A TWO-LEVEL CBIR PLATFORM WITH APPLICATION TO BRAIN MRI RETRIEVAL

John Moustakas[1,2], Kostas Marias[2], Socrates Dimitriadis[3] and Stelios C. Orphanoudakis[1,2]

[1] Department of Computer Science, University of Crete, Heraklion, Hellas
[2] Institute of Computer Science, Foundation for Research and Technology, ICS-FORTH, Hellas
[3] Department of Cognitive and Linguistic Sciences, Brown University, Providence RI, USA

ABSTRACT

This paper presents a novel platform for image retrieval based on a two-level architecture inspired from human cognitive mechanisms. These two levels provide both generic similarity and semantic information related to special characteristics. Although the proposed architecture can be customised for any content-based image retrieval application, our work is focused on medical images and more specifically on brain MRI data. Our main motivation is that in medical applications, it is crucial to be able to combine in an efficient way, image similarity with specific, semantics related to pathology in order to provide the user with relevant cases and aid diagnosis. A description of the architecture and function of the proposed CBIR platform is presented, as well as specific details for the application to brain MRI retrieval.

1. INTRODUCTION

The already large and continuously increasing volume of multimedia data in the internet dictates the development of efficient methods and tools for retrieval. Content Based Image Retrieval (CBIR) has drawn the attention of the Computer Vision research community, while it may also play an important role as a decision support tool for clinicians [1]. Most frequently, retrieval is based on the computation of features, such as colour, texture, shape, edges in combination with several transformations and similarity/distance measures (see [2] for a comprehensive review). Consequently, the problem is reduced to the retrieval of images with similar 'visual' characteristics.

However, the semantic content of images is subjective, and depends on the specific application. For this reason, such approaches often fall into the 'semantic gap' problem, meaning that the computed features can't always describe well the real characteristics of the image. As a remedy to this problem, it is often suggested to incorporate prior knowledge and/or adopt an 'intelligent' approach. This trend has also led to an important challenge of CBIR; to draw inspiration and model mechanisms of biological perception and vision.

This paper presents a novel CBIR platform that features a two-tier architecture. Our work is inspired by the human cognitive architecture. The key idea underlying our platform emanates from psychological and neuroscientific studies which indicate that the human visual system processes information at several stages:

- According to the classic view of Feature Integration Theory in attention [3], the visual system retains independent retinotopic maps for different primitive visual features (color, form, etc.). In the pre-attentive or early stage of vision, the processing on these feature maps is undertaken in parallel and independently, whereas in the subsequent attentive stage the visual modalities engage in a co-operative work. In other words, the pre-attentive tier decomposes the optical scene in its primitive characteristics which are – to a large extent – independently, in parallel and autonomously processed.
- After the first stage of fixed-time pre-attentive processing the primitive features are organised in such a way that the semantic entities of the scene can be re-composed (e.g. the Gestalt rules [4]). In order to perceive and comprehend the scene, the human visual system performs a serial and selective examination of semantic objects that draw the subject's attention – that is the attentive level of perception.

Neuroscientific evidence supports a similar organization and structure as well [5]. The primary visual cortex, V1, segregates signals related to different attributes of the visual scene and distributes them out to different areas, V2-V6, for further independent, and specialized processing.

Based on these concepts we developed a two-tier CBIR platform featuring a pre-attentive and an attentive level of retrieval. This work is an extension of our previous work on a generic, agent-based, single-tier platform for CBIR [6]. The current platform though is designed for a specific domain, i.e. brain MRI image retrieval. In the next section we explain the proposed architecture.

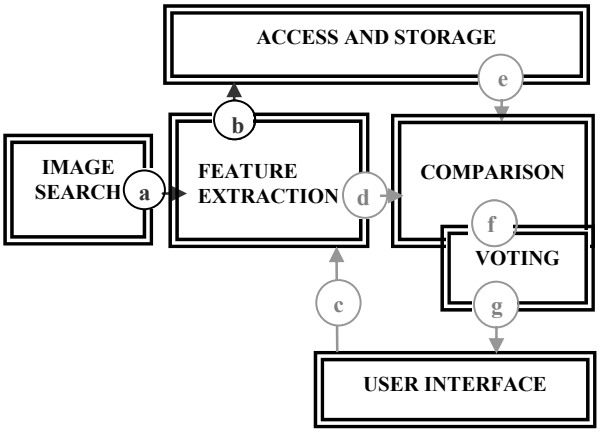


Fig. 1: Schematic view of the platform's function

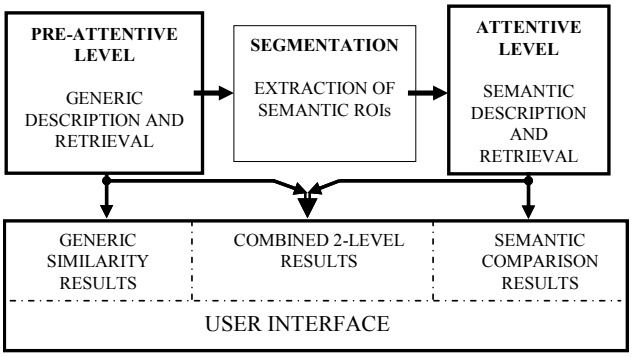


Fig. 2: The 2-tier retrieval architecture of the platform

2. A NOVEL 2-TIER CBIR PLATFORM

2.1. General Description

Fig. 1 is a schematic overview of the function of the CBIR platform which is designed to be fully automated and unbiased (i.e. no relevance feedback was used). Two main functional stages can be identified:

i) The import stage where an autonomous agent searches for images and imports them in the DB (process **a**) followed by feature extraction and storage (process **b**)

ii) The retrieval stage where the user request (**c**), is followed by the feature extraction from the Query Image (**d**), the access to the DB and feature comparison (**e**), then to the voting stage (**f**), and finally the presentation of the results to the user (**g**). The retrieval result is updated as more images from the DB are analysed, until the user is satisfied or all the DB images are compared to the query.

In order to avoid the 'semantic gap' problem, the proposed CBIR platform is designed to retrieve medical images (brain MRI) using a combination of both 'pre-attentive' and 'attentive' criteria, as was discussed in the introduction. This two-tier functional specialization takes place only in processes **b**, **d** and **e**.

The platform's computing modules are autonomous agents [7]. An application-specific agent (brain MRI in our case), continuously searches for images and imports them in a database. In order to serve the two-tier architecture, the platform comprises both pre-attentive (computing generic image similarity) and attentive agents (computing and analyzing semantic features). Each agent is capable of extracting, retrieving and storing information at his personal DB space. The analysis and retrieval-decision tasks are divided in asynchronous parallel processes, each one managing features independently. The fusion of the partial results is based on a voting scheme. The voting weights are computed exclusively by the agent's methodology regardless of the selected voting scheme. If P is the number of agents participating in a particular decision group, and α the feature distance vector for a specific query image, then the similarity score S is calculated using the following equation:

$$S = \sum_{i}^{P} w_i \frac{a_i}{\sigma(a_i)} \qquad (1)$$

where w_i is the agents' weight vector, and $\sigma(\cdot)$ the variance of the respective values in the retrieval set, used to normalize the computed values among the different agents. However, each agent's task is different and the ranges of values produced are often difficult to compare across agents. In our platform, we incorporated processes for 'lifetime' update of several statistical measurements, as for example the fluctuation of the comparison results. This way, while the platform acquires experience via successive retrievals, such measurements converge, reflecting the actual range of values for each agent in image-data of a particular application.

These agents support the two-tier architecture of the platform that is described in the next section.

2.2. The two-tier Retrieval Architecture

Fig. 2 is a schematic representation of the proposed two-tier retrieval architecture. The first pre-attentive tier (parallel feature map generation), is concerned with a 'general' description of the images, while in the second, a serial analysis of specific regions of interest is performed. The user can chose to use one of the two levels for retrieval, or a combination of them. This way, the two-tier architecture enhances flexibility and allows the user to determine the relative effect of generic similarity analysis in the pre-attentive level, and semantic analysis (e.g. presence of disease in brain MRI), in the attentive level. Additionally, the pre-attentive level may be used as a rejection-filter (excluding several potential 'matches') for computational efficiency.

Fig. 3 illustrates the function of each tier. The pre-attentive layer (Fig. 3a), produces independent, parallel feature maps (A, B, C), each one coding an independent visual feature. During the retrieval stage, each agent

compares the computed values between the query and each database image. The comparison scores from all relevant agents are driven to the voting system resulting to the final score for the candidate retrieval image (according to equation 1).

Fig. 3b illustrates the attentive layer for semantic description and retrieval. This layer is application specific, and in our platform a segmentation algorithm has been designed in order for the agents to receive, one by one, the regions of interest (ROIs). Then, a specialized group of 'attentive' agents carefully examines ROIs in a serial fashion. As is shown in Fig. 3b, first region '1' is analyzed by the agents and compared to the single region of the other image. Only after this comparison ends, region '2' is also compared to the same region. In our implementation, the attentive similarity of a given pair of MRI images is defined as the similarity of their 'closest' pair of ROIs.

3. APPLICATION TO BRAIN MRI CBIR

The human brain exhibits a remarkable degree of symmetry with respect to the mid-sagittal plane and the identification of regions of asymmetry is often indicative of diseases such as schizophrenia, epilepsy, and Alzheimer. In addition, pathologies such as stroke and tumors, often exhibit significant asymmetries in the two hemispheres of the brain. Therefore, asymmetry analysis in brain images is a valuable *semantic* characteristic for brain MRI CBIR. This particular application was chosen in order to prove the usefulness of this concept in developing a tool for medical decision support. Medical image retrieval is a very demanding application since it is expected to provide the clinician decision support for diagnosis by retrieving 'relevant' cases.

3.1. Asymmetry detection and segmentation of ROIs

In order to apply the proposed CBIR architecture to the brain MRI retrieval problem it is necessary to automatically identify visual properties that are important for interpreting such images. In our work we:

i. Propose a new measure for brain asymmetry detection that allows us to compute binary maps of asymmetry that are used as a generic similarity measure in the pre-attentive level.
ii. Segment and analyse asymmetrical regions for the 'attentive' level of CBIR.

A complete account of these medical image analysis tasks is out of the scope of this paper; for this reason a summary of the several processing steps is provided:

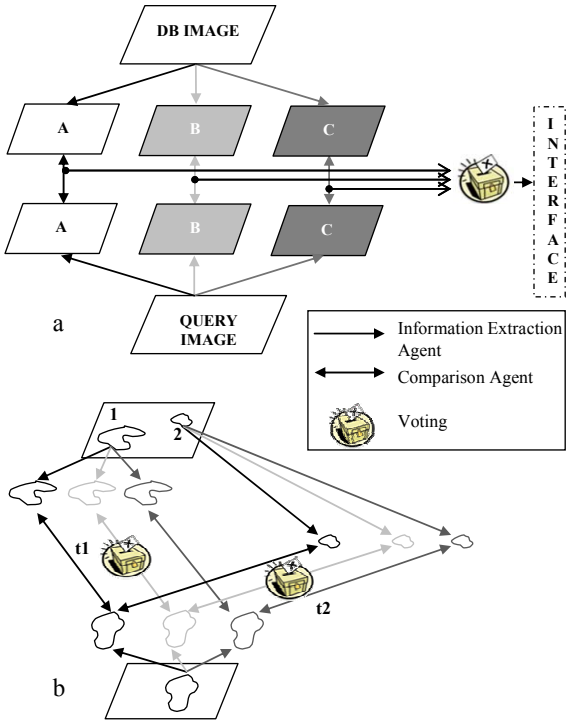


Fig. 3: a) The pre-attentive level architecture; Parallel extraction and retrieval of visual information (A, B, C) for the whole of the image – comparison via voting. b) The attentive level architecture features a serial examination and comparison of significant ROIs.

a) Following the minimal mapping criteria of proximity and similarity suggested by Ullman [8] we propose a novel measure for asymmetry (with respect to the mid-sagittal axis of symmetry) detection that significantly reduces false positives due to edges:

$$C(i,j) = \left| \min_{k,l} \left(I(i,j) - I(w-i+k, j+l) \right) \right| \qquad k,l \in [-N \ldots N] \quad \textbf{(2)}$$

for all k,l of each ($2N+1$ x $2N+1$) window (N is scale dependent and in our experiments we used N=2 for 200x200x256 sized MRI data). This way, a topographical map of asymmetry (Fig. 4b) that reduces 'normal' structural asymmetry, is computed by a pre-attentive agent.

b) From the pre-attentive asymmetry map we first process each image using a morphological filter and clustering of neighboring segments (Fig. 4c). This step produces several 'symmetrical' regions. However, in most cases only one of these pairs is a 'true' positive (regarding asymmetry).

c) Lastly, we introduce further criteria of homogeneity to separate true positives from false positives (Fig. 4d). This step is essential since the asymmetry map often produces a 'mirror reflection' of true positives (see Fig. 4b). Most frequently, such regions contain heterogeneous tissues and can be identified on the basis of low homogeneity.

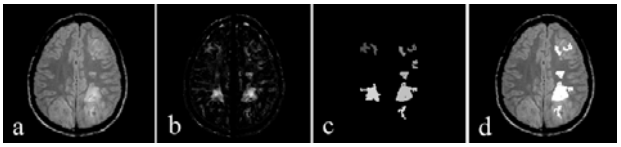


Fig. 4: a) Original image, b) The proposed measure produces reduced-edge, map of asymmetry, c) Image processing to reduce the number of asymmetrical regions, d) Homogeneity criteria separate true from false positives.

5.1. Agents used and results

In the pre-attentive level the agents compute: Mutual information (45%), intensity histogram (20%), eccentricity (20%) and topographical asymmetry map (15%). In the attentive level they compute: Location of ROI (40%), intensity histogram (25%), size (25%) and eccentricity (10%). This way, as is shown in Fig. 5a, by using the pre-attentive agents only, the platform retrieves images that exhibit general similarity and correspond to similar slides of the 3D MR volume. On the other hand, based on the pre-attentive agents only, the platform retrieves cases where an abnormality is present in the same location as in the query image, regardless of the overall similarity.

While it is difficult to assess the effectiveness of the platform in this application (the authors aim to publish a broader report of this work with more results, at a later stage), initial results are very encouraging. Fig. 5b shows initial retrieval results from a 120-image database (taken from [9]), comprising of 70 pathological images (from 7 different cases in [9]) and 50 normal. In 10 retrieval experiments of a 'cancer' query image, the fraction of the 10 retrieved images that are also pathological, is assessed as well as the fraction (again in the 10 first hits) of cases where the pathological ROI is similar (visual assessment of location and shape). Retrieved images of the same case were excluded from the results while the preattentive-attentive tiers were equally weighted. It should be noted that although this is an initial experiment, the results are promising, especially since the images used had different characteristics (T1, T2 and PD MRI scans were used).

3. DISCUSSION

We presented a two-tier, CBIR architecture with special application to brain MRI data. It is important to note that in clinical decision support system design, it is recommended that a CBIR platform should retrieve perceptually 'similar' cases, without suggesting a possible diagnosis to the user (to avoid bias). Our implementation conforms to this requirement while it allows the user to weight the relative effect of 'pre-attentive' vs. 'attentive' retrieval. The semantics were defined for the specific application since it remains hard to define important semantics suited for any application. Decomposing an image to its primitive visual features, and extracting the true underlying semantics is an endeavor that puzzles the cognitive science community for many years. Until new experimental evidence shed light on this process, we have no other means but to keep on defining manually the most promising agents to participate in retrieval.

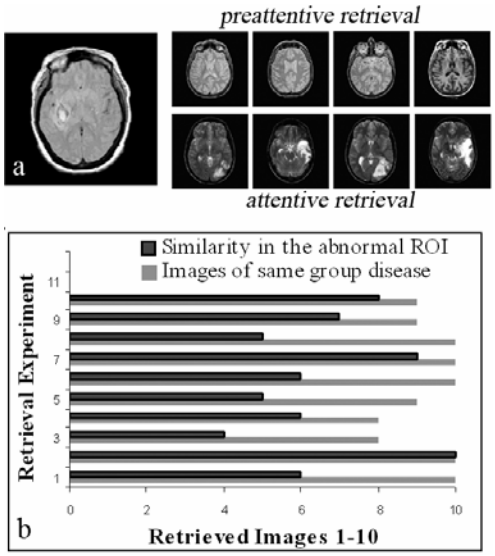


Fig. 5: a) Preattentive and attentive retrieval of a diseased query MRI image, b) Fraction of the 10 first hits that belong to the same group of diseases (cancer) and that exhibit ROI similarity.

5. REFERENCES

[1] D. Comaniciu1, P. Meer, D.J. Foran, "Image-guided decision support system for pathology", *Machine Vision and Applications,* 11: 213–224, 1999

[2] R.C. Veltkamp, H. Burkhardt and H.-P. Kriegel, *State of the Art in Content Based Image and Video Retrieval*, Kluwer Academic Publishers, 2001

[3] A.M. Treisman, and G.A. Gelade, "Feature integration theory of attention", *Cognitive Psychology*, 12, 97-136, 1980

[4] S.E. Palmer and I. Rock, "Rethinking perceptual organization: The role of uniform connectedness", *Psychonomic Bulletin & Review*, 1 (1), 29-55, 1994

[5] Semir Zeki. *A Vision of the Brain*, Blackwell Scientific Publications, 1994

[6] S. Dimitriadis, K. Marias and S.C. Orphanoudakis, "A Versatile Image Retrieval Platform based on a Multiagent Architecture", *6th International Conference on Visual Information Systems*, Miami, 2003

[7] Pattie Maes, "Modeling Adaptive Autonomous Agents", *Artificial Life Journal*, edited by C. Langton, Vol. 1, No. 1 & 2, pp. 135-162, MIT Press, 1994

[8] S. Ullman, *The interpretation of Visual Motion*, MIT Press, Cambridge, MA, 1979

[9] "The Whole Brain Atlas" www.harvard.edu