# Image Geometric Structure representation using a clustering approach for Content Based Image Retrieval

Ahmad Nishat (조선대학교, 박사과정)
Muhammad Riaz (조선대학교, 석사과정)
Jong An Park (조선대학교, 교수)

Key Words : CBIR; Clustering, Geometric Structure

목 차



## Ⅰ. Introduction and related work

Research in the area of 'Content based image retrieval' (CBIR) and CBIR systems is getting more and more attention due to ever increasing digital multimedia contents and the user friendliness of humans to interact with multimedia instead of text at many levels. CBIR holds tremendous prospects for strengthening the human computer's relationship. There are excellent reviews covering the progress made so far and discussing the open ends for future research [1~3].

The geometric structure of an image exhibits fundamental information. Various approaches have been used in the past to describe the image structure based on shapes or taking local approaches.

In the literature there are various approaches used to use image structure as, Li and Shapiro (2001) [4] used consistent line clusters for building recognition in CBIR but the algorithm does not use the geometrical properties of line segments; rather it uses color-consistent clusters based on color pairs from around line segments. The approach is different from the proposed algorithm as it does not consider the shape making semantic arrangements of line segments.

The paper sequence is organized conventionally, discussing prior related work, proposed idea and its various aspects with a discussion on the experimental results with concluding remarks in the end followed by relevant references consulted while preparing this paper.

## Ⅱ. Proposed idea/algorithm

As a first step we need to find out the line segments. Many approaches have been used in the past for line segment detection such as proposed by Etemadi [5] and Burn's straight lines detector [6]. The Hough transform [7] can be efficiently used to search the straight lines in the images using the parameterized line equation (1).

$$\rho = x\cos\theta + y\sin\theta \quad (1)$$

Each line in the image can be associated with a couple $(\rho, \theta)$ which is unique if $\theta \in [0, \pi]$ and $r \in R$ or if $\theta \in [0, 2\pi]$ and $r \geq 0$. The $(\rho, \theta)$ plane is sometimes referred to as Houghspace. From the Hough space the lines can be found using the inverse Hough transform[8].

For further processing, we have discarded lines by setting a threshold on the line lengths, so that only prominent lines are considered and the rest which mostly provide object details are discarded.

The approach we used is basically a linking approach. For using a linking approach we calculate pair wise distance between the detected line segments angles and lengths. The number of calculations involved is $m(m-1)/2$, where m is the number of detected lines. This way, we construct Euclidean Distance Matrices (EDM) for line segment lengths and angles. By definition the EDM is invariant to rotation and translation.

For grouping the detected line segments we used the single linkage clustering which is an agglomerative scheme. Single linkage defines the distance between any two clusters as the minimum distance between them, i.e. the distance between the two closest points

$$d(k, l) = \min(dist(x_{ki}, y_{lj})) \quad (2)$$

$k \in (1, \ldots, n_i)$, $1 \in (1, \ldots, n_j)$, where $n_i$ and $n_j$ are the number of objects in cluster $k$ and $l$, respectively. $x_{ki}$ denotes the $i-th$ object in cluster $k$ and $y_{lj}$ the $j-th$ object in cluster $l$. This method tends to produce elongated clusters, which is known as chaining effect.

The algorithm uses the above calculated formulae for calculating minimum distance between line segments and generates the hierarchical linkage structure of line segments. The result of the hierarchical clustering is presented in the form of a dendrogram, as shown in Figure 3.

In order to determine the final clusters we have to cut or partition the hierarchy. Two groups of approaches exist, where the first one is the cutting of the dendrogram at a given height, that is, the distance between the nodes in the graph. The second method is to prune the dendrogram by a manual or automatic selection of clusters at various distances. Since the algorithm is based on the minimum line segment distances, we use the node distances to partition the hierarchy to form line segment groups. The line segment groups obtained this way represent some basic level semantic groupings.

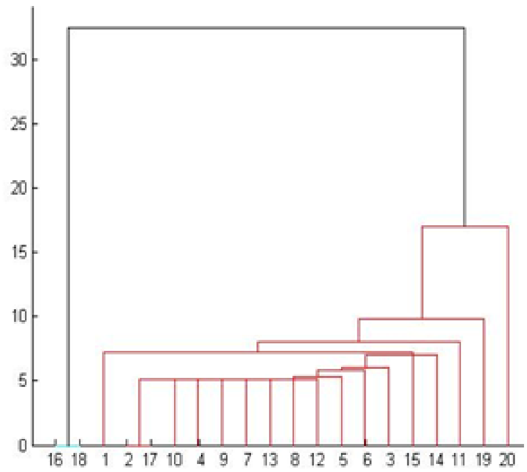


Figure 1. Dendogram of the line segment groups

We draw histograms representing the members count in each cluster. Thus the final feature vector is a histogram giving the number of members in each cluster.

## III. Experiments and results

Content-based image retrieval calculates visual similarities between a query image and images in a database instead of exact matching. Many similarity measures have been developed for image retrieval based on empirical estimates of the distribution of features in recent years. Different similarity/distance measures will affect retrieval performance of an image retrieval system significantly.

In order to evaluate the affect of different similarity measures on the algorithm we used three known similarity measures for histogram matching. As each image contains different number of groups, so the matching algorithm compares each group and then over all similarity between the images is decided in terms of percentage of bins matched. Using two histograms (segment lengths and angles) for one group of line clusters plays a significant disambiguation role, as both must be matched for a similar group.

### 1. Euclidean distance measure

The most commonly used Euclidean distance is given as:

$$d_{eucl}(H, H') = \sqrt{\sum_{i=1}^{n} (H_i - H'_i)^2} \quad (3)$$

### 2. Relative histogram deviation measure

Relative histogram deviation measure gives the deviation between two histograms as:

$$d_{rd}(H, H') = \frac{\sqrt{\sum_{m=1}^{M} (H_m - H'_m)^2}}{\frac{1}{2}\left(\sqrt{\sum_{m=1}^{M} H_m^2} + \sqrt{\sum_{m=1}^{M} H'^2_m}\right)} \quad (4)$$

### 3. Quadratic distance measure

Quadratic Forms are capable of considering the similarities between different bins by incorporating a matrix $A = A_{m,n}$ with $A_{mn}$ denoting the dissimilarity between the bins $m$ and $n$ [9]. Let $H$ and $H'$ be the histograms represented as vectors, the Quadratic form can be calculated as:

$$d_{qd}(H, H') = \sqrt{(H - H')^T . A . (H - H')} \quad (5)$$

### 4 Data set

For testing the proposed idea, we used the Wang dataset [10]. The database is a subset of 1,000 images of the Corel stock photo database which have been manually selected and which forms 10 classes of 100 images each. The 10 classes are used for relevance estimation: given a query image, it is assumed that the user is searching for images from the same class, and therefore the remaining 99 images from the same class are considered relevant and the images from all other classes are considered irrelevant.

### 5. Retrieval results

Since content based image retrieval is all about visual information retrieval, in order to discuss various aspects of experiments carried out, results from the three distance measures used are displayed from the visual class buses. The results displayed are the first 25 results obtained in a random query.

Figure 2. Query Image

Figure 2 displays the query image used for the results shown and discussed in the succeeding paragraphs.

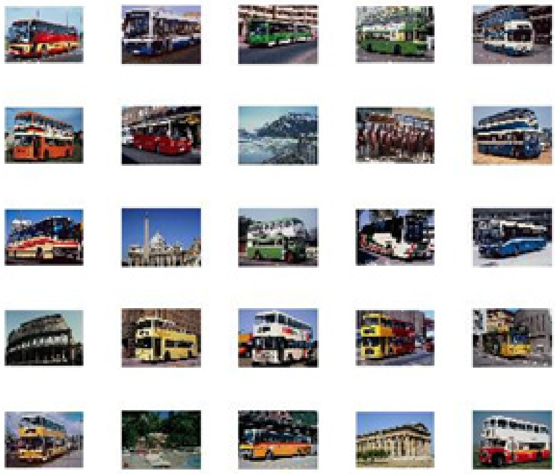

Figure 3. Euclidean distance measure

For figure 3~5, we can see that different distance measures have different performance. However with different databases, the behavior of distance measure changes. One distance measure performing well for one data base may not be giving that much accurate result in another. So the choice of similarity measure still remains an open ended research question.

In figure 3 of Euclidean distance measure, there are 6 false positives with the first false positive at 8, with a total precision of 76% for 25 results.

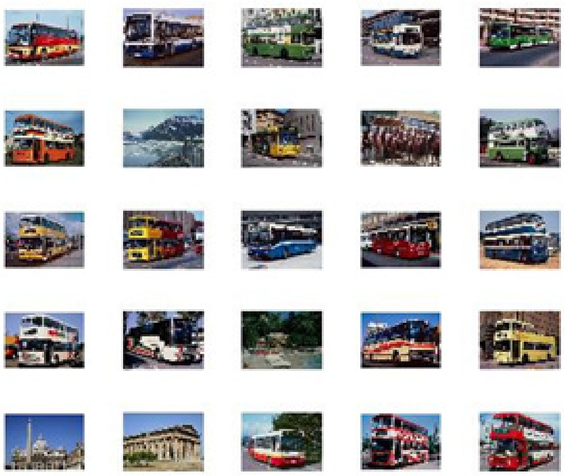

Figure 4. Relative histogram deviation measure

In figure 4 of relative histogram deviation measure, first false positive is at image 7, total of 5 false positives in 25 results, with precision of 80%. While comparing figure 5 and 6, we can see that the first false positive in both is the same image; however it has been assigned different weights by the two distance measures.

Figure 5. Quadratic distance measure

Figure 5 is showing results for the quadratic distance measure, there are total of 9 false positives, giving us a precision of 64%.

## 6. Performance evaluation

The two traditional measures for retrieval performance in the information retrieval literature; precision and recall have been used. To evaluate the overall retrieval performance (precision and recall) of a class, first, the database is queried with each of the images in a visual class, then average precision and recall percentages are computed for the entire class. To rank-order the database images, distance measures discussed above are used.

Figure 6 shows the averaged precision and recall for each class of the database for Euclidean distance measure, Relative histogram deviation measure and Quadratic distance measure being better performer than other measures tested.

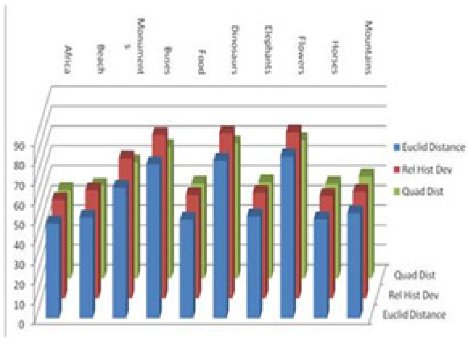


Figure 6. Performance evaluation

Another aspect of ambiguity in defining visual classes has also been highlighted through figure 6. Few generic class are semantically correct but in abstraction very vague. Like for example class food, Africa and even monuments are vague in terms of lower level features. On the other hand classes describing animals like elephant and horse are expected to have a strong background consisting of vegetation due to which the object becomes a non entity in the global context causing low precision. That's why the classes 'Horses' 'Africa' and 'Monuments' have low precision and 'Buses', Dinosaurs have high precision.

## Ⅳ. Conclusion

A new approach for image retrieval based on the concept of semantic line groupings using proximity between line segments has been presented. The proposed feature set is significantly smaller in size and thus comparisons are computationally efficient. The algorithm has been implemented and tested using a publically available database for checking of image retrieval algorithms. The results obtained using the algorithm supplement the idea that line segments when placed at a small distance from each other in certain order form a basic semantic structure, which can be exploited for image retrieval. Furthermore, strengths and weaknesses of similarity matching approaches have been investigated.

## Acknowledgements

This study was supported by Ministry of Culture & tourism and Culture & Content Agency in Republic of Korea.

## References

[1] Arnold W. M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, Ramesh Jain, "Content Based Image Retrieval at the end of the early years", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 22, No. 12, pp 1349-1380, 2000.

[2] Michael Lew, et al., "Content based Multimedia Information Retrieval: State of the Art and Challenges", ACM Transactions on Multimedia Computing, Communications, and Applications, pp. 1-19, 2006.

[3] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of the New Age", ACM Computing Surveys (ACM-CSUR 2008), 2008.

[4] Li, Y. and L. G. Shapiro, "Consistent Line Clusters for Building Recognition in CBIR", Proceedings of the International Conference on Pattern Recognition, pp. 952-956, 2002

[5] Etemadi, "Robust Segmentation of Edge Data", Proceedings of the IEEE International Conference on Image Processing and its Applications, 1992, Volume , Issue , 7-9 Apr 1992 Page(s):311 - 314.

[6] J. B. Burns, A. R. Hanson, and E. M. Riseman, "Extracting straight lines", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, no. 4, pp. 425-455, 1986.

[7] R.O. Duda, P.E. Hart, "Use of the Hough transform to detect lines and curves in pictures", Communications ACM, Vol. 15, pp 11-15, 1972.

[8] Anastasios L. Kesidis & Nikos Papamarkos, "On the Inverse Hough Transform" IEEE trans. Pattern analysis & machine intelligence, vol 21, No. 12 Dec 1999

[9] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, W. Equitz. "Efficient and Effective Querying by Image Content" Journal of Intelligent Information Systems, Vol. 3, No. 3/4, pp. 231 - 262, July 1994.

[10] Jia Li and James Z. Wang, "Automatic Linguistic Indexing of Pictures by a Statistical Modeling Approach", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 9, pp. 1075-1088, 2003. http://wang.ist.psu.edu/docs/related.shtml