Literature Survey to Improve Image Retrieval Efficiency by Visual Attention Model

T. Venkata Ramana, K. Venugopal Rao and G. Shyama Chandra Prasad

Abstract Now a day's CBIR is facing several performance issues because of the growth of digital world. To overcome the issues of CBIR, one challenging task is using the simulation of the visual attention model. To implement visual attention model several factors to be considered like similarity measures, Saliency model. Whereas the traditional CBIR focuses on image features. This paper presents analysis of different concepts which are used to improve the image retrieval efficiency. After analyzing it was understood that there exists some gap to concentrate in increasing the effectiveness of image retrievals. In accomplishing the gap we are presenting a kind of scope which improvises the performance issues in image retrievals.

Keywords Visual attention model · Saliency model · CBIR · Similarity measures · Image retrieval

1 Introduction

Digital libraries are growing in rapid pace in increasing the volumes of electronic data day by day, which is due to the electronic gadgets like cell phones, web cameras and cam carders etc. To satisfy the needs of the user an expert system is needed to have the effective retrieval of similar images for the given query image

T.V. Ramana (✉)
JNTU, Hyderabad, India
e-mail: meetramana_12@yahoo.co.in

K.V. Rao
CSE Department, Narayanamma Engineering College, Hyderabad, India
e-mail: kvgrao@gmail.com

G.S.C. Prasad
CSE Department, Matrusri Engineering College, Hyderabad, India
e-mail: gscprasad@gmail.com

S.C. Satapathy et al. (eds.), *Information Systems Design and Intelligent Applications*, Advances in Intelligent Systems and Computing 434,
DOI 10.1007/978-81-322-2752-6_11

[1]. CBIR system is one of such experts systems that highly rely on appropriate extraction of features and similarity measures used for retrieval [2]. The area has gained wide range of attention from researchers to investigate various adopted methodologies, their drawbacks, research scope, etc. [3–11]. This domain became complex because of the diversification of the image contents and also made interesting [2].

The recent development ensures the importance of Visual information retrieval like images and videos, in the meantime so much of research work is going on in implementing in many real world applications such as Biological sciences, medical sciences environmental and health care, in digitally stored libraries and social media such as twitter, LinkedIn, etc. CBIR understands and analyzes the visual content of the images [12]. It represents an image using the renowned visual information like color, texture, shape, etc. [13, 14]. All these features are often referred as basic features of the image, which undergoes lot of variations according to the need and specifications of the image [15–17]. Since the image acquisition varies with respect to illumination, angle of acquisition, depth, etc., it is a challenging task to define a best limited set of features to describe the entire image library.

Similarity measure is another processing stage that defines the performance of the CBIR system [18]. In this stage, the similarity between the query image and the images in the database is determined using the distance between their feature vectors [19]. The database images that exhibit least distance with the input query are identified as required images and they are accessed from the database. Despite numerous distance measures are reported in the literature [4], Euclidean distance and other norm-based distance measures gain popularity due to its simplicity. Nevertheless, these norm-based distance measures often fail due to the global variances and practical dynamics of the image acquisition technology [18]. From the above Kullback divergence is identified as most promising one to rectify the aforesaid problem [20], still it is facing some issues like compatibility with frequent domain representations such as wavelets [18].

The first initiative step has taken by the IBM company towards image retrieval by proposing query-by image content (QBIC), the other name of it is content based image retrieval. Basic features considered in CBIR includes color (distribution of color intensity across image), texture (Homogeneity of visual patterns), shape (boundaries, or the interiors of objects depicted in the image), spatial relations (the relationship or arrangement of low level features in space) or combination of above features were used. In order to add to the image databases every image has to undergo the feature extraction process, so that all low level features are extracted before it gets added to the database. While performing the process of feature extraction, like extracting the image color, texture or shape are extracted in feature extraction stage. If the user wants to retrieve an image, he provides a sample image and applying extraction process and then the similarity measurement engine is accountable in estimating the similarities between the query image and the images present in the database. After completion of the similarity measure, ranking is given by considering the similar features of the given query image. Approaches for feature extraction include histogram and color movement (Niblack et al.), color

histogram (Sahwney et al.), region histogram (Chad carson et al.), Fourier transforms, Gabor filter and statistical methods. There are many sophisticated algorithms are there to which will be used to describe low level image features like color, shape, texture and spatial features approaches. But these algorithms did not gave the satisfaction, clarity and comfort to visualize it.

This is because of not having the basic features of image in describing high level concepts in the users mind such as searching of a image of a baby boy who is laughing a lot. The only way a machine is able to perform automatic extraction is by extracting the low level features that represented by the color, texture, shape and spatial from images with a good degree of efficiency. Till now there are different CBIR systems which are implemented are available, still there exists the problem in retrieving the images by considering the pixel content. It is not solved till now. To have the Semantic image retrieval, to identify the higher level concepts it requires a feedback mechanism with the human intervention. Out of the different methods, the most common method for comparing two images in content-based image retrieval (typically an example image and an image from the database) is using an image distance measure. The functionality of distance measure is to have the comparison of two images to know the similarity of two images in different dimensions like shape, color, texture and others. After applying the distance measure method if the result is a distance of 0, then it means an exact match with the query, with respect to the dimensions that were considered. If the result of distance measure function is greater than 0, then it indicates various degrees of similarities between the images. Search results then can be sorted based on their distance to the queried image. Color histograms are used to know the color similarity in Distance measures. Apart from that segment color proportion by region and by spatial relationship among several color regions are also used.

In all the methods of retrieval two issues are ignored. They are Performance and usability. There is high variability of interpretation in judging the success of an algorithm. To judge the Performance of an algorithm, evaluation can be done through precision and recall but still it is not meeting the requirements of the user. Queries to incorporate complete context are critical to the success of CBIR.

2 Literature Review

(A) Murala et al. [1] have used local tetra patterns (LTrPs) to construct an image indexing and retrieval system. In contrast to the local binary pattern (LBP) and local ternary pattern (LTP), LTrPs determines the first-order derivatives in both horizontal and vertical directions to identify the pixel direction and hence to represent the reference pixel and the neighbors. They have also introduced a methodology to determine the nth order LTrP using $(n - 1)$th order horizontal and vertical derivatives. The experimental results have demonstrated that LTrP based CBIR system outperforms LBP, LTP and local derivative patterns based CBIR systems. Though the methodology is proved for its computational

efficiency and good performance for diverse texture patterns, the feature descriptor is falling under low level category. The extracted features consider only the pixel relationship and not the pixel relevance with the respective image category. As a result, these feature descriptors probably fail to handle different classes of images, but with similar visual patterns.

(B) Su et al. [2] have introduced a CBIR system in which retrieval accuracy is defined by the users feedback. They have addressed that the conventional relevance feedback based CBIR system has entertained iterative process to ensure the relevance between the retrieval results and the query input. To overcome this, they have proposed navigation—pattern based relevance feedback (NRPF) method that substantially minimizes the number of iterations incurred to ensure the retrieval precision. The method has also been facilitated by three query refinement strategies such as Query Point Movement (QPM), Query Reweighting (QR), and Query Expansion (QEX). Performance demonstration is given by them over conventional systems. The relevance feedback model has been entertained well in the recent days. But now a days this is not that much effective because of huge growth of data. However, these semi-automatic retrieval systems do not cope well with practical scenarios, where retrieval efficiency plays the major role rather than the retrieval precision. These methods are insensitive to similarity measures as the relevance feedback plays crucial role. Under the scenario of worst case performance, the methodology tends to move towards manual retrieval process.

(C) Guo et al. [21] have proposed an indexing method based on Error-Diffusion Block Truncation Coding (EDBTC). Here, the image features are extracted by vector quantization (VQ) on color quantized and bit mapped images obtained from EDBTC. To know the image similarities between the images of the query and the images present in the data base, authors had introduced bit pattern features and color histogram.

(D) Guo et al. [22] have exploited low complexity ordered dither block truncation coding (ODBTC) for CBIR system. ODBTC has been used to compress the image as quantizes and bitmap image. Here, the image is represented as color co-occurrence feature (CCF) and bit pattern features (BPF). ODBTC is shown high performance on BTC based image retrieval systems in CBIR. In [21], the adopted scheme (EDBTC) for extracting features such as color histogram feature (CHF) and bit pattern histogram feature (BHF) were promising in the image compression scheme. In the same way [22] has also adopted ODBTC to extract CCF and BPF features. Any way out of the extracted features, 4 are low level, since they consider only the pixel intensity and the structure of pixel neighborhood. Likewise [1], the retrieval efficiency will be reduced when trying to retrieve visually similar images of dissimilar classes. By this, the practical relevance is found to be less while working on with such feature descriptors.

(E) Liu et al. [23] had introduced a computational vision attention model, known as saliency structure model, to support CBIR system. This saliency model is implemented in three steps. Initially at the beginning stage, they have detected

saliency regions using color volume and edge information rather than basic image features. After getting the output of the first step next step is global suppressing of maps has been performed using the energy feature of the gray level co-occurrence. Finally, the last step is used to construct the saliency structure histogram to represent the image. The saliency structure model has been experimentally proved for its performance over bag of visual words (BOW) model and micro-structure descriptor. Even after having precise information with the help of descriptor, similarity measure is considered which can improve the efficiency of image retrieval. However, they have not provided adequate significance to the similarity measure and hence it is quite complex to decide upon the performance accomplishment.

(F) Stella Vetova et al. [24], had proposed a novel algorithm which is used for extracting the image features by using Dual-Tree Complex Wavelet Transform (DT CWT) in the CBIR system. In their experiment they have shown the results of algorithm which satisfies the conditions of feature extraction rate, feature vector length and high information concentration is necessary for CBIR system. Authors has discussed Dual-Tree Complex Wavelet Transform as a filter bank (FB) structure, running process, conditions for shift-invariance and applications. This mechanism is suitable for edge and surface detection in image processing. By using this algorithm authors achieved high accuracy in comparison among the feature sets.

3 Scope of Research

To develop the CBIR system which has to cope up with the human visual system, a lot more has to done with visual attention model. To implement the human visual system a simulation model is needed which is termed as computational visual model. The high level challenges reside in this area has attracted the world wide researchers to work on it. The existing global features are not promising to build the visual attention model. For instance, the feature descriptors [1, 21, 22] are still under the category of basic features and hence they does not represent the image in the human perception. Semi-automatic methodologies can rectify the issue, but it may tend to be manual retrieval under worst case scenarios [2].

Recently, saliency model is found to be suitable for image representation [25, 26]. However, constructing visual attention models are still challenging [23]. Despite successful development has been made in [23] to construct the model as per the requirement, there is no adequate significance given for exploiting suitable similarity measure. These are the primary research gaps to be considered in my proposal for further development of an effective CBIR system.

Advantages of the Proposal: In the proposed CBIR system, three major components play key roles. They are Saliency model, higher order statistics and decision rules library. In Saliency model, we will be giving an input image which has to

be subsampled for extracting the visual features of the image like color, shape, intensity respectively. These features are used in fine tuning of images by using some of the decision rules, so that visually similar images are extracted. While implementing this mechanism we can check which rules and image features are to be considered so that we can achieve highest relevant image for the given query. Upon this different other methods like similarity measures are applied to have the effective image retrieval.

4 Conclusions

In order to fulfill the gap between CBIR and Visual attention model, each should be implemented then the performance of both should be investigated. To know the performance of each comparison should be made for each of the methodology and applying some other techniques like similarity measures and even we can apply saliency model to know the efficiency. The performance of the methodology will be quantified using renowned performance metrics such as efficiency, precision, recall and F-scores. Even we can use MATLAB to implement these methodologies.

References

1. Murala, S.; Maheshwari, R.P.; Balasubramanian, R., "Local Tetra Patterns: A New Feature Descriptor for Content-Based Image Retrieval", IEEE Transactions on Image Processing, Vol. 21, No. 5, pp. 2874–2886, 2012.
2. Ja-Hwung Su; Wei-Jyun Huang; Yu, P.S.; Tseng, V.S., "Efficient Relevance Feedback for Content-Based Image Retrieval by Mining User Navigation Patterns" IEEE Transactions on Knowledge and Data Engineering, Vol. 23, No. 3, pp. 360–372, 2011.
3. Y. Ruiand, T.S. Huang, "Imageretrieval: Current techniques, promising directions and open issues," J. Visual Commun. Image Represent", vol. 10, no. 1, pp. 39–62, Mar. 1999.
4. A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 12, pp. 1349–1380, Dec. 2000.
5. M. Kokare, B.N. Chatterji, and P.K. Biswas, "A survey on current content based image retrieval methods," IETE J. Res., vol. 48, no. 3&4, pp. 261–271, 2002.
6. Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics," Pattern Recogn., vol. 40, no. 1, pp. 262–282, Jan. 2007.
7. G. Qiu, "Color Image Indexing Using BTC," IEEE Trans. Image Processing, Vol. 12, No. 1, Jan. 2003.
8. C. H. Lin, R. T. Chen, and Y. K. Chan, "A smart content-based image retrieval system based on color and texture feature," Image and Vision Computing, vol. 27, no. 6, pp. 658–665, May 2009.
9. N. Jhanwar, S. Chaudhuri, G. Seetharaman and B. Zavidovique, "Content based image retrieval using motif co-occurrence matrix," Image and Vision Computing, vol. 22, pp. 1211–1220, Dec. 2004.
10. P. W. Huang and S. K. Dai, "Image retrieval by texture similarity," Pattern Recognition, vol. 36, no. 3, pp. 665–679, Mar. 2003.

11. C. C. Lai, and Y. C. Chen, "A user-oriented image retrieval system based on interactive genetic algorithm," IEEE Trans. Inst. Meas., vol. 60, no. 10, October 2011.
12. K. Seetharaman, "Image retrieval based on micro-level spatial structure features and content analysis using Full Range Gaussian Markov Random Field model", Engineering Applications of Artificial Intelligence, Vol. 40, pp. 103–116, 2015.
13. K. Vu, K.A. Hua, and N. Jiang,"Improving Image Retrieval Effectiveness in Query-by-Example Environment," Proc. 2003 ACM Symp. Applied Computing, pp. 774–781, 2003.
14. W. Xingyuan and W. Zongyu, "The method for image retrieval based on multi-factors correlation utilizing block truncation coding," Pattern Recog., vol. 47, no. 10, pp. 3293–3303, 2014.
15. M. Unser, "Texture classification by wavelet packet signatures," IEEE Trans. Pattern Anal. Mach. Intell., vol. 15, no. 11, pp. 1186–1191, Nov. 1993.
16. B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," IEEE Trans. Pattern Anal. Mach. Intell., vol. 18, no. 8, pp. 837–842, Aug. 1996.
17. M. Kokare, P. K. Biswas, and B. N. Chatterji, "Texture image retrieval using rotated wavelet filters," Pattern Recogn. Lett., vol. 28, no. 10, pp. 1240–1249, Jul. 2007.
18. Missaoui, R.; Sarifuddin, M.; Vaillancourt, J., "Similarity measures for efficient content-based image retrieval", IEE Proceedings—Vision, Image and Signal Processing, Vol. 152, No. 6, pp. 875–887, 2005.
19. Hua Yuan; Xiao-Ping Zhang, "Statistical Modeling in the Wavelet Domain for Compact Feature Extraction and Similarity Measure of Images", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 20, No. 3, pp. 439–445, 2010.
20. M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback–Leibler distance," IEEE Trans. Image Process., vol. 11, no. 2, pp. 146–158, Feb. 2002.
21. Jing-Ming Guo; Prasetyo, H.; Jen-Ho Chen "Content-Based Image Retrieval Using Error Diffusion Block Truncation Coding Features", IEEE Transactions on Circuits and Systems for Video Technology, vol. 25, No. 3, pp. 466–481, 2015.
22. Jing-Ming Guo; Prasetyo, H., "Content-Based Image Retrieval Using Features Extracted From Halftoning-Based Block Truncation Coding", IEEE Transactions on Image Processing, Vol. 24, No. 3, pp. 1010–1024, 2015.
23. Guang-Hai Liu, Jing-Yu Yang, ZuoYong Li, "Content-based image retrieval using computational visual attention model", Pattern Recognition, In Press, Corrected Proof, 2015.
24. Stella vetova, Ivan ivanov., " Image Features Extraction Using The Dual-Tree Complex Wavelet transform",- Advances in Applied and Pure Mathematics, 978-960-474-380-3.
25. A. Toet, "Computational versus psychophysical bottom-up image saliency: a comparative evaluation study", IEEE Trans. Pattern Anal. Mach. Intell. 33 (11) (2011) 2131–2146.
26. A. Borji, L. Itti., "State-of-the-art in visual attention modeling", IEEE Trans. Pattern Anal. Mach. Intell. Vol. 35, No. 1, pp. 185–207, 2013.