

Texture analysis based image segmentation using optimized unsupervised learning

Cheng, Shun

(出版者 / Publisher)

法政大学大学院情報科学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 情報科学研究科編 / 法政大学大学院紀要. 情報科学研究科編

(巻 / Volume)

9

(開始ページ / Start Page)

39

(終了ページ / End Page)

44

(発行年 / Year)

2014-03

(URL)

<https://doi.org/10.15002/00010519>

Texture analysis based image segmentation using optimized unsupervised learning

Shun Cheng

Graduate School of Computer and Information Sciences

Hosei University

Email: shun.cheng.4v@stu.hosei.ac.jp

Abstract—This paper proposed a new high accuracy method using textural feature analysis and optimized unsupervised classification to segment images, which contained multiple objects with own textures. In the proposed method, some spatial filtering processes were applied to extract the texture features as multi-dimensional vectors. Following area segmentation was done by applying K-means clustering to the feature vectors with possible number of clusters. The AIC was adopted for selecting the optimal number of clusters. The method was successfully applied to simple and complex natural images and the result of experiments indicated that the segmentation accuracy rate was achieved 75.67%.

Keywords—local texture analysis; image segmentation; AIC; optimized method.

1. INTRODUCTION

Nowadays, the most common method in the field of image retrieval search engine is still Text-Based Image Retrieval (TBIR), which generally speaking, is converting visual information from arbitrary images into text information manually, exactly as labeling images. However, the biggest flaw of TBIR is that the main kernel of image understanding part is finished by human, manual work on every image cannot adapt to the image number with such an explosive growth in the era of information.

As this situation, Content-Based Image Retrieval (CBIR) technology naturally becomes hotspot in image retrieval field. Meanwhile, leading searching engines, like Google and Baidu have already designed some interesting application named "Search by image". This kind of application is a great

improvement; still it is image matching rather than real CBIR, real CBIR should possess the ability to recognize arbitrary images automatically and that leads to a great demand of image segmentation.

Image segmentation is an important technology in many applications of image recognition [1][2]. As a main direction of image segmentation, current texture analysis based image classification method normally can reach a high accuracy coping with images with single texture [3][4]. However, in real images case, textures it contains are much more complex.

The first difficulty of texture based image analysis is choosing or designing a proper texture which can distinguish different texture characteristics, as shown in Fig. 1. To cope with this problem, we extract three radically different texture feature types motivated by 1) statistical, 2) psychological and 3) color information which are discriminative and sufficient for the description of the objects.

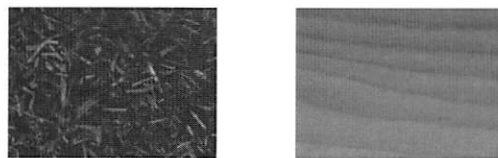


Fig. 1. Texture from chips and wood

Human visual perception is quite unpredictable, and the second difficulty is designing a scientifically optimized method to determine the number of objects in the source image. We need to design model selection method [5][6] to cope with this difficulty, in this thesis, we propose one Occum's razor principle [7] and one Akaike Information Criterion (AIC) [8] based optimized methods to solve the problem.

The method we proposed is specially aiming at coping with difficulties in the process of image segmentation by information from the input image itself. This unsupervised and automatic method can supply indispensable object number and distribution information within the input image for the image recognition application.

2. APPROACH

The proposed method contains four sub-processes; feature-extraction for extracting appropriate low level features which separate dissimilar objects, area segmentation for clustering pixels, optimized method for determining the final segmentation result and post processing for quality improvement, Fig. 2 shows flow of the image segmentation.

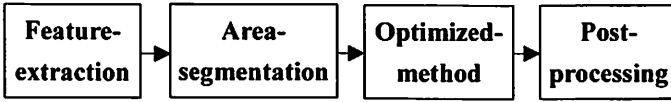


Fig. 2. Flow of images segmentation

A. Feature extraction

Most of the former papers devised and evaluated texture features by testing the texture classification accuracy of a certain texture image database. However, real images are not made up of uniform texture patchworks, in this thesis, we use features combination to couple with natural images with much more complex textures.

For separating an object from background, local features are widely used and optimally selected ones are powerful tools for extracting complex objects. In the point of view of efficiency, we adopt the following features.

1. Grey level Co-occurrence Matrix features:

The co-occurrence matrix [9] [10] of an $N \times N$ -pixel area I , comprises the probabilities $P_{d,\theta}(i, j)$ of the transitions from a grey-level i to a grey-level j in a given direction at a given inter sample spacing d [11]:

$$P_{d,\theta}(i, j) = \frac{C_{d,\theta}(i, j)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} C_{d,\theta}(i, j)}. \quad (1)$$

where $C_{d,\theta}(i, j)$ means the occurrence frequency of pixel $(m, n), (u, v) \in N \times N$, $f(m, n) = j$, $f(u, v) = i$. $|(m, n) - (u, v)| = d, \angle((m, n) - (u, v)) = \theta$.

Denotes the number of elements in the set, $f(m, n)$ and $f(u, v)$ correspond to the grey-levels of the pixel located at (m, n) and (u, v) respectively, and N_g is the total number of grey-levels in the image. Within this thesis, we choose the three most commonly used GLCM features, listed in Table 1.

TABLE I. FEATURES CALCULATED FROM THE NORMALISED GLCM

Feature	Formula
Energy	$\sum_i \sum_j P_{d,\theta}(i, j)^2$
Entropy	$\sum_i \sum_j P_{d,\theta}(i, j) \log P_{d,\theta}(i, j)$
Homogeneity	$\sum_i \sum_j \frac{P_{d,\theta}(i, j)}{1 + i - j }$

2. Tamura features:

Tamura, Mori and Yamawaki [12] designed six basic textural features, namely, coarseness, contrast, directionality, line likeness, regularity and roughness to simulate visual perception. Within this thesis, we choose the two most commonly used Tamura features.

Coarseness: This feature has a direct relationship to scale and repetition rates. Coarseness aims to identify the largest size of texture existing in this image. Computationally one first takes averages at every point over neighborhoods the linear size of powers of 2. Set the average over the neighborhood of size $2^k \times 2^k$ at the point (x, y) is

$$A_k(x, y) = \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} f(i, j) / 2^{2k}. \quad (2)$$

In both horizontal and vertical orientations, at each point one takes differences between average pairs corresponding to non-overlapping neighborhoods on opposite sides. In horizontal case it is

$$E_{k,h}(x, y) = |A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y)|. \quad (3)$$

For each pixel, picks the best size which gives the highest $E_{k,h}(x, y)$ value, where k maximizes E in both directions. Then coarseness is the average of $S_{opt}(x, y) = 2^{k_{opt}}$.

Contrast: This feature is designed to capture the grey levels dynamic range in an image; it is derived from the following four factors: dynamic range of levels of grey, polarization of the white-black distribution in the grey level histogram, edge sharpness and frequency of repeated patterns

$$F_{con} = \sigma / (\alpha_4)^n \quad (4)$$

where $\alpha_4 = \mu_4 / \sigma^4$, and μ_4 is the fourth moment about the mean and σ^2 is the deviation of the values of grey in the image. Experimentally, it can get the closest agreement to human visual measurements when $n=1/4$.

3. Color information:

In both textural and natural images, color information surely contains a lot of object information which can be used for our purpose. In grey level images, since the original three channels has already been linear converted into one, in our experiments, we use average values within multiple size regions around reference pixel directly.

B. Area-segmentation

After extracting all above features for each pixel, we use Linde-Buzo-Gray algorithm (LBG algorithm) [13] to cluster feature vectors. LBG algorithm takes sets of input feature vectors $S = \{x_i \in R_d | i=1, 2, \dots, n\}$ as input and generates a representative subset of vectors $C = \{c_j \in R_d | j=1, 2, \dots, K\}$ with a user specified value K as output according to the similarity measure, where S is the feature vector space, d is the dimensions of feature vector space, c_j is the center of clustering j (codebook j), and K is the number of total clusters.

1. Input feature vectors $S = \{x_i \in R_d | i=1, 2, \dots, n\}$.
2. Initiate the codebook $C = \{c_j \in R_d | j=1, 2, \dots, K\}$, the initial position of c_j is the average of all x_j adding a small random variable.
3. Classify the n training vectors into K clusters according to $x_i \in S_q$, if $\|x_i - c_q\|_p \leq \|x_i - c_j\|_p$ for all $j \neq q$.
4. $\Delta C_{jx} = \Delta t * \sum_{i=1}^K \sum_{i=1}^n (x_{ix} - c_{jx})$, $\Delta C_{jy} = \Delta t * \sum_{i=1}^K \sum_{i=1}^n (x_{iy} - c_{jy})$.
5. Set $k \leftarrow k+1$, compute the distortion $D_k = \sum_{j=1}^K \sum_{x_i \in S_j} |x_i - c_j|$.
6. If $(D_{k-1} - D_k) / D_k > \epsilon$, repeat steps 4 to 5.
7. With $(D_{k-1} - D_k) / D_k < \epsilon$, clustering is finished.

Since the cluster number K cannot be decided, we can get multiple clustering results with multiple K value.

C. Optimized method

The cluster number K cannot be determined automatically, we propose two estimation method, one is based on Occam's razor principle, the other one is based on AIC to cope with this problem. In the experiment section, both these methods are estimated through experiment.

1. Occam's razor optimized method

Occam's razor principle states that among competing hypotheses, we should select the hypothesis with the fewest assumptions. Applying to our purpose, what we need is simple structure with a moderate outcome. In segmentation, the complexity of structure corresponds to the number of clusters; the degree of outcome does to the total error. To obtain the balanced case of complexity and total error, we designed an index E_{occ} to evaluate multiple choices and the case with minimum E_{occ} is determined as the final output.

$$E_{occ} = w_1 \sum_{i=1}^n \sum_{j=1}^K d_{ij} + w_2 * K. \quad (5)$$

where K is the number of codebooks, n is the number of vectors, $\sum_{i=1}^n \sum_{j=1}^K d_{ij}$ is the distances between every codebook and the sample vectors belonging to that codebook, w_1 and w_2 are weights for adjusting two term's contribution.

2. AIC optimized method

AIC is based on the Kullback-Leibler divergence between the statistical probability distribution model and the true probability distribution [14], which can be estimated by the maximum likelihood estimator. AIC is defined as

$$AIC = (-2)(\text{maximum logarithmic likelihood}) + 2(\text{degree of freedom}) \quad (6)$$

Applying to our purpose, AIC based optimized method is expressed as

$$E_{AIC} = -2 \sum_{j=1, N} n_j \log(n_j) + 2 M \log(M) + KM \log(2\pi) + \sum_j \sum_i (n_j \sum_{k=1, K} \log(s_{jk})) + 2(N-1) + 2KN + KM \quad (7)$$

where N is the number of codebooks, n_j is the number of sample vectors belonging to the codebook j , q_j is the probability that a sample pattern belongs to cluster C_j (note that $\sum_{j=1, N} q_j = 1$ must be satisfied) and $p = i_j x / s_j$ is the probability that a sample pattern in cluster C_j is located at $i_j x$

which is calculated by the probability density function of C_j with the parameter s_j (variance).

D. Post-processing

The purpose of the post-processing is to process the preliminary result of section C to a final segmentation result of objects. The filter we use is a particular case of dilation template, if the majority pixels in the detecting pixels' adjacent area are belonging to the same cluster, the detecting pixel will be considered as a part of that cluster.

3. EXPERIMENT

In this chapter, we implement the feature extraction and area segmentation on test images, calculate the accuracy of two optimized methods, and evaluate segmentation accuracy. In our experiments, surface texture images are from Outex texture database [15].

A. Image database

- 1) Grey level Material surface images: Each test image contains 16 (4×4) textures; in order to evaluate the optimized method, type numbers of texture are varied from 2 to 8. The original source images are color images; in this test section we transform them into grey level images.

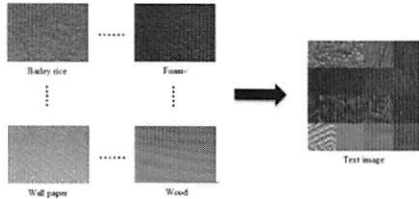


Fig. 3. Material surface image (4×4)

- 2) Grey level Simple natural images: Simple images we use are natural images with low resolution (256×256) and uncomplicated background and foreground from an open source image database. The subjects of these images contain book, car, indoor, house, etc.



Fig. 4. Simple natural image

- 3) Grey level Complex natural images: Complex images we use are natural images with high resolution, complicated background and foreground. The subjects of these images contain natural, elephant, city, dessert, etc.

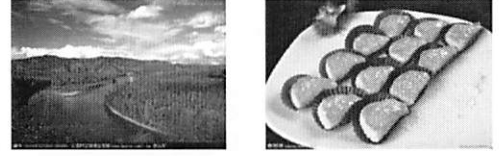


Fig. 5. Complex natural image

- 4) Color images: color images we use are the original color images of part 1 and part 3 (original source images of part 2 are grey level images).

B. Feature extraction and Area segmentation

In feature extraction, we extract different features with varying parameters for each pixel, storing these quantified features into different images for observation, and building a multi-dimensional feature space for clustering.

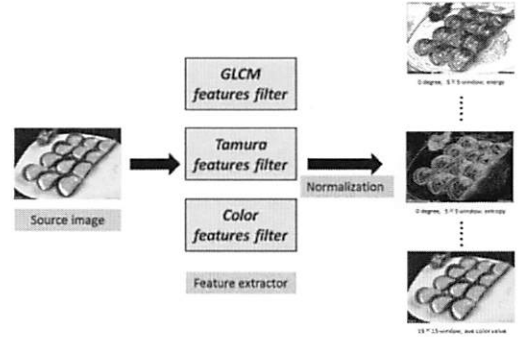


Fig. 6. Feature extraction process

For each codebook number we will get a segmentation result with i clusters, as shown in Fig. 7.

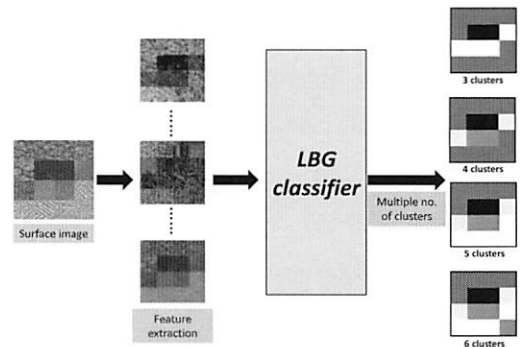


Fig. 7. Area segmentation process

C. Optimized method

In this part, we test two optimized methods mentioned in section B to determine one optimized result from all the segmentation results we got above. For each of these two methods, we design an evaluation index, as shown in Fig. 8, the result with minimum evaluation index is determined as the final output, as shown in Fig. 9.

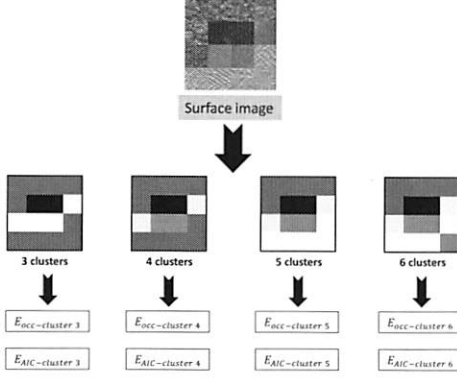


Fig. 8. Evaluation index

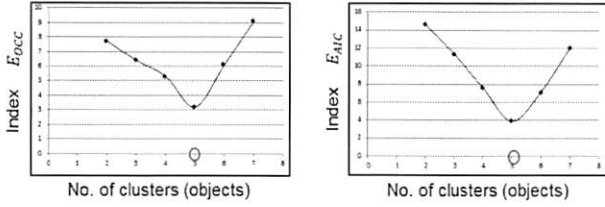




Fig. 9. Evaluation index (material surface 3 with 5 types of texture)

TABLE II. OPTIMIZED METHOD EVALUATION

<i>Input image</i>		
	Material surface 1	Material surface 4
<i>Types number</i>	Six	Six
<i>Result of E_{occ}</i>	Six	Five
<i>Result of E_{AIC}</i>	Six	Six

In the experiments, twenty images are tested; part of the results is shown in Table 2. The accuracy of Occam's razor principle based optimized method is 75% and the accuracy of AIC based optimized method is 85%, and the average error is about one object in wrong cases.

D. Segmentation accuracy

TABLE III. SEGMENTATION SAMPLES OF SIMPLE IMAGES



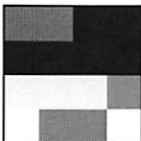

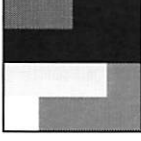






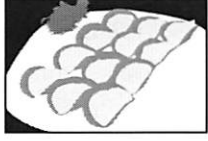
<i>Input image</i>		
	Grey Material surface	Grey simple natural
<i>Segmentation result</i>		
<i>GT image</i>		

TABLE IV. SEGMENTATION SAMPLES OF COMPLEX IMAGES

<i>Input image</i>		
	Grey complex natural	Color complex natural
<i>Result</i>		
<i>GT image</i>		

The segmentation accuracy of grey level material surface images is about 86.7%, accuracy of grey level simple natural images; grey level complex natural images and color complex natural images are 78.6 %, 70.32% and 75.67 %, respectively.

E. Discussion

The experiment results showed the validity of our proposed method with the ability of solving the main difficulties in the segmentation process.

For material surface images, the original usage of these images is testing how well a new designed texture descriptor can classify them. The-state-of-art technologies like local binary pattern texture descriptor can reach about 95% recognizing accuracy, but normally this kind of research focuses on distinguishing images with only one kind texture in each of them. In the experiment, multiple textures images surely pose some difficulties for us; still the performance of our multiple features combination is acceptable. Since this kind of images contains no global information, classification accuracy ties up with efficiency of textures directly.

For simple and complex natural images, the segmentation accuracy is about 75%. The main reason of error segmentation consists of two parts. The first one is the extracted features cannot represent different textures perfectly. The second reason is lack of global information, and these flaws raise parts of our future work.

4. CONCLUSIONS

In this thesis, we proposed a method to cope with texture analysis based image segmentation. The method consists of four sub-processes, feature extraction for extracting low level features which separate dissimilar objects, area segmentation for clustering pixels, optimized method for determining the final segmentation result and post processing for quality improvement.

This method can cope with different kinds of images, including one and three-channel material surface images, simple natural images, and complex natural images. The experiments achieve about 85% of the optimized method accuracy, 70.32% and 75.67% segmentation accuracy of grey-level complex images and color complex images respectively. With its ability to partition a digital image into

multiple objects, provide object number and distribution information, the proposed method can be effectively applied to the image recognition applications.

Our future work includes: 1) the features we extracted in this paper haven't been proved to be the best choice; therefore, other local texture features should be examined. 2) To deal with the high dimensional feature space, we should consider dimensionality reduction or factor analysis methods. 3) In the optimized method part, other information criterion based model selection method should be checked. 4) The way we deal with three channel images is relatively rough, designing an analysis strategy aiming at color images segmentation should be considered.

REFERENCE

- [1] Jianbo Shi and Jitendra Malik, "Normalized Cuts and Image Segmentation," *IEEE T Transactions on pattern analysis and machine intelligence*, 22(8), 888-905, 2000
- [2] R. Peter M. and Winter Martin "Survey of Appearance-Based Methods for Object Recognition", Technical Report, Inst. for Computer Graphics and Vision, Graz University of Technology, Austria; 2008.
- [3] Schwarz, Gideon E. "Estimating the dimension of a model". *Annals of Statistics* 6 (2): 461-464. MR 468014, 1978.
- [4] Ando, Tomohiro. "Bayesian predictive information criterion for the evaluation of hierarchical Bayesian and empirical Bayes models". *Biometrika* 94 (2): 443-458. 2007
- [5] P. Howarth and Stefan Ruger, "Evaluation of Texture Features for Content-Based Image Retrieval," *Proceedings of the CIVR 2004*, 326-334, 2004.
- [6] Savvas A. Chatzichristofis and Yiannis S. Boutalis, "Compat composite Descriptors for Content Based Image Retrieval," *VDM Verlag Dr. Muller*, 2011
- [7] Paul Vitányi and Ming Li, "Minimum Description Length Induction, Bayesianism and Kolmogorov Complexity." *IEEE Transactions on Information Theory*, 46(2), 446-464, 2000
- [8] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on AC*, 19 (6): 716-723, 1974
- [9] M. Haralick, K. Shanmugam and Itshak Dinstein, "Textural features for image classification," *IEEE Transaction on SMC*, 3(6), 610-621, 1973.
- [10] P. Howarth and Stefan Ruger, "Evaluation of Texture Features for Content-Based Image Retrieval," *Proceedings of the CIVR 2004*, 326-334, 2004.
- [11] Savvas A. Chatzichristofis and Yiannis S. Boutalis, "Compat composite Descriptors for Content Based Image Retrieval," *VDM Verlag Dr. Muller*, 2011
- [12] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *IEEE Transaction on SMC*, 8(6), 460-472, 1978.
- [13] Linde, Y., Buzo, A., Gray, R.M., An Algorithm for Vector Quantizer Design, *IEEE Transactions on Communications*, vol. 28, 84-94, 1980.
- [14] H. Sako, "Optimization of the Distance-Based Neural Network by Akaike's Information Criterion," *Artificial Neural Network, 2*, Elsevier Science Publishers B.V. 1127-1133, 1992
- [15] T. Ojala, M. Pietikäinen, "Multiresolution Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, 971-986, 2002.