Image Classification for Digital Archive Management

Cheng-Hung Li, Chih-Yi Chiu, Hsiang-An Wang

Institute of Information Science, Academia Sinica,
128. Sec. 2. Academia Road, Nankang, Taipei 115, Taiwan
{chli, cychiu, sawang}@iis.sinica.edu.tw

Abstract. As tools and systems for producing and disseminating image data have improved significantly in recent years, the volume of digital images has grown rapidly. An efficient mechanism for managing such images in a digital archive system is therefore needed. In this study, we propose an image classification technique that meets this need. The technique can be employed to annotate and verify image categories when gathering images. The proposed method segments each image into non-overlapping blocks from which color and texture features can be extracted. Support Vector Machine (SVM) classifiers are then applied to train and classify the images. Our experimental results show that the proposed classification mechanism is feasible for digital archive management systems.

1 Introduction

There has been an enormous growth in the number of digital image files in digital libraries in recent years; however managing these files has become increasingly difficult. To reduce the cost of digital image management, an automatic tool for digital libraries is required. In digital image management, research activities include image retrieval [1], digital rights management [2], and image annotation [3]. These research activities are very much related to image classification techniques. In addition, if there is a large volume of images, an efficient classification technique is very useful for maintaining the quality and correctness of the digital library. Take, for example, a set of naïve images in which no metadata has been embedded. An image classification technique can be applied to annotate the images and store them in appropriate categories. Furthermore, we can use the classification technique to verify the correctness of manually labeled image categories and dubious images can be identified and double-checked to maintain the quality of the digital library.

In this paper, we integrate an image classification technique with a digital archive management system. Image classification in our system comprises two components: off-line image training and on-line image classification. In the first component, each image is segmented into non-overlapping blocks. A color and texture histogram of each image block is then extracted as its feature representation, and the Support Vector Machine (SVM) technique is applied to train several image category classifiers. In on-line classification, we extract the features of the image and use the trained SVM classifiers to store gathered images in appropriate categories.

The remainder of this paper is organized as follows. In the next section, we discuss related works. In Section 3, we address SVM training and classification, together with image feature extraction. Section 4 describes the implementation platform and experimental results. Finally, we present our conclusions and discuss future research directions in Section 5.

2 Related Work

Many content-based image retrieval (CBIR) systems have been developed since the early 1990s. CBIR for large image databases [5] has been studied extensively in relation to image processing, information retrieval, pattern recognition, and database management. Most CBIR research efforts have focused on finding effective feature representations for images. First, the histogram method has been widely used for various image categorization problems. Szummer and Picard [4] and Vailaya et al. [6] use color histograms to classify indoor and outdoor images. Chapelle et al. [7] apply SVM to classify images containing a generic set of objects based on color histogram features. The above works show that histograms can be computed at a low cost and are effective for certain classification cases. However, a major shortcoming of general histogram representation is that spatial structural information is missed. Hence, other features such as texture and shape have been proposed to solve this problem [8-11].

A number of methods that extract features by dividing an image into blocks have been proposed. Gorkani and Picard [12] first divide an image into 16 non-overlapping, equal-sized blocks, and then compute the dominant orientations for each block. Finally, the image is classified as a city or suburb scene by the major orientations of the blocks. Wang et al. [13] develop an algorithm for classifying graphs or photographs. The classifier divides an image into blocks, and classifies each block into one of two categories based on the wavelet coefficients in high frequency channels. If the percentage of blocks classified as photograph is higher than a pre-defined threshold, the image is marked as a photograph; otherwise, it is marked as a graph. Yu and Wolf [14] propose a Hidden Markov Model (HMM) classifier for indoor/outdoor scene classification that is trained on vector quantized (VQ) color histograms of image blocks. Li and Wang [15] suggest that a particular category of images can be captured by a two-dimensional multi-resolution HMM trained on the color and texture features of the image blocks. Yanai [16] describe a generic image classification system that uses images gathered automatically from the World Wide Web as training images for generic image classification, instead of using image collections gathered manually. Murphy et al. [17] propose four graph models that associate the features of image blocks with objects and perform joint scene and object recognition.

3 Image Classification

Image classification in our system comprises two components: off-line image training and on-line image classification. In the training component, we extract features from images sourced by the archive system and link them to categories through SVM training. Next, in the classification component, we classify unknown images into one of the categories using SVM classifiers. Figure 1 shows the workflow of the image training and classification modules in our digital archive management system.

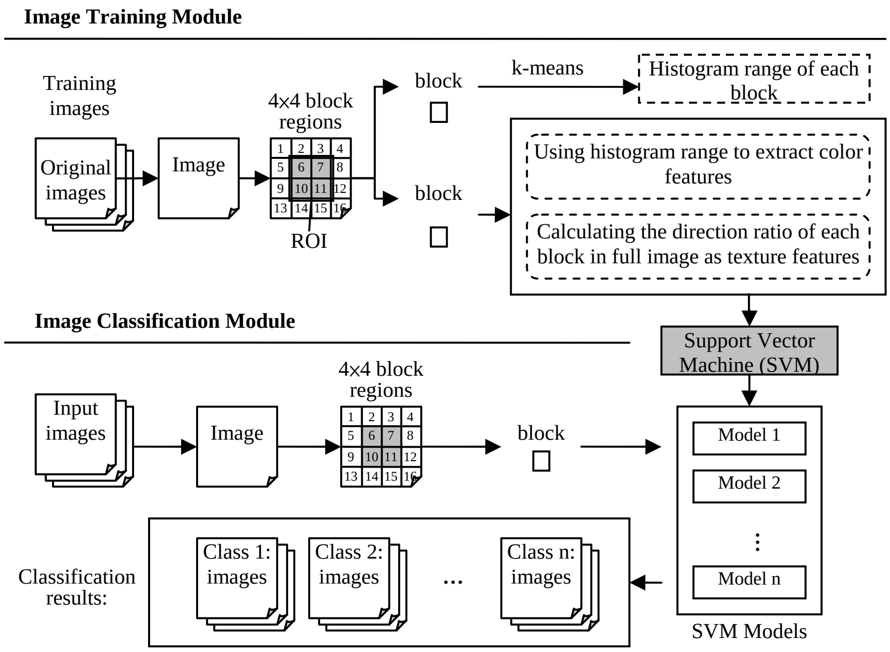

Figure 1. An overview of the image training and classification components.

3.1 Feature Extraction

Since archive images in the same category have similar characteristics, an image's content, such as its color and texture, can be used to represent the characteristics of the category. For example, an image of Archaeological Data usually has an ancient object in the center of the image, while a Han Wooden Slip image usually exhibits natural wood colors, but the use of ultra red rays may reveal detail not visible in the chromatic image. In the following, we describe the color and texture feature extraction process.

To obtain color features, we first normalize the size of a training image into 96×120 pixels, and segment them into 4×4 non-overlapping blocks (see Figure 1). Next a color histogram vector for each block is computed. A mean vector is also computed by averaging the color histogram vectors of the sixteen blocks. We then calculate the color histogram intersection between each block's histogram vector and the mean vector. Finally we have sixteen similarity values as our color feature representation.

For texture feature extraction, we apply Canny's edge detection algorithm [18] to each block. First, we use a Gaussian mask to smoothing the image's noise, and use a pair of 3x3 Sobel convolution masks to estimate the gradient in the *x*-direction (*Gx*) and *y*-direction (*Gy*). Then we use the gradient of *Gx* and *Gy* to compute the pixel direction as follows:

$$Direct = 180 - ((\tan^{-1}(Gy/Gx)) * (180/\pi)); \quad (1)$$

For quantization, an edge direction is given a setting based on the range it falls within. An edge direction in the range (0~22.5 & 157.5~180 degrees) is set to 0 degrees. Any edge direction falling in the range (22.5~67.5 degrees) it is set to 45 degrees; in the range (67.5~112.5 degrees) it is set to 90 degrees; and in the range (112.5~157.5 degrees) is set to 135 degrees. We calculate the direction histogram of four region-of-interest (ROI) blocks (blocks 6, 7, 10, and 11 in Figure 1) in the image to determine its texture feature representation.

3.2 SVM Training and Classification

The Support Vector Machine (SVM) technique utilizes the minimal structured risk principle in machine learning theory, and generally performs effectively on pattern classification problems without incorporating domain knowledge [19]. It is based on a binary classification method, and looks for a hyper-plane to divide objects into two classes. It can ensure the minimal error rate of classification. The main advantage of SVM is that it can process linearly inseparable cases by training and analyzing the given data to generate support vectors. After removing extreme data from the training data, it builds the support vectors into a model. It then classifies any existing test data into the appropriate classes.

The color and texture features are used for training and classification. We normalize the features of training images as follows:

[Label] [Index1]:[Value1] [Index2]:[Value2] ...

For example, a feature vector of the first image could be:

1 1:0.268169 2:0.332564 3:0.246752 4:0.101631 5:0.202088 6:…etc

where [Label] is the correct category number of the image, and [Index] is series of feature values of the image. In this study, we employ the radial basis function as the kernel function.

4 Experimental Results

Our image collections contain Archaeological Data, Buddhist Rubbings, Han Dynasty Stone Relief Rubbings, Fu-Ssu Nien Ancient Books, and Han Wooden Slips. An image database, built by Redhat Linux 7.3, is divided into two components: the multimedia center web system (MMC) and the image processing daemon (IPD). These two components share the kernel function module layer and the database and file storage layer. The MMC executes the Servlet through JavaBeans invoked and compiled by the Application Server. The IPD uses the API through the kernel function module layer to communicate with the database and file storage layer. The classification module was developed by Java-Servlet and SVM tools [20]. Figure 2 shows some of the class images from the digital archive management system.

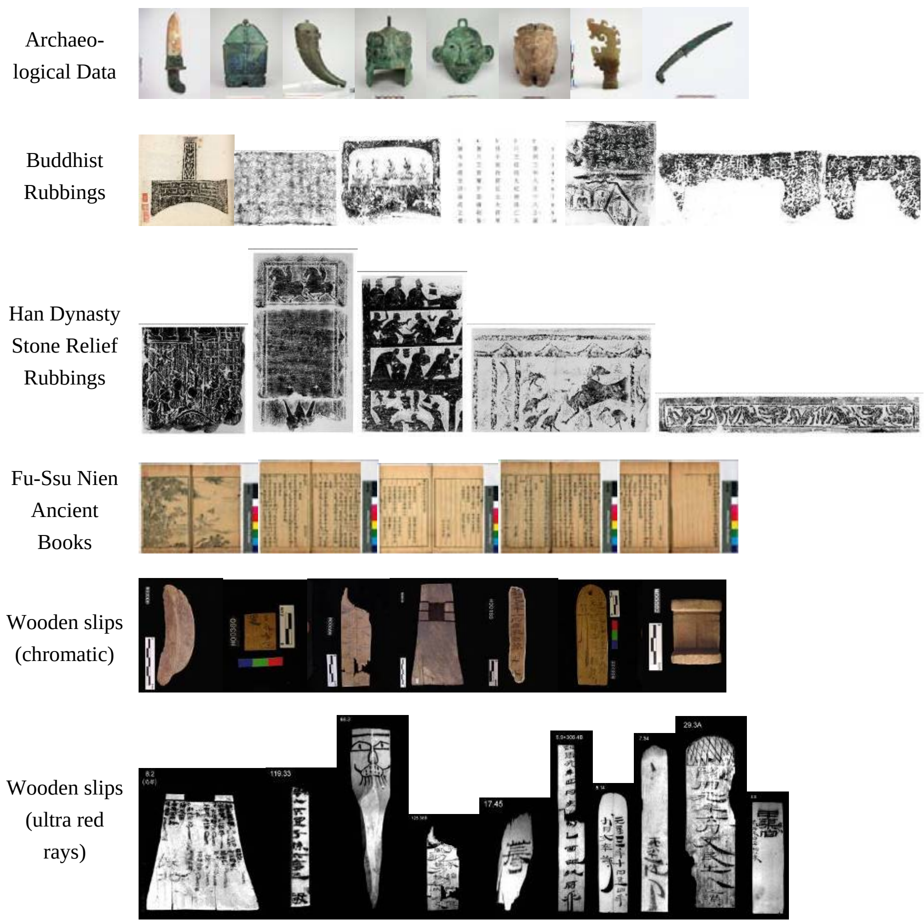

Figure 2. Part of each class of images from the digital archive management system.

Table 1 lists the image collections. There are 2716 collected images (# CI) 385 training images (# TRI), and 2331 testing images (# TI). In the training stage, we use a different cross-validation parameter (# CV) in each class to increase the accuracy ((%) CVA). The recall rate, precision rate, and F-measure are used to evaluate the classification performance. Table 2 shows the accuracy rates derived from experiments that evaluated the proposed classification technique. The recall rate (Rec. (%)) is denoted by TI_R / TI_T, the precision rate (Pre. (%)) is defined by TI_R / (TIr + TIw), and the F-measure is the harmonic mean of the recall and the precision rates.

TI_R, TIw, and TI_T are the number of correctly clustered images, the number of incorrectly clustered images, and the number of test images for each class, respectively. All values are represented in percentages. A test image estimated by each SVM classifier returns a decision value, and the classifier with the highest decision values is assigned as the category of the test image. ADV is average of the decision values in the classifier. We obtained an average F-measure of 86.67%, as shown in Table 2.

Table 1. Image data collections.

Class	**Archaeo-logical Data**	**Buddhist Rubbings**	**Han Dynasty Stone Relief Rubbings**	**Fu-Ssu Nien Ancient Books**	**Wooden slips (chromatic)**	**Wooden slips (ultra red rays)**	total / avg
# CI	154	498	499	498	569	498	2716
# TRI	36	72	72	72	71	62	385
# TI	118	426	427	426	498	436	2331
# CV	5	50	40	5	5	15	
(%) CVA	98.96	90.65	89.87	100	98.96	94.55	

Table 2. Experimental results of image classification.

Class	**Archaeo-logical Data**	**Buddhist Rubbings**	**Han Dynasty Stone Relief Rubbings**	**Fu-Ssu Nien Ancient Books**	**Wooden slips (chromatic)**	**Wooden slips (ultra red rays)**	total / avg
ADV	**1.571186**	**1.007611**	**0.958000**	**0.989020**	**0.903878**	**0.982123**	**1.0728**
Rec. (%)	114/118 **96.61**	339/426 **79.58**	281/427 **65.81**	421/426 **98.83**	498/498 **100.00**	352/436 **80.73**	**86.93**
Pre. (%)	114/122 **93.44**	339/457 **74.18**	281/358 **78.49**	421/422 **99.76**	498/574 **86.76**	352/398 **88.44**	**86.85**
F (%)	**95.00**	**76.78**	**71.59**	**99.29**	**92.91**	**84.41**	**86.67**

Figures 3 and 4 show the classification results of Archaeological Data and Fu-Ssu Nien Ancient Books in the digital archive management system. The results of Archaeological Data, Fu-Ssu Nien Ancient Books, and Han Wooden Slips (ultra red rays) are superior to the Buddhist Rubbing, Han Dynasty Stone Relief Rubbings, and Han Wooden Slips (chromatic). We assume that the color tones of the first three categories are similar. The image sizes in the last three categories are irregular so that the images have different textures. From our experimental results, it is clear that the average F-measure is extremely effective for image classification.

Figure 3. The classification results of Archaeological Data

Figure 4. The classification result of Fu-Ssu Nien Ancient Books

5 Conclusions and Future Work

In this paper, we have integrated an image classification technique with a digital archive management system. Images are segmented into blocks to extract their color and texture histograms as their feature representations, which are then trained to generate Support Vector Machine (SVM) classifiers. Our experiments show very promising results for the proposed technique. We believe the proposed image classification technique would be very useful for automatic annotation and validation in large image databases.

In our future work, we will investigate and compare the effectiveness of other image features (shape, spatial layout, etc.) and classification techniques (Bayesian networks, decision trees, etc.). In addition, we will build a multi-level semantic annotation mechanism to facilitate image management and searching.

Acknowledgements

This work was partially supported by the following grants: NSC 94-2422-H-001-006, 94-2422-H-001-007, and 94-2422-H-001-008. The authors would like to thank Dr. L. F. Chien and various professional parties in the NDAP Research & Development of Technology Division for their contributions to this paper.

Reference

1. Saux, B.L., Amato, G.: Image recognition for digital libraries. Proceedings of the 6th ACM SIGMM International Workshop on Multimedia Information Retrieval, (2004) 91-98
2. Liu, Q., Safavi-Naini, R., Sheppard, N.P.: Digital rights management for content distribution. Proceedings of the Australasian Information Security Workshop Conference on ACSW Frontiers, 2 (2003)
3. Cheng, P.-J., Chien, L.-F.: Effective Image Annotation for Search Using Multi-Level Semantics. International Journal of Digital Libraries: Special Issue on Asian Digital Libraries, (2004) 258-271
4. Szummer, M., Picard, R.W.: Indoor-outdoor image classification. In Proc. IEEE Int'l Workshop on Content-Based Access of Image and Video Databases, (1998) 42–51
5. Kim, D.-H., Chung, C.-W.: QCluster: Relevance Feedback Using Adaptive Clustering for Content-based Image Retrieval. In Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data, (2003) 599-610
6. Vailaya, A., Figueiredo, M.A.T., Jain, A.K., Zhang, H.-J.: Image classification for content-based indexing. IEEE Transaction on Image Processing, Vol. 10, No. 1, (2001) 117-130
7. Chapelle, O., Haffner, P., Vapnik, V.N.: Support vector machines for histogram-based image classification. IEEE Transactions on Neural Networks, 10(5) (1999) 1055–1064
8. Do, M.N., Vetterli, M.: Rotation Invariant Texture Characterization and Retrieval Using Steerable Wavelet-Domain Hidden Markov Models. IEEE Transactions on Multimedia, 4 (2002) 517-527
9. Jing, F., Li, M., Zhang, H.-J., Zhang, B.: An Effective Region-based Image Retrieval Framework. In Proceedings of the 10th ACM International Conference on Multimedia, (2002) 456-465

10. Safar, M., Shahabi, C., Sun, X.: Image Retrieval by Shape: A Comparative Study. In Proceedings of IEEE International Conference on Multimedia and Expo (ICME'00), (2000) 141-144
11. Zhang, D., Lu, G.: Generic Fourier Descriptors for Shape-based Image Retrieval. In Proceedings of IEEE International Conference on Multimedia and Expo (ICME'02), 1 (2002) 425-428
12. Gorkani, M.M., Picard, R.W.: Texture orientation for sorting photos "at a glance". In Proceedings of the 12th IAPR International Conference on Pattern Recognition,, (1994) 459-464
13. Wang, J.Z., Li, J., Wiederhold, G.: SIMPLIcity: Semantics-sensitive integrated matching for picture libraries. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(9) (2001) 947–963
14. Yu, H.H., Wolf, W.H.: Scenic classification methods for image and video databases. In Proc. SPIE Int'l Conf. on Digital Image Storage and archiving systems, 2606 (1995) 363-371
15. Li, J., Wang, J.Z.: Automatic linguistic indexing of pictures by a statistical modeling approach. IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(9) (2003) 1075-1088
16. Yanai, K.: Generic Image Classification Using Visual Knowledge on the Web. In Proceedings of the Eleventh ACM International Conference on Multimedia, (2003) 167-176.
17. Murphy, K., Torralb, A., Freeman, W.T.: Using the forest to see the trees: a graphical model relating features, objects, and scenes. In Advances in Neural Information Processing Systems 16. Cambridge, MA:MIT Press, (2004)
18. Canny, J.: A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(6) (1986) 679-698
19. Moghaddam, B., Yang, M.-H.: Gender Classification with Support Vector Machines. IEEE Int. Conf. on Automatic Face and Gesture Recognition (FG), (2000) 306-311
20. Chang, C.-C., Lin, C.-J.: LIBSVM: a library for support vector machines. (2001). Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm