Data-Guided Image Retrieval System under Big Data Environment: Design and Implementation

HUANG Chao[1, a]

[1]College of Information Science and Engineering, Zaozhuang University, Zaozhuang 277160,China

[a] hch1618@126.com

Keywords: Image Retrieval System; Data Mining and Analysis; Image Representation; Big Data Environment; Saliency Map; Graph Learning.

Abstract. In Bag-of-Words based image retrieval, the SIFT visual word has a low discriminative power, so false positive matches occur prevalently. The research on data-guided image retrieval system is a hot topic. Its innovation is using PCNN and ICM in image feature extraction with translation, rotation, scale and distortion invariance and good resistance to noise, the PCNN and ICM extracted features as the image texture feature is applied to image retrieval system. The main idea is to use the PCNN and the ICM process images, get corresponds to different gray level values of binary image sequence, the sequence of the entropy of each image sequence, the one dimensional feature vector as the texture feature; Then using Euclidean distance similarity calculation. The experimental results show that the method not only has strong robustness to noise, at the same time can reduce eigenvector dimension, scale, translation and rotation invariance, and can get higher retrieval rate. In the future, we decide to do more comparison experiment to verify the effectiveness of the method.

Introduction

With the development of computer technology, multimedia technology and Internet technology, people are increasingly exposed to a lot of image information. How to make quickly and efficiently from the massive image database retrieval required image, has become the current retrieval is an important and challenging research subject. Content-based image retrieval (CBIR) technology is to solve how to effectively retrieve the relevant images from image database problems and put forward. In recent years, CBIR has become the focus of widespread attention at home and abroad, and in many fields such as medical image, multimedia is widely used, has made a lot of research results, and in some prototype and business get successful application in the system. With database technology, the rapid development of multimedia technology and network technology, more and more people come into contact with a lot of digital image database, in order to effectively manage image library, and efficient image retrieval system is needed for people. The traditional image retrieval method based on text already cannot satisfy the actual needs. In order to overcome the limitations of traditional retrieval methods, content-based image retrieval technology arises at the historic moment. So content-based image retrieval technique has rapidly become one of the hot research topics in image database technology, and received wide attention and research field of information science and technology personnel at home and abroad [1-5].

Content-based image retrieval is mainly use the image color, texture, shape and spatial location relationship between the underlying characteristics of retrieval. Color is the most direct and the simple feature in the image, it about the size of the image itself, direction, and rotational less dependence, so the use of color features to retrieve image is commonly used in the content-based image retrieval technology of the basic method. But color features is more sensitive to changes in brightness, and color histogram as feature does not contain any color space layout information, and the shape characteristic, as one of the essential feature of the object, it is not over and brightness changes in the environment around the change, is characteristic of stable object, relative to the color and texture feature is more intuitive, and can carry a certain space layout information advantages of shape features itself makes it possible to separate the images of the different target zone, so adopt the

method of combining color and shape features for image retrieval, and make up for the deficiencies of single feature image retrieval. In addition, the current image semantic annotation results failed to reflect the user more easily by the fact that natural language to express the retrieval intention, the existing main retrieval way or use the form of images or keyword, namely by the user to enter a picture or a group of key words, the system to retrieve the relevant images. In order to overcome this limitation, Zhu and others made a meaningful exploration and implements a prototype system and the user can input text to retrieve images. The prototype system through natural language processing techniques to extract the user input the key words in short sentences, and then retrieve the relevant images presented to the user after combination. Although in this way the user can be expressed by a richer semantics to describe to retrieve images, but due to the system by extracting phrase of the keyword to retrieve the image, so the semantic is hurt, or essentially similar to the traditional retrieval based on keyword. We consider the statement text is the most direct in daily life people to express the semantic of a carrier, has the ability to express rich semantic if it can be directly in the text and image characteristics of the underlying semantic space which establish an effective mapping relationship between contained not only can better express the image of semantics, and can be naturally support in the form of text statements for image retrieval, this is a problem worthy of attention and research.

Therefore, we conduct research on data-guided image retrieval system under big data environment. The sparse expression mechanism is introduced to study the correlation between visual features and semantic space. On this basis, the directly uses the way of image retrieval based on text statements, the users expressed in natural language retrieval intention projection into the semantic space, and through the text semantics and the mapping relationship between image features for image retrieval. This is actually will use a form of media representation of semantic conversion to use another form of media to express, so in essence, is a way of cross-media retrieval. The detailed analysis will be conducted in the following parts.

Our Designed Algorithm and System

The Feature Selection Methodology. In CBIR system, image description indexes through its visual characteristics, these visual features mainly including color, texture and shape etc. Due to the characteristics of the image content is difficult to accurately extract and description, the characteristics of the similarity calculation and there was some difference of perception of the human eye, CBIR has many problems to be solved. As one of the most significant characteristics of visual image, the texture is always an important research content in the field of image retrieval; through the visual characteristics of the texture, the user can query in the database to have the same model of image texture. But because of the complexity of the phenomenon of texture and diversity, and the texture information in content-based image retrieval is not very widely used. Color is one of the most notable features of the image, relative to other traits, because it is very stable, i.e. and for rotation, translation and scaling change has very strong robustness, and simple calculation, thus become the most widely used in the existing retrieval system characteristics. In a lot of color space model is commonly used in actual applications which choose what kind of color model had a great influence on the retrieval results. RGB color model is the most basic, the most commonly used color model, digital image is generally the RGB color model, but does not conform to the RGB color model of color perception in the mind, and HSV color model can better conform to the human eye perception characteristics, can be respectively to extract the image at the same time tonal V, H and brightness and saturation S information, so this article use the HSV color space model. On the edge of each pixel corresponds to an edge gradient direction here, edge can be regarded as composed, has the certain direction of edge pixels of edge pixels gradient direction information statistics to better characterize the shape of the object. The formula 1 defines this property.

$$n(i) = \sum_{x,y} \begin{cases} 1, & if A_s f(x,y) \in [i, j] \\ 0, & others \end{cases} \tag{1}$$

In order to make the edge gradient direction histogram of image illumination change changes, translation, scale invariance, with the total number of edge pixels finally get the edge gradient direction histogram normalization. The L value, the greater the angle of quantitative unit number, the more will increase the amount of calculation, but some image edge information in image library. For digital image, each pixel only around eight adjacent points, it will be a plane is divided into eight sectors, so that each point can be used as a gradient in the direction of only eight directions, considering both must include enough information in order to improve the retrieval accuracy, and to make the edge gradient direction histogram feature vector data as little as possible in order to improve the retrieval speed finally choose L=8. The primary reason lies in that feature fusion works better for features with low correlation, such as SIFT and color. In feature matching, complementary information may be of vital importance. For example, given two key-points quantized to the same SIFT visual word, if the coupled color features are largely different, they may be considered to be a false match. The match map is shown in the figure 1. Then, we adopt the basic concepts of deep neural network which is shown in the figure 2.
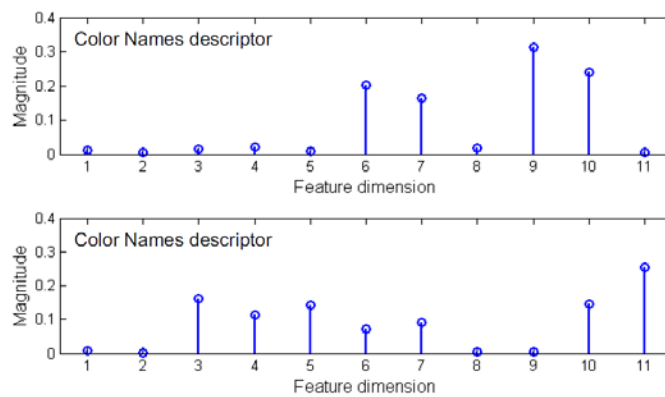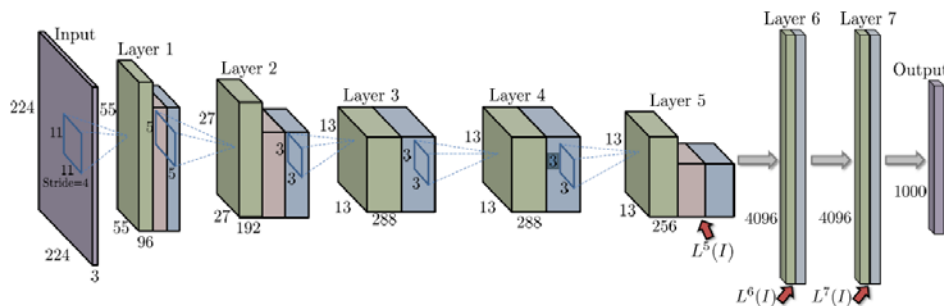
Figure 1.The Feature Description Map.

Figure 2.The Organization of Deep Neural Network

Although CCA has been widely used, and other statistical analysis algorithm, and factors related to CCA algorithm to get is the linear combination of all the characters (variable), the result can be interpreted. This article hopes to learn the training set, from the original high-dimensional feature in the collection to select the most significant correlation with performance feature subset, of linear combination of these characteristics, forming a great semantic relevance of sparse. This not only can make the results can be interpreted, also can eliminate noise in the correlation analysis of the influence of variables, which improve the stability of the model, effectively prevent appeared the situation of the fitting. Directly into the objective function of CCA after adding a penalty term, expanded the Sparse eigenvalue problem solving calculation for solving Sparse CCA, expressed in the music association between semantics and words, so as to extract the most appropriate some words to mark the music. But the algorithm will be two sets of features into a set of feature selection, not two sets of characteristics respectively set up their own sparse expression. The formula 2 shows this.

$$H(P) = -P_1 lb(P_1) - P_0 lb(P_0) \qquad (2)$$

The information entropy in information theory is used to measure the amount of information of a physical quantity. An image can bring a certain amount of information, so information entropy can

also be used to represent the amount of information of image. Digital image is composed of pixels, different grey value of pixels to occupy the space of different, form different distribution, make the image presents different texture feature. Therefore, the information entropy is not the same, different image contains information entropy to describe the texture feature can be used. Of course, the information entropy can also describe color characteristics of the binary image.

Theoretical Analysis and Implementation. Bag image visual word said is based on a code on the basis of this, for all the feature points on each image from a training set to extract feature vector for clustering, to define each cluster center as a visual words, all of the visual word form a code. The images that will be handling to map all of the feature points to visual words are formed on the image of the word bag said. In terms of form and the semantic word bag image representation and text vector model are similar, so naturally you can text classification, lingo righteousness analysis technology used in image processing, etc. At the same time, through the word bag said can make the text and images have a relatively unified expression, in this way, the corresponding relationship between text and related image is similar to the semantic association between the multilingual texts. Based on this consideration, this paper tries to establish image visual words and the mapping relationship between text semantic find intrinsic connection between the text and images, which attempted to retrieve the associated by text image. The framework is shown in the formula 3.

$$\min_{a,b} \left\| Xa - Yb - y_k \right\|^2 + \mu_1 \left\| a \right\|_1 + \mu_2 \left\| b \right\|_1, s.t. \left\| b \right\|_\infty \leq 1 \tag{3}$$

Method, this paper proposes an image retrieval method based on text statement. For a given text statements, can according to this text semantic find related image from the image library. The whole process is composed of two process training and retrieval: like (image, text) training data is a set of binary group, training process using Sparse CCA algorithm, according to the corresponding relationship between the text and images on the training set, study out a semantically related space, images and text semantic space and feature space to the bottom semantically related space transformation; Retrieval can be given to retrieve text into semantically related space, and then to find and retrieve text in semantically related space adjacent points, these points are semantically related to retrieve text image in semantically related space projection. The system model can be roughly divided into storage and query the two stages. In storage stage, the first MPEG standards recommended by the layout, color, color edge histogram and shape descriptor based on region extraction of image base color, texture and shape features and saved to the vision in the library; Then use algorithm to build more class classifier, and draw some pictures from visual feature library as the training sample for training; Finally with trained more than class classifier to classify image library, image semantic labels and belonging to different semantic classes, and deposited in the semantic repository. In the stage of the query, the first is provided by the user query images; System then queries the low-level features of images, and trained more than before class classifier to classify the query image, obtain the query image semantic labels and belonging to different semantic classes; According to the similarity calculation method to calculate the degree of similarity between the query image and the image library, and returns the similarity S not 0 similar to query image the first N picture, if all of S is zero, it returns with the query graph N in front of the low-level features the most similar picture. The distance measure metric is defined in the formula 4.

$$D = \sqrt{\sum_{i=1}^{n} \left(f_i - f_{qi} \right)^2} \tag{4}$$

Content-based image retrieval mainly USES the similarity retrieval method, namely through the similarity between feature vector to determine the similarity between two images, then obtain some image with the most similar to query image. This paper USES color histogram to represent the image color features, edge gradient direction histogram to represent the image shape feature, so the image by comparing the similarity between color and shape similarity measure between the feature vectors.

Experiment and Simulation Result

The Robust Detection. For a basic CBIR system, first of all need to certain pretreatment of image data, including image format conversion, neat image size, and image enhancement and de-noising: then is the extraction of image characteristics, this paper extract image texture feature; Database information including images of the original database and extraction of characteristic database, specific retrieval is based on the data; Image similarity measure is image matching, the use of the characteristics of the image to quantify the similarity of images; Finally according to output the retrieval results, because the content based image retrieval is still in the experimental exploration stage, so use some more research in the study of image database, such as Corel image database, Brodatz texture image library, etc., the focus of the research focuses on the feature extraction, similarity measure, etc. The robust detection result is shown in the figure 3.
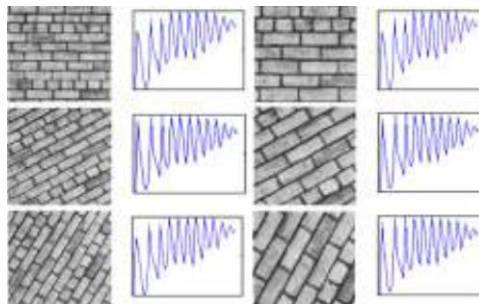

Figure 3.The Robust Detection Result

After standard PCNN image database input or ICM, extraction of information entropy as the texture feature of image, composition standard feature library; Rotation, distorted or add the image of the query image noise, using PCNN or ICM, information entropy as the texture feature extracting of feature library, then follow the standard feature similarity calculation, finally get the retrieval rate. The experimental results show that the method not only has strong robustness to noise, at the same time can reduce eigenvector dimension, scale, translation and rotation invariance, and can get higher retrieval rate; And using ICM than PCNN retrieval rate is higher, and the retrieval speed faster, it is also the advantages of the ICM. Retrieval speed faster, because five equations of PCNN contains three oscillator and two convolution, and only two of the ICM oscillator and a convolution, convolution is the only more time-consuming term in the equation, so the ICM is faster than the PCNN retrieval; Using ICM is PCNN retrieval rate is higher, because ICM is PCNN with no connections between neurons of a special case; Compared with PANN output, the output of the ICM regularity is better, also may be the cause of the parameter Settings. In addition to the main reference the PCNN model ICM, at the same time absorb the advantages of other neural model.

Numerical Analysis and Simulation. In the experiments, the first thing we SIFT features were extracted from training focus all the images, then SIFT feature extraction using k - means clustering method, the clustering center as a visual words. Because we use the Sparse CCA semantically related algorithm learning space, so we created 1 000 visual words, so I can let the Sparse CCA algorithm to study in a larger scope and choose the right visual words. The first set of experiment's result is shown in the figure 4. The others are shown in the figure 5.
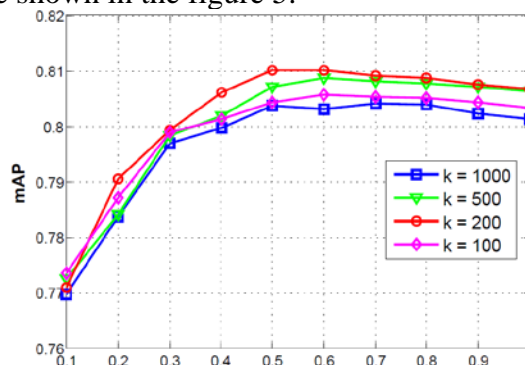

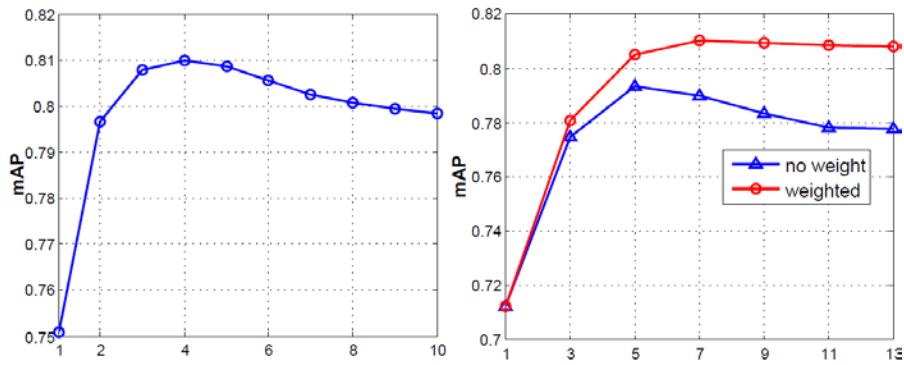Figure 4.The Experimental Result of the First Set of Experiment

Figure 5.The Other Sets of Experiment

Conclusion and Summary

Image retrieval is a broad application prospect and challenging research subject. This paper proposes a texture image retrieval method based on pulse coupled neural networks. Its innovation is using PCNN and ICM in image feature extraction with translation, rotation, scale and distortion invariance and good resistance to noise, the PCNN and ICM extracted features as the image texture feature is applied to image retrieval system. The main idea is to use the PCNN and the ICM process images, get corresponds to different gray level values of binary image sequence, the sequence of the entropy of each image sequence, the one dimensional feature vector as the texture feature; Then using Euclidean distance similarity calculation. The experimental results show that the method not only has strong robustness to noise, at the same time can reduce eigenvector dimension, scale, translation and rotation invariance, and can get higher retrieval rate. This is a lot of content based image retrieval system. In this paper, we design a sparse CCA algorithm, proposes and implements a can carry on the image retrieval method based on text statements, allows users to directly use a small piece of text to describe his retrieval intention, retrieval system can according to the semantic retrieval query text contains to the relevant images. This method is based on Sparse CCA algorithm, the proposed algorithm can learn from the training sample out part of the characteristics of the linear combination of meaningful construct a semantically related space, the query text and images are mapped to the space at the same time, and effectively depict the semantic relation between them. The result is satisfactory.

References

[1] Gong, Yunchao, et al. "Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval." Pattern Analysis and Machine Intelligence, IEEE Transactions on 35.12 (2013): 2916-2929.

[2] Wu, Zhong, et al. "Scalable face image retrieval with identity-based quantization and multireference reranking." Pattern Analysis and Machine Intelligence, IEEE Transactions on 33.10 (2011): 1991-2001.

[3] H. Wang and J. Wang, "An effective image representation method using kernel classification," in Tools with Artificial Intelligence (ICTAI), 2014 IEEE 26th International Conference on, Nov 2014, pp. 853–858.

[4] Kekre, H. B., et al. "Performance evaluation of image retrieval using energy compaction and imagetiling over DCT row mean and DCT column mean." Thinkquest~ 2010. Springer India, 2011. 158-167.

[5] Zhang, Lining, Lipo Wang, and Weisi Lin. "Semisupervised biased maximum margin analysis for interactive image retrieval." Image Processing, IEEE Transactions on 21.4 (2012): 2294-2308.