SPATIAL NEIGHBORING HISTOGRAM FOR SHAPE-BASED IMAGE RETRIEVAL

Noramiza Hashim [1, 2], Patrice Boursier [1]
[1] *Laboratoire Informatique,Image et Interaction (L3iI), Universite de La Rochelle, 23 Avenue Albert Einstein 17071 La Rochelle Cedex 9, France*

Hong Tat Ewe [2]
[2] *Faculty of Information Technology, Multimedia University, Jalan Multimedia, 63100 Cyberjaya, Selangor, Malaysia*

Keywords: Building recognition, shape histogram, spatial neighboring.

Abstract: Man-made object recognition from ground level image requires a fast and efficient approach especially in a large image database. Our work focuses on recognizing buildings based on a shape-based histogram descriptor. A 2-dimensional histogram is generated from gradient direction information of edge pixels and local spatial analysis of its neighbors. The edge direction histogram is a global representation of edge pixels. The neighborhood structure is coded in a 4-bit binary representation which offers a simple and efficient way to incorporate local spatial data into the histogram. We find that the proposed spatial neighboring histogram increases the retrieval precision by approximately 10% compared to other shape-based histogram methods.

1 INTRODUCTION

Images can be exploited to various purposes. In an urban scene, images of building can be used to get additional information on the building or the surrounding environment. The problem of recognizing buildings from aerial images has been extensively studied throughout the years. Although building recognition from ground level images has not been as widely researched as its counterpart, it has gained more interest from the research communities due to the rapid development of digital imaging around the world.

In content based-image retrieval, one of the earlier approaches in this area may include the use of edge direction based features (Vailaya et al., 1998), which was found to produce optimum result in distinguishing city versus landscape images. Man-made objects in city scenes usually have strong vertical and horizontal edges compared to non-city scenes where the edges are randomly distributed in various directions. This property might be useful to identify building in images. In (Iqbal & Aggarwal, 1999), retrieval by classification was investigated using perceptual grouping to extract structures containing man-made object. These structures include straight and linear lines, junctions, graph and polygons. In Consistent Line Clusters (Li & Shapiro, 2002), the lines extracted from images were grouped into clusters. Inter-cluster and intra-cluster relationships were exploited to recognize complex object, in particular buildings. A hierarchical approach was employed for recognition of building in (Zhang & Kosecka, 2005). A localized color histogram, constructed based on pixels whose orientation complies with main vanishing directions, was used in combination with a keypoint descriptor.

In our work, we employed a shape descriptor for building recognition. The objective is to implement fast and simple methods for recognizing man-made objects. For this reason, histogram-based method was chosen as a basis for the method we developed.

The use of histogram is acknowledged as a powerful tool in image retrieval systems. The edge direction histogram is used in (Vailaya et al., 1998) as a shape-based attributes but it ignores the relationship between edge pixels. To overcome this kind of limitation, (Chalechale & Mertins, 2002) integrates spatial information into a histogram in Edge Pixel Neighboring Histogram or EPNH. Information about the neighbors of edge pixels is

Hashim N., Boursier P. and Tat Ewe H. (2008).
SPATIAL NEIGHBORING HISTOGRAM FOR SHAPE-BASED IMAGE RETRIEVAL.
In *Proceedings of the Third International Conference on Computer Vision Theory and Applications*, pages 256-259
DOI: 10.5220/0001075302560259

obtained in form of codes which are used to produce the neighboring histogram. The correlation between edge pixels was added to the edge direction histogram in Edge Orientation Autocorrelogram or EOAC (Mahmoudi et al, 2003).

Our method involves finding and representing significant pixels in a histogram and associates a local image descriptor to each edge pixel. This descriptor is built based on the gradient direction of the edge pixels and the positions of the surrounding edge pixels. We define the latter as the spatial neighboring property. The analysis of the surrounding pixels is based on the local binary pattern or LBP (Ojala et al, 1999). The LBP operator was developed as a gray-scale invariant texture measure in images. It generates a binary code that describes the local texture pattern. We have adapted the LBP operator to be used with the edge orientation information in order to describe the spatial structure of the local edge pixels. Our method combines both the low level edge feature (i.e. gradient direction) and the middle level edge feature (i.e. spatial information).

The next section contains further description of our method. In section 3, an explanation of the experimental setup and results for image retrieval process is presented followed by discussion and recommendation for future works in the last section.

2 SPATIAL NEIGHBORING HISTOGRAM

The spatial neighboring histogram is a two dimensional histogram comprising the gradient direction in one dimension and the spatial neighboring property in the other dimension.

It is constructed in three stages. The first stage calculates the edge direction using the Canny edge detector. The second stage is the analysis and coding of the neighborhood pixels' pattern. The last stage combines this information to construct the neighboring structure two-dimensional histogram.

2.1 Edge Direction

The directions of edge pixels can capture the general shape information and can be used for discriminating cue especially in the absence of color information (Veltkamp, 2001). We use the Canny edge detector to find the edge map of an image.

The edges are then quantified into a fixed number of bins according to their direction. This constructs the original edge direction histogram or EDH. It is invariant to translation; the positions of the objects in the image have no effect on the edge directions. The histogram is normalized by the total number of edge pixels to achieve scale invariance.

2.2 Spatial Neighboring Property

After the edge direction histogram is constructed, an analysis of the surrounding edge pixels is performed inside a 3 by 3 window. This local analysis will associate the edge pixels to its corresponding spatial neighboring property.

An edge pixel can have zero neighbors (i.e. solitary pixel) and up to 8 neighbors. We define four main directions with respect to the position of the neighbors, which is shown in figure 1. X is the center edge pixel and N1 to N9 are the neighbor edge pixels.

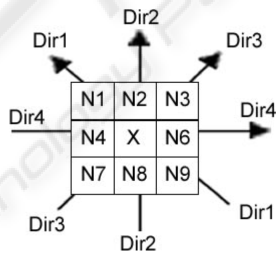

Figure 1: Neighbor pixels and the four main directions.

Each direction is associated with a type of neighbor, named T1, T2, T3 and T4 respective of the four directions. For example, the neighbor pixel N1 and N9 are both in the direction Dir1 and they are the neighbors of type T1.

We consider a type of neighbor as present if there is at least one neighbor pixel (belonging to the type) found. For example, if an edge pixel has 3 neighbors at the position N1, N2 and N8, it will be classified as having the two types of neighbor present i.e. type 1 for N1 and type 2 for N2 and N8.

An edge pixel can have zero type of neighbor present to all the fours types of neighbor present. All the possible combination of present and absent edge pixels creates sixteen distinctive patterns. These combinatorial patterns can be coded in 4 bits with each bit representing a type of neighbor. Thus, the coding of the different patterns becomes a simple binary number representation ranging from [0000] to [1111].

2.3 Histogram Construction

In this last stage, the spatial neighboring histogram is generated. This histogram is a two-dimensional histogram; one dimension represents the edge direction information, in n equally spaced bins and another dimension represents the 16 possible combinations of the neighboring edge pixels. We have chosen n equals 16 bins; therefore, the histogram matrix has 16 rows and 16 columns.

Each $<i,j>$ element of this matrix ($1\leq i \leq 16$ and $1\leq j \leq 16$) represents the number of edge pixels in a direction i with the combinatorial pattern j. An example of the spatial neighboring histogram is shown in figure 2.

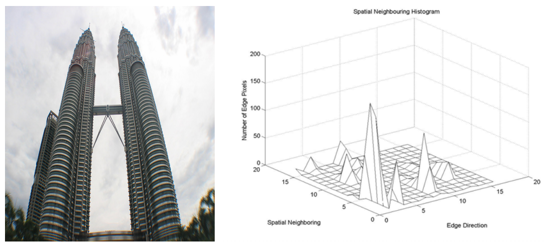

Figure 2: An image of a building and its corresponding Spatial Neighboring Histogram.

3 RETRIEVAL PROCESS

The methodology used for evaluating the performance of our method is presented in this section. The similarity measure, image databases and explanation of our retrieval evaluation are described in subsection 1, 2 and 3 respectively. The last subsection present the retrieval results obtained from our experiment.

3.1 Similarity Measure

For image retrieval, instead of employing exact image matching, a similarity measure is calculated between a query image and images in the database.

The result is a list of images rank by their similarity to the query image. This result depends on the type of distance or similarity measurement used during retrieval process. The Histogram Intersection property is used to calculate the similarity between two histograms.

3.2 Image Database

To evaluate the performance of the retrieval system, we compare the retrieval result of our method against the two alternative methods described in the previous sections applied to the same image database.

The image database used for this experiment contains two different image sets; training set and query set. Each set contains 50 images of 10 classes of building with 5 different acquisitions for each class. The images are chosen such that they feature the standard frontal view of a building taken at ground level. The images also have minimal occlusions and rotations.

For the training set, the histogram matrix of the images is extracted offline and stored in a database. For the test phase, we use the query set. This query set contains the same classes of building but with different acquisitions. The histogram intersection distance will be between histograms of the query image and of each image in the database. The results will then be sort from the closest match to the farthest match.

3.3 Performance Evaluation

The retrieval performance is analyzed in term of retrieval accuracy and the average normalized modified retrieval rank (ANMRR) proposed in MPEG-7 (Manjunath et al, 2002).

The retrieval accuracy concerns two metrics; recall and precision rates. Recall is defined as the proportion of relevant images in the database that are retrieved in response to a query. Precision is defined as the proportion of the retrieved images that are relevant to a query.

The ANMRR combines the precision and recall information as well as the rank information among the retrieved images. ANMRR are in the range of [0,1]. The smaller the ANMRR, the better the retrieval performance is.

3.4 Results

The retrieval accuracy evaluation, we have calculated the average value of precision and recall rates for all 50 query images. We compare the performance of our method against three other similar methods: Edge Direction Histogram (EDH), Edge Orientation Auto-Correlogram (EOAC) and Edge Pixel Neighboring Histogram (EPNH).

The EDH is chosen as an edge orientation based method while EPNH acts as a basis for spatial one dimensional histogram based on LBP. The EOAC is chosen in order to compare our method with another edge-based technique using two-dimensional orientation-based histogram.

Figure 3 shows the precision versus recall plot for comparing the performance of methods. From the plot, it is noted that our method (SNH) outperformed the other three methods with higher precision for the same recall rate. For top rank match, SNH obtains the highest precision rate of 0.86 followed by EOAC, EPNH and EDH with 0.74, 0.72 and 0.70 respectively.

For ANMRR evaluation, we exploit the top 5 retrieved images for a query images and calculate the NMRR. Then, we calculate the ANMRR for all 50 query images. The result obtained is shown in table 1.

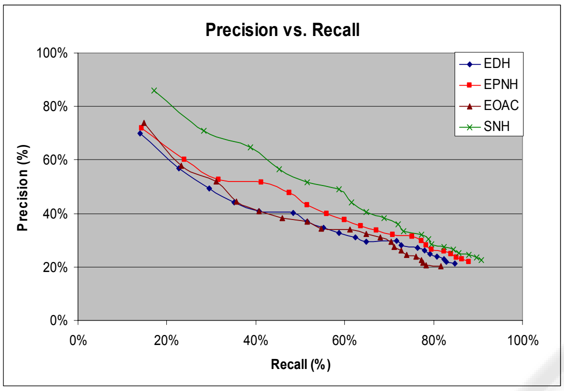

Figure 3: Performance comparison between SNH (our method), EDH, EPNH, and EOAC.

The experimental results have indicated that our method is capable of obtaining an acceptable performance in terms of ANMRR. SNH has the lowest ANMRR followed by EPNH. Lower value of ANMRR shows that our method has better result than the other three methods. The results also show that, for EOAC and EDH, their performance is at approximately the same level. These results have indicated that our method is efficient and capable of producing a good performance.

Table 1: Comparison of ANMRR for SNH, EPNH, EOAC and EDH.

	SNH	EPNH	EOAC	EDH
ANMRR	0.4500	0.4935	0.5630	0.5625

4 CONCLUSIONS

Our study has shown that integrating spatial neighborhood information into a histogram can increase the retrieval system performance. The separate use of the edge and LBP information produces good retrieval result. In our work, we have shown that combining both properties can further improve a system's performance. For images of man-made structure such as buildings, SNH produces better results when compared to other similar methods.

Although our method is simple and straightforward, the experimental results have shown that it is capable of improving the retrieval precision. For future work, further tests with large-scale image database are expected. We also plan to integrate other features to the histogram in order to improve its efficiency.

REFERENCES

Chalechale, A & Mertins, A 2002, 'Semantic Evaluation and Efficiency Comparison of the Edge Pixel Neighboring Histogram in Image Retrieval', *Proceedings of the First Workshop on the Internet, Telecommunication and Signal Processing*, vol. 1, pp. 179–184.

Iqbal, Q & Aggarwal, J.K. 1999, 'Applying perceptual grouping to content-based image retrieval: Building Images', *Proceedings of the IEEE International Conference of Computer Vision and Pattern Recognition*, pp. 42-48.

Li, Y & Shapiro, LG 2002, 'Consistent line clusters for building recognition in CBIR', *Proceedings of the 16th International Conference of Pattern Recognition*, vol.3, pp. 952-956.

Manjunath, BS, Salembier, P & Sikora, T 2002, 'Introduction to MPEG-7', *John Wiley & Sons Ltd.*

Mahmoudi, F, Shahbenzadeh, J, Eftekhari-Moghadam, A-M & Soltanian-Zadeh, H 2003, 'A new non-segmentation shape-based image indexing method', *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol.3, pp. III-17-20

Ojala, T, Pretikainen, M & Harwood, D 1999, 'A Comparative Study of Texture Measures with Classification Based on Feature Distribution', *Pattern Recognition*, vol. 29(1), pp 51-59.

Vailaya, A, Jain, A & Zhang, HJ, 1998, 'On Image Classification: City vs. Landscape', *Proceedings of the IEEE Workshop on Content-Based Access of Image and Video Libraries*, pp. 3 -8.

Veltkamp, RC 2001, 'Shape Matching: Similarity Measures and Algorithms', *Proceedings of Shape Modelling International*, pp 188-197.

Zhang, W & Kosecka, J 2005, 'Localization Based on Building Recognition', *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 3, p. 21.