A POWERFUL AUTOMATED IMAGE INDEXING AND RETRIEVAL TOOL FOR SOCIAL MEDIA Sample

Nemir Ahmed [1], Ameer Morad [1], Furat Nidhal Tawfeeq [2]

[1] Department of Mechatronics Engineering, Al Khwarizmi College of Engineering
[2] Information Engineering Iraqi National Cancer Research CenterUniversity of Baghdad, Iraq

---***---

Abstract - *The Internet image retrieval is an interesting task that needs efforts from image processing and relationship structure analysis. In this paper, has been proposed compressed method when you need to send more than a photo via the internet based on image retrieval. First, face detection is implemented based on local binary patterns. The background is notice based on matching global self-similarities and compared it with the rest of the image backgrounds. The propose algorithm are link the gap between the present image indexing technology, developed in the pixel domain, and the fact that an increasing number of images stored on the computer are previously compressed by JPEG at the source. The similar images are found and send a few images instead of huge images with the same parameters. This provides users in social media with efficient tools to search their images and establish a database or a collection of special images to their interests. Experiments using face detection and background notice based indexing techniques support the idea that the proposed algorithm achieves significantly better results in terms of computing cost. This technology will support control the explosion of media-rich content by offering users an efficient automated image indexing and retrieval tool for compressed images on the Internet.*

***Key Words*: local binary patterns, face detection and recognition, image indexing technology, image retrieval, matching global self-similarities.**

1. INTRODUCTION

The fast developments of social media network technologies have chief to an explosive growth of users in recent years. Therefore, efficient search technologies for sent user images are of great importance. The general media search that heavily depend on the contextual text information, such as titles, and substitute texts on web pages, and content-based multimedia retrieval [1], social media data are commonly associated with user generated tags that define the images, provide meta information (i.e. date, location, etc.). These tags can be used to index the multimedia data to facilitate their search [2]. However, recognizing individuals from Meta information is inflexible problem, particularly when the images are same or wrong Meta information. Mainly, tags annotated to an image may not essentially describe its visual content. For example, tags assigned to images may describe the time (e.g., 2012) and location (e.g., Asia, India) where the photos are taken. Furthermore, even the preparation by manually labelling and tags images, to enhanced search system becomes burdensome. In some individual cases in social media they need to send images for same person with different location with compressed form (without redundant). This paper argues that by designing automatic search system in personal pc to recognize the face and background. This difficulty is exacerbated in large online photo collections in which hundreds of millions of individuals might appear; the difference in appearance between individuals becomes very small relative to the appearance variation of any particular individual. In the present day, efficient content-based search of images over World Wide Web is becoming ever more desirable due to the needs motivated by various applications. In recent years, there has been a huge interest in image content-based indexing in the research community and commercial applications. The first system to attract considerable attention was the Query by Image Content (QBIC) system from IBM Almaden Research Center [3]. This system, which has continued to evolve, allowed the user to find images similar to a given example image using low-level cues such as color and texture similarity. Many research systems approach image retrieval by analysing images in terms of visual properties such as color and texture. However, results of usability studies call into question the usefulness of image searching according to low-level visual properties [4]. Face detection is an important step to all image analysis algorithms. Various approaches to face detection are discussed based on local binary patterns [5], fuzzy transform, wavelet transform [6], nonsubsampled contourlet transform [7-8] and machine learning [9]. The goal of face detection is to calculate the existence of faces in the image and extent of each face. K. Rodden, et al [4], performed a features global image properties and spatial layout of image regions to determine the similarity. Schettini et al., 2001 [10], used the color properties of an image to characterize by the probability distribution of the colors in the image. These images distribute nearly none of the usual photometric properties such as colour, texture or edges and similarities in shape. The capability to equivalent these standard shapes can be considered an main sub-task for object class recognition [11]. The difficulty of matching such shapes is addressed by the descriptor of Shechtman and Irani [12]. However, their work concentrates on matching templates at similar scales and does not address the problems of false positive matches or retrieval in large datasets [13]. Global shape deformations are modelled using the Implicit Shape Model of Leibe et al [14].

Ximena Olivares et al. [15], studied in content- based image retrieval, where the objective is to include the visual characteristics of an image into the search process. Using the query by image content (QBIC) search paradigm similar images are retrieved for a given sample image by extracting visual features from all the images in the collection. The narrow internal layouts of self-similarities are shared by these images E.g., people wear the same clothes in consecutive frames and same backgrounds.

In this paper we propose a method that powerful automated image indexing to enhance the retrieval performance of keyword-based queries. Using notes users can highlight a certain region in the photo and associate a tag (label) with the region for example face or background. The proposed algorithms are based on local binary patterns to detect the face and matching global self-similarities to detect the background. The propose algorithm are link the gap between the present image indexing technology, developed in the pixel domain. The results on retrieval experiments show considerable improvements when you need to send more than a photo via the internet based on image retrieval.

2 FACE DESCRIPTION WITH LOCAL BINARY PATTERNS

Face detection and recognition has been considered for decades in the computer vision literature [16-17]. The innovative LBP operator, introduced by Ojala et al. [18], is a powerful means of texture description. The operator labels the pixels of an image by thresholding the 3x3-neighbourhood of each pixel with the center value and considering the result as a binary number. After that the histogram of the labels can be used as a texture descriptor. Figure 1 shows the illustration of the basic LBP operator. The original LBP operator was defined to only deal with the spatial information. LBP operator was improved by using neighborhoods of different sizes [19]. Later, it was extended to a spatiotemporal representation for dynamic texture analysis by using a circular neighborhood. For neighborhoods we will use the notation (P,R) which means P sampling points on a circle of radius of R. See Figure 2 for an example of the circular (8,1) neighborhood. Another extension to the original operator uses so called uniform patterns [19]. A Local Binary Pattern is called uniform if it contains at most two bitwise.

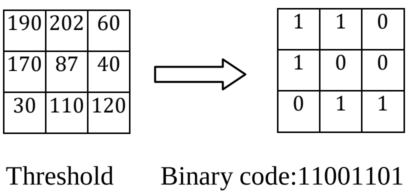

Fig. 1: The basic LBP operator.

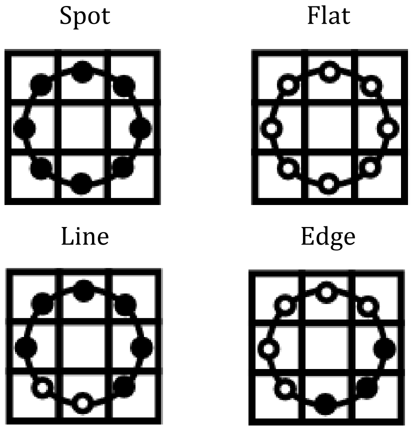

Figure 2. The circular (8,1) neighbor hood.

We derive the operator for a general case based on a circularly symmetric neighbour set of P members on a circle of radius R, denoting the operator as $LBP_{P,R}^{riu2}$. The value of the center pixel (gc) from the values of the circularly symmetric neighbourhood gp (p=0,...,P-1) by assigning a binomial factor 2p for each signs (gp – g0) giving:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_0)\, 2^P \tag{1}$$

After the LBP pattern of each pixel is identified, a histogram is built to represent the texture image:

$$H_i = \sum\nolimits_{x,y} I\big(LBP_{P,R}(x,y), k\big)\,, k \in [0, K], \tag{2}$$

where n is the number of different labels produced by the LBP operator and

$$I\{x,y\} = \begin{cases} 1, & x = y \\ 0, & otherwisw \end{cases} \tag{3}$$

This histogram contains information about the distribution of the local micro-patterns, such as edges, spots and flat areas, over the whole image (figure 2). For efficient face representation, one should retain also spatial information. For this purpose, the image is divided into regions $R_0, R_1, \ldots R_{m-1}$ and the spatially enhanced histogram is defined as:

$$H_{i,j} = \sum\nolimits_{x,y} I\{f_l(x,y) = i\}\,, I\{(x,y) \in R_j\} i = 0, \ldots, n-1, j = 0, \ldots, m-1 \tag{4}$$

In this histogram, we successfully have a description of the face on three different levels of locality: the labels for the histogram contain information about the patterns on a pixel-level, the labels are summed over a small region to produce information on a regional level and the regional histograms are concatenated to build a global description of the face.

From the pattern classification point of view, a usual problem in face recognition is having a excess of classes and only a few, possibly only one, training samples per class. For this reason, more sophisticated classifiers are not needed but a nearest-neighbour classifier is used.

When the image has been divided into regions, it can be expected that some of the regions contain more useful information than others in terms of distinguishing between people. For example, eyes seem to be an important sign in human face recognition [20]. In this study, a test sample S was assigned to the class of the model M that maximized the log likelihood statistic:

$$L(S,M) = \sum_{b=1}^{B} S_b \log M_b \tag{5}$$

where B is the number of bins, and Sb and Mb correspond to the sample and model probabilities at bin b, respectively. To take advantage of this, a weight can be set for each region based on the importance of the information characterizes the spatial structure of the local image texture Chi square statistic (χ2) becomes:

$$\chi^2_w(S,M) = \sum_{i,j} w_j \frac{(S_{i,j} - M_{i,j})^2}{S_{i,j} + M_{i,j}} \tag{6}$$

which wj is the weight for region j.

3 GLOBAL SELF-SIMILARITIES DESCRIPTOR REVIEW

Local Self-Similarity as proposed by Shechtman et al. [12] captures self-similarities surrounded by relatively small (40x40 pixel) regions. The local binary patterns (LBP) is one of the greatest carrying out texture descriptors and it has been commonly used in various applications[21]. The Global Self-Similarity proposed by [22] which can be computed by directly extending LSS which is much faster to compute and uses far less memory. To compute the Global Self-Similarity tensor S_I for image I, we correlate t_p for each pixel $p \in I$ with the entire $H \times W$ image resulting in $H \times W$ different correlation surfaces C_p . S_I is a 4D tensor collecting the C_p's :

$$S_I(p,p') = C_{p(p')} \quad \forall p,p' \in I \tag{7}$$

Determine the $N \times N$ correlation surface C_p of the $w \times w$ patch t_p with the surrounding $N \times N$ region R_p. Both R_p and t_p are centered on p. $C_p(x)$ is the correlation of t_p with a patch t_x centered on x:

$$C_p(x) = \exp\left(-\frac{Self\ Similarity\ Descriptor\ (t_p,t_x)}{\sigma}\right) \tag{8}$$

The suggestion proposed by [12] is to not use the image appearance directly but instead to generate a correlation surface of local self-similarities from intensity patterns across the image. Although not mentioned in [12], an easy

speedup is to compute convolutions using the Fast Fourier

Transform (FFT. To create $\hat{S}_I$ we quantize the patches t_p according to a codebook Θ of prototype patches θ . Then, we define that two patches t_p and $t_{p'}$ are similar if they are assigned to the same prototype θ.

4 PROPOSED METHOD

The collection contains 2500 images, based on a set of tags, which corresponds with the topics that are used for the experiment. As a result we obtained a set of images that at least had one of the tags, (No. of similar Face and No. of similar Background). The next step is indexing and decision and finally sent images without repeating. Figure 3, shows proposed implementation.

5 EXPERIMENTAL RESULTS

Web pages have grown considerably in just the last four years in both the number of requests per page and the size of those requests. Fortunately, many graphics can be cached for future requests. Relative to other content types, static images account for 60% of bytes transferred as shown in Figure 4.

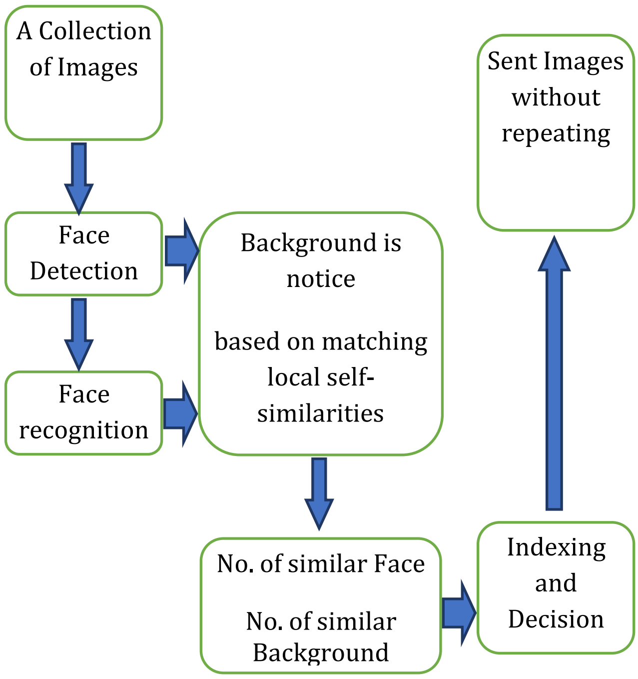

Fig. 3: The proposed implementation.

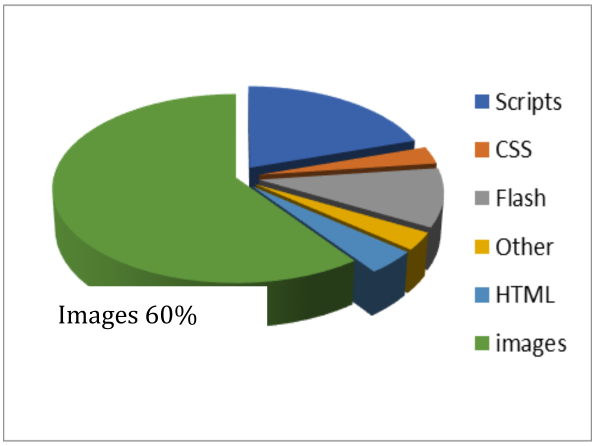

Fig. 4: Average bytes, Images account for 60% of total byte transfer per page, on average.

All of the algorithms are implemented in MATLAB 7.4 and executed on the same computer. The images are registered as described in [23]. In the implementation process, the proposed are used as a filter to kick out images that do not seem promising or redundant so that the number of images to be sent can be greatly reduced. Figure 5 shows the reduction in bytes.

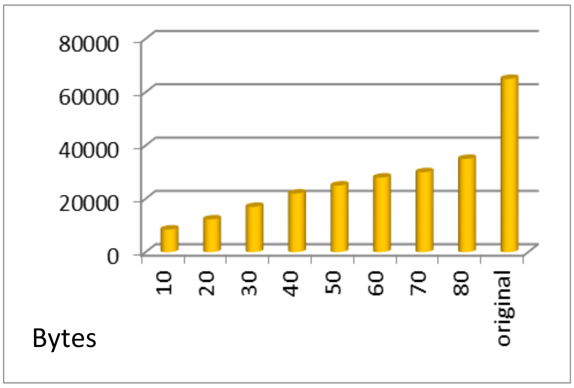

Fig. 5: The percentage number of images to be sent and reduction in bytes.

Along with recognition rates at manual tag, statistical measures (the mean recognition rate with a 95 percent confidence interval) are used to compare the performance of the algorithms. There are several parameters that can be selected to optimize the performance of the LBP-based algorithm. These include choosing the style of the LBP operator, division of the images into regions, selecting the distance measure for the nearest neighbour classifier and finding the weights w_j.

Using a web browser, we design a system which allows a user to control between images. One of the images is a high-quality reference image (original), and the other repeated images as shown in Figure 6. A LBP value is calculated for this center pixel and stored in the output 2D array with the same width and height as the input image. Figure 7 and 8 is an example of computing and visualizing a full LBP 2D array.

A main advantage of LBP implementation is that we can capture extremely fine-grained details in the image. The number of uniform prototypes in a Local Binary Pattern is totally dependent on the number of point's *p*. Before we get started extracting Local Binary Patterns from images and using them for classification, we need to create a temporary dataset of textures.

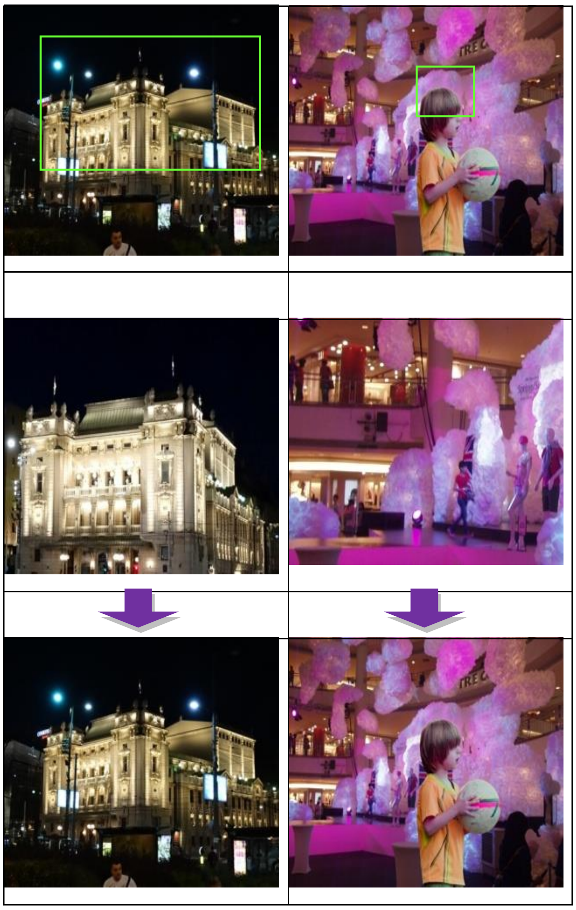

Fig. 6: Sample of reduced multiple image with same sense to single image after face detection and recognition with background indexing techniques.

(a) (b)

Fig. 7: An example of computing the LBP representation *(right)* from the original input image *(left)*.

Original image

Efficient LBP image

Fig. 8: Results of sample of Local Binary Patterns.

Here, we compare LBP and the successful gets feature with different groups of (P, R) in texture analysis. From Table 1 we can see that the accuracy of LBP is providing an appropriate approach in real tasks. The experimental results show that the proposed method yields excellent performance on face verification test and optimum number of images to be sent in social media.

Table -1: Accuracies (mean±std%) of LBP

P, R	(16, 1)	(16, 2)	(8, 1)	(8, 2)
LBP	92.2±0.75	87.9±0.93	91.1±0.64	87.4±0.87

6 CONCLUSIONS

In this paper, we have studied the problem of send more than a photo via the internet (social media) based image retrieval on a various image collection, such as typically found in the personal PC. The results of the retrieval performance experiment clearly showed that the optimum of the results significantly improves when applying proposed procedure based on LBP operator and global self-similarities. Moreover, the results of our overall visual search show a good improvement when compared with the tags only run. When estimate the visual search with a textual filter on the tags we can further bound our search space, and show another major boost in retrieval performance in terms of precision.

REFERENCES

[1] J. Davidson, et al., "The YouTube video recommendation system," 2010, pp. 293-296.

[2] D. Liu, et al., "Semi-automatic tagging of photo albums via exemplar selection and tag inference," IEEE Transactions on Multimedia, vol. 13, pp. 82-91, 2011 2011.

[3] M. Flickner, et al., "Query by image and video content: The QBIC system," Computer, vol. 28, pp. 23-32, 1995.

[4] K. Rodden, et al., "Does organisation by similarity assist image browsing?," in Proceedings of the SIGCHI conference on Human factors in computing systems, 2001, pp. 190-197.

[5] T. Ahonen, et al., "Face description with local binary patterns: Application to face recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, pp. 2037-2041, 2006.

[6] K.-C. Kwak and W. Pedrycz, "Face recognition using fuzzy integral and wavelet decomposition method," IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 34, pp. 1666-1675, 2004.

[7] Y. Cheng, et al., "Robust face recognition based on illumination invariant in nonsubsampled contourlet transform domain," Neurocomputing, vol. 73, pp. 2217-2224, 2010.

[8] N. A. Al-Azzawi, et al., "MR image monomodal registration based on the nonsubsampled contourlet transform and mutual information," in International Conference on Computer Applications & Industrial Electronics, ICCAIE 2010, Kuala Lumpur, Malaysia, 2010, pp. 493-497.

[9] K.-K. Sung and T. Poggio, "Example-based learning for view-based human face detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, pp. 39-51, 1998.

[10] R. Schettini, et al., "A Survey on Methods for Colour Image Indexing and Retrieval in Image Database," in

Color Imaging Science: Exploiting Digital Media, E. R. L. a. L. MacDonald, Ed., 1 ed: John Wiley, 2001.

[11] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol. 60, pp. 91-110, 2004.

[12] E. Shechtman and M. Irani, "Matching local self-similarities across images and videos," in Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on, 2007, pp. 1-8.

[13] K. Chatfield, et al., "Efficient retrieval of deformable shape classes using local self-similarities," in Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on, 2009, pp. 264-271.

[14] B. Leibe, et al., "Combined object categorization and segmentation with an implicit shape model," in Workshop on statistical learning in computer vision, ECCV, 2004, p. 7.

[15] X. Olivares, et al., "Boosting image retrieval through aggregating search results based on visual annotations," in Proceedings of the 16th ACM international conference on Multimedia, 2008, pp. 189-198.

[16] Y. M. Chen and J.-H. Chiang, "Face recognition using combined multiple feature extraction based on Fourier-Mellin approach for single example image per person," Pattern Recognition Letters, vol. In Press, Corrected Proof.

[17] S. Gundimada, et al., "Face recognition in multi-sensor images based on a novel modular feature selection technique," Information Fusion, vol. In Press, Corrected Proof.

[18] T. Ojala, et al., "A comparative study of texture measures with classification based on feature distributions," Pattern Recognition, vol. 29, pp. 51-59, 1996.

[19] T. Ojala, et al., "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, pp. 971-987, 2002.

[20] S. Gong, et al., Dynamic Vision, From Images to Face Recognition. London: Imperial College Press, 2000.

[21] T. Ojala, et al., "A comparative study of texture measures with classification based on featured distributions," Pattern Recognition, vol. 29, pp. 51-59, 1996.

[22] T. Deselaers and V. Ferrari, "Global and efficient self-similarity for object classification and detection," in Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, 2010, pp. 1633-1640.

[23] N. Al-Azzawi, et al., "Fast free-form registration based on kullback- leibler distance for multimodal medical image," in IEEE Symposium on Industrial Electronics and Applications, ISIEA 2010, Penang, Malaysia, 2010, pp. 484-489.