Proceedings of the International MultiConference of Engineers and Computer Scientists 2010 Vol II,
IMECS 2010, March 17 - 19, 2010, Hong Kong

Latent Semantic Analysis for Classifying Scene Images

Chu-Hui Lee and Kun-Cheng Chiang

***Abstract*—With the increase of digital images, how to classify digital images efficiently becomes an important research topic. Semantic analysis that is used in the scene image classification has received much attention. The digital images are represented by low-level features often. It is a valuable research to reduce semantic gap between the high-level semantic and low-level features. This study proposes a new scene image classification method. The latent semantic analysis is used to select the important low-level features of the scene. Then scene similarity measurement is used to classify images. Experimental results show that the proposed method in the classifying scenes images has better correct classification ratio than frequency feature analysis.**

***Index Terms*—image classification、latent semantic analysis、low-level features、scene image.**

I. Introduction

Information technology has brought convenience due to rapid development of digital images. Image database becomes larger and larger over time so that efficiently using or searching images becomes an important topic. Content-based Image Retrieval (CBIR) is an important technology in image retrieval. An image contains many objects and is represented by low-level features that have color, texture, shape, spatial relationships and etc. There is the gap between the high-level semantic content and the low-level image feature. This is called the semantic gap problem [10][15]. The semantic gap problem is mainly due to low-level image feature that can not map to the high-level semantic content smoothly. Thus identifying and classifying high-level semantic content in an image is very difficult. Therefore, how to reduce the gap between high-level semantic content and low-level features in an image is still an interesting issue.

The good image classification greatly improves the effectiveness of content-based image retrieval systems and solves semantic gap [3][13]. Thus, many schemes have been proposed many relevant researches [2][8][9]. Such as Tsai et al. [12](2005) used low-level feature combined with support vector machine (SVM) to image classification. Chow and Rahman [4](2007) used tree structure combined with global features and regional features of image to image classification.

In general, latent semantic analysis has good results in text data. In our paper, we use latent semantic analysis to find the correlations between low-level features and high-level semantics and reduce semantic gap. We propose the scene image classification based on the semantics of image features. We use statistical histogram of color features and use latent semantic analysis to find the latent semantic features in the images. Then, we use the latent semantic feature to classify various scene images. Experimental results show that the proposed method in the classifying scenes images has better correct classification ratio than frequency feature analysis.

Framework of this paper as follows: In section II describes the latent semantic analysis (LSA) and low-level feature. In section III describes how to use latent semantic analysis to get important features and calculate scene similarity. Experimental results are provided in section IV. The conclusion of this study and future research directions are given in Section V.

II. Previous work

In this section, we will explore the latent semantic analysis (LSA) and the low-level features of images.

A. Latent Semantic Analysis (LSA)

Latent semantic analysis is a vector space model for index and retrieval of information technology, this method mainly uses singular value decomposition (SVD) and reduces dimension as the basis of the theory of modules to find implicit concept in document [1][5][7]. Some of the latent semantic analysis researches are used in text or web pages. SVD is a decomposition technology and uses the singular value decomposition to reduce the size of the high-dimensional matrix [6]. Through dimension reduced can extracted important information in the semantic space. After the decomposition, new matrix and original matrix of the features are similar and new matrix can more accurately describe the matrix of the hidden semantic concept.

Let T denote a matrix with size $m \times n$. After singular value decomposition, matrix T will be transfer to three matrices, matrix S represents the semantic space for a diagonal matrix, matrix A and matrix B^T represent the semantic space of keywords and document for an orthogonal matrix, that is $T = ASB^T$, as shown in Fig. 1. After the singular value decomposition, matrix S of singular value will be sorted by large to small, that re-ordering of the matrix S will require re-ordering the columns of A and the rows of B^T. Singular values corresponding to the feature of larger are more important. We will use dimension reduction to popup the important scene features, that is $T' = A'S'B^{T}{}'$, as shown in Fig. 2. Matrix T can be expressed as:

$$T = \sum_{i=1}^{r} a_i s_i b_i^T \tag{1}$$

B. Low-level image features

There are plenty low-level features in the image content, such as color, texture, shape. In general, color is one of the most widely used low-level features. A lot of space of colors can be used such as RGB, LAB, HSV and CMYK. HSV is

This work was supported by National Science Council under Grant NSC 98-2221-E-324-042.

Chu-Hui Lee Author is with the Chaoyang University of Technology, Wufong Township Taichung County, 41349 Taiwan (R.O.C.). Phone: 886-4-23323000 ext 4388; fax: 886-4-23742337; e-mail: chlee@ cyut.edu.tw.

Kun-Cheng Chiang is with the Chaoyang University of Technology, Wufong Township Taichung County, 41349 Taiwan (R.O.C.). E-mail: s9714627@cyut.edu.tw.

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

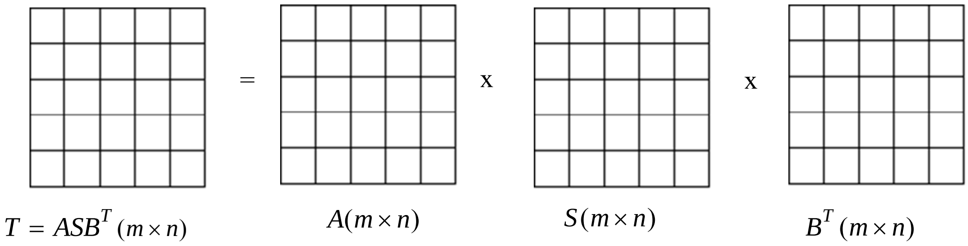

Fig.1. Latent Semantic Analysis [7]

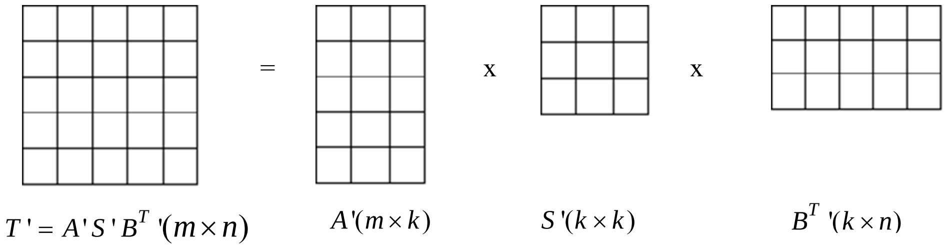

Fig.2. Reduced Dimension [7]

closest to human vision effect among the color features. HSV color space is shown in Fig. 3.The value of H refers to hue between 0 and 360, the value of S that represents the saturation is between 0 and 1 and the value of V means the brightness is between 0 and 1. The three elements make up the color space.

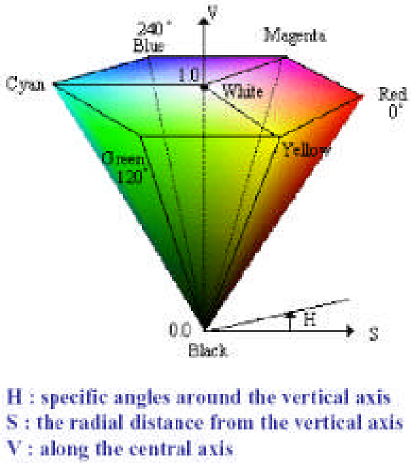

Fig.3. HSV color attribute model [14]

Among the color features technologies, the color histogram proposed by Swain and Ballard is the most popular [11]. The color space is quantified into several bins. Each pixel in the image is distributed into the corresponding bin according the color feature. Values for every bin are recorded to acquire the color histogram of the image. The reason of the popularity is that histogram simply considers the frequency of each color. It can also keep the invariant after rotation of image.

III. Scene Image Classification

Most of scene images contain a lot of semantic information, using semantic information to classify each image to the appropriate category is an important challenge. In this paper, we extracted color feature and applied latent semantic analysis to popup the semantic features. Then, the system calculated the scene similarity for each image. The image would be classified to the suitable scene based on voting mechanism. The proposed process is shown in Fig. 4:

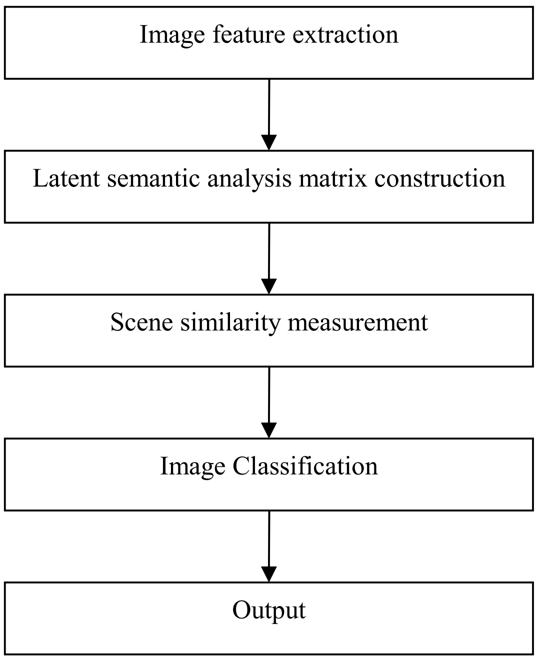

Fig.4. Scene image classification process

A. Image feature extraction

The HSV color attribute model are used to present color features. Among the color features technologies, the color histogram proposed by Swain and Ballard is the most popular [11]. The color space is quantified into several bins. Each pixel in the image is distributed into the corresponding bin according the color feature. Values for every bin are recorded to acquire the color histogram of the image. The reason of the popularity is that histogram simply considers the frequency of each color. It can also keep the invariant after rotation of image. The HSV color feature is quantified into several bins and consists the color histogram to get the color feature vector $c = (f_1, f_2, \cdots, f_k)$。

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

$$A=\begin{bmatrix} sf_{11} & sf_{12} & \cdots & sf_{1j} \\ sf_{21} & sf_{22} & \cdots & sf_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ sf_{i1} & sf_{i2} & \cdots & sf_{ij} \end{bmatrix} \; S=\begin{bmatrix} s_{11} & 0 & \cdots & 0 \\ 0 & s_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_{ij} \end{bmatrix} \; B^T=\begin{bmatrix} sc_{11} & sc_{12} & \cdots & sc_{1j} \\ sc_{21} & sc_{22} & \cdots & sc_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ sc_{i1} & sc_{i2} & \cdots & sc_{ij} \end{bmatrix}$$

$$\Big\downarrow \text{ Reduced dimension}$$

$$A'=\begin{bmatrix} sf_{11}' & sf_{12}' & \cdots & sf_{1k}' \\ sf_{21}' & sf_{22}' & \cdots & sf_{2k}' \\ \vdots & \vdots & \ddots & \vdots \\ sf_{i1}' & sf_{i2}' & \cdots & sf_{ik}' \end{bmatrix} \; S'=\begin{bmatrix} s_1' & 0 & \cdots & 0 \\ 0 & s_2' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_k' \end{bmatrix} \; B^T{}'=\begin{bmatrix} sc_{11}' & sc_{12}' & \cdots & sc_{1j}' \\ sc_{21}' & sc_{22}' & \cdots & sc_{2j}' \\ \vdots & \vdots & \ddots & \vdots \\ sc_{k1}' & sc_{k2}' & \cdots & sc_{kj}' \end{bmatrix}$$

Fig.6. Singular value decomposition and reduced dimension

B. Latent semantic analysis matrix construction

A set of images in the same scene will be extracted feature to constitute a matrix H_{ij}, as shown in Fig. 5, where i_j represents the *j*'th image, f_i represents the *i*'th color bin and w_{ij} represents frequency of *j* feature in the image *i*. Therefore, each column represents the feature vector of each image and each row represents the frequency of each feature in all images. We want to find distinguishing features for a specific scene, therefore, the matrix for a specific scene is constructed. Then latent semantic analysis of singular value decomposition is used to analysis the matrix. Re-ordering the matrix of *S* will require re-ordering the columns of *A* and the rows of B^T. The feature corresponding with larger singular values is more important, therefore, after the reorganization of the matrix will be accordance with the features of importance in scene by large singular values to small. After reduced dimension, the pre-k important features for specific scene are selected, as shown in Figure 6.

$$\begin{array}{c c} & \begin{matrix} i_1 & i_2 & \cdots & i_j \end{matrix} \\ \begin{matrix} f_1 \\ f_2 \\ \vdots \\ f_i \end{matrix} & \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1j} \\ w_{21} & w_{22} & \cdots & w_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ w_{i1} & w_{i2} & \cdots & w_{ij} \end{bmatrix} \end{array}$$

Fig.5. A scene matrix H_{ij}

C. Scene similarity measurement and image classification

In the LSA, the feature corresponding to larger singular values is more importance. We use this characteristic to find out pre-*k* suitable features for each scene. If there are *r* scenes, then *r* times LSA analysis are applied to find the suitable features for *r* kinds of scenes. Scene similarity measurement for an image in scene *i* is defined as (2).

$$C_i=\frac{(\log s_1^{i}{}')^n}{\log s_{total}}f_1^{i}{}'+\frac{(\log s_2^{i}{}')^n}{\log s_{total}}f_2^{i}{}'+\cdots+\frac{(\log s_j^{i}{}')^n}{\log s_{total}}f_j^{i}{}'+\cdots+\frac{(\log s_k^{i}{}')^n}{\log s_{total}}f_k^{i}{}' \qquad (2)$$

The specific feature *j* is determined in the LSA process, where $s_j^{i'}$ is the singular value corresponding to feature *j* and $s_{total}=(\log s_1^{i}{}')^n+(\log s_2^{i}{}')^n+\cdots+(\log s_k^{i}{}')^n$. Let *n* denote the magnified factor. Different scene has its own magnified factor *n* to enhance the important ratio. The $f_j^{i}{}'$ denote frequency of specific feature *j* in an image, where *j*=1 to k. (2) measures the degree of an image that belongs to scene *i*. Each image can calculate the *r* degrees for different scene, that is $C_1, C_2, \cdots, C_r$. Therefore, we can classify the image to the scene *s* when C_s is the maximum among *r* degree values. It is a majority function as denoted as:

$$Max(C_1, C_2, \cdots, C_k) = C_s \qquad (3)$$

IV. Experiment

The experiment images were downloaded from http://wang.ist.psu.edu/docs/related/. We choice five different classifications that are coast、flower、sky、forest、sunset. Those classifications contain 99、100、79、100 and 100 color images respectively. All of the images have dimensions of 100 x 100 pixels. We use HSV color attribute model and quantify into 40 bins which composed of 10 bins for H value、2 bins for S value and 2 bins for V value respectively.

The performances of our method are measured by correct classification ratio (C.C.R.) as (4). In order to assess the proposed method's validity, we compared our method with feature frequency analysis. In the frequency analysis, we select the top six color features of high frequency in a group of the same scene images. The selected features are used for classification. The average performance of results is shown in Table 1. The average of each scene correct classification rates is between 24.05% and 75%. The overall classification rate is 54.39%.

$$C.C.R = {Correct\ classification\ image} \big/ {Test\ image} \qquad (4)$$

In our method, we use the LSA to select the top six important features and to calculated scene similarity measurement in (2) with the magnification factor n=5、5、15、1 and 7 of coast、sunset、sky、forest and flower respectively. The average performance of results is shown in table2. The average of each scene correct classification rates is between 64% and 88%. The overall classification rate is 72.17%. For the table1 and table2, our method is superior to feature frequency analysis; this indicated that the method used in the image classification is efficiency.

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Table1. The average performance of frequency analysis

Frequency analysis			
Scenes	Test image	Correct	C.C.R
Coast	99	73	73.73%
Sky	79	19	24.05%
Sunset	100	75	75%
Forest	100	45	45%
Flower	100	48	48%
Total	478	260	54.39%

Table2. The average performance of LSA

LSA			
Scenes	Test image	Correct	C.C.R
Coast	99	74	74.74%
Sky	79	51	64.55%
Sunset	100	88	88%
Forest	100	64	64%
Flower	100	68	68%
Total	478	345	72.17%

V. CONCLUSION

In this paper, we adopt latent semantic analysis to selected important semantic features in scenes and classify the scene image. In general, latent semantic analysis has good results in text data. We modify the model to apply on image data to find latent semantic and develop a suitable scene similarity measurement. Experimental results indicate that our method indeed has better performance in scene image classification. Future work will be to explore how to integrate other low-level features such as texture and shape to raise the retrieval accuracy further.

REFERENCES

[1] Vinay Bhat, Tim Oates, Vishal Shanbhag, and Charles Nicholas, "Finfing aliases on the web using latent semantic analysis," *Data and Knowledge Engineering*, vol.49, 2004, pp.129-143.

[2] Chuan-Yu Chang, Hung-Jen Wang and Chi-Fang Li, "Semantic analysis of real-world images using support vector machine," *Expert Systems with Applications*, vol.36, 2009, pp.10560-10569.

[3] Ya-Chun Cheng and Shu-Yuan Chen, "Image classification using color, texture and regions," *Image and Vision Computing,* vol.21, 2003, pp.759–776.

[4] Tommy W.S. Chow and M.K.M. Rahman, "A new image classification technique using tree-structured regional features," *Neurocomputing*, vol.70, 2007, pp. 1040-1050.

[5] S. Deerwester, Susan Dumais, G. W. Furnas, T. K. Landauer and, R. Harshman, "Indexing by Latent Semantic Analysis," *Journal of the American Society for Information Science*, vo.41, no.6, 1990, pp.391–407.

[6] Gene H. Golub and Charles F. Van Loan, Matrix Computations, *Johns-Hopkins, Baltimore, Second Edition*, 1989.

[7] April Kontostathis and William M. Pottenger, "A framework for understanding Latent Semantic Indexing (LSI) performance," *Information Processing and Management*, vol.42, 2006, pp.56-73.

[8] Chu-Hui Lee and Chi-Hung Hsu, "Scene Classification Using Efficient Low-level Feature Selection," *The 2008 IAENG International Conference on Imaging Engineering*, 2008, Hong Kong.

[9] Chu-Hui Lee and Yi-Ling Yin, "A Semantic Consumer Image Classification System Based on Multi-layer Color Analysis," *Graphics and Image Processing,* 2007, Japan.

[10] Ying Liu, Dengsheng Zhang, Guojun Lu and Wei-Ying Ma, "Asurvey of content-based image retrieval with high-level semantics," Pattern Recognition, vol.40, 2007, pp.262 – 282.

[11] M. J. Swain and D. H. Ballard, "Color indexing," International Journal of Computer Vision, Vol. 7, No. 1, pp. 11-32.

[12] Chih-Fong Tsai, "Training support vector machines based on stacked generalization for image classification," *Neurocomputing* , vol.64, 2005, pp.497–503.

[13] A. Vailaya, A.K. Jain, H.J. Zhang, "On image classification: city images vs. landscapes," *Pattern Recognition,* vol,31, 1988, pp.1921–1935.

[14] J. Wales, HSV Color Space, *Wikipedia encyclopedia*, 2002.

[15] Rong Zhao and W.I. Grosky, "Negotiating the semantic gap: from feature maps to semantic landscapes," Pattern Recognition, vol.35, 2002, pp.593–600.

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)