Chapter 81
Multimodal Indexing and Information Retrieval System Based on Mammographic Image Analysis: Gabor Model in Digital Learning

Sahbi Sidhom
University of Lorraine, France

Noureddine Bourkache
University of Mouloud Mammeri, Algeria

Mourad Laghrouche
University of Mouloud Mammeri, Algeria

ABSTRACT

In this chapter, we propose a new indexing approach on medical "image scanner" databases combining the analysis process of the texture characteristics with the descriptive information. The proposed model is based on the digital image components using the characteristics vector. This vector represents the morphological processing result on image texture. It is linked to image semantic attributes using the annotations of medical professionals. Our application context is based on "Mammographic Image Analysis" (MIAS) in databases. The first aspect concerning the morphology processing on images called the "numerical signature" vector. In this approach, the texture analysis of the image is based on the Gabor Wavelets (or Filters) Theory. In offline processing for each image in MIAS databases, the Gabor Wavelets determine all numerical signatures: image characteristics as multi-index vectors. In online, the query processing by image in real-time defines the query signature (or image-query vectors) and determines all similarities by multi-index matching with images in databases. The similarities are built between the image-query and images in MIAS databases using the same Gabors' algorithms implementation. In order to evaluate the robustness of our system (based on multi-index, semantic attributes, query and information retrieval by image), we experiment with a controlled database of 320 mammographies. The efficient results show a set of successful criteria in image representations based on the Gabor's Wavelets, semantic attributes and significant ratios in the system recall and precision. The objective is to design an intelligent application to assist medical professionals in the decision-making on tumor dignosis based on mammography scanner.

DOI: 10.4018/978-1-5225-0571-6.ch081

INTRODUCTION

The development of imagery electronic devices in various forms has contributed to the creation of trillions of stored images on the internet and new needs have been expressed for image analysis in medical contexts especially in medicine. Actually, the medical field generates significant billions of images for therapeutic interest. In the medical context, images come mainly from x-rays (or scanner), magnetic resonance image (or MRI), ultrasonic imagery, microscopic nuclear medicine imagery (Singh & Mazumdar, 2010),(Zheng, Kuai, Liu, Teng, & Zhang, 2011). This dynamic in information sources comes from practical variability with extreme difficulty to analyze and to exploit images in quantitative, qualitative and objective ways (Sivakamasundari, Kavitha, Natarajan, & Ramakrishnan, 2014),(Candemir et al., 2014).

In this referring works to medical image, a lot of implemented systems were based on text representation of images (Text-Based Image Retrieval: TBIR). For these kind of systems, the image indexing process is based only on text descriptors (as significant words in text, index terms using specific thesaurus in the domain, keywords in descriptions or annotations, user or expert semantic tags, semantic terms in bibliographic records, etc.). Today, this way of image processing is the case of major search engines in the Web. However, it was proved that these processes in image indexing are not sufficient particularly concerning the indexing of significant images in sensitive areas like medicine, mechanical engineering, biometrics, photo-satellites, etc. other than those found in social networks. Especially since our objective is to design an intelligent application to assist medical professionals in decision-making on tumor dignosis from mammography scanner (Oussalah, 2008),(Sinha, Ahmed, & Greenspan, 2014),(Zhang, Zhang, Yao, Pu, Knudson, Weinstock, & Krol, 2005),(Wang, Wang, Li, Liu, Li, & Zhang, 2007). Therefore, it will be a significant contribution to develop other kinds of systems based on digital components and numerical characteristics of images. Since the 90s, scientific communities started implementations of image indexing systems based on query by contents (Content-Based Image Retrieval System: CBIR system) (Akgül, Rubin, Napel, Beaulieu, Greenspan, & Acar, 2011),(Florea, Rogozan, Bensrhair, Dacher, & Darmoni, 2005),(De Oliveriaa, Machadob, & Chaveza, 2010),(Ramos, Kockelkorn, Ramos, Ramos, Ginneken, Viergever, Campilho, & Grutters, 2015), (Rajalakshmi & Minu, 2014).

Today, we find some kinds of CBIR Systems which make the information retrieval from images over the web possible. As examples, we can list *tineye (*www.tineye.com*), cydral (*www.cydral.fr*), gazopa (*www.gazopa.com*),* etc. These image recognition systems using the web of images make it possible to carry out an information image search by URL or by a direct remote loading of an image query stored in a local machine. Several indexing techniques and information image search approaches were elaborated (Harbaoui, Ghenima, & Sidhom, 2009),(Sidhom & David, 2007). However observations in this type of image representation show limits to remain faithful to the analyzed image (Rangayyan, Mudigonda, & Desautels, 2000),(Tong, Xie, & Weng, 2014),(Marino & Tarchi, 2014). Certainly an image file contains a set of digital components characterized by sets of colors, textures and forms (Ravani, Mirali, & Baniasadi, 2010), instead of text representations. For some of these reasons, the indexing systems and image-query (or query by image) engines have been under the study since the 90s (Ravani, Mirali, & Baniasadi, 2010),(Giro-i-Nieto, Ventura, Jordi Pont-Tuset, & Marques, 2010).

We begin the chapter with a review in scientific literature on content-based image retrieval (CBIR), contrasting their respective technologies and applications, and following with related work to present a set of approaches, as the approach using the text semantic attributes and as the approach using the