Leukocyte Recognition Using EM-Algorithm

Mario Chirinos Colunga[1], Oscar Sánchez Siordia[2], and Stephen J. Maybank[1]

[1] Department of Computer Science and Information Systems, Birkbeck, University of London, Malet Street, London WC1E 7HX
[2] Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET) Cuernavaca, Morelos, México

Abstract. This document describes a method for classifying images of blood cells. Three different classes of cells are used: Band Neutrophils, Eosinophils and Lymphocytes. The image pattern is projected down to a lower dimensional sub space using PCA; the probability density function for each class is modeled with a Gaussian mixture using the EM-Algorithm. A new cell image is classified using the maximum a posteriori decision rule.

Keywords: Leukocyte recognition, EM-Algorithm, PCA.

1 Introduction

Blood cell analysis is an important diagnostic tool because it can help to detect a width range of diseases. Two types of blood cell analysis are performed: complete blood count and differential blood count. In complete blood count the numbers of erythrocytes (red cells), leukocytes (white cells) and platelets in the peripheral blood (the blood in the circulatory system) are counted to obtain a concentration of cells per unit volume. In a differential blood count the different classes of leukocytes in the peripheral blood or bone marrow are counted to provide a more detailed diagnosis.

The cell counts for peripheral blood are obtained manually or by an automatic flow cytometer, whereas the counts for bone marrow are always obtained manually. The manual counting is done via visual inspection by technicians under a microscope. It has the advantage of being accurate, in that the counts are repeatedly close to the correct values, with little bias; On the other hand it has the disadvantage of being a highly time consuming process. Only a few hundred cells are counted. This reduces the precision of the counts of cells that have a low probability of occurring. The precision of the low cell counts can also be affected by subjectivity, different results can be obtained by different technicians or by the same technician in a second count. In the automatic flow cytometer the cells pass one by one through a channel where sensors measure volume, conductivity and scattered light. This method is accurate and precise for normal blood samples because it can count up to 80 times more cells than for manual counting. However it cannot be used for pathelogical blood samples. In such cases it is necessary to revert to manual counting

A. Hernández Aguirre et al. (Eds.): MICAI 2009, LNAI 5845, pp. 545–555, 2009.

For some time researchers have been developing automatic visual inspection systems using computer vision and pattern recognition methods. This make it possible to obtain cell count automatically from images, reducing the cost of analysis [5]. The process of blood cell recognition described in this document involves three steps: segmentation, feature selection and classification. A classifier is implemented by modelling the probability distribution of each class of cells with a mixture of Gaussian functions. The parameters of the Gaussian functions are obtained using the EM-Algorithm on the data from each class. The EM-Algorithm is initialized using K-Means clustering and the data is projected down into a lower dimensional space using PCA. The class conditional densities are found using Bayes' theorem. Classification is done by choosing the class with the highest probability given the data.

The document is structured as follows: section 2 contains the background theory: K-Means clustering, EM-Algorithm and principal component analysis. Section 3 describes the method for cell classification and its implementation. Section 4 presents the experimental results and section 5 offer some conclusions about the experiment.

2 Background Theory

2.1 EM-Algorithm for Gaussian Mixtures

The expectation maximization algorithm or EM-algorithm is a method to find the maximum likelihood solution for models with latent variables. The name was given in [4] which points out that the method had been proposed many times by other authors. The model is a Gaussian mixture distribution (1) and the latent variables z_k are labels that indicate from which component k of the mixture the measurements came from [6].

The probability density function for a mixture model with K Gaussian distributions $N(x|\mu_k, \Sigma_k)$ with means μ_k and covariance matrices Σ_k is defined by:

$$p(x) = \sum_{k=1}^{K} \pi_k N(x|\mu_k, \Sigma_k) \tag{1}$$

The parameters $\pi_k \in [0, 1]$ are the mixing coefficients for the Gaussian densities, they are subject to the restriction:

$$\sum_{k=1}^{K} \pi_k = 1 \tag{2}$$

The conditional probability of z_k given x can be seen as the responsibility that the component k takes for explaining the observation x [2]. This conditional probability is given by:.

$$p(z_k|x) = \frac{\pi_k N(x|\mu_k, \Sigma_k)}{\sum_{j=1}^{K} \pi_j N(x|\mu_j, \Sigma_j)} \tag{3}$$

The log of the likelihood function for a data set $X = (x_1, .., x_N)^T$ with N data points under the Gaussian mixture model is given by:

$$\ln\, p(X|\pi, \mu, \Sigma) = \sum_{n=1}^{N} \ln \left(\sum_{k=1}^{K} \pi_k N(x_n|\mu_k, \Sigma_k) \right) \tag{4}$$

There is no closed form solution for the values of π, μ and Σ which maximize the log likelihood in (4).

We can find an estimates of π, μ and Σ by first setting to zero the derivative of $\ln\, p(X|\pi, \mu, \Sigma)$ with respect to the mean μ_k of the Gaussian component.

$$\frac{\partial}{\partial \mu_k} \ln\, p(X|\pi, \mu, \Sigma) = 0 \tag{5}$$

$$\sum_{n=1}^{N} \frac{\partial}{\partial \mu_k} \ln \left(\sum_{j=1}^{K} \pi_j N(x_n|\mu_j, \Sigma_j) \right) = 0 \tag{6}$$

Using the chain rule and equations (1), (3) and (4), we have:

$$\sum_{n=1}^{N} \frac{(\pi_k N(x_n|\mu_k, \Sigma_k))}{p(x_n)} \left(-\Sigma_k^{-1}(x_n - \mu_k) \right) = 0 \tag{7}$$

$$-\Sigma_k^{-1} \sum_{n=1}^{N} p(z_k|x_n)(x_n - \mu_k) = 0 \tag{8}$$

therefore:

$$\sum_{n=1}^{N} p(z_k|x_n)(x_n - \mu_k) = 0 \tag{9}$$

which yields:

$$\mu_k \sum_{n=1}^{N} p(z_k|x_n) = \sum_{n=1}^{N} p(z_k|x_n) x_n \tag{10}$$

If we define the effective number of points N_k assigned to cluster k as:

$$N_k = \sum_{n=1}^{N} p(z_k|x_n) \tag{11}$$

we find that μ_k is given by:

$$\mu_k = \frac{1}{N_k} \sum_{n=1}^{N} p(z_k|x_n) x_n \tag{12}$$

Now we set the derivative of $\ln\, p(X|\pi, \mu, \Sigma)$ with respect to the covariance matrix Σ_k to zero,

$$\frac{\partial}{\partial \Sigma_k} \ln\, p(X|\pi, \mu, \Sigma) = 0 \tag{13}$$

thus:

$$\sum_{n=1}^{N} \frac{\partial}{\partial \Sigma_k} \ln \left(\sum_{j=1}^{K} \pi_j N(x_n|\mu_j, \Sigma_j) \right) = 0 \tag{14}$$

which using (1) yields:

$$\sum_{n=1}^{N} \frac{\pi_k}{p(x_n)} \frac{\partial N(x_n|\mu_k, \Sigma_k)}{\partial \Sigma_k} = 0 \tag{15}$$

defining $M(x)^2 \equiv (x-\mu_k)^T \Sigma^{-1}(x-\mu_k)$ and $\alpha \equiv \frac{1}{(2\pi)^{D/2}}$, where D is the dimension of the vector x, we have:

$$\sum_{n=1}^{N} \frac{\pi_k \alpha}{p(x_n)} \frac{\partial}{\partial \Sigma_k} \left(\frac{e^{\left(-\frac{1}{2}M(x_n)^2\right)}}{|\Sigma_k|^{1/2}} \right) = 0 \tag{16}$$

Where $|\Sigma_k|$ is the determinant of the matrix Σ_k. Equation (16) yields:

$$\sum_{n=1}^{N} \frac{\pi_k \alpha}{p(x_n)} \left(\left(\frac{\partial}{\partial \Sigma_k} \frac{1}{|\Sigma_k|^{1/2}} \right) e^{\left(-\frac{1}{2}M(x_n)^2\right)} + \frac{1}{|\Sigma_k|^{1/2}} \frac{\partial}{\partial \Sigma_k} e^{\left(-\frac{1}{2}M(x_n)^2\right)} \right) = 0 \tag{17}$$

On using $\frac{\partial}{\partial X}|X| = |X|X^{-T}$, where X is a square matrix with a non-zero determinant. Equation (17) yields:

$$\sum_{n=1}^{N} \frac{\pi_k \alpha}{p(x_n)} \left(\frac{-|\Sigma_k|\Sigma_k^{-T} e^{\left(-\frac{1}{2}M(x_n)^2\right)}}{2|\Sigma_k|^{3/2}} - \frac{e^{\left(-\frac{1}{2}M(x_n)^2\right)}}{|\Sigma_k|^{1/2}} \frac{\partial}{\partial \Sigma_k} \frac{1}{2} M(x_n)^2 \right) = 0 \tag{18}$$

On using $\frac{\partial}{\partial X} a^T X^{-1} b = -X^{-T} a b^T X^{-T}$, where a and b are vectors and X is any invertible matrix, (18) yields:

$$-\sum_{n=1}^{N} \frac{\pi_k \alpha}{p(x_n)} \left(\frac{\Sigma_k^{-1} e^{\left(-\frac{1}{2}M(x_n)^2\right)}}{2|\Sigma_k|^{1/2}} - \frac{\Sigma_k^{-1} e^{\left(-\frac{1}{2}M(x_n)^2\right)}}{2|\Sigma_k|^{1/2}} (x_n-\mu_k)(x_n-\mu_k)^T \Sigma_k^{-1} \right) = 0 \tag{19}$$

Factorizing and multiplying both sides of the equation by Σ_k:

$$\sum_{n=1}^{N} \frac{\pi_k \alpha |\Sigma_k|^{-1/2} e^{\left(-\frac{1}{2}M(x_n)^2\right)}}{p(x_n)} (I - (x_n-\mu_k)(x_n-\mu_k)^T \Sigma_k^{-1}) = 0 \tag{20}$$

It follows from (3) and (20) that:

$$\sum_{n=1}^{N} p(z_k|x_n)(I - (x_n-\mu_k)(x_n-\mu_k)^T \Sigma_k^{-1}) = 0 \tag{21}$$

thus:

$$N_k I = \sum_{n=1}^{N} p(z_k|x_n)(x_n-\mu_k)(x_n-\mu_k)^T \Sigma_k^{-1} \tag{22}$$

from which follows that Σ_k is given by:

$$\Sigma_k = \frac{1}{N_k} \sum_{n=1}^{N} p(z_k|x_n)(x_n - \mu_k)(x_n - \mu_k)^T \tag{23}$$

Let λ be the Lagrage multiplier associated with the constraint (2) on the π_k and define V by:

$$V = \ln p(X|\pi, \mu, \Sigma) + \lambda \left(1 - \sum_{j=1}^{K} \pi_j \right) \tag{24}$$

it follows that:

$$\frac{\partial V}{\partial \pi_k} = -\lambda + \sum_{n=1}^{N} \frac{\partial}{\partial \pi_k} \ln \left(\sum_{j=1}^{K} \pi_j N(x_n|\mu_j, \Sigma_j) \right) \tag{25}$$

On setting $\frac{\partial V}{\partial \pi_k} = 0$ it follows that:

$$\sum_{n=1}^{N} \frac{1}{p(x_n)} \frac{\partial \pi_k N(x_n|\mu_k, \Sigma_k)}{\partial \pi_k} = \lambda \tag{26}$$

thus:

$$\sum_{n=1}^{N} \frac{1}{p(x_n)} N(x_n|\mu_k, \Sigma_k) = \lambda \tag{27}$$

which yields:

$$\pi_k = \frac{1}{\lambda} \sum_{n=1}^{N} \frac{\pi_k N(x_n|\mu_k, \Sigma_k)}{p(x_n)} = \frac{1}{\lambda} \sum_{n=1}^{N} p(z_k|x_n) = \frac{N_k}{\lambda} \tag{28}$$

Substituting into the constraint (2):

$$\frac{1}{\lambda} \sum_{k=1}^{K} N_k = 1 \tag{29}$$

It follows from (28) and (29) that:

$$\pi_k = \frac{N_k}{N} \tag{30}$$

This results do not constitute a closed form solution for $\ln p(X|\pi, \mu, \Sigma)$, since these parameters appear in $p(z_k|x)$, however a solution for the maximum likelihood problem can be obtained by iteratively updating the parameters.

The EM – Algorithm begins by first choosing initial values for π_k, μ_k and Σ_k and then iterates two steps to update their values. It is common to run a

K-Means algorithm in order to find suitable initial values. In the first step called E-step (expectation) the current values of the parameters are used to evaluate the a posteriori probabilities (3). Then these probabilities are used in the M-step (maximization) to re-estimate the parameters π_k, μ_k and Σ_k. In each iteration the log likelihood is guaranteed to increase [2]. The iterative process stops when the change in the log likelihood or in the parameters falls below a threshold.

There are some difficulties with the EM-Algorithm. It is inclined to get stuck in a local maximum, and there is no guarantee that it converges to a the global maximum [6]. Another problem happens if one of the Gaussians collapses to a single point [2]. This severe overfitting can be avoided using a Bayesian approach in which a prior density $p(\pi, \mu, \Sigma)$ is defined. A MAP (maximum a posteriori) solution is found [2].

2.2 K-Means

K-Means algorithm is one of the most popular clustering methods. It is also called c-means, iterative reallocation or basic iso data [16]. The K clusters are formed such that the distances between the points in a cluster are relatively small compared to the distances to data outside the cluster. The aim is to find an assignment of the data to clusters that minimizes the sum of squared distances of each data point to its closer prototype or cluster mean μ_k.

Data assignment is indicated via a matrix of flags (1-of-K coding scheme) with elements r_{nk} that describe to which of the K clusters the element $\mathbf{x_n}$ is assigned. If $\mathbf{x_n}$ is assigned to cluster k then $r_{nk} = 1$ and $r_{nj} = 0\ \forall j \neq k$. The goal is to find the values for r_{nk} and μ_k that minimize an objective function or distortion J [2].

This is done by a two steps that are iterated. The steps correspond to successive optimizations of r_{nk} and μ_k. In the first step each data point is assigned to the cluster which has the closest prototype. The first set μ of cluster centres is chosen randomly before the iteration begins. In the second step new μ_k are calculated using the assignation of data points to clusters obtained in the previous step. The process stops when no movement of data from a cluster to another reduces the objective function J or a maximum number of iterations is reached. The objective function is the sum of squared distances for each data point $\mathbf{x_n}$ to its corresponding cluster prototype μ_k.

$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} \|\mathbf{x_n} - \mu_k\|^2 \tag{31}$$

Since in each step the objective function is reduced the convergence is assured, nevertheless it can converge into a local minimum instead of a global minimum, and it may even produce empty clusters. This means that the result is a deterministic function of its initial parameters. One way to deal with this problem is to run K-means many different times and chose the best result [6].

Variants of K-Means algorithm are obtained by replacing the Euclidean distance with another measure of distance. In [13] a distance based on symmetry is used successfully to produce spherical, elliptical and ring shaped clusters. The detection of ring shaped clusters in digital images is important in industrial applications.

2.3 Lower Dimensional Space Mapping Using PCA

Working in high dimensional space leads to many problems. As the dimension of the feature space grows the number of sample points needed to reliably estimate the parameters of a classifier grows exponentially. This is known as the curse of dimensionality. It causes a poor generalization ability and thereby reduces the accuracy of a classifier. However, real data will often be confined to a region of the feature space having a lower effective dimension (intrinsic dimensionality) [2]. It is desirable to project the data to a feature space with a dimension equal to the effective dimension of the data. Feature extraction methods determine an appropriate subspace of dimension M in the original space of dimension D where $M \leq D$.

PCA is perhaps the most used technique for feature extraction. It involves finding an orthogonal projection P of the data X onto a lower dimensional linear space, where the variance of the projected data is maximized [2]. This means that the projection $Y = P^T X$ diagonalize the covariance matrix [3]:

$$\Sigma_Y = P^T \Sigma_X P \tag{32}$$

The covariance matrix Σ_X is a symmetric matrix. Symmetry guarantees that all eigenvalues of Σ_X are real and that there is an orthonormal basis of eigenvectors [7]. Then the selected projection is a real orthogonal matrix of eigen vectors Φ that diagonalize the covariance matrix:

$$\Phi^T \Sigma_X \Phi = \Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_D) \tag{33}$$

By ranking the eigenvectors by their corespondent eigenvalues and selecting the first M principal components, is generated an orthogonal decomposition of the vector space $\Re^D$ into two complementary and orthogonal subspaces, the principal subspace $F = \{\Phi_i\}_{i=1}^{M}$ and its orthogonal complement $\overline{F} = \{\Phi_i\}_{i=M+1}^{D}$. The components in $\overline{F}$can be interpreted as noise present in the signal [9].

PCA is a partial Karhunen-Loève transformation, a term that include all the transformations based on the covariance matrix [16], which extract a subspace of lower dimension, $\mathbf{y} = \Phi_M^T \mathbf{x} : \Re^D \rightarrow \Re^M$, corresponding to the maximum eigenvalues [8].

Two of the most wide spread applications of PCA in computer vision are face detection and face recognition. A. Turk and P. Pentland [14] introduced for the first time the term "eigenfaces", by projecting the space of the sample faces in a subspace whose axes or egienfaces, are the principal components of the training data set.

3 Leukocyte Recognition Implementation

The leukocyte images used for this experiment were acquired from "Hemosurf - an Interactive Hematology Atlas" [15] with permission of the authors to use them for research proposes. The images were taken with a Sony DXC-3000A 3CCD video camera and a Zeiss Axioskop microscope using a 100X objective and a 10x eyepiece. Judging by the size of the erythrocytes, which normally exhibit high uniformity in diameter [1], the resolution of the images is approximately 160nm/pixel.

A "ground truth" for each one of the images was created manually by labeling the regions of interest: nucleus, cytoplasm and background, figure 1, a similar manual segmentation has been created in [10] using 431 images to evaluate image segmentation accuracy; in [11] a manual selection of regions of interest is used to learn the parameters of a image segmentation. Leukocytes are extracted from the images by centring a square window at the center of mass of the cytoplasm. The window size is set at the diameter of the largest cell, in this case the neutrophil with a maximum diameter of 104 pixels. The numbers of leukocytes in each image and the corresponding pixels for each leukocyte are found using connected components, a description of some connected components algorithms can be found in [3]. After the extraction each window containing a leukocyte is dawn sampled and saved as a independent image.

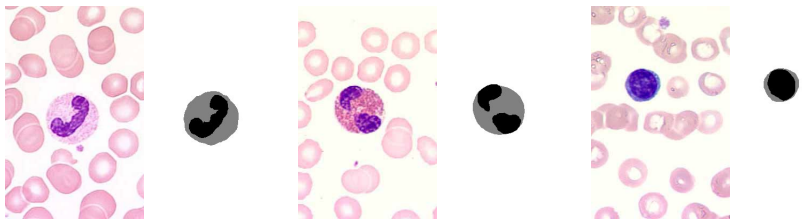


Fig. 1. "Ground truth" images manually created to label the regions of interest: nucleus, cytoplasm and background. a) The original image of a Neutrophil and its correspondence "ground truth" image. b) Eosinophil c) Lymphocyte.

The class conditional density function $p(x|\omega_c)$ for each leukocyte class ω_c is modelled using a mixture of K Gaussian densities. To find the parameters of each mixture, the means μ_{ck} and the covariances Σ_{ck}, the EM-Algorithm is applied independently to each class. The parameters for the EM-algorithm are initialized using the best result from a set of K-Means clustering trials.

The training set to learn the class conditional densities is created as follows: each cell example is rotated at a fixed number of different angles to generate more images, figure 2; in this way the classifier is made invariant to rotation i.e. the classifier learn each example at different rotations. In other works the cells are rotated to have a uniformly aligned training and test set. In [12] the cells are

rotated based on the vector created from the center of mass of the cytoplasm and the center of mass of the nucleus and in [17] nucleus centroids are aligned with a reference image and the image is rotated until the absolute difference between the images is a minimum.

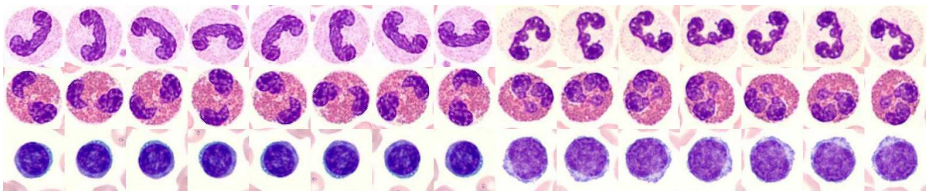


Fig. 2. Each image is rotated several times to make the classifier invariant to rotation

After training, a new leukocyte image is classified using the maximum posteriory rule, i.e. the new image is assigned to class ω_i if $p(\omega_i|x) > p(\omega_j|x) \forall\, i \neq j$. The posterior probabilities are found using the Bayes' theorem. Both the training and test images have been contrast stretched.

4 Experimental Results

Three classes of leukocytes (band neutrophils, eosinophils and lymphocyte) were used in the experiment. They were chosen because are the three most dissimilar classes. At this point the data set consists only of 15 examples per class. To create the training set each image is rotated [0, 45, 90, 135, 180, 225, 270, 315] degrees and then down sampled at scale $\sigma = 0.25$. The data set is projected down into a lower dimensional space using PCA and the first 7 principal components, figure 3; the number of principal component used was found empirically which is the maximum number of principal components that allow to have nonsingular covariance matrices for all the mixture components. The number of Gaussian functions that form the mixture of each class was set arbitrarily to $K = 4$. The parameters of the EM-Algorithm were initialized using the best result from 10 trials of the K-Means clustering. The parameters obtained for each one of the Gaussian mixtures can be seen in figure 4; it can be seen that the means correspond to different representation of a cell class as variation in rotation, shape or size. The classifier is tested using cross validation leave one out method.

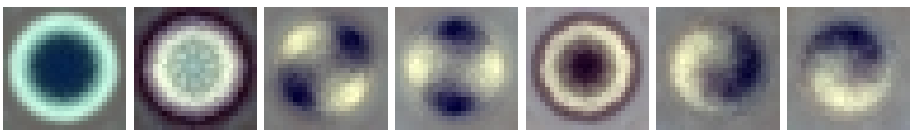


Fig. 3. First seven principal components for the leukocyte images data

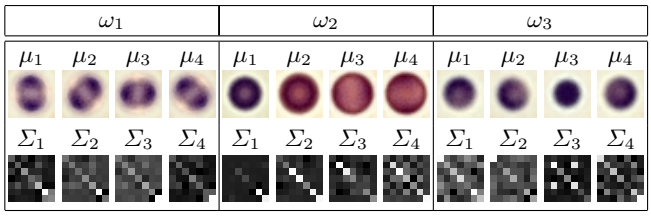
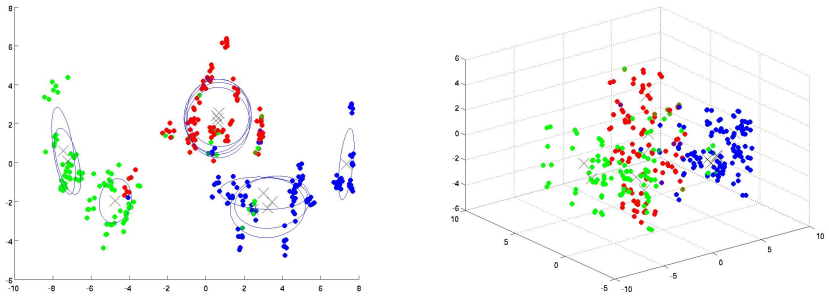


Fig. 4. Means and covariance matrices for the four components Gaussian mixture. a) neutrophils class parameters. b) eosinophils class parameters. c) lymphocyte class parameters.



a b

Fig. 5. a) Cell data plotted using the first two principal components, red dots indicate neutrophils samples, green correspond to Eosinophils and blue to lymphocytes; ellipses correspond to the influence of each Gaussian. b) Same data plotted using the three principal components.

Confusion matrix for 15 training sets			
	ω_1	ω_2	ω_3
ω_1	0.82	0.06	0.01
ω_2	0.05	0.89	0.05
ω_1	0	0	1

15 test sets using leave one out cross validation			
	ω_1	ω_2	ω_3
ω_1	0.80	0.013	0.06
ω_2	0	0.86	0.13
ω_1	0.06	0.06	0.86

Fig. 6. Confusion matrices for the test and training sets.

Fifteen different training sets were created by using 14 cells per class and rotated them to obtain 112 examples per class. 15 different test sets were created by using the remaining one example per class without rotation. Figure 4 shows the confusion matrix for the training and test sets.

5 Conclusions

This was my first test using blood cell images; at this time just tree of the six different types of leukocyte in peripheral blood were used, neutrophils, eosinophils

and lymphocytes. Just 15 samples per class were available in this experiment. The lack of samples meant that only a few principal component could be used to represent the data and that the number of Gaussian per mixture was small. Currently 787 images of blood slides containing 957 leukocytes have been acquired and have to be manually labeled and segmented to evaluate classification and segmentation performance.

A extension to this experiment could be to use a diferent feature extraction method to improve class separability, use a method to find the optimal numbers of components for each Gaussian mixture and use a different image descriptor

References

1. Hoffbrand, A.V., Pettit, J.E.: Essential Haematology., 3rd edn. Blackwell Science, Malden (1993)
2. Bishop, C.M.: Pattern Recognition and Machine Learning (Information Science and Statistics). Springer, Heidelberg (2006)
3. Davies, E.R.: Machine Vision: Theory, Algorithms, Practicalities. Morgan Kaufmann Publishers Inc., San Francisco (2004)
4. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the em algorithm. Journal of the Royal Statistical Society. Series B (Methodological) 39(1), 1–38 (1977)
5. Fluxion Biosciences: Microfluidic flow cytometer with stop-flow capability provides enhanced cellular imaging. Technical report, Fluxion Biosciences, QB3- Mission Bay; 1700 4th St., Suite 219; San Francisco, CA 94143 (2005)
6. Forsyth, D.A., Ponce, J.: Computer Vision: A Modern Approach. Prentice Hall Professional Technical Reference, Englewood Cliffs (2002)
7. Golub, G.H., Van Loan, C.F.: Matrix Computations, 3rd edn. Johns Hopkins University Press (October 11, 1996)
8. Moghaddam, B., Pentland, A.: Probabilistic visual learning for object detection. In: ICCV 1995, vol. 1, pp. 786–793 (1995)
9. Moghaddam, B., Pentland, A.: Probabilistic visual learning for object representation. In: PAMI, vol. 19, pp. 696–710 (July 1997)
10. Theera-Umpon, N.: Patch-Based White Blood Cell Nucleus Segmentation Using Fuzzy Clustering. Transactions on Electrical Eng., Electronics, and Communications, ECTI-EEC 3, 5–19 (2005)
11. Sabino, D.M.U., da Fontoura Costa, L., Rizzatti, E.G., Zago, M.A.: A texture approach to leukocyte recognition. Real-Time Imaging 10, 205–216 (2004)
12. Sanei, S., Lee, T.M.K.: Cell Recognition Based on PCA and Bayesian Classification. In: Fourth International Symposium on Independent Component Analysis and Blind Signal Separation. ICA, April 1 (2003)
13. Su, M.C., Chou, C.H.: A modified version of the k-means algorithm with a distance based on cluster symmetry. IEEE Trans. Pattern Anal. Mach. Intell. 23(6), 674–680 (2001)
14. Turk, M.A., Pentland, A.P.: Face recognition using eigenfaces, pp. 586–591 (1991)
15. Woermann, U., Montandon, M., Tobler, A., Solemthaler, M.: HemoSurf - An Interactive Hematology Atlas. Division of Instructional Media, Bern (2004)
16. Webb, A.R.: Statistical pattern recognition. John Wiley & Sons, Chichester (1999)
17. Yampri, P., Pintavirooj, C., Daochai, S., Teartulakarn, S.: White Blood Cell Classification based on the Combination of Eigen Cell and Parametric Feature Detection. In: IEEE (ed.) Industrial Electronics and Applications (May 24, 2006)