JOURNAL OF KOREA MULTIMEDIA SOCIETY, VOL. 7, NO. 12, DECEMBER 2004(pp. 1657-1664)

Classification of Man-Made and Natural Object Images in Color Images

Chang-Min Park[†], Kyung-Mo Gu[††], Sung-Young Kim[†††], Min-Hwan Kim[††††]

ABSTRACT

We propose a method that classifies images into two object types man-made and natural objects. A central object is extracted from each image by using central object extraction method[1] before classification. A central object in an images defined as a set of regions that lies around center of the image and has significant color distribution against its surrounding. We define three measures to classify the object images. The first measure is energy of edge direction histogram. The energy is calculated based on the direction of only non-circular edges. The second measure is an energy difference along directions in Gabor filter dictionary. Maximum and minimum energy along directions in Gabor filter dictionary are selected and the energy difference is computed as the ratio of the maximum to the minimum value. The last one is a shape of an object, which is also represented by Gabor filter dictionary. Gabor filter dictionary for the shape of an object differs from the one for the texture in an object in which the former is computed from a binarized object image. Each measure is combined by using majority rule tin which decisions are made by the majority. A test with 600 images shows a classification accuracy of 86%.

Keywords: Image classification, content-based image retrieval, object of interest, object extraction, man-made object, natural object

1. INTRODUCTION

Content-based image retrieval (CBIR) is to find all images in a given database depicting scenes or objects of some specified type by users. In CBIR, images are automatically indexed by low level features such as color, texture, or shape. Although the low level features are easily measured and computed, there is an obvious semantic gap between what user-quires represent based on the features and what the users think.

Automatic classification of scenes into general types such as indoor/outdoor or city/landscape[2-4] tried to extract semantic information directly from images to overcome the semantic gap. A successful indexing of images through the classification greatly enhances the performance of CBIR systems by filtering out irrelevant images and substantially reducing the computational time required to retrieve the images. Szummer *et al.*[3] developed a method for inferring high level image properties such as indoor vs. outdoor by classifying low-level features such as color and texture. Oliva *et al.*[5] classified images into artificial image and natural image through its power spectrum. All of them tried to classify only scenery images, in which object(s) of interest usually do not exist. An object can play an important role in the image retrieval

※ Corresponding Author : Sung-Young Kim, Address : (730-701) 188, Shinpyung-dong, Gumi, Gyeongbuk, Korea, TEL : +82-55-380-9489, FAX : +82-55-380-9249
E-mail : sykim@kumoh.ac.kr
Receipt date : Oct. 14, 2004, Approval date : Dec. 23, 2004
[†] School of Multimedia Engineering YoungSan Univ. (E-mail : cmpark@ysu.ac.kr)
[††] Dept. of Computer Engineering Pusan National Univ. (E-mail : kmkool@pusan.ac.kr)
[†††] School of Computer Engineering Kumoh National Institute of Technology
[††††] Dept. of Computer Engineering Pusan National Univ. (E-mail : mhkim@pusan.ac.kr)
※ This work has been supported by "Research Center for Future Logistics Information Technology" hosted by the Ministry of Education in Korea.

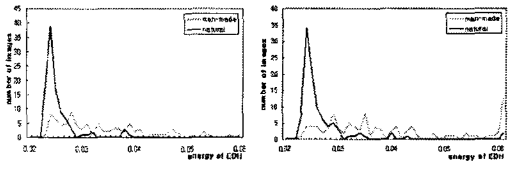
because users generally want to search for the images containing particular *object(s) of interest*. Thus the key to effective CBIR performance lies in the ability to access images at the level of objects. Object / non-object image classification[6] can improve retrieval performance by filtering out images that are classified as another class. The classification can be also utilized for the pre-process of object-based applications, such as the extraction of objects from object class images[1,7] and the classification of object types to improve image retrieval performance[8]. Park *et al.*[8] tried to classify object types into pre-defined 30 classes (for example, dog, frog, car, boat etc) based on texture features from wavelet coefficient and neural network classifier.

In this paper, we propose a method that automatically classifies objects into the man-made and natural object image classes. Classes that we try toclassify in this paper provide more generic information than those in ref.[8]. For example, dog and frog classes can belong to natural class and car and boat to man-made class. So classifying the generic classes requires higher-level knowledge than classifying specific classes because common attributes among some specific classes must be extracted to represent more generic class. To classify an object type, first of all, the objects are extracted from the images[1]. Generally man-made objects in the images usually have many straightened edges, whereas natural objects tend to have edges randomly distributed in various directions. Also man-made objects consist of regularly and repeatedly arranged textures but natural objects have complicated irregular textures. Thus, we defined three measures based on the characteristics of each object class. The first measure is Energy of edge direction-histogram that indicates how direction of edges in the object is regular. The second measure is defined as regularity of textures in the objects. We used Gabor filter[9] to extract texture information in the objects. In Gabor filter dictionary maximum and minimum energy along directions are selected and then the energy difference, which is second measure, is computed as the ratio of the maximum to the minimum value. Sometimes shape feature of an object is not suitable for image classification. Also the shape of the automatically extracted objects is incomplete. Nevertheless, this shape information can be useful in the classification into man-made and natural object because a natural object has significantly different characteristics on the shape from a man-made object. A natural object has a rounded boundary but a man-made object has a straight one. We use Gabor filter dictionary for shape information of an object. The Gabor filter dictionary is computed from the binarized object image.

2. EXTRACTION OF CENTRAL OBJECTS FROM COLOR IMAGES

Before we classify images based on their objects, we extract central objects from images using the central object extraction method[1]. A central object in a color image is defined as a relatively big object that has different color characteristics against its surrounding regions and is located around the center of the image. The center regions are expected to represent contents of the image more efficiently than the boundary regions, because people tend to locate the most interesting object at the center of a picture when they take the picture. A very small object cannot be a central object, because they may be background objects or the scene itself.

To extract a central object, first of all, significant pixels in color and texture are determined by using the difference between the correlogram for the center area of an image and the one for the surrounding region. Then two types of core regions are determined from the segmented image to characterize the foreground and the background in the image. One is the core object region that has a lot of significant pixels and the other is the core background region that is adjacent to the corners or borders of the image. The core object region is extended through merging its unlabeled neighbor regions, if they have similar color distribution to it though they are dissimilar to the core background region in color distribution. The final merging result, a set of regions connected to each other, is the central object.



(a) (b)

Fig. 1. The first image shows EDH-Energy distribution curves of man-made and natural objects in the various types of images. The second image shows the EDH-Energy distribution curves after circular edge pixels are removed.

3. CLASSIFICATION OF NATURAL AND MAN-MADE OBJECTS

3.1 Measures for Classification

3.1.1 Energy of Edge Direction Histogram

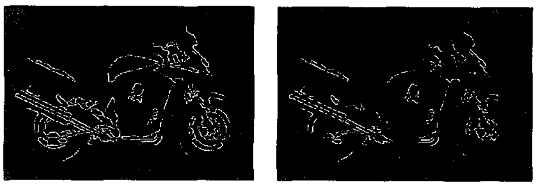
The distribution of edge pixels in a man-made object forms in a straight-line, whereas a natural object does not. In this paper, we represent the distribution of edge pixels in an object with edge direction histogram (EDH) as in Eq. 1 where $H(i)$ and Ne are the count in bin i of the edge direction histogram and total number of edge pixels detected in the object, respectively. The Canny edge detector [10] is used to extract the edges from an image. The EDH, however, is not rotation-invariant, and so the EDH may be changed when the image is rotated. As we are only interested in how many edge pixels belong to the same direction, we use energy of EDH such as in Eq. 2 instead of EDH as the first measure for classification.

$$EDH(i) = H(i) / N_e \quad (1)$$

$$E = \sum_i EDH^2(i) \quad (2)$$

Fig. 1(a) shows the EDH-Energy distribution curves for man-made and natural objects in various types of images. The energies of natural objects tend to concentrate at a lower value. On the other hand, the energies of man-made objects are distributed at larger values than those of natural objects.

On the one hand, sometimes a man-made object includes components with circular edges, for example, tires of a motorcycle. Such circular edges decrease EDH-Energies for the man-made objects. Consequently, discrimination ability between the man-made and the natural objects gets lower. To eliminate the circular edges, we label each edge segment in an edge image, which is determined by using 8 neighbored connected component labeling.



(a) (b)

Fig. 2. Energy of the edge direction histogram can be increased by eliminating circular edge pixels. Left and right images show edge pixels before and after elimination of the edge pixels, respectively. The energy of left and right image is 0.026 and 0.037, respectively.

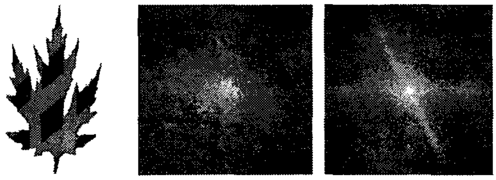
Then we determine circular edge pixels from the edge segments by using Hough transform[11]. Each edge segment is treated independently to prevent the disconnected and independent edge pixels from creating a circle. When the number of circular edge pixels is more than 25% of the circumference, the circular edges will be eliminated. Fig. 2 shows the increased the energy after elimination of circular edge pixels. Also, as we can see in Fig. 1(b), the energy in the man-made objects with extremely low energy is shifted to higher values.

3.1.2 Energy Difference Ratio along Directions in Gabor Filter

While textures in man-made objects show repetitive regularity, texture in natural objects gives complicated irregularity. Man-made objects have some strong energy along the specific angles. In contrast, natural objects have randomly distributed energy. Thus energy difference along directions can be used as a measure of distinguish man-made objects from natural objects. In this paper, we use the Gabor filter dictionary[9] to compute the energy difference along directions. The texture feature with Gabor filter is represented as $t(Oi, Sj)$ where Oi and Sj indicate i-th orientation and j-th scale, respectively. The energy difference ratio (EDR) is defined as in Eq. 3 where M and N are the number of scales and orientation that are used in the filter dictionary, respectively.

$$
\begin{aligned}
&EDR = E_{O_{\max}} / E_{O_{\min}} \\
&E_{O_{\max}} = \max(E_{O_0}, E_{O_1}, \cdots, E_{O_N}), \\
&E_{O_{min}} = \min(E_{O_0}, E_{O_1}, \cdots, E_{O_N}) \\
&E_{O_r} = \left\{ \sum_{i=0}^{M} t(O_r, S_i) \right\}^2
\end{aligned}
\tag{3}
$$

On the one hand, we try to classify objects that are separated from background. After extraction of objects, background is filled with constant color. So, there exist very clear boundaries between the objects and their background. This strong edginess may change the Fourier spectrum of the objects. We smooth the boundaries between objects and their background. As we can see in Fig. 3(b), the spectrum does not characterize the texture in the object (leaf) because of the strong edginess in the object boundary. However, after smoothing the object boundary, as Fig. 3(c), the spectrum represents well the characteristic of the texture.



(a) (b) (c)

Fig. 3. The spectrums of the object: (a) original image (b) the spectrum of the image before smoothing the object boundary (c) the spectrum of the image after smoothing the boundary.

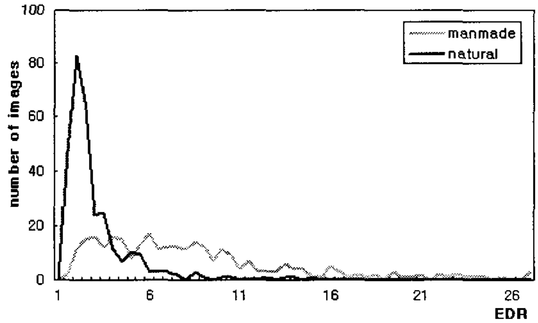


Fig. 4. Energy difference rate curves between maximum and minimum energy.

Fig. 4 shows the distribution curves of the energy difference ratio between man-made and natural objects. The ratios of natural objects tend to concentrate at the low value. On the other hand, the ratios of man-made objects are distributed at larger values than those of natural objects.

3.1.3 Gabor Filter for an Object Shape

In general, the boundary of a man-made object can be represented as a polygon consisting of straight line and can be represented in symmetrical and relatively simple form. In contrast, a natural object has complex form with irregular curves. This shape characteristic can be observed through Fourier spectrum of a binarized object image. Fig. 5 shows the binary representation of extracted objects in natural and man-made object images and their Fourier spectrum. The spectrum of man-made object shows strong directivity because the boundary of object consists of straight lines. In contrast, the spectrum of a natural object shows

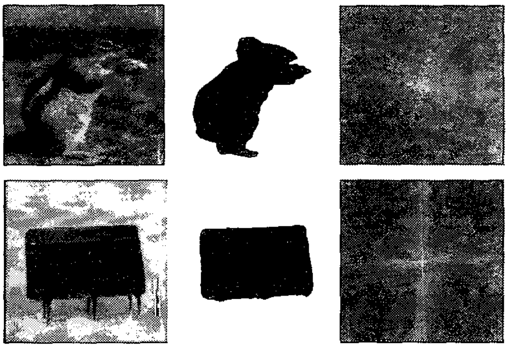


Fig. 5. The upper row represents a natural object image, the extracted object region and its Fourier spectrum. The spectrum has no directivity. The lower row represents a man-made object image, the extracted object and its Fourier spectrum. The spectrum gives strong directivity.

almost no directivity. We represent the shapeof an object using Gobor filter dictionary.

4. CLASSIFICATIONS AND EXPERIMENTAL RESULTS

The suggested method is evaluated on 300 man-made images and 300 natural images that are, selected from Corel Gallery photo-CD and Web. The classification accuracy depends on a training set. Therefore, to increase the reliability of classification result, we randomly chose 400 from 600 images as a training set and the rest 200 images as a test set. This procedure is repeated 100 times and the final accuracy is determined by the average.

We first classify object images based on each measure. When using energy of EDH and EDR in Gabor filter to train the classifier, thresholds are determined by minimizing the classification errors. Then images in test set are compared to these thresholds to determine their classes. To classify objects based on the shape feature vector which is represented Gabor filter dictionary we use the K-nearest neighbor (K-NN) classifier. For Gabor filter dictionary we extract a 24-dimentional feature vector which is made of 4-scales and 6-orientations in anobject region. In many approach, 24-dimentional vector is used to represent texture feature. However, this is very computationally intensive because we must compare each unknown example with every example in the training set. So we first classify the training data into the small number of groups by using K-mean algorithm. To determine the number of groups, we examine the classification error rate according to the number of groups. As we can see in Fig. 6, there is no significant difference in classification error when K=11 and more. Other training sets show the same result. We classify an object by using K-NN classifier that calculates distance between shape feature vector of the object and each center vector of 11 clusters. A feature vector is classified into a cluster that is of the closest distance.

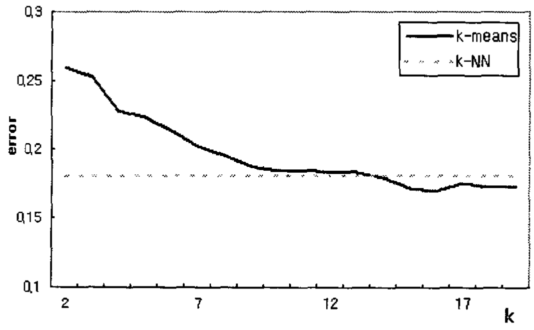


Fig. 6. Distribution of classification error rates according to K in K-mean clustering method on the shape feature vectors.

However, as we can see in Fig. 1 and Fig. 4, the distribution curves for the energy of edge direction histogram and the energy difference ratio in Gabor filter are not clearly separated. Thus each measure does not provide very good performance by itself. So each classified result is combined by majority rule in which decisions are made by voting with a majority determining the position of the entire group. Table 1 shows classification accuracy based on the precision, recall and F-measure. Table 2 shows the classification results for each measure. The measure on the energy difference ratio in

Gabor filter shows the best accuracy. The results reveal that energy of edge direction histogram and energy difference ratio in Gabor filter show the similar accuracies and characteristics but the former represents straightened edginess in the objects and the latter shows how many edges go along the same directions.

Fig. 7 shows a representative subset of the misclassified man-made and natural objects. The main reason for the misclassification of man-made objects is their streamlined shapes and smooth textures. They don't have any strong edginess or textures. The reason for the misclassification of the natural objects is directional patterns in the objects. They have smooth shape but have many straightened lines and also indicate same direction.

Table 1. Evaluation of the classification result by using majority rule

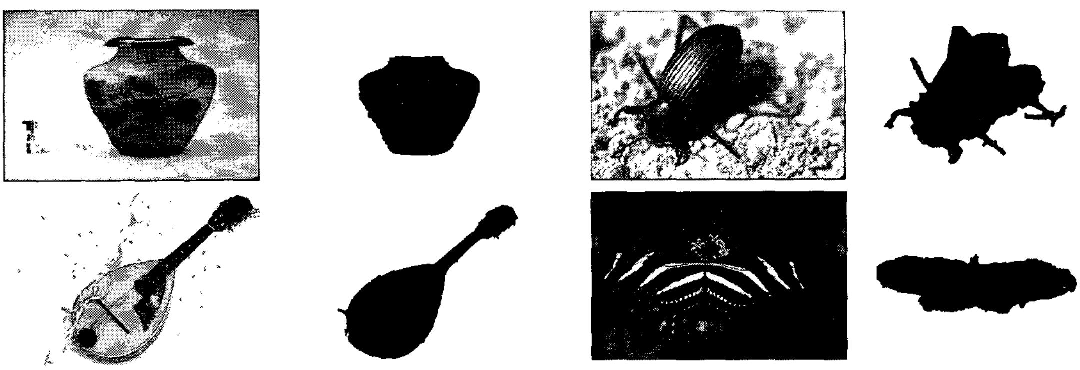
	Man-made	Natural
Precision	0.87	0.85
Recall	0.85	0.87
F-measure	0.86	0.86

Table 2. Evaluation of the classification results for each measure

		Energy of Edge Direction Histogram	Energy Difference Ratio in Gabor Filter	Gabor Filter for an Object Shape
Object	Precision	0.88	0.85	0.77
	Recall	0.77	0.86	0.78
	F-measure	0.82	0.86	0.78
Non-Object	Precision	0.79	0.86	0.78
	Recall	0.89	0.85	0.77
	F-measure	0.84	0.85	0.77

5. CONCLUSIONS

Our method of image classification is proven to have 86% accuracy. For this classification we proposed three measures (energy of edge direction histogram, energy difference ratio along directions in Gabor filter, and Gabor filter for an object shape) and used the majority rule to combine them. The proposed method was useful in classifying object images into man-made and natural object image although the measures were relatively simple. It must not be a trivial work to represent common attributes of generic classes such as man-made and natural object classes. In our future work, we will design more effective measures hoping to improve the classification accuracy. The classified objects are expected to be effectively used in image indexing for CBIR because they can provide higher level knowledge about the images in image databases.



(a) (b)

Fig. 7. A subset of the misclassified (a) man- made and (b) natural objects.

6. REFERENCES

[1] S.Y. Kim, S.Y. Park, and M.H. Kim, "Central Object Extraction for Object-Based Image Retrieval," *Int'l Conf. on Image and Video Retrieval* (*CIVR2003*), pp. 39-49, 2003.

[2] A. Vailaya, A.K. Jain, and H.J. Zhang, "On Image Classification: City images vs. landscape," *Pattern Recognition*, Vol. 31, No. 12, pp. 1921-1936, 1998.

[3] M. Szummer, and R.W. Picard, "Indoor-outdoor image classification," *IEEE Int'l Workshop Content-Based Access Image Video Databases*, pp. 42-51, 1998.

[4] A. Vailaya, M.A.T. Figueiredo, A.K. Jain, and H.J. Zhang, "Image Classification for Content-Based Indexing," *IEEE Trans. on Image Processing*, Vol. 10, No. 1, pp. 117-130, 2001.

[5] A. Oliva, A. Torralba, A.G. Dugue, and J. Herault, "Global Semantic Classification of Scenes using Power Spectrum Templates," *CIR'99*, Newcastle UK, 1999.

[6] S.Y. Kim, S.J. Park, and M.H. Kim, "Image Classification into Object / Non-object Classes," *Int'l Conf. on Image and Video Retrieval* (*CIVR2004*), pp. 393-400, 2004.

[7] J.R. Serra, and J.B. Subirana, "Texture Frame Curves and Regions of Attention Using Adaptive Non-cartesian Networks," *Pattern Recognition*, Vol. 32, pp. 503-515, 1999.

[8] S.B. Park, J.W. Lee, and S.K. Kim, "Content-based image classification using a neural network," *Pattern Recognition Letter*, Vol. 25, pp. 287-300, 2004.

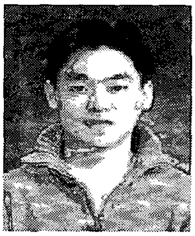
[9] T.S. Lee, "Image representation using Gabor wavelet," *IEEE Trans. on Pattern Analysis Machine Intelligent*, Vol. 1, No. 18, pp. 959-971, 1996.

[10] J. Canny, "A computational approach to edge detection," *IEEE Trans. on Pattern Analysis Machine Intelligent* (*PAMI*), Vol. 8, pp. 679-698, 1996.

[11] D.D. Ballard, "Generalizing the Hough Transform to Detect Arbitrary Shapes," *Pattern Recognition*, Vol. 13, pp. 111-122, 1981.

Chang-Min Park

He received B.S. and M.S. degrees in computer engineering from Dortmund National University, Dortmund, Germany in 1991 and 1995, respectively. Since 1998, he has been a Ph.D. candidate in the Department of computer engineering at Pusan National University. Since 2003, he has been an assistant Professor in the School of multimedia engineering, Young San University, Pusan, Korea. His research interests include Color Image Segmentation, Content based Image Retrieval, Image Processing.



Kyung-Mo Gu

He received B.S. and M.S. degrees in computer engineering from Pusan National University, Pusan, Korea in 2002 and 2004, respectively. He is currently working toward the Ph.D. degree in computer engineering at Pusan National University. His research interests include Color Image Segmentation, Content based Image Retrieval, Image Processing.

Sung-Young Kim

He received B.S., M.S., and Ph.D. degrees in computer engineering from Pusan National University, Pusan, Korea in 1994, 1996 and 2003, respectively. Since 2004, he has been an assistant Professor in the School of computer engineering, Kumoh National Institute of Technology, Gumi, Gyeongbuk, Korea. His research interests include Color Image Segmentation, Content based Image Retrieval, Image Processing.

Min-Hwan Kim

He received B.S. degree in electrical engineering from Seoul National University, Seoul, Korea in 1980. M.S., and Ph.D. degrees in computer engineering from Seoul National University, Seoul, Korea in 1983 and 1988, respectively. Since 1986, he has been an Professor in the Department of computer engineering, Pusan National University, Pusan, Korea. His research interests include Image Processing & Understanding, Color Engineering.