (1) It preserves the original topology; (2) It works well even though the original space has a high number of dimensions; (3) It incorporates the selection feature approach; (4) Although one class has few examples they are not lost; (5) It provides an easy way to show data; (6) It is organized in an autonomous way to be adjusted better to data. Moreover, Kohonen Maps are a Soft Computing technique that allows the management of uncertain, approximate, partial truth and complex knowledge. These capabilities are useful in order to manage real domains, which are often complex and uncertain. On the other hand, the drawbacks are that it is influenced by the order of the training samples, and it is not trivial to define how many clusters are needed.

```
input : Let R be the set of retrieved elements from SOMCBR; let D+ be the set of
        relevant elements; let D− be the set of non-relevant elements; let I be the actual
        element that is being evaluated;
Function Full Strategy is
    D−=∅, D+={initial mammography image}
    forall I de D+ do
        //Retrieve the most similar cases from the X most similar model
        R=R+(SOMCBR(I)−D−)
    //Show the results to the user
    if human expert finds what he is looking for then
        End the execution
    else
        //The user marks the positive and negative images/cases
        D+=<Relevant elements>
        D−=D− ∪ <Non-relevant element>
```

Fig. 1. Relevance Feedback strategy using the SOMCBR framework

They have successfully been used in a variety of clustering applications for CBIR. Zhang was the first in used them to filter images according to the colour and texture [20]. Next, Han and Myaeng [6] used them to define the outline of objects. Also, they have been used to develop search engines as in PicSOM [11] or WEBSOM [9].

SOM allows the projection of the original n-input data space into a new shorter m-output space in order to highlight the more important data features. This property allows the defining of clusters represented by a vector, which models certain patterns. These clustering capabilities are used in SOMCBR [4] to do a selective retrieval. This way, CBR only compares with cases belonging to the X most similar clusters instead of comparing with all cases from the case memory. The definition of the X value depends on the relation between time execution and error rate desired, because it determines the number of cases used in the comparison. Finally, figure 1 describes the SOMCBR integration in the Relevance Feedback strategy.

4 Experiments and Results

It is difficult to measure the improvement of Relevance Feedback into HRIMAC. There is not a standard benchmarking because evaluation is completely related to domain, its complexity, and the points of view of the expert. However, the introduction of an expert into the retrieval process allows system to obtain more accurate results under the perception of the expert, and consequently, its integration can be considered as positive. However, we can evaluate how SOMCBR works instead of the CBR without clustering (CBRWC) into the Relevance Feedback strategy. It means, the impact of using less cases in the retrieve phase on the error rate. Because we only study the measure capabilities, both systems are studied using a 1-NN in the retrieve phase.

Our main goal is to reduce the case retrieval mean time, which is related with the cases used in the retrieve phase. The number of models used determines this value as we can see in the figure 1. Let I be the number of interactions; let D be the results marked by expert as positives from CBR in each interaction; let K be the map size; let S be the size of case memory; and considering the mean number of elements by cluster as S/K^2, the number of comparisons in retrieve phase can be modeled in equation 1 and 2. They show how the number of operations in SOMCBR is smaller than in CBRWC for short X values, and how the operations are incremented for bigger values of X. Also, SOMCBR and CBRWC work similar when X is equal to the number of cases.

$$number\ of\ operations\ in\ CBRWC = (1 + D \cdot (I - 1)) \cdot S \tag{1}$$

$$number\ of\ operations\ in\ SOMCBR = (1 + D \cdot (I - 1)) \cdot (K^2 + X \cdot S/K^2) \tag{2}$$

Next, we study these equations over several datasets (see table 1) from HRIMAC with the aim studying this in a more quantifiable way. The μCa dataset [12] contains samples from Trueta Hospital (in Girona), while DDSM [8] and MIAS [18] are public mammographic images datasets, which have been studied and preprocessed in [15,14] respectively. The μCa dataset contains samples of mammographies previously diagnosed by surgical biopsy, which can be benign or malign. DDSM and MIAS-Bi classify mammography densities, which was found relevant for the automatic diagnosis of breast cancer. Experts classify them either in four classes (according to BIRADS [16] classifications) or three classes (classification used in Trueta Hospital). Therefore, all this information is used with to aims: (1) Detecting abnormalities or espicular lesions in shape of μCa, and analysing whether they are benign or malign; (2) Defining the density of tissue to improve the mammographic interpretation.

The charts in figure 2 show the evolution of the number of operations. They have been build using the equation 2, and supposing common values for I (5) and D (3 and 5) variables. Also, we have test situations with few (K=2) and many clusters (K=8). It is obvious that SOMCBR strategy drastically reduces the number of comparisons required, and consequently, SOMCBR is better than CBRWC in terms of execution time.

The next step is to evaluate the influence of the reduction in the number of comparisons on the precision of results. We have three parameters to tune: the similarity function, the X value, and the map size. The similarity function in CBR used is Minkowski with r=1 because it provides the best error rate. Also, we focus on the worst situation for SOMCBR, that is, when X value is 1. Finally, we do not know the optimal number of clusters. For this reason, we study several map sizes in order to represent situations defined by many clusters

Table 1. Description of datasets from HRIMAC

Dataset	**Attributes**	**Class distribution**
μCa	22	benign (121), malign (95)
DDSM	143	b1(61), b2(185), b3(157), b4(98)
Mias-Bi	153	b1(128), b2(78), b3(70), b4(44)
Mias-3C	153	fatty(106), dense(112), glandular(104)

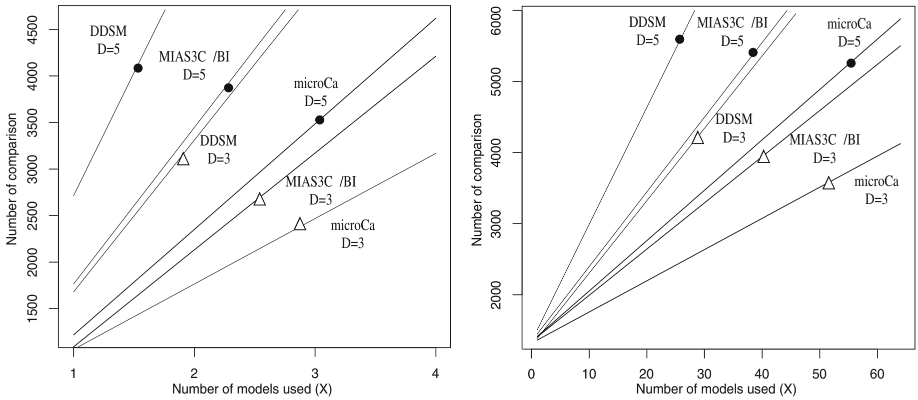

Fig. 2. Evolution for the number of operations in Relevance Feedback strategy using SOMCBR configured with 2 × 2 and 8 × 8 respectively, applying the equation 2. The 'Δ' and '•' symbols represent the configurations for D=3 and 5 respectively.

with few cases, and defined by few clusters with many cases. The configurations tested are: 2 × 2 (4 clusters), 4 × 4 (16 clusters), and 8 × 8 (64 clusters).

Table 2 summarizes the error rates, their standard deviation, and the case retrieval mean time in milliseconds for the CBRWC and for several SOMCBR configurations. All computes have been done applying a 10-fold stratified Cross Validation. Comparing the error rates, we can observe that strategy based on clustering the case memory maintains the error rate for the datasets, and it provides equivalent results if we apply a t-student at 95% of confidence level. This is produced because the Soft Computing capabilities of SOM allows the management of uncertain, approximate, partial truth and complex knowledge. Thus, the SOMCBR strategy is able to manage these complex data, and it does not affect negatively the error rate. Also, the results show the improvement of time, which is significantly better when the CBR case memory is clustered, that is, in SOMCBR strategy.

Therefore, we can conclude that the application of SOMCBR as kernel of our Relevance Feedback strategy is positive, because we drastically reduce the number of comparisons in retrieve phase, and consequently the execution time, without negatively affecting the error rate.

Table 2. Summary of error rates, their standard deviation, and the case retrieval mean time in milliseconds for CBR and SOMCBR using several map sizes ($K \times K$), X=1, and Minkowski (r=1) as similarity function

Code	**CBRWC**		**SOMCBR - 2×2**		**SOMCBR - 4×4**		**SOMCBR - 8×8**	
	%AR(std.)	Time	%AR(std.)	Time	%AR(std.)	Time	%AR(std.)	Time
μCa	31.02(10.5)	0.1	34.26(9.5)	0.04	33.80(7.7)	0.02	34.26(8.1)	0.01
DD	55.49(5.8)	1.9	53.49(5.6)	1.30	53.89(5.2)	1.18	54.29(4.3)	1.10
MB	30.94(11.4)	1.5	29.69(5.5)	0.69	31.25(7.4)	0.61	32.19(5.8)	0.59
M3	29.81(6.4)	1.5	29.19(6.2)	0.69	29.81(6.2)	0.61	31.06(8.5)	0.59

5 Conclusions and Further Work

We are developing a tool called HRIMAC for retrieving mammographic images depending on certain typology from several public databases. The original kernel of HRIMAC is based on a CBR approach, which looks for the most similar cases in comparison with the input case using a similarity function. One of the main difficulties is the definition of the similarity function because the information available from the domain is complex and uncertain. Also, it is not trivial incorporate capabilities for detecting concepts or abstraction inside the similarity function. Relevance Feedback strategies allow the introduction of human experience into the retrieval process in order to reduce the differences in similarity concepts between human and machine through an iterative interaction process. Thus, the retrieval process can benefit of the experts abilities for creating concepts and relationships that systems usually can not detect. For this reason, we want to integrate this strategy into HRIMAC in order to make the task of retrieving mammographic images easier.

On the other hand, the interaction needs to be fast because it is done in real time. The CBR retrieve phase is the bottle neck because it has to compare with all the cases from the case memory. One way of improving it is by means of a selective retrieval, in which CBR only compares with potentially useful cases. For this reason, we have proposed the use of SOMCBR as kernel of our Relevance Feedback strategy because it is a CBR with a case memory clustered by SOM. SOM is a clustering algorithm that projects the original space into other more reduced with the aim of highlighting the more important features. This property is used to build clusters that model the information. Thus, CBR can organize the case memory to improve the retrieval time. Also, SOM has Soft Computing capabilities that allow the management of uncertain, approximate, partial truth and complex knowledge and, consequently, improve the management capacity. The experiments done show that error rate has not been negatively influenced by the reduction of the information used. Therefore, SOMCBR strategy is better than CBR in these datasets because it improves the execution time without negatively affecting the error rate.

The further work is focused on improving the Relevance Feedback strategy allowing semantic content by means of the introduction of keywords, which represent concepts of experts, in order to improve the capacity of retrieving results more accurately. This goal requires two previous steps: (1) Experts have to mark all the mammographic image in order to set high level relations, (2) We need to define a similarity function capable of managing and measuring distances between concepts.

Acknowledgements

We would like to thank the Spanish Government for the support under grants TIC 2002-04160-C02-02 and TIN 2005-08386-C05-04, and the *Generalitat de*

Catalunya (DURSI) for the support under grants 2005SGR-302 and 2006FIC-0043. Also, we would like to thank *Enginyeria i Arquitectura La Salle* of Ramon Llull University for the support to our research group.

References

1. A. Aamodt and E. Plaza. Case-based reasoning: Foundations issues, methodological variations, and system approaches. *IA Communications*, 7:39–59, 1994.
2. Z. Chen, L. Wenyin, C. Hu, M. Li, and H. Zhang. ifind: a web image search engine. In *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*, page 450. ACM Press, 2001.
3. I.J. Cox, T.P. Minka, and T.V. Papathomas. The bayesian image retrieval system, pichunter: Theory, implementation, and psychophysical experiments. *IEEE Transaction on Image Processing – special issue on digital libraries*, 9:20–37, 2000.
4. A. Fornells, E. Golobardes, D. Vernet, and G. Corral. Unsupervised case memory organization: Analysing computational time and soft computing capabilities. In *8th European Conference on Case-Based Reasoning*, 2006. In press.
5. E. Golobardes, X. Llorà, M. Salamó, and J. Martí. Computer aided diagnosis with case-based reasoning and genetic algorithms. *Journal of Knowledge Based Systems*, 15:45–52, 2002.
6. K. Han and S. Myaeng. Image organization and retrieval with automatically constructed feature vectors. *SIGIR Forum*, special issue:157–165, 1996.
7. S. Haykin. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, 2nd edition, 1999.
8. M. Heath, K. Bowyer, D. Kopans, R. Moore, and P.J. Kegelmeyer. The digital database for screening mammography. *Int. Workshop on Dig. Mammography*, 2000.
9. S. Kaski, T. Honkela, K. Lagus, and T. Kohonen. Websom: Self-organizing maps of document collections. *Neurocomputing*, 21(1):101–117, 1998.
10. T. Kohonen. *Self-Organization and Associative Memory*, volume 8 of *Springer Series in Information Sciences*. Springer, Berlin, Heidelberg, 1984. 3rd ed. 1989.
11. J. Laaksonen, M. Koskela, and E. Oja. Picsom: Self-organization maps for content-based image retrieval. In *Proceedings of Int. Joint Conference on NN*, 1999.
12. J. Martí, J. Español, E. Golobardes, J. Freixenet, R. García, and M. Salamó. Classification of microcalcifications in digital mammograms using case-based reasonig. *International Workshop on Digital Mammography*, 2000.
13. H. Müller, Wolfgang Müller, Stéphane Marchand-Maillet, and Thierry Pun. Strategies for positive and negative relevance feedback in image retrieval. *International Conference on Pattern Recognition*, 1:1043–1046, 2000.
14. A. Oliver, J. Freixenet, A. Bosch, D. Raba, and R. Zwiggelaar. Automatic classification of breast tissue. In *Iberian Conference on Pattern Recognition and Image Analysis*, pages 431–438, 2005.
15. A. Oliver, J. Freixenet, and R. Zwiggelaar. Automatic classification of breast density. *IEEE International Conference on Image Processing*, 2:1258–1261, 2005.
16. T. H. Samuels. *Illustrated Breast Imaging Reporting and Data System BIRADS*. American College of Radiology Publications, 3rd edition, 1998.
17. Stan Sclaroff, Leonid Taycher, and Marco LaCascia. Imagerover: A content-based image browser for the world wide web. Technical Report 5, 1997.

18. J. Suckling, J. Parker, and D.R. Dance. The mammographic image analysis society digital mammogram database. In A.G. Gale et al., editor, *Proceedings of 2nd International Workshop on Digital Mammography*, pages 211–221, 1994.
19. D. Winfields, M. Silbiger, and G. Brown. Technology transfer in digital mamography. *Report of the Joint National Cancer Institute, Workshop of May 19-20, Invest Radiololgy*, pages 507–515, 1994.
20. H. Zhang and D. Zhong. A scheme for visual feature based image indexing. In *Storage and Retrieval for Image and Video Databases III*, volume 2420, 1995.