Recognition of Blood and Bone Marrow Cells using Kernel-based Image Retrieval

Chen Pan†, Xiangguo Yan††, and Chongxun Zheng††

† The School of Mathematics & Computer Engineering, Ningxia University, Yin chuan, 750021, P. R. China.
†† Key Laboratory of Biomedical Information Engineering of Education Ministry, Xi'an Jiaotong University, Xi'an, 710049, P.R. China

Summary
This paper presents a novel cell classification method based on image retrieval by learning with kernel. Cell image is firstly segmented into cytoplasm and nucleus regions in order to keep more spatial information. RGB color histogram of cell and two intensity histograms corresponding to those local regions compose feature vector represents the cell image. Kernel principal component analysis (KPCA) is utilized to extract effective features from the feature vector. The weight coefficients of features are estimated automatically using relevance feedback strategy by linear support vector machine (SVM). Classification depends on the decision distance obtained by SVM and the nearest center criterion. Experimental results on the ten-class task of 400 cells from blood and bone marrow smears show a 90.5% classification accuracy of the method when combined with standardized sample preparation and image acquisition.

Key words:
Classification, KPCA, SVM, feature extraction, blood and bone marrow cells

1. Introduction

The analysis of blood and bone marrow smears is a powerful diagnostic tool for the detection of leukemia. It is a classical and challenging pattern recognition task that always includes two stages: one is object detection/segmentation and the other is object recognition/classification. In this paper, we focus on the solution of the later.

There are many types of cell with different lineage and maturity level in bone marrow. Most of them only have subtle visible differences. It is difficult to achieve a consistent diagnose during microscopic evaluation by subjective impressions of observers. Computer-assisted morphologic cell classification can improve accuracy, objectivity and reproducibility in diagnosis [1][2][3]. Yet, the classification performance extremely lies on the strategy of image feature selection and extraction.

Large numbers of literatures existed for feature extraction or translation of the features given by physicians [2][4]. It leads to the application of known classification methods (e.g. Bayesian classifier) using traditional morphological characteristics (such as geometric, texture, form oriented, color and combined features). In order to discriminate similar classes, much more specific features need be offered and measured. Although economical and efficient in achieving data reduction and insensitive to variations in illumination and viewpoint, such features rely heavily on the extraction of cell features. However, feature detection and measurement techniques developed to date have not been reliable enough to cater to this need [5].

Another interesting way is to consider every pixel as a feature, in which neural network [6] generally operate directly on an image-based representation (i.e. pixel intensity array with rotational invariant features as input). Because the detection and measurement of cell features are not required, this kind of methods has been more practical and reliable compared to the concrete feature extraction. However, training of the classifier is very crucial. Since training samples are often limited, classification with good generalization property is most appealing.

Apart from mentioned above, we notice that differences between two similar classes' images could be discriminated easily by comparison each other. This fact can help us to find those features with obvious otherness between similar classes. Content-based image retrieval (CBIR) is just a comparison technique actually. The main purpose of CBIR is to find the relevant image from a database comparing with user provided query image using only the image information. Many image classification problems can be thought of as image retrieval in a high similarity database [7]. In biomedical applications, recently, David [2] presented an image-retrieval based system for discrimination among malignant lymphomas and two subclasses of leukemia, and Cecilia [3] also proposed same idea for recognition of infected blood cells by malaria parasites. Good results had been shown in their papers. Although these systems are only suited to peripheral blood with few types of cell, they point out a

Manuscript received October 10, 2006.
Manuscript revised October 25, 2006.
.

reasonable direction for cell recognition.

In this paper, we present a CBIR-based method to classify complex cells from blood and bone marrow smears. Color and intensity histograms as image-based feature representation are mainly adopted in our method, and kernel principal component analysis (KPCA) [8] is used to reduce the high dimensionality of the data representation and deal with the nonlinear distribution of features. In order to achieve satisfied retrieval accuracy and generalization performance, a relevance feedback strategy based on linear SVM is presented for training of classifiers. We analyze 10-class task involving monocytic and granulocytic series of white blood cells with new method. Experimental results show that over 90% accuracy could be achieved.

This paper is organized as follows. In section 2, the framework of our method is presented. An efficient and robust image-based representation is discussed in detail. KPCA and linear SVM are briefly introduced and their training procedures are proposed. In section 3, the experimental results are described. Finally discussion and conclusion is given in section 4.

2. The Proposed Method

2.1 Image and its representation

Successful retrieval in high similarity image database needs more restrictions put on the acquisition procedures. Retrieving entries based on subtle differences requires the control of all influencing parameters to reduce unnecessary variation that could affect the retrieval performance. That is to say, the sample preparation and presentation procedures have to be standardized and well defined. And image acquisition should also be standardized by color and geometric calibrations. It is important to notice that the standardizations applied in the acquisition procedures will affect the choice of features in such a way that a good standardization will simplify the features [7].

In our study, nucleated cells in an original microscopic image have been detected and pre-segmented to a series of sub-images by image segmentation algorithm, which was presented in [9]. Each sub-image represents an individual cell. See Fig.1b. Every individual cell image will be regarded as query image to find the relevant images from a ground truth database and decide its class according to similarities.

Two problems should be solved in this work: one is how to construct a ground truth database; the other is how to define an efficient and robust image descriptor. A simple way to solve the former problem is made of two steps: (1) use image segmentation algorithm to obtain individual cell [9], (2) select appropriate samples and label them by pathologist. The later problem is discussed in detail as following.

As mentioned in section 1, feature selection and extraction is key to retrieval. Here, we prefer image-based feature representation rather than concrete features extraction. In such case, it should be noticed that the features translate from pixels could potentially involving information about object's color, texture and shape.

Color histogram is the most traditional way of describing low-level color properties of objects. It provides an efficient representation of color content, and their computation is quite trivial. It is invariant to translation and rotation of objects, partially occlusions, and normalizing the histogram with respect to the object area leads to scale invariance. Histogram is also well suited to describe the content of textured region. A frequently used approach for texture analysis is based on statistical properties of the intensity histogram [10][11]. The shortage of histogram is which does not include any spatial information. Errors perhaps occur when different images have same histograms. Therefore spatial features along with color have been intensively researched [12][13][14]. The better way is to extract spatial/shape information by segmenting the object within the image. By proper choice of region segmentation it is possible to describe some of the spatial relations [7]. Thus, in our retrieval, cell image is segmented into cytoplasm and nucleus regions. It's not a difficult task because individual cell image always has deep stained nucleus region and shallow stained cytoplasm region. A bimodal histogram is often obtained by intensity. Histogram thresholding by classical Otsu's method could get binary regions corresponding to cytoplasm and nucleus, respectively.

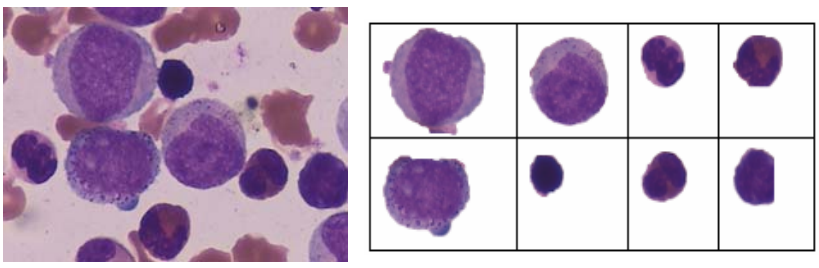


Fig.1. (a)bone marrow cells (b)pre-segmentation results

Motivated by mentioned above reasons, a mixed image representation is adopted in our work, which is composed with three normalized histograms. There are:

- RGB color histogram of individual cell image, $\boldsymbol{h_{3dRGB}}$,
- Intensity histogram of cytoplasm, $\boldsymbol{h_{cytoplasm}}$,
- Intensity histogram of nucleus, $\boldsymbol{h_{nucleus}}$.

Here, these histograms can be assembled to a high dimensional vector $\mathbf{h}=(\boldsymbol{h_{3dRGB}}, \boldsymbol{h_{cytoplasm}}, \boldsymbol{h_{nucleus}})^T$, namely mixed histogram, which gives a standardized image representation. A simplest possible nonlinear remapping of the input that does not affect the dimension is adopted to improve robustness with respect to changes in scale [15],

$$\mathbf{h} = \mathbf{h}^{a} \quad \text{, with } 0 \le a \le 1 \tag{1}$$

As is known, the histogram of image is very sparse. For a high dimension vector with sparse entries, dimensionality reduction could benefit to computation. So a feature extractor to deal with high dimensional data is introduced in the following.

2.2 Kernel feature analysis

Principal component analysis (PCA) is a very popular technique for dimensionality reduction. It's optimal in the sense that it minimizes the mean square error of the L2-distance between the original and the approximated vectors. PCA is a linear method, however, image features is more likely to be nonlinear. Kernel PCA is viewed as a nonlinear extension of the PCA with kernel trick [8]. The idea is that we could implicitly map input data into high dimension feature space via a nonlinear function:

$$\begin{aligned} \Phi &: X \to F \\ \mathbf{x} &\mapsto \phi(\mathbf{x}) \end{aligned} \tag{2}$$

And a similarity measure is defined from the dot product in F as follows:

$$k(\mathbf{x}, \mathbf{x}') \equiv < \phi(\mathbf{x}), \phi(\mathbf{x}') > \tag{3}$$

where the kernel function $k(\cdot,\cdot)$ should satisfy Mercer's condition [16].

Given a set of $\{\mathbf{h}_i\}_{i=1}^{m} \in R^{d}$, $\mathbf{h}_i$ is the mixed histogram of the ith sample in ground truth database with m samples, using the nonlinear mapping and kernel trick defined as in (2) and (3), Kernel PCA mainly depends on the eigen-decomposition problem on a Gram matrix.

$$m\lambda\alpha = \mathbf{K}\alpha \tag{4}$$

where $\alpha = (\alpha^{1}, ..., \alpha^{m})^{T}$ is the expansion coefficient, λ is the eigen-value and $\mathbf{K}$ is the $m \times m$ Gram matrix with element $K_{ij} = < \phi(\mathbf{h}_i), \phi(\mathbf{h}_j) >$.

To extract nonlinear features $\boldsymbol{f}$ from a test point $\mathbf{h}$, KPCA computes dot product between $\phi(\mathbf{h})$ and the n^{th} eigenvector v^{n} in feature space to obtain the projection $\mathbf{f} = (f^{1}, ..., f^{p})$.

$$f^{n} = < v^{n}, \phi(\mathbf{h}) > = \sum_{i=1}^{m} \alpha_i^{n} k(\mathbf{h}_i, \mathbf{h}) \quad n = 1, ..., p \tag{5}$$

where $\mathbf{p}$ $\mathbf{(p<m)}$ is the dimension of extracted features space $\boldsymbol{F}$. KPCA is an efficient nonlinear method for feature extraction. The advantage of it is that data with complex structure could be well clustered in feature space.

2.3 SVM learning for image retrieval

Since the large gap exists between high-level concepts and low-level features, the accuracy of image retrieval is not often satisfied to users' expectations. An interactive learning mechanism called 'relevant feedback' was presented [17]. The basic idea is to build a model according to the relevance information, feedback by users to indicate which images he or she thinks are relevant to the query, and to do retrieval again for better result. With the feedback technique, the available information is not only the feature space itself, but also the relevance information given by users. Obviously, relevant feedback changes the distance metric by re-weighting to features to make relevant images closer. Traditional relevant feedback is interactive mode, here an automatic approach using linear SVM is presented.

SVM is a novel type of learning machine based on statistical learning theory[16]. Let the training set $\boldsymbol{D}$ be $\{(\mathbf{x}_i, y_i)\}_{i=1}^{N}$, with input $\mathbf{x}_i$ and $y_i = \{\pm 1\}$. When the data is linearly separable in F, linear SVM constructs a hyperplane $\mathbf{w} \cdot \mathbf{x}_i + b$ for which separation between the positive and negative examples is maximized. It can be shown that $\mathbf{w} = \sum_{i=1}^{N} \alpha_i y_i \mathbf{x}_i$, where $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_N)$ can be found by solving the following quadratic programming problem:

$$\min \frac{1}{2} \boldsymbol{\alpha}^{T} \mathbf{Q} \boldsymbol{\alpha} - \boldsymbol{\alpha}^{T} \mathbf{I} \tag{6}$$

subject to $\boldsymbol{\alpha} \ge 0$ and $\boldsymbol{\alpha} \cdot \mathbf{y} = 0$.

where, $\mathbf{y} = (y_1, ..., y_N)^{T}$, $\mathbf{I} = (1, 1, ..., 1)$ and $\mathbf{Q}$ has entries $y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j$, where $\mathbf{x}_i \cdot \mathbf{x}_j$ is dot product. When the training set is not separable in F, the SVM algorithm introduces non-negative slack variable $\xi_i \ge 0$. The resultant problem becomes

$$\min \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i}^{n} \xi_i \tag{7}$$

subject to $y_i(\mathbf{w}\cdot\mathbf{x}_i+b)\geq 1-\xi_i$. C is a regularization parameter controlling the tradeoff between model complexity and training error.

The $\mathbf{x}_i$ for which $\alpha_i \neq 0$ are defined as the support vectors, since they determine the optimal hyperplane, the hyperplane with maximal margin. Geometrically, the support vectors correspond to the closest to the optimal hyperlane. The decision function is

$$f(\mathbf{x}) = sign(\sum_{i=1}^{N}\alpha_i y_i < \mathbf{x}_i, \mathbf{x} > +b) \quad (8)$$

The formula (8) shows SVM is neural network essentially, but automatically chooses the net centers $\mathbf{x}_i$ by weights α_i and threshold b in training. Thus, we utilize SVM to weight the features and construct classification model. In order to deal with several classes, a binary classifier with "one against the others" strategy, which compares a given class with all the others put together, is adopted here. In such algorithm, *n* hyperplanes are constructed, where *n* is the number of classes. Each hyperplane separates one class from the other classes.

To training classifiers, images in the ground truth database can be marked as either relevance or irrelevance. We carry out two-class learning algorithm by SVM, and construct a classifier suitable to represent same class's query. Sorting images according to their distance to the hyperplane, the rank of relevance images can be obtained in the result. Repeat such procedure until almost every sample belongs to relevance class is in the top rank. The process is improved from [17] and described below.

①Mark top ***N1*** images into two classes: relevance set I^+ and irrelevance set I^-.

② Prepare for SVM the training data $(\mathbf{f}_i, y_i)$, $\mathbf{f}_i$ represents the extracted features from the mixed histogram $\mathbf{h}_i$ by KPCA.

$$\mathbf{f}_i \in I^+ \cup I^-, y_i = \begin{cases} +1, if\ \mathbf{f}_i \in I^+ \\ -1, if\ \mathbf{f}_i \in I^- \end{cases} \quad (9)$$

③Construct decision function using SVM algorithm

$$d(\mathbf{f}) = \sum_i \alpha_i y_i < \mathbf{f}_i, \mathbf{f} > + b \quad (10)$$

④Calculate the score for each image Ii in the database.

$$score(I_i) = d(\mathbf{f}_i) \quad (11)$$

⑤Sort all images by score, if top ***N2*** rank include almost all samples belong to one class, stop the training; else repeat until an iterative limit arrived.

Due to SVM has good generalization performance, the repeat times of feedback is not over 3 in our study when the parameters be well selected.

2.4 Similarity evaluation

Compared formula (8) with (10), we ignored the function ***sign(.)*** in later. So $d(\mathbf{f})$ is a decision distance from each image to the separating hyperplane. Moreover, in relevance feedback, we concern only one class samples rather than all the others classes' samples, so the two-class SVM used here is more likely one-class SVM [18]. The separating hyperplane shapes a smallest enclosing sphere which encloses almost samples belong to one class in the feature space. We can test the cell images one by one, and assign each vector $\mathbf{h}_i$ to the class which minimizes $d(\mathbf{f}_i)$, i.e. the nearest center criterion is used here.

2.5 Parameters adjusting for linear SVM

Only one parameter ***C*** govern the performance of linear SVM classifier (see formula (7)), i.e. control the shape of the enclosing sphere (separating hyperplane). The size of ***C*** determines the number of outliers. As ***C*** is decreased, the number of outliers increases. While the number of support vectors also depends on ***C***. We observed that as ***C*** is decreased, the number of SVs decreases since the increased number of outliers makes the shape smoother, and thus easier to describe with less support vectors.

Therefore, we prefer such case, which is every relevance sample can be included in an enclosing sphere and the enclosing sphere can be constructed with smallest support vectors. In this case, the optimal size of ***C*** that corresponds to smallest SVs can be selected automatically.

3. Experimental results

In our classification system, the ground truth database is made of 10 classes' white blood cell from bone marrow smears involving monocytic and granulocytic series, and each class includes 15 samples labeled by pathologist. That is to say, SVM classifiers will be trained only with 150 prototypes. Fig.2 shows a part of prototypes. Over one hundred images include 400 cells were acquired under similar condition compose the test dataset. Total images in this paper vary a little in color and intensity due to the inconsistent staining technique, variation of the thickness of smear and illumination.

In our experiment, the color histogram is quantized from ***256*256*256*** to ***32*32*32*** bins and intensity histogram is quantized from 256 to 64 bins. Three parameters need to be adjusted respectively according to experiments in our system. They are: (1) ***a***, the coefficient of nonlinear remapping of the input, i.e. the exponentiation of each component of the input vector; (2) ***d***, degree of the polynomial kernel $K(\mathbf{x}, \mathbf{x}') = (\mathbf{x}\cdot\mathbf{x}' + 1)^d$; (3) ***p***, the

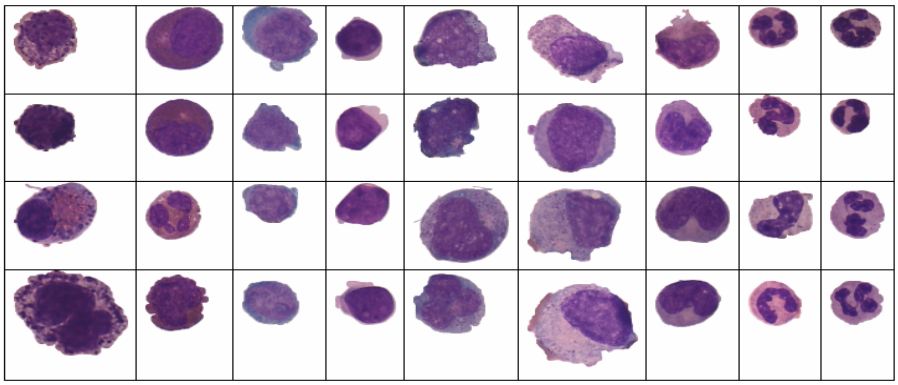
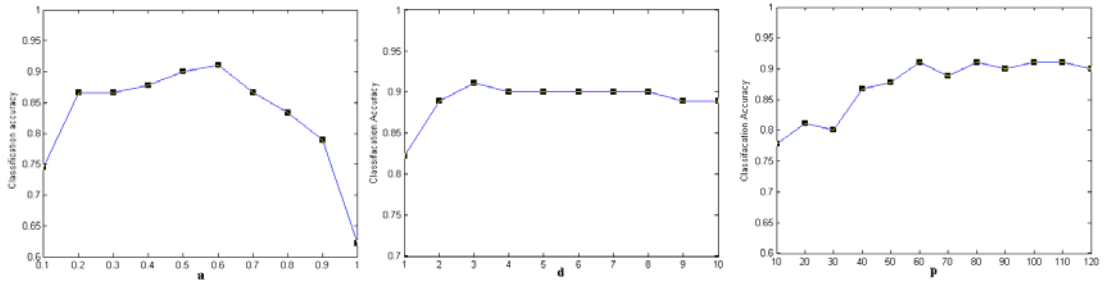


Fig.2 A part of prototypes of nine-class white cells in ground truth database



(a)The effect of ***a*** ***(d=3,p=100)*** (b)The effect of ***d*** ***(a=0.6,p=100)*** (c)The effect of ***p*** ***(a=0.6,d=3)***

Fig.3. The effect of three parameters

dimension of extracted features by KPCA (see formula (5)).

In order to estimate reliable parameters, a five-fold cross-validated classification was implemented by randomized splitting of the ground truth data into 5 equal sub-test-sets. We investigated the effect of these parameters, Fig.3 (a), (b) and (c) show the average accuracy curves of classification corresponding to single variation of those parameters, respectively. The classification accuracy (CA) is defined as follows:

$$\text{classification accuracy} = \frac{\sum N_{correct}}{N_T} \qquad (12)$$

where, $N_{correct}$ is the number of samples to be classified correctly in test set, and N_T is the total number of samples in test set.

Fig.3(a) shows a proper ***a*** can improve the performance of classification. The following gives us reasons to believe that ***a***-exponentiation may improve robustness with respect to changes in scale. If the histogram component x_{col} is caused by the presence of color *col* in some object, we increase the size of the object by some scaling factor ***s***, the number of pixles is multiplied by s^2, and x_{col} is multiplied by the same factor. The ***a***-exponentiation could lower this quadratic scaling effect to a more reasonable s^{2a}, with $a<1$ [15].

Fig.3(b) shows when the ***d*** increases, the CA is not improved as direct proportion of ***d***. It just preserved in a high level. That suggests us that a suitable lower ***d*** could obtain good performance rather than making the dimension of feature space even higher. The computations may be faster and the performances comparable.

Fig.3(c) shows when the dimension of the efficient features increase, CA may decrease. We think a few feature components of noise may be introduced in training when the ***p*** is higher, so overfitting may appear in SVM training and influence the result of classification. However, from the overview, the higher ***p*** corresponds to higher CA. Here, optimal parameters ***d=3***, ***p=100*** and ***a=0.6***.

Table.1 presents the class-confusion matrix corresponding to the use of the optimal parameters mentioned above and linear SVM. 400 cells were test in this experiment. The classification accuracy of each class is shown in Fig.4. So the average classification accuracy is 90.5%.

The meanings of the class labels shown in Table.1 are explained in following:

s1: Basophilic granulocyte
s2: Eosinophilic granulocyte
s3: Monoblast

s4: Myeloblast
s5: Promylocyte
s6: Neutrophilic myelocyte granulocyte
s7: Neutrophilic metamyelocyte granulocyte
s8: Neutrophilic band（stab） granulocyte
s9: Neutrophilic segmented granulocyte
s10: others (include shivers of cell, impurity, a few of lymphocyte and erythrocytic)

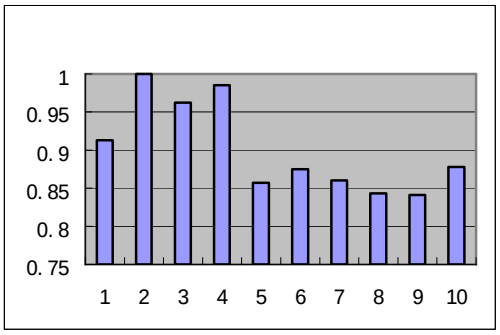


Fig.4. The classification accuracy of each class

Table 1. Class-Confusion Matrix

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
S1	21						1	1		
S2		25								
S3			51		1	1				
S4				67			1			
S5			1		24	2	1			
S6	1					21	1	1		
S7	1			1			37	2	1	1
S8		1					4	43	3	
S9	1							5	37	1
S10	1		2	1			1			36

The best CA appears in class of s2 where the color of it is obviously different from others'. The most common confusions happen among s7, s8 and s9, which are consistent since those cells belong to granulocytic series with closer maturity level. The good performance is result from the superior generalization ability of SVM in efficient feature spaces and nonlinear transformation applied to the histogram bin values.

Pure SVM for histrogram-based image classification presented by Olivier Chapelle [15], in which nonlinear SVM is used to not only map the input data to high-dimension feature space, but also to extract and weight features. However, on one hand, nonlinear SVM need adjust more parameters (>2) simultaneously, that may result in high computation and time cost in training. On the other hand, color histogram in [15] is quantized to a low resolution (***16*16*16***). It's too coarse to distinguish high similarity images. Although SVM is not sensitive to high dimension data, the cost in computation and memory is too intensive as the resolution increased. Our method avoids these disadvantages, which maps original data and extracts features using kernel PCA and weights them by linear SVM, respectively. KPCA used here can greatly enhance the efficiency of the method in which the dimensionality of mixed histogram is sharply reduced from (***32*32*32+2*64***) to 100, and thereby decrease the disturbance of noise exists in training set.

4. Conclusion

We have introduced a novel method based on image-retrieval to classify cell image from high similarity image databases. The cell image is represented with combined global and local histogram features. KPCA is used to reduce the high dimensionality of the data representation and deal with the nonlinear distribution of features. Linear SVMs are presented to construct clustering centers (separating hyperplanes) of multiclass. Similarity measure depends on decision distances from each image to the separating hyperplanes. The method has proven to have a high performance on cell images.

The image-retrieval based system brings two obvious advantages. One is to enable consulting physicians and experts engage in interactive diagnosis easily. The other is to automatically search pathology image records to support reliable decision in detecting and discriminating. This is an open system, with the addition of temporal data to the ground truth database, the number of test classes will be extended and new purposes may be produced. In order to facilitate the expansion of ground truth databases, standardized imaging methodology, cell selection, and digitization need be developed in the future work. The system might be useful tool for clinical diagnosis and pathology studies.

Acknowledgment

This work was supported by the Natural Science Foundation of Ningxia Province under grant No. NZ0610 .

References

[1] H. Hengen, S. Spoor, M. Pandit. Analysis of Blood and Bone Marrow Smears using Digital Image Processing Techniques. SPIE Medical Imaging, San Diego, Vol. 4684, Feb.2002, pp. 624-635.

[2] J.F. David, C. Dorin, M. Peter. Computer-Assisted Discrimination Among Malignant Lymphomas and Leukemia Using Immunophenotyping, Intelligent Image Repositories, and Telemicroscopy. IEEE trans. on information technology in biomedicine, Vol.4, No.4, 12,2000.

[3] D.R. Cecilia, D. Andrew, K. Shahid, et al. Analysis of infected blood cell images using morphological operators. Image and Vision Computing 20 (2002), pp.133-146

[4] I. Baumann, R. Nenninger, H. Harms et al. Image analysis detects lineage-specific morphologic markers in leukemic blast cells. Hematopathology, 1996,105(1), pp.23-30

[5] G. Guodong, Z.L. Stan, L.C. Kap. Support vector machines for face recognition. Image and Vision Computing 19(2001), pp.631-638

[6] M. Beksac , M.S. Beksac, V.B. Tipi et al . An artificial intelligent diagnostic system on differential recognition of Hematopoietic cells from microscopic images. Cytometry, 1997, 30, pp.145-150

[7] E.H. Michael, M.C. Jens. Color-Based Image Retrieval form High-Similarity Image Databases. J.Bigun and T.Gustavsson (Eds.): SCIA 2003, LNCS 2749, pp.1098-1105, 2003.

[8] B. Schölkopf, A. Smola, and K. Müller. Nonlinear component analysis as a kernel eigenvalue problem. Neural Computation, vol. 10, pp. 1299-1319, 1998.

[9] Chen Pan, Yi Fang, Xiangguo Yan, and Chongxun Zheng. Robust segmentation for low quality cell images from blood and bone marrow. International Journal of Control, Automation, and Systems. 2006, vol.4, No.5, pp.637-644

[10] C.R. Gonzalez, E. R. Woods. Digital Image Processing. (Second Edition), Pearson Education, 2003.

[11] D.A. Bimbo, M. Mugnaini, P. Pala et al. Visual querying by color perceptive regions, Pattern Recognition 31 (9) 1998.

[12] J. Huang, S. R. Kumar, M. Mitra, et al. Image indexing using color correlograms, Proc. Conf. Computer Vision Pattern Recognition, 1997, pp. 762-768.

[13] Ying Deng, B.S. Manjunath, Charles Kenney, et al. An efficient color representation for image retrieval. IEEE trans. on Image Processing, vol.10, 1.2001, pp.140-147

[14] Chad Carson, Serge Belongie, Hayit Greenspan, et al. Blobworld: Image Segmentation Using Expectation-Maximization and its Application to Image Querying. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.24. No.8, 2002, pp.1028-1038

[15] Olivier Chapelle, Partick Haffner, Vladimir N. Vapnik. Support Vector Machines for Histogram-Based Image Classification. IEEE Trans. on Neural Networks, vol.10, No.5, Sept.1999, pp.1055-1064

[16] V. Vapnik. Statistical learning theory. New York, John Willey &Sons, 1998.

[17] Lei Zhang, Fuzong Lin, Bo Zhang. Support Vector Machine Learning for Image Retrieval. 2001 International Conference on Image Processing, vol.2, 7-10 Oct. 2001 pp.721 - 724

[18] A. Ben-Hur, D. Horn, H.T. Siegelmann, V. Vapnik. A support vector clustering method. 15th International Conference on Pattern Recognition, 2000. vol.2, 3-7 Sept 2000, pp.:724 – 727

Chen Pan received the Ph.D. degree in Biomedical Engineering from Xi'an Jiaotong University in 2005. He is currently an associate professor of Computer Science at Ningxia University. His research interests include medical image analysis, pattern recognition and machine learning.

Xiang-guo Yan is a professor of Biomedical Engineering at Xi'an Jiaotong University. His research interests primarily are centered with medical signal processing and image processing.

Chong-xun Zheng is a Professor of Biomedical Engineering at Xi'an Jiaotong University. He received the diploma in Electrical Engineering from Xi'an Jiaotong University, China, in 1962. His research interests include biomedical signal processing and applications, image processing, cardiac electrophysiology, neurophysiology, noninvasive bio-signal extracting.