Modified Semantic Classification for very large image database

Chang Min Park

School of Undeclared Majors, Youngsan University, Busan, Korea
cmpark@ysu.ac.kr

Abstract: Semantic indexing of images is an ongoing study. Recently many researchers in the field of image classification for indexing of the very large image databases are interested in object(s) in an image. Thus, for efficient image matching use the semantic gap between higher concept of users and low-level image features. In this paper, we present a modified semantic classification method of the object(s) into natural/artificial classes, which can be of great interest for semantic indexing applications processing very large image databases. We first show that the modified segmentation method of the color images with irregular textures based on the block homogeneity. These works based on small-size blocks the color histogram of each of which is computed preliminarily once. Thus it works fast but provides rough segmentation. And using the segmented image, we show that dominant orientation features in modified Gabor orientation energy map of artificial objects are very useful for discriminating them from natural objects. Such Gabor energy maps for artificial objects tend to have dominant orientation features through analysis of Gabor filtering results for many object images. Then a modified sum of sector power differences is proposed as a classification measure, which shows a classification accuracy of 90.7% in a test with 3,000 object images.
[Chang Min Park. **Modified Semantic Classification for very large image database.** *Life Sci J* 2014;11(7):661-665]. (ISSN:1097-8135). http://www.lifesciencesite.com. 94

Keywords: Semantic indexing; classification; image data base; segmentation; dissimilarity

1. Introduction

In very large image database (VLID), images are automatically indexed by summarizing their visual contents, and are searched and matched usually based on low-level features such as color, texture, shape, and spatial layout. Usually a successful indexing of database images through appropriate classification (Vailaya *et al*., 1998; Vailaya *et al*., 2001; Zhou *et al*., 2010; Yang *et al*., 2009; Y. Yang and S. Newsam., 2011; Lowe, 2004) greatly enhances the performance of VLID systems by filtering out irrelevant images. Automatic segmentation (Deng *et al*., 1999; Carson *et al*., 2002; Handayani *et al*., 2012; Ibaa *et al*., 2012) method for VLID has to segment an image into the regions of interest and it is more important to segment an image into the regions of the objects as well. Additionally fast computation time is required.

On the one hand, many researchers believe that the key to effective VLID performance lies in the ability to access images at the level of objects because users generally want to search for the images containing particular object(s) of interest. Thus several methods (Huang *et al*., 1995; Manjunath, B.S. and Ma, W.Y, 1996, Liu *et al*., 2010, M. Muja and D. G. Lowe., 2009) that extract object(s) of interest from object images are studied. And no perfect pixel wise segmentation (T. Brox and J. Malik,. 2010, R. Arandjelovi´c and A. Zisserman., 2011, Arbelaez et al., 2011, Boix *et al*., 2012, A. Ion, J. Carreira, and C. Sminchisescu, 2011) is required, because shape information is less useful than color and texture information. Instead, a fast and unsupervised segmentation method is needed, even though segments are not disjunctive or their boundaries are a little incorrect.

First, a rough and fast segmentation method of color images is proposed which uses dissimilarity of local color distribution to measure discontinuity of colors and textures. The dissimilarity of local color distribution is computed using a modified color histogram intersection technique that enables us to measure the discontinuity of colors at boundaries of objects efficiently and to lessen the effect of discontinuity at finely textured regions. This modified segmentation method is based on the block homogeneity and then, using the segmented image (*D*-image).

In this paper show that the Gabor orientation energy for artificial objects varies more drastically than one for natural objects does. The Gabor orientation energy is the sum of Gabor energies that have the same orientation in the Gabor energy map. Then, as a discriminating measure, the sum of sector power differences is proposed in this paper, which can represent the drastic change of the Gabor orientation energy well. Those measure show a classification accuracy of 90.7% in a test with 3,000 object images.

2. Dissimilarity of local color distribution

In VLID, it is more important to segment a few salient regions in short processing time than to segment many regions with pixel-wise exact

boundaries. Thus an image is divided into blocks of 4×4 pixels and the dissimilarity of local color distribution at each block is computed using a modified color histogram intersection technique. There are three types of dissimilarity. One is the color dissimilarity that represents discontinuity of color distribution in a block against its neighbors. The other is texture dissimilarity for discontinuity of texture distribution. The dissimilarity in a block can be considered as boundary strength in the block. A block that consists of 16 pixels is not enough to compute color distribution information of the block. Thus we define the super-block S for a block B that consists of the block B and 8-neighbor blocks of B (Fig. 1). The color histogram of each super-block, H_{SB} can be easily computed by just adding the color histograms of its nine blocks, H_{Bi}'s each of which is computed preliminarily once.

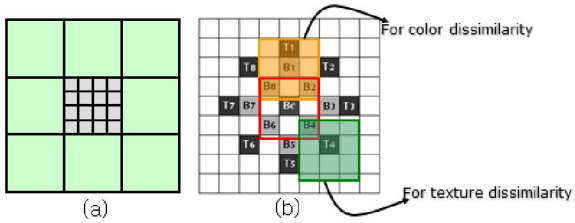

Figure. 1 Definition of a 4x4 pixels block(a) and neighbor blocks(b)

2.1 Color dissimilarity

The color dissimilarity D_C at a block B_c is defined as described in Eq. (1) by using a modified color histogram intersection.

$$D_C = M - \sum_{k=1}^{N} H_C(k) = M - \sum_{k=1}^{N} \min\{H_{I1}(k),...,H_{I8}(k)\},$$
$$where\ H_{Ii}(k) = \min\{H_{SBc}(k), H_{SBi}(k)\} \quad ---(1)$$

The neighbor blocks B_i's for color dissimilarity are selected as shown in Fig. 1(b). The M and N are total number of pixels in a super-block and the number of quantized colors, respectively. The color dissimilarity represents discontinuity of color distribution at B_c in eight directions. Even if only one of neighbors of B_c has different color distribution with B_c, the color dissimilarity at B_c will be increased.

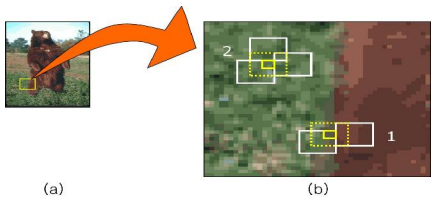

Figure. 2 An example of the color dissimilarity, (a): original image, (b): a part image of (a).

Blocks with the higher color dissimilarities can be located near the region boundaries with a strong probability. Fig. 2 shows an example of the color dissimilarity. The dotted area 1 in Fig. 2(b) contains the boundary of the object and background. This boundary appears as vertical line. Therefore, we consider only left and right neighbor blocks. We first compute the color histogram intersection between center block and blue block. And then center block and red block. As shown in the Figure, there is no similar color between the left half of the center block and right half of the blue block. In the same way, there is no similar color between the rights half of the center block and left half of the red block. Therefore, the color dissimilarity becomes higher value. But, area 2 contains the region with fine texture or uniform color, so its color dissimilarity becomes a lower value.

Fig. 3(b) shows the color dissimilarity results for the image with various texture sizes in Fig. 3(a). In the regions with fine textures or uniform colors, the color dissimilarity becomes a lower value. However, in the regions with the coarse textures, the color dissimilarity is increased and so it is hard to discriminate the boundaries of the objects and the coarse textures. The coarse and large texture boundaries are apt to be misunderstood as the object boundaries based on the only this color dissimilarity value. Therefore, we define another measure to discriminate texture boundaries with coarse patterns and object boundaries.

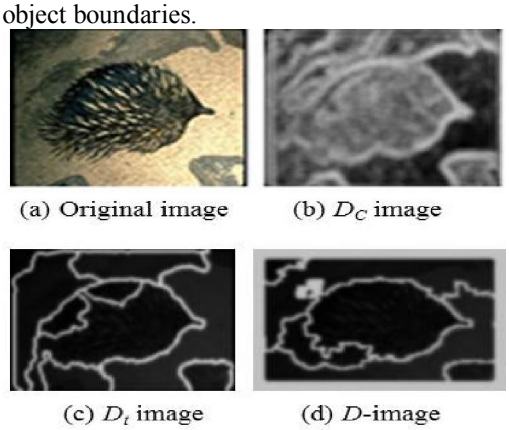

Figure. 3 Result images using boundary dissimilarity

2.2 Texture dissimilarity

For the region with the coarse and large texture, the color dissimilarity is not useful. In Fig. 4, center-block and red-block are overlapped with white color, center-block and blue block as well. Therefore, color dissimilarity becomes higher value. So, the coarse and large texture boundaries are apt to be misunderstood as the object boundaries. We define the texture dissimilarity. Texture dissimilarity is designed to weaken the strength of false boundaries at

sparse and coarse texture patterns. The texture dissimilarity D_T at a block B_c is defined as described in Eq. (2) by subtracting the averaged sum of usual color histogram intersection between the super-block histogram of the B_c and one of each T_i from total number of pixels in a super-block. The neighbor blocks T_i's for texture dissimilarity are determined as shown in Fig. 1(b). Note that they are located farther from the B_c than the B_i's are. Fig. 4 shows an example of the texture dissimilarity.

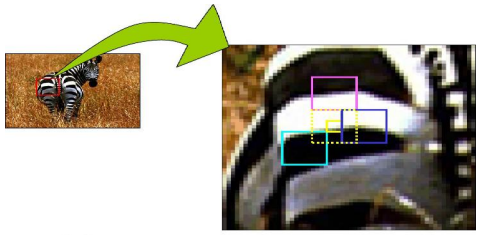

Figure. 4 An example of the texture dissimilarity, (a): original image, (b): a part image of (a).

$$D_T = M - \frac{1}{8}\sum_{i=1}^{8}\left[\sum_{k=1}^{N}\min\{H_{SB_c}(k), H_{ST_i}(k)\}\right] \quad ---(2)$$

Actually the second term in Eq. (2) is the averaged sum of color similarity between nearly overlapped super-blocks, so the texture dissimilarity represents color dissimilarity in relatively large-scale. Thus the texture dissimilarity weakens the dissimilarity at coarsely textured regions than the color dissimilarity. In the Fig. 3(c), we can see that boundaries of regions with sparse and coarse texture have low texture dissimilarity value.

2.3 Generation of dissimilarity images

The dissimilarity D at a block is defined as the weighted sum of color dissimilarity D_C and texture dissimilarity D_T of the block (Eq. (3)). The weight factor α can be appropriately adjusted in each application. The image that consists of D's of all the blocks is called the D-image in this paper.

$$D = \alpha D_C + (1-\alpha) D_T \quad ---(3)$$

Fig. 3(d) shows the dissimilarity image for the original image in Fig. 3(a), when the weight factor is set to 0.5. We can see that dissimilarities at the boundaries of the four textured regions are greater than those at the texture boundaries. Thus the boundaries of the regions can be easily discriminated from the texture boundaries.

3. Analysis and Clustering of Gabor Characteristics

3. 1 Gabor Energy Map

Given an D-image (Fig. 5(a)), we pass it through a bank of 24 Gabor filters as shown in Fig. 5(b). The filter bank is similarly designed to the Gabor filter dictionary in [13]. The number of scales, the number of orientations, the lower center frequency and the upper center frequency of interest are set by 4, 6, 0.1, and 0.4, respectively. The filter corresponding to i-th scale and j-th orientation is denoted by F_{ij} and the filtered image through F_{ij} is denoted by f_{ij}. The Gabor energy e_{ij} for f_{ij} is defined as the sum of magnitude squares over all complex pixel values in f_{ij}, as shown in Eq. 4. Fig. 5(c) shows the Gabor energy map M for the object image in Fig. 5(a). The value of $D(i,j)$ is the Gabor energy e_{ij}. The Gabor energy map in Fig. 5(c) represents texture feature of the given D-image in Fig. 5(a) well. We call a Gabor energy map for an object image its Gabor characteristics.

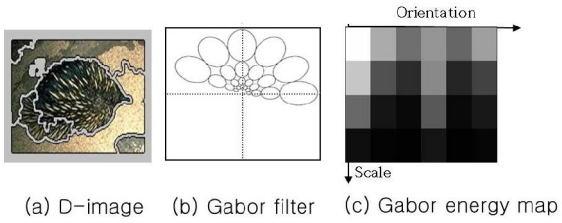

Figure. 5 . Gabor filtering and Gabor characteristics: (a) an D-image, (b) a Gabor filter bank designed with parameters, the number of scales = 4, the number of orientations = 6, the lower center frequency of interest = 0.1, and the upper center frequency of interest = 0.4, (c) a Gabor energy map

$$e_{ij} = \sum\sum |f_{ij}(x,y)|^2 \quad ----(4)$$

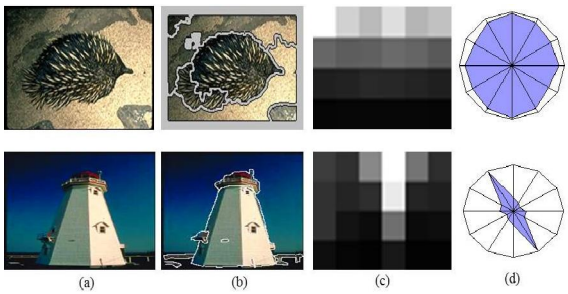

Figure. 6. An example of constructing Gabor energy diagrams: (a) original image (b) D-image, (c) Gabor energy map, (d) Gabor energy diagrams

3. 2 Diagram of Gabor orientation energy Map

Fig. 6(d) shows the Gabor energy diagram corresponding to the Gabor energy map in Fig. 5(c). Gabor energy diagrams of artificial class look to be

sharp-pointed, while those of natural class tend to have round shapes.

4. Modified semantic classification using Natural/artificial images

4.1 Sum of Sector Power Differences

The dominant orientation discussed in section 3 can be considered as a texture feature of artificial images. This texture feature can be analyzed well in Fourier power spectrum by measuring the sector powers. Let $F(u,v)$ be the discrete Fourier transform of an image $f(x,y)$. The power spectrum $|F(u,v)|^2$ is defined by the magnitude of the spectral components squared. A sector power is measured by summing the power over range of corresponding sector. In this paper, six sectors are defined over half the power spectrum, as shown in Fig. 7(a). Very low frequency powers are excluded from computing the sector powers in order to make dominant orientations clear. The Gabor energy diagrams of sector power (Fig. 7(b)) for the Lighthouse image in Fig 6 is very similar to that of Gabor orientation energy as shown again in Fig. 7(c).

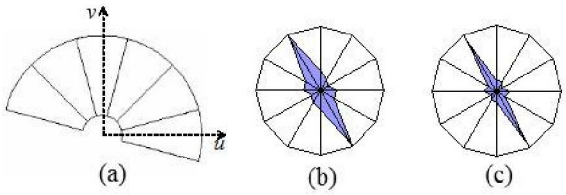

Figure. 7. Comparing Gabor energy diagram of sector power with that of Gabor orientation energy for the Lighthouse image in Fig. 6: (a) sector definition used in this paper, (b) Gabor energy diagram of sector power, (c) Gabor energy diagram of Gabor orientation energy

Let the six sector powers be SP_i (i = 0,1,…,5). The sum of sector power difference (SSPD) can be computed as in Eq. 3, which is very useful for representing existence of dominant orientations of any direction in Gabor energy diagram of sector power.

$$SSPD = \sum_{j=0}^{5} \left| SP_{(j+1) \bmod 6} - SP_j \right| \quad --- (5)$$

The SSPD has a great value when there is abrupt change between neighboring sector powers. Thus SSPDs for artificial object images have greater values than those for natural object images. Fig. 8 shows distribution of the SSPDs for 1,500 artificial object images and 1,300 natural ones.

5. Experimental Results and Discussions

The proposed classification method is evaluated on 3,000 object images (1,500 natural images and 1,500 artificial images) selected from the Corel Gallery photo-CD. Fig. 9 shows examples of object images in several categories.

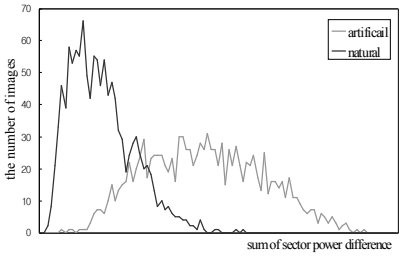

Figure. 8. Distribution of SSPDs (sum of sector power differences) for 1,300 artificial object images and 1,300 natural object images. The SSPDs for natural object images tend to have lower values than those for artificial object images

To compensate for different object sizes, each object is scaled for longer one between its width and its height to be about 210 pixels and is centered in a (256 x 256) black background. These training images are randomly chosen from the 3,000 object images. Object classification is performed on the remaining object images. This procedure is repeated 50 times to reduce dependence of classification on the training set of object images. Average classification accuracy is 90.7%.

Figure. 9 Examples of object images in several categories

Table 1 shows classification accuracy based on precision, recall and F-measure. We can see in the classification of artificial images that the recall is high and the precision is low. This means that many natural images are misclassified into artificial class even though almost all of artificial images are classified correctly. We can also see that small number of natural images is correctly classified. In case of K-means classification, misclassification rates are almost equal to in two object classes.

Table 1. Evaluation of the classification results for each measure by using 6-fold cross-validation

		Proposed Method	K-means Classification
Artificial image	Precision	0.77	0.86
	Recall	0.93	0.78
	F-measure	0.84	0.82
Natural image	Precision	0.91	0.80
	Recall	0.73	0.87
	F-measure	0.81	0.83

6. Conclusions

We first proposed a block-based color image segmentation method in this paper, which can classify textured color patterns as a region instead of separating them several regions. For this purpose, we proposed the Dissimilarity of local color distribution method to extract the closed boundary of region as possible. And we showed that the dominant orientation(s) of artificial images in their energy diagrams of Gabor orientation energy was very useful for discriminating artificial image class from natural image class. As classification measures, the sum of sector power differences in Fourier power spectrum was proposed, which represented existence of the dominant orientations in any direction well. They showed classification accuracy of 90.7% on a test with 3,000 object images. Our work can be applied to improving the performance of semantic-based image indexing.

Acknowledgements:

This paper was supported by research funding in 2013 from Youngsan University, Busan, Korea.

Corresponding Author:
Park Chang Min
School of Undeclared Majors
Youngsan University, Busan, Korea
E-mail: cmpark@ysu.ac.kr

References

1. Vailaya, A., Jain, A.K., and Zhang, H.J.: On Image Classification*: City Images vs. Landscape. Pattern Recognition*. **31**(**12**) (1998) 1921-1936
2. Vailaya, A., Figueiredo, M.A.T., Jain, A.K., and Zhang, H.J.: Image Classification for Content-Based Indexing. *IEEE Trans. on Image Processing*. **10**(**1**) (2001) 117-130
3. X .Zhou, K. Yu, T. Zhang, and T. S. Huang. Image classification using super-vector coding of local image descriptors. In *ECCV*, page 141-154, 2010
4. J. Yang, K. Yu, Y. Gong, and T. Huang. Linear spatial pyramid matching using sparse coding for image classification. In *CVPR*, page 1794-1801, 2009.
5. Y. Yang and S. Newsam. Spatial pyramid co-occurrence for image classification. In *ICCV*. Page 1465-1472, 2011.
6. D. Lowe, Distinctive image features from scale invariant features. *International Journal of Computer Vision,* 60:91-110, 2004
7. Y. Deng, B. S. Manjunath, and H. Shin, "Color Image Segmentation," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 446-451, 1999
8. C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: Image segmentation using Expectation-Maximization and its application to image querying," *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002
9. Handayani T, Ari W, Nanik S. Optic Nerve Head Segmentation Using Hough Transform and Active Contours. *TELKOMNIKA Indonesia Journal of Electrical Engineering*. 2012; 10(3): 531-536.
10. Ibaa J, Akram UM, Anam T. Retinal Image Preprocessing: Background and Noise Using Adaptive Non-cartesian Networks. Pattern Recognition. **32** (1999) 503-515
11. Huang, Q., Dom, B., Steels, D., Ashely, J., and Niblack, W.: Foreground / Background Segmentation of Color Images by Integration of Multiple Cues. Int'l Conf. on Image Processing. **1** (1995) 246-249
12. Manjunath, B.S. and Ma, W.Y.: Texture Features for Browsing and Retrieval of Image Data. IEEE Trans. on Pattern Analysis and Machine Intelligence. **18**(**8**) (1996) 837-842
13. T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H. Shum. Learning to detect a salient object. Published by the IEEE Computer Society, 2010.
14. M. Muja and D. G. Lowe. Fast approximate nearest neighbors with automatic algorithm configuration. In International Conference on Computer Vision Theory and Application
VISSAPP'09), pages 331–340. INSTICC Press, 2009.
15. T. Brox and J. Malik, Object Segmentation by Long Analysis of Point Trajectories. LNCS. Pp282-295, 2010
16. R. Arandjelovi´c and A. Zisserman. Smooth object retrieval using a bag of boundaries. In *Proc. ICCV*, 2011.
17. P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. *IEEE Trans. on PAMI*, 2011.
18. X. Boix, J. M. Gonfaus, J. van de Weijer, A. D. Bagdanov, J. Serrat,and J. Gonz´alez. Harmony Potentials Fusing Global and Local Scale for Semantic Image Segmentation. *International Journal of Computer Vision*, 96(1):83–102, December 2012.
19. A. Ion, J. Carreira, and C. Sminchisescu. Image segmentation by figure-ground composition into maximal cliques. In *Proc. ICCV*, 2011.

5/26/2014