Expert Systems with Applications 42 (2015) 4927–4936

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Image classification and retrieval using optimized Pulse-Coupled Neural Network

Mona Mahrous Mohammed [a,*], Amr Badr [b], M.B. Abdelhalim [a]

[a] College of Computing and Information Technology (CCIT), Arab Academy of Science and Technology and Maritime Transport (AASTMT), Cairo, Egypt
[b] Department of Computer Science, Faculty of Computers and Information, Cairo University, Cairo, Egypt

ARTICLE INFO

Article history:
Available online 24 February 2015

Keywords:
Content-Based Image Retrieval (CBIR)
Image classification
Visual features
Pulse-Coupled Neural Network (PCNN)
Image signature
K-Nearest Neighbor
Genetic algorithm

ABSTRACT

Content-Based Image Retrieval (CBIR) has become a powerful tool that is used in many image applications and search engines. Thus, many techniques and approaches for CBIR were developed in literature. The CBIR approach works on the visual features of the image rather than a descriptive text. Therefore, it provides more effective and efficient retrieval. On the other hand, PCNN has proved its efficiency as an image processing tool for various tasks such as image segmentation and recognition, feature extraction, edge and object detection. This article introduces a technique for content-based image classification and retrieval using PCNN. The proposed technique uses an optimized Pulse-Coupled Neural Network (PCNN) to extract the visual features of the image in a form of a numeric vector called image signature. An optimization mechanism was applied to the PCNN parameters in order to improve the signature quality. Thus improving the classification and retrieval results. Additionally, it employs the K-Nearest Neighbor (K-NN) algorithm for classification and matching. By applying classification before retrieval, the number of images in the search space is optimized to include one category instead of multiple categories. Moreover, we developed a CBIR prototype to validate our technique. The results show that our technique can retrieve and classify images efficiently. Furthermore, we evaluated our prototype against one of the widely used techniques and it was proven that the proposed technique can enhance the search results and improve the accuracy by 3.5%.

1. Introduction

In recent years, there has been a popularization in the use of digital images in computers and over the internet, especially since the obvious increase of digital image facilities, such as scanners, digital cameras, high quality printers, and so on. The explosion of inexpensive digital equipment and storage devices make users easily own and access huge amount of digital images. More recently the fast growth of the Internet has encouraged the real time processing and retrieval of digital images as one of the most important communication media for daily life.

On the other hand, in contrast with the availability and fast development of hardware, the availability of software for managing large image collections is still limited. The obvious problem is how to allow users to retrieve some images that contain specific contents from a large image database. In general, a common way to search for images is to store them with keywords. These keywords are used to classify them into specific categories, perhaps accompanied by the use of traditional database management systems, which mostly rely on manual operations. However, most likely those similar methods will not support the complex tasks of seeking images. Therefore, motivated by emergent requirements, this research was proposed to explore and develop advanced techniques to overcome these problems.

Image retrieval (IR) is considered one of the most extensive and richest research areas in the field of computer vision. Many techniques and algorithms have been studied and implemented for image search and retrieval (Datta, Joshi, Li, & Wang, 2008). The basic and most popular approach for image retrieval is the annotation-based approach, in which each image associated with a set of captions or keywords that describe the image content. This method is widely used in many commercial applications and especially in online image search engines such as Google, Yahoo, and so on (Yonekawa & Kurokawa, 2012). Annotation-based approach is subjective to the culture, the knowledge and the perspective of the annotator. Therefore, it is difficult to evaluate the accuracy of the annotations that may lead to poor search results.

* Corresponding author.
E-mail addresses: mona.mahrouse@gmail.com (M.M. Mohammed), a.badr.fci@gmail.com (A. Badr), mbakr@ieee.org (M.B. Abdelhalim).

http://dx.doi.org/10.1016/j.eswa.2015.02.019
0957-4174/

The difficulties and drawbacks founded in the annotation-based technique was the motivation to start working on the content of the image itself rather than its description. This new approach called Content-Based Image Retrieval (CBIR). The CBIR analyzes and extracts the visual features from the images such as color, texture, and shape for retrieval.

There are many CBIR systems developed for more than 20 years ago (Chang & Fu, 1980). IBM's QBIC (Query By Image Content) is the start of Content-Based Image Retrieval commercial system (Flickner et al., 1995). In addition to commercial systems, there are many systems developed and introduced by the academia. These systems include for example Candid, Photobook and Netra which all use simple color and texture characteristics to describe the image content (Veltkamp & Tanase, 2000). Furthermore, using higher level of extracted features such as color, texture, location and shape was developed in Blobworld System by Computer Science Division, the University of California, Berkeley (Carson et al., 1999). On the other hand, PicHunter was developed to match an exact image in the database through user feedback (Cox et al., 2000).

Although there are various approaches for developing CBIR, the efficiency of retrieval is strongly dependent on features' representations of an image. This article introduces a new technique for Content-Based Image Retrieval and classification. The proposed technique represents the visual features of an image using simple signature vector generated from PCNN and then classified and matched through K-Nearest Neighbors (K-NN) algorithm. In addition, this technique is optimized using the genetic algorithm. We aim to improve the retrieval performance by applying classification first and then retrieving the matched images from the defined class. Furthermore, we implemented a prototype that validates our technique and proves its ability in classifying and retrieving images.

The rest of the article is organized as follows: Section 2 is surveying the related work. Section 3 is describing the proposed image classification and retrieval technique. Section 3 shows our experimental results while these results are discussed in Section 4. Section 5 concludes the article and introduces possible future work.

2. Related work

The basic technology of developing CBIR is based on the feature extraction (Kumaran & Bhavani, 2012; Xavier & Mary, 2011). The various methods to achieve the CBIR had been proposed in literature (da Silvaa, Falcaeob, & Magalhaesa, 2011; Rao & Rani, 2012). Despite of the variance among the different CBIR techniques, they follow the same procedure. This procedure composes of two phases, one for the feature representation and the other for similarity measure. Agarwal, Verma, and Singh (2013) introduced a technique that uses image texture and shape for retrieval. The proposed technique combines Discrete Wavelet Transform (DWT) and Edge Histogram Descriptor (EHD) feature of MPEG-7 to extract the feature vector. In addition, Choudhary et al. (2014) proposed an integrated approach for image retrieval. The technique integrates color moment (CM) to represent the color feature and local binary pattern (LBP) to represent the texture. Also the research study in (Shrinivasacharya & Sudhamani, 2013) presented a novel method that integrates texture information and modified Block Truncation Coding (BTC) mechanism to represent the color feature vector of an image. Furthermore, a simple approach that uses localized multi-texton histogram is described by Qazi and Farid (2013) in order to define a unique feature vector.

Unlike the mentioned approaches, there are other techniques that gave color advantage over shape. Elumalaivasan et al. (2013) proposed a color based image retrieval technique that represents the color feature vector using RGB, HSV and CIE. Also, authors in Luszczkiewicz-Piatek and Smolka (2015) used Gaussian mixtures model to represent the color information of images. This methodology overcomes the color distortions – caused by lossy compression – by evaluating the color homogeneity using Dijkstra algorithm.

On the other hand, The Pulse Coupled Neural Network has been a powerful processor for most of image processing tasks such as segmentation, feature extraction, recognition, edge and object detection and so on (Kurokawa, Kaneko, & Yonekawa, 2009; Lindblad & Kinser, 2005; Masato & Kurokawa, 2011; Mingliang, 2011; Yonekawa & Kurokawa, 2009). Object identification was one the functions in which PCNN were used and proposed in (Johnson, 1994; Lindblad & Kinser, 2005). Additionally a new method of an automatic parameter adjustment of Pulse Coupled Neural Network was introduced by Yonekawa and Kurokawa (2009) for Image Segmentation. Another study by Kurokawa et al. (2009) is proposing a PCNN with inhibitory connections and which used in color image segmentation. Moreover, a method for pattern recognition using image signature had been stated by Mureşan (2003). This study applied Fourier transforms on an image signature that had been generated from PCNN to provide object recognition. In addition to this study, Tolba, Abdel-Wahab, Aboul-Ela, and Samir (2010) built their work on image signature. They introduced a new technique for generating the image signature using PCNN and using it for Arabic sign language recognition. Likewise, Xiao-hua, Juan-juan, Zhong-hua, and Yang (2014) developed a License plate recognition (LPR) technique by employing PCNN to the gray image segmentation and characters' extraction from the plate.

The most important benefit of image signature is representing image content in a condensed manner. This has a great impact on saving storage and reducing processing cost. The simplest way of generating an image signature from PCNN was proposed by Johnson as the summation of pulsed neurons over time intervals (Johnson & Padgett, 1999). On the other hand, one of the main features of the PCNN that was demonstrated by early experiments was the invariant result of shifted, rotated, scaled, and skewed images (Lindblad & Kinser, 2005). Because of these merits, it was effective to use PCNN for image retrieval and matching.

3. Materials and methods

In this section, we present the PCNN model and the K-NN classifier. In addition, we describe our proposed CBIR technique. The technique operates in two modes: the optimization mode and the retrieval mode. In optimization mode, we adjust the PCNN parameters to enhance the generated signature which will be illustrated in the next subsections. On the other hand, the retrieval mode is responsible for taking an input image, then classifies and searches for similar images in the database.

3.1. Pulse-Coupled Neural Network model

In this section, we demonstrate the proposed PCNN model. Fig. 1 shows that the PCNN is composed of two main components, one for linking and the other one for feeding. In addition, a threshold calculator and step function are employed. The PCNN is working in iterative manner where the previous neuron values affect the next ones.

As described in Eq. (1), the linking component is working on the linking value of the neuron and its output value from previous iteration.

$$L_{ij}(n) = L_{ij}(n-1) * e^{\frac{-1}{AL}} + VL * Y_{ij}(n-1) \quad (1)$$

where:

L_{ij} is the linking value associated with neuron (i, j)