Content-Based Image Retrieval Using Moments

Prashant Srivastava[1], Nguyen Thanh Binh[2], and Ashish Khare[1(✉)]

[1] Department of Electronics and Communication, University of Allahabad, Allahabad, Uttar Pradesh, India
prashant.jk087@gmail.com, ashishkhare@hotmail.com
[2] Faculty of Computer Science and Engineering, Ho Chi Minh City University of Technology, Ho Chi Minh, Vietnam
ntbinh@cse.hcmut.edu.vn

Abstract. Due to the availability of large number of digital images, development of an efficient content-based indexing and retrieval method is required. This paper proposes a moment based image retrieval method. Image is divided into blocks and geometric moment of each block is calculated followed by computation of distance between block moments of query image and database images. Then threshold, using distance values, is applied to retrieve images similar to the query image. Performance of the proposed method is compared with other state-of-the-art image retrieval methods on the basis of results obtained on Corel-1000 database. The comparison shows that the proposed method gives better results in terms of precision and recall as compared to the other image retrieval methods.

Keywords: Content-based image retrieval · Moments · Euclidean distance

1 Introduction

With the advent of numerous digital image libraries, containing huge amount of different types of images, it has become necessary to develop the systems that are capable of performing efficient browsing and retrieval of images. Pure text based images are prevalent but are unable to retrieve visually similar images. Also, it is practically very difficult to annotate manually large number of images. Hence pure text based retrieval is insufficient.

Content-Based Image Retrieval (CBIR) - the retrieval of images on the basis of features present in the image, is an important problem in Computer Vision. Content-based image retrieval, instead of using keywords and text, uses visual features such as colour, texture and shape to search an image from large database [1, 2]. These features form a feature set and act as an indexing scheme to perform search in an image database. These feature sets of query images are compared with database images to retrieve visually similar images.

Early image retrieval systems were based on primitive features such as colour, texture and shape. The field of image retrieval has witnessed substantial work on colour features. Colour is a visible property of an object and is a powerful descriptor.

P.C. Vinh et al. (Eds.): ICCASA 2013, LNICST 128, pp. 228–237, 2014.
DOI: 10.1007/978-3-319-05939-6_23, © Springer International Publishing Switzerland 2014

CBIR systems based on colour feature use conventional colour histogram to perform retrieval. Texture is another feature that has been used extensively for image retrieval. Texture feature represents structural arrangement of a region and describes characteristics such as smoothness, coarseness, roughness of a region. These features are used for classification of images into different categories. Content-based retrieval methods based on shape feature are being used extensively for image retrieval. Shape does not mean shape of whole image but shape of a particular object or a region in the image. Shape features are generally used after segmentation of objects from images unlike colour and texture [3]. Since segmentation is a difficult problem, therefore, shape features have not been exploited much. But, still shape is considered as a powerful descriptor. Moment is a measure of shape of object. In the present work we have proposed a method for image retrieval based on geometric moment. Image is divided into blocks of sizes and geometric moment of each block is computed. Euclidean distance between moments of image blocks of query image and database image is computed. This is followed by computation of threshold value, which is used to find images similar to the query image.

Rest of the paper is organized as follows. Section 2 discusses some of the related work in the field of image retrieval. Section 3 describes fundamentals of image moments and its properties. Section 4 of this paper is concerned with the proposed method. Section 5 discusses experimental results and Sect. 6 concludes the paper.

2 Related Work

Over a past few decades the field of image retrieval has witnessed a number of approaches to improve the performance of image retrieval. Text-based approaches are still in use and almost all web search engines follow this approach. Early CBIR systems were based on colour features. Later on, colour based techniques saw use of colour histograms.

Zhand et al. [4] proposed a region based shape descriptor, namely, Generic Fourier Descriptor (GFD). Two dimensional fourier descriptor was applied on polar raster sampled shape image in order to extract GFD, which was applied on image to determine the shape of the object. Lin et al. [5] proposed a rotation, translation and scale invariant method for shape identification which is also applicable on the objects with modest level of deformation. Yoo et al. [6] proposed the concept of histogram of edge directions, called as edge angles to perform shape based retrieval. Combination of gabor filter and Zernike moments has been proposed in [7]. Gabor filter performs texture extraction while Zernike moment performs shape extraction. This method has been applied for face recognition, fingerprint recognition, shape recognition.

Subrahmanyam et al. [8] proposed two new features, namely Local Tetra Patterns (LTrP) and Directional Local Extrema Pattern (DLEP) [9], based on the concept of Local Binary Pattern (LBP) as features for image retrieval. Liu et al. [10] proposed the concept of Multi-texton Histogram (MTH) which is considered an improvement of Texton Co-occurrence Matrix (TCM) [11]. The concept of MTH works for natural images. The concept of Micro-structure Descriptor (MSD) has been described in [12]. This feature computes local features by identifying colours that have similar edge orientations.

Wang et al. [13] incorporated colour, texture and shape features for image retrieval. Colour feature has been exploited by using fast colour quantization. Texture feature is extracted using filter decomposition and finally, shape feature has been exploited using pseudo-Zernike moments. Li et al. [15] proposed the use of phase and magnitude of Zernike moment, for image retrieval. Pan et al. [19] proposed image segmentation method on the basis of cultural algorithms. Optimal threshold values have been computed and selected to perform segmentation. This approach can be used to perform segmentation of objects from an image and then perform retrieval.

3 Moments in Image Analysis

Image moment is a certain particular weighted average of the image pixels' intensities or a function of such moments, usually chosen to have some attractive property or interpretation. Image moments are useful to describe objects after segmentation. The $(p + q)^{th}$ order geometric moment M_{pq} of a gray-level f(x, y) is defined as

$$M_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)dxdy \tag{1}$$

In discrete cases [14], the integral in the Eq. (1) reduces to summation and Eq. (1) becomes

$$M_{pq} = \sum_{x=1}^{n}\sum_{y=1}^{m} x^p y^q f(x,y) \tag{2}$$

where $n \times m$ is the size of image and f(x, y) denotes gray-level image.

When a complete object has been identified in an image, it can be described by a set of moments $m_{pq.}$ The (p, q) - moment of an object $O \subseteq \mathbb{R}^2$ is given as

$$m_{pq} = \int_{(x,y)} \in_o x^p y^q dxdy \tag{3}$$

In terms of pixels in a binary $[1, n] \times [1, m]$ image f(x, y):

$$m_{pq} = \sum_{x=1}^{n}\sum_{y=1}^{m} x^p y^q f(x,y) \tag{4}$$

where background pixels have value zero, and the object pixels have the value one. The infinite sequence of moments, p, q = 0, 1, 2 … uniquely determines the shape and vice versa.

Simple properties of image which are found via image moments include area, its centroid and information about the orientation.

Moment features are invariant to geometric transformations. Such features are useful to identify objects with unique shapes regardless of their shape, size and orientation. Being invariant under linear coordinate transformations, the moment invariants are useful features in pattern recognition problems.

3.1 Properties of Image Moments

Image moments hold following properties useful for image retrieval-

1. Moment features are invariant to geometric transformations.
2. Moment features provide enough discrimination power to distinguish among objects of different shape.
3. Moment features provide efficient local descriptors for identifying the shape of objects.
4. Infinite sequence of moments uniquely identifies objects.

4 The Proposed Method

The proposed method consists of three steps:

(a) The first step is concerned with division of image into blocks and computation of moments of each block.
(b) In the second step, we compute distance between the block moments of query image and database images.
(c) Threshold is computed to perform retrieval in third step.

The schematic diagram of the proposed method has been shown in Fig. 1.

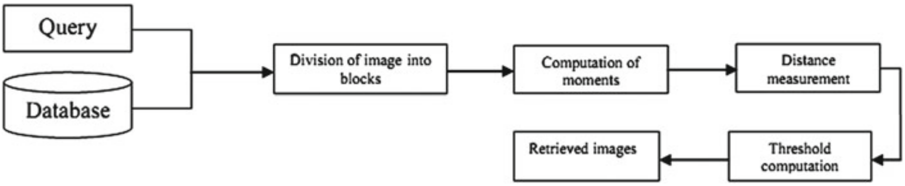

Fig. 1. Schematic diagram of the proposed method

4.1 Computation of Moments

The algorithm for the computation of moments is as follows:

1. Convert the image into gray scale.
2. Rescale the image to 256×256.
3. Divide the image into blocks of different sizes and compute the moments of each block using Eq. (2) taking values of p and q from 0 to 15.
4. Compute the distance between moments of image blocks of query image and database image.

4.2 Distance Measurement

Let the block moments of query image be represented as $m_Q = (m_{Q1}, m_{Q2}, \ldots, m_{Qn})$. Let the moments of blocks of database images be represented as $m_D = (m_{Di1}, m_{Di2}, \ldots, m_{Din})$.

Then the Euclidean distance between block moments of query image and database image is given as

$$D(\mathrm{m_Q}, \mathrm{m_D}) = \sqrt{(\mathrm{m_Q} - \mathrm{m_D})^2} \tag{5}$$

4.3 Computation of Threshold

Threshold is used to perform retrieval. Use of threshold improves the retrieval results as compared to the retrieval result obtained without using threshold. The basic idea behind threshold computation is to find the range of distance values which return images similar to the query image. The Euclidean distance values computed using Eq. (5) are sorted in ascending order so that images are arranged according to similarity to query image. That is, the most similar image first and others after that. The index of similar images is stored along with their distance values to identify minimum and maximum values of range. This determines the range of similarity to a query image. This procedure is repeated for every image of database to find the range of similarity. Finally, the minimum and maximum of all range of values is determined. These values determine threshold of the entire category of similar images. This is done for all categories of images in database. To compute threshold, let

1. *N* be total number of relevant images in database and NDB be total number of images in the database.
2. *sortmat* be the sorted matrix (ascending order) of distance values and minix be first N indices of images in sortmat matrix.
3. *start_range* and *end_range* be the range of relevant images in the database.
4. *maxthreshold* and *minthreshold* are respectively the maximum and minimum distance values of each query image.
5. *mthreshmat* be the maximum of all the values of maxthreshmat.

   Then the algorithm to compute threshold is given below:

```
1.  For u ← 1 to N
    1.1 if (minix(u) >= start_range and minix(u) <= end_range) then
        1.1.1   mthresh ← sortmat(u);
    1.2 endif
2.  endfor
3.  maxthreshmat ← mthresh
4.  mthreshmat ← max(maxthreshmat);
5.  for u ← 1 to NDB
    5.1 if (sortmat(u) >= minthresh and sortmat (u) <= maxthresh) then
        5.1.1 if(ini(u) >= start_range and ini(u) <= end_range) then
              5.1.1.1 freq ← freq + 1;
        5.1.2 endif
    5.2 endif
6.  endfor
7.  end
```

Fig. 2. Sample images from Corel-1000 database

5 Experiment and Results

To perform experiment using the proposed method, images from Corel-1K database [20] have been used. The images in this database have been classified into ten categories, namely, Africans, Beaches, Buildings, Buses, Dinosaurs, Elephants, Flowers, Horses, Mountains, Food. Each image is of size either 256×384 or 384×256. Each category of image consists of 100 images. Each image has been rescaled to 256×256 to ease the computation. Sample images from each category are shown in Fig. 2.

Each image of this database is taken as query image. If the retrieved images belong to the same category as that of the query image, the retrieval is considered to be successful, otherwise the retrieval fails.

5.1 Performance Evaluation

The performance of the proposed method has been measured in terms of precision and recall. Precision is defined as the ratio of total number of relevant images retrieved to the total number of images retrieved. Mathematically, precision can be formulated as

$$P = \frac{I_R}{T_R} \tag{6}$$

where I_R denotes total number of relevant images retrieved and T_R denotes total number of images retrieved. Recall is defined as the ratio of total number of relevant images retrieved to the total number of relevant images in the database. Mathematically, recall can be formulated as

$$R = \frac{I_R}{C_R} \tag{7}$$

where I_R denotes total number of relevant images retrieved and C_R denotes total number of relevant images in the database. In this experiment, $T_R = 10$ and $C_R = 100$.

5.2 Retrieval Results

For the experimentation purpose, each image is divided into blocks of various sizes such as 32 × 32, 64 × 64, 128 × 128 and 256 × 256. Moments of each block have been computed and distance between block moments of query image and database

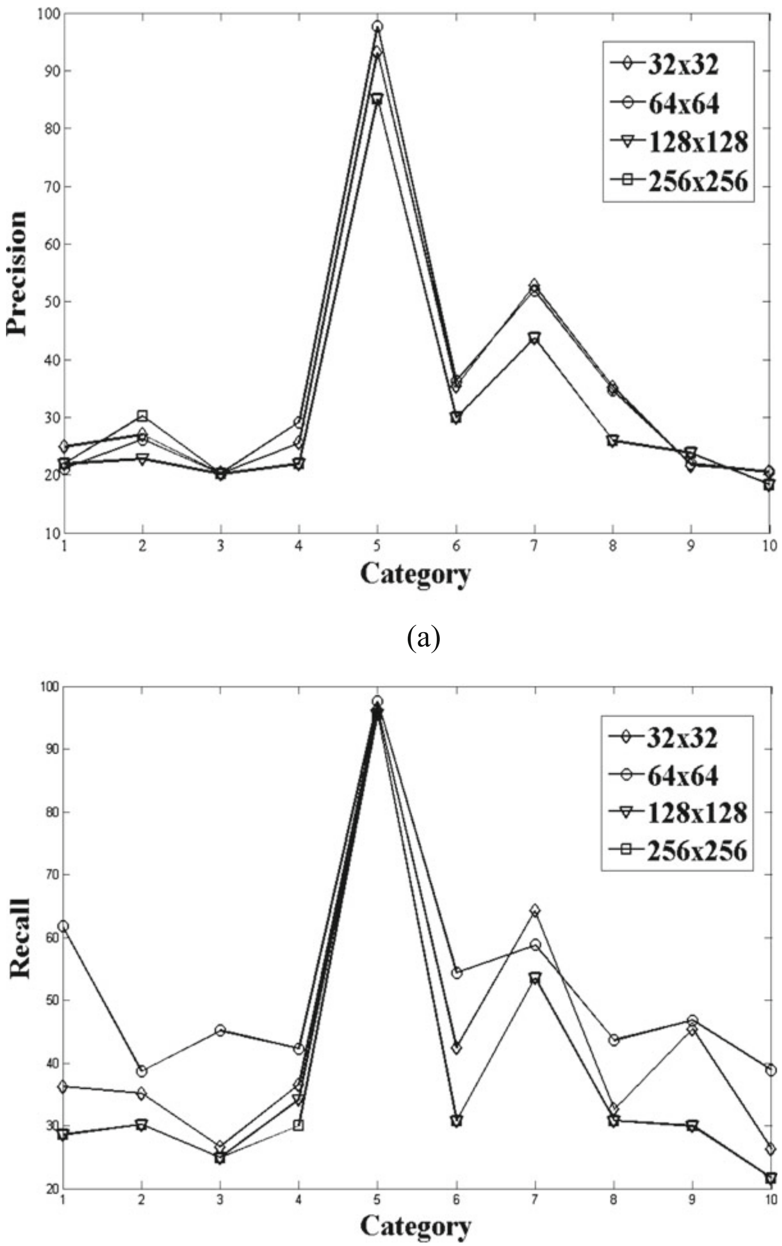

Fig. 3. (a) Precision vs. Category plot (b) Recall vs. Category plot

Table 1. Average precision and recall for each block size

Category	R_{avg} (%)	P_{avg} (%)
Africans	61.76	21.10
Beaches	38.68	26.20
Buildings	45.12	20.20
Buses	42.28	29.10
Dinosaurs	97.66	97.60
Elephants	54.36	36.30
Flowers	58.76	51.90
Horses	43.62	34.60
Mountains	46.78	21.90
Food	38.90	20.50
Average	**52.79**	**35.94**

Table 2. Performance comparison of the proposed method with other methods

	Edge histogram [16]	Gabor vector [18]	Block based LBP [17]	Proposed method
Precision (%)	26.5	23.7	23	**35.94**
Recall (%)	5.3	NA	NA	**52.79**

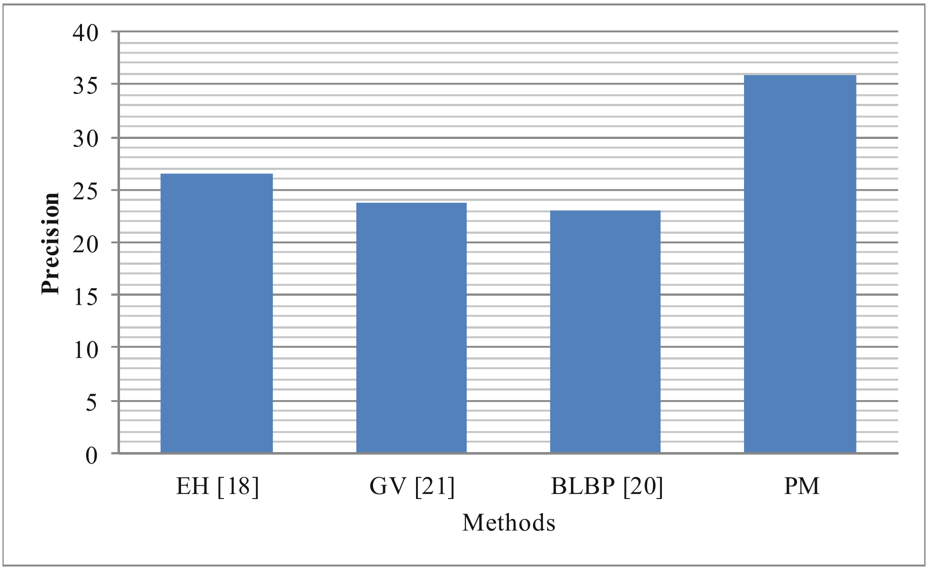

Fig. 4. Comparison of the Proposed Method with other methods

image has been determined. Then retrieval is performed using threshold computed using threshold algorithm. The performance of retrieval has been measured in terms of precision and recall. Precision and recall have been computed for each category.

Retrieval is considered to be good if the values of precision and recall are high. Figure 3 shows the plot between recall and precision and values for different block sizes for different image categories. From Fig. 3 it is observed that the performance of block size 64 × 64 is better than the rest of the block size. Therefore, we have used block size 64 × 64 for comparison of the proposed method with other methods. Table 1 shows the performance of 64 × 64 block size for each category of image in terms of precision and recall.

Table 2 shows the performance comparison of the proposed method with other state-of- the-art methods in terms of precision and recall. Figure 4 shows method vs. precision plot for the proposed method and various other methods [16–18]. From Table 2 and Fig. 4 it is clearly observed that the proposed method outperforms other state-of-the-art methods [16–18].

6 Conclusion

In this paper, we have presented a moment based image retrieval method. Moment based method divides image into blocks of different size and computes geometric moment of each block. The method then computes distance between blocks of query and database images and retrieval is performed on the basis of threshold. Performance of the method was measured in terms of precision and recall. The proposed method outperformed some of the other image retrieval methods such as Edge Histogram, Gabor Vector and Block based LBP. The proposed method took sequence of moments from 0 to 15. Results can be improved by computing moments in more number of sequences at the cost of time. Also, segmentation of object from image and computation of moments for individual objects may produce accurate results.

References

1. Long, H., Zhang, H., Feng, D.D.: Fundamentals of content-based image retrieval. Multimedia Information Retrieval and Management. Springer, Berlin (2003)
2. Rui, Y., Huang, T.S., Chang, S.: Image retrieval: current techniques, promising directions, and open issues. J. Vis. Commun. Image Represent. **10**, 39–62 (1999)
3. Khare, M., Srivastava, R.K., Khare, A.: Moving object segmentation in daubechies complex wavelet domain. Sig. Image Video Process. doi: 10.1007/s11760-013-0496-4, Springer (2013)
4. Zhang, D., Lu, G.: Fourier descriptor for shape-based image retrieval. Sig. Process.-Image Commun. **17**, 825–848 (2002)
5. Lin, H., Kao, Y., Yen, S., Wang, C.: A study of shape-based image retrieval. In: 24th International Conference on Distributed Computing Workshops (2004)
6. Yoo, H., Jang, D., Jung, S., Park, J., Song, K.: Visual information retrieval via content-based approach. J. Pattern Recogn. Soc. **35**, 749–769 (2002)
7. Fu, X., Li, Y., Harrison, R., Belkasim, S.: Content-based image retrieval using gabor-zernike features. In: 18th International Conference on Pattern Recognition, vol. 2, pp. 417–420 (2006)

8. Murala, S., Maheshwari, R.P., Balasubramanian, R.: Local tetra patterns: a new descriptor for content-based image retrieval. IEEE Trans. Image Process. **21**, 2874–2886 (2012)
9. Murala, S., Maheshwari, R.P., Balasubramanian, R.: Directional local extrema patterns: a new descriptor for content-based image retrieval. Int. J. Multimedia Inf. Retrieval **1**, 191–203 (2012)
10. Liu, G., Zhang, L., Hou, Y., Yang, J.: Image retrieval based on multi-texton histogram. Pattern Recogn. **43**, 2380–2389 (2008)
11. Liu, G., Yang, Y.: Image retrieval based on texton co-occurrence matrix. Pattern Recogn. **41**, 3521–3527 (2008)
12. Liu, G., Li, Z., Zhang, L., Xu, Y.: Image retrieval based on microstructure descriptor. Pattern Recogn. (2011)
13. Wang, X., Yu, Y., Yang, H.: An effective image retrieval using color, texture and shape features. Comput. Stan. Interfaces **33**, 59–68 (2011)
14. Flusser, J.: Moment invariants in image analysis. Acad. Sci. Eng. Technol. (2005)
15. Li S., Lee M.C., Pun C.M.: Complex zernike moments shape-based image retrieval. IEEE Trans. Syst. Cybern. Part A: Syst. Humans, **39**, 227–237 (2009)
16. Takala, V., Ahonen, T., Pietikäinen, M.: Block-based methods for image retrieval using local binary patterns. In: Kalviainen, H., Parkkinen, J., Kaarna, A. (eds.) SCIA 2005. LNCS, vol. 3540, pp. 882–891. Springer, Heidelberg (2005)
17. Yuan, X., Yu, J., Qin, Z., Wan, T.: A SIFT-LBP based image retrieval model based on bag-of-words features. In: International Conference on Image Processing, pp. 1061–1064 (2011)
18. Deselaers, T., Keysers, D., Ney, H.: Features for image retrieval: an experimental comparison. Inf. Retrieval **11**, 77–107 (2008)
19. Pan, Z., Chen, Z., Zhang, G.: Threshold segmentation using cultural algorithms for image analysis. In: International Symposium on Photoelectronic Detection and Imaging: Technology and Applications, pp. 662512–662512 (2007)
20. http://wang.ist.psu.edu/docs/related/