Cell Image Segmentation for Diagnostic Pathology

Dorin Comaniciu

Imaging Department

Siemens Corporate Research

Princeton, NJ 08540

Peter Meer

ECE Department

Rutgers University

Piscataway, NJ 08854

1   Introduction

The colors associated with a digitized specimen representing peripheral blood smear are typically characterized by only a few, non Gaussian clusters, whose shapes have to be discerned solely from the image being processed. Nonparametric methods such as mode based analysis [10], are particularly suitable for the segmentation of this type of data since they do not constrain the cluster shapes. This chapter reviews an efficient cell segmentation algorithm that detects clusters in the $L^*u^*v^*$ color space and delineates their borders by employing the gradient ascent mean shift procedure [8, 9]. The color space is randomly tessellated with search windows that are moved till convergence to the nearest mode of the underlying probability distribution. After the pruning of the mode candidates, the colors are classified using the basins of attraction. The segmented image is derived by mapping the color vectors in the image domain and enforcing spatial constraints.

The segmenter is the core module of the Image Guided Decision Support (IGDS) system [14, 13] which is discussed next. The IGDS architecture supports decision making in clinical pathology and provides components for remote microscope control and multiuser

visualization. The primary and long term goal of the IGDS related research is to reduce the number of false negatives during routine specimen screening by medical technologists. The Decision-Support component of the system searches remote databases, retrieves and displays cases which exhibit visual features consistent to the case in question, and suggests the most likely diagnosis according to majority logic. Based on the Micro-Controller component the primary user can command a robotic microscope from the distance, obtain high-quality images for the diagnosis, and authorize other users to visualize the same images. The system has a natural man-machine interface that contains engines for speech recognition and voice feedback.

Section 2 of the chapter concentrates on the segmentation algorithm. In Section 3 we underline the idea of diagnosis support through the presentation of relevant cases and show how the IGDS system was developed based on the robust handling of the image features.

2 Segmentation

We explain first the advantages of nonparametric methods for the analysis of feature spaces. An iterative, gradient ascent procedure for mode seeking is presented next. The section concludes with the description of the cell segmentation algorithm and segmentation examples.

2.1 Feature Space Analysis

Feature space analysis is a widely used tool for solving image understanding problems. An image feature is defined as a local, meaningful, and detectable part of the image [41, p.68].

Given the input image, feature vectors representing color, texture, or parameters of curves or surfaces, are extracted from local neighborhoods and mapped into the spaced spanned by their components. Significant features in the image then correspond to high density regions in this space.

The feature space provides an intrinsic tolerance to noise but a reduced sensitivity to features represented by only a few points. The space the features originate is often used to compensate for allocation errors. For example, the image domain is used during the task of color segmentation to compensate for errors resulted from the analysis of the color space. The feature space can be regarded as a sample drawn from an unknown probability distribution. Analyzing this distribution based on a parametric model (e.g., Gaussian mixture) will introduce severe constraints since then the shape of the delineated clusters is predefined. By contrast, nonparametric cluster analysis uses the *modes* of the underlying probability density to define the cluster centers and the *basin of attraction* to define the boundaries separating the clusters.

The nonparametric methods, however, require *multidimensional range searching* [38, p.373], that is, the search for the data points falling into a given neighborhood. Since the optimization of this task is difficult [38, p.385] the nonparametric analysis of large data sets is computationally expensive, with a complexity proportional to the square of the number of data points. Therefore, a practical algorithm for cluster delineation should involve the tessellation of the feature space and selective processing of data points. Following this idea, we have developed a general technique for the analysis of multimodal data based on the *mean shift* property, first described in [19], and more recently discussed in [6].

2.2 Mean Shift Procedure

Let $\{\mathbf{x}_i\}_{i=1...n}$ be a set of n points in the d-dimensional Euclidean space R^d, i.e., the feature space. Define the hypersphere $S_h(\mathbf{x})$ of radius h centered on $\mathbf{x}$ and containing $n_{\mathbf{x}}$ data points. It can be shown [10] that the sample mean shift is

$$M_h(\mathbf{x}) = \frac{1}{n_{\mathbf{x}}} \sum_{\mathbf{x}_i \in S_h(\mathbf{x})} \mathbf{x}_i - \mathbf{x} \sim \frac{\hat{\nabla} f(\mathbf{x})}{\hat{f}(\mathbf{x})}, \tag{1}$$

where $\hat{f}(\mathbf{x})$ is the local density estimate and $\hat{\nabla} f(\mathbf{x})$ is the local density gradient estimate.

The expression (1) indicates that a local estimate of the normalized gradient can be obtained by computing the sample mean shift and that the mean shift vector alway points towards the direction of the maximum increase in the density. Moreover, the procedure defined by recursively moving the hypersphere by the mean shift vector defines a path leading to a local density maximum, i.e., to a mode of the density. For a more detailed discussion, as well as proof of convergence of the mean shift procedure see [10].
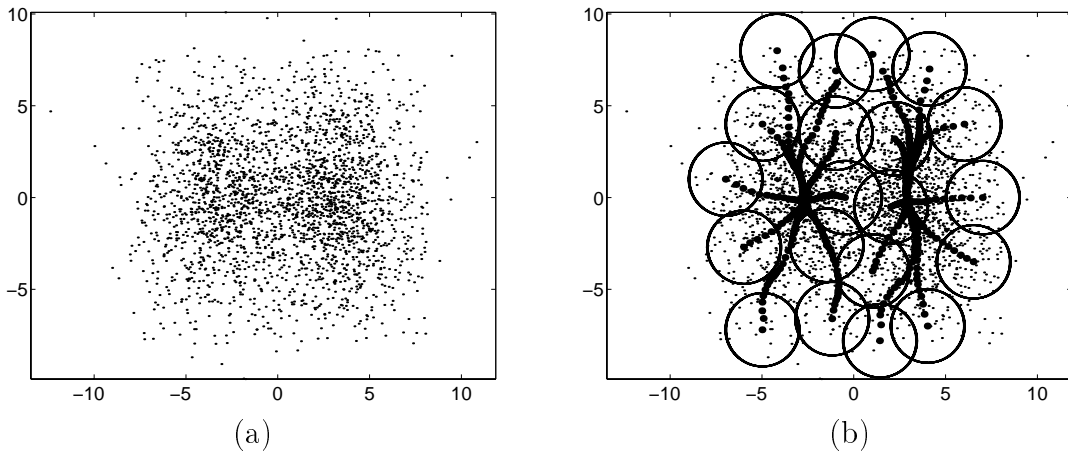

Figure 1: 2-D mode seeking through mean shift iterations. (a) Input data. (b) Trajectories of 20 mean shift procedures.

4

As an example, Figure 1b presents the trajectories of 20 mean shift procedures started from random locations and applied in parallel on the data shown in Figure 1a. The entire process results in two convergence points that correspond to the two modes of the underlying density. In addition, by associating the points with the modes, the structure of the data (number of clusters and their shapes) is revealed. The analysis needs only one parameter, the radius h of the searching sphere and no other a priori knowledge. It is simple and straightforward to implement, being based on the iterative shifting of a fixed size window to the average of the data points within. However, a practical method must handle artifacts such as poor convergence over the low density regions or plateaus without a clear local maximum.

2.3 Cell Segmentation

This section briefly describes the cell segmentation based on the mean shift procedure (see [8, 9] for details). Figure 2 shows a summary of the algorithm.
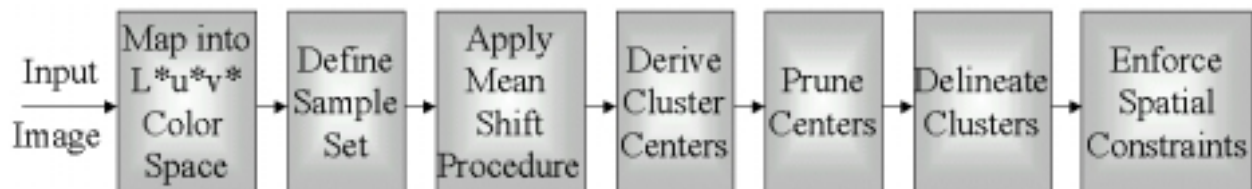

Figure 2: The processing flow of the segmentation algorithm.

The RGB input vectors are first transformed into $L^*u^*v^*$ vectors to obtain a perceptually uniform color space [42, Sec. 3.3.9].

Then, a set of m points called the *sample set* is randomly selected from the data. Distance and density constraints are imposed on the points retained in the sample set, fixing the sample set cardinality. The distance between any two neighbors should be larger than h,

the radius of a searching sphere $S_h(\mathbf{x})$, and the sample points should not lie in sparsely populated regions. Whenever the number of points inside the sphere is below a threshold T_1, a region is regarded as sparse.

The mean shift procedure is applied to each point in the sample set, and the resulting convergence points define m *cluster center candidates*. Since a local plateau in the color space can prematurely stop the mean shift iterations, each cluster center candidate is perturbed by a random vector of small norm and the mean shift procedure is let to converge again.

The candidates are then pruned to obtain $p \leq m$ cluster centers. The mean of any subset of cluster center candidates which are less than h close to each other defines a *cluster center*. In addition, the presence of a valley between each pair of cluster centers is tested [10] and if no valley is found, the tested center of lower density is removed.

Cluster delineation has two stages. First, each sample point is allocated to a cluster center based on the trajectory of its initial window. Then, each data point is classified according to the majority of its k-nearest sample points. Finally, small connected components containing less than T_2 pixels are removed, and region growing is performed to allocate the unclassified pixels.

2.4   Segmentation Examples

Three parameters control the segmentation: the searching sphere radius h, the threshold T_1 which enforces the density constraint, and the threshold T_2 which determines the minimum connected component size. The results presented in here were obtained with $h = 4$, $T_1 = 50$, and $T_2 = 1000$.

A typical leukocyte image is shown in Figure 3a. The segmented image is presented in pseudocolors in Figure 3b. Identical pseudocolors are used to display the cluster delineation in Figure 4e, where the clusters are shifted for better visualization. Figures 4a, b, c, and d present, respectively, the corresponding $n = 6005$ color vectors, sample set with $m = 73$, cluster center candidates, and $p = 4$ cluster centers.
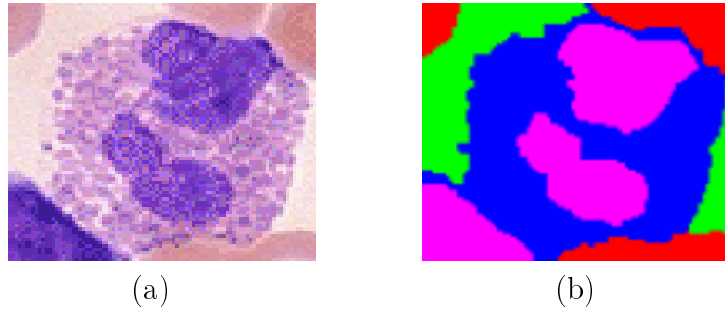

(a)   (b)

Figure 3: (a) Original image. (b) Segmented.

(a)   (b)

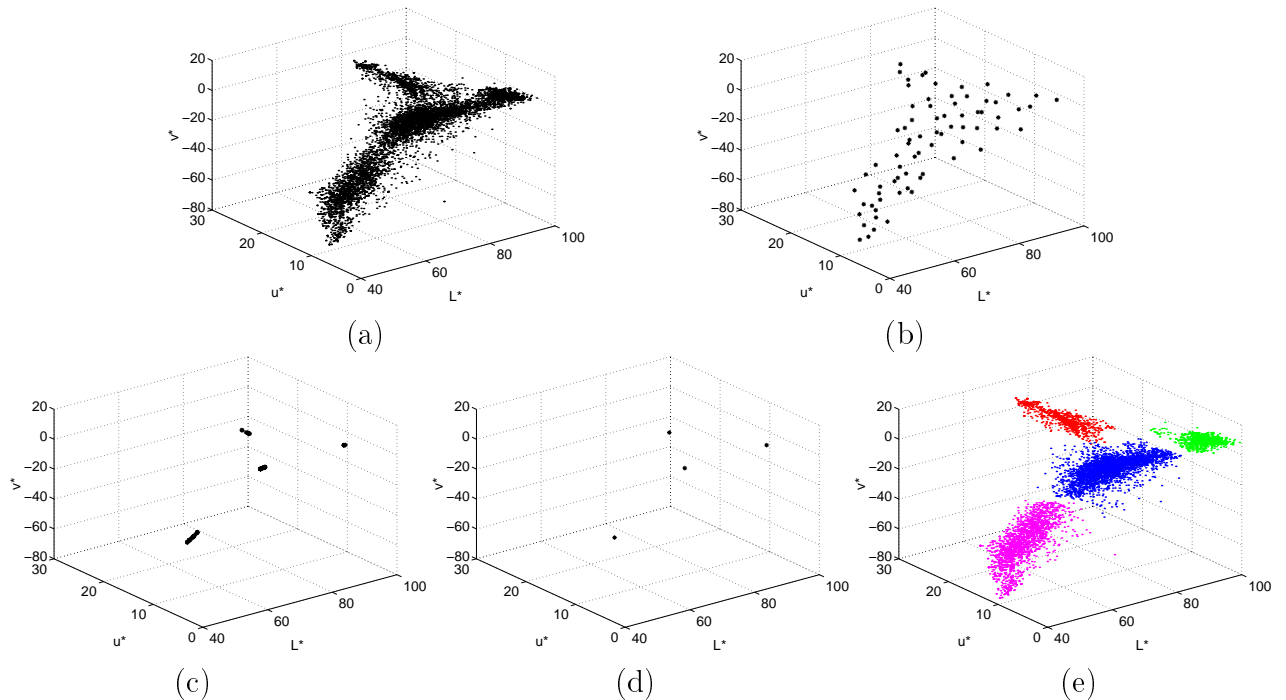
(c)   (d)   (e)

Figure 4: (a) Color vectors. (b) Sample set. (c) Cluster center candidates. (d) Cluster centers. (e) Delineated clusters. The color vectors are randomly sampled yielding the sample set which converges to cluster center candidates from which the cluster centers are extracted.

7

The resulting decomposition is meaningful in both the spatial domain and the color space, the cytoplasm texture, for example, being classified as one cluster.

The segmentation quality can also be evaluated from Figure 5 which shows the original, contour, and segmented images of a stained specimen of Mantle Cell Lymphoma (upper) and two stained specimens of Chronic Lymphocytic Leukemia (lower) [5].

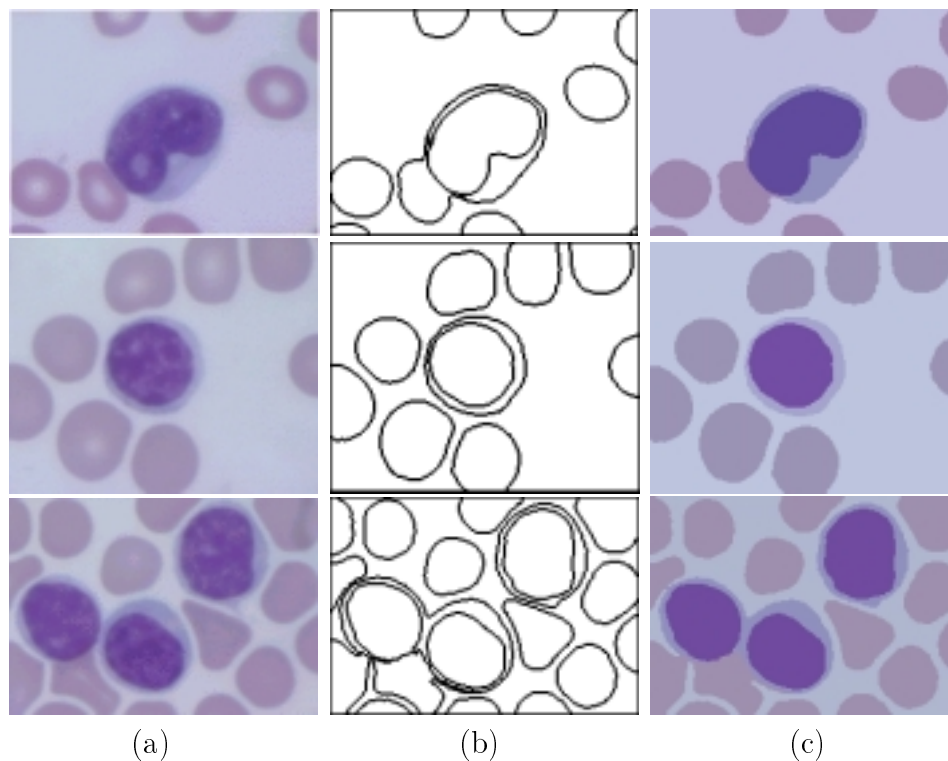| (a) | (b) | (c) |

Figure 5: (a) Original images. (b) Contours. (c) Segmented.

Additional results are presented in Figure 6 where the nucleus of each cell was delineated using the algorithm from above and no further postprocessing was necessary. The data set contains images of different color, sharpness, contrast, noise level and size (for convenience, they are displayed at the same size).

The algorithm running time is linear with the number of pixels in the image. It takes fractions of a second to segment a 256×256 pixel image on a standard PC / workstation.
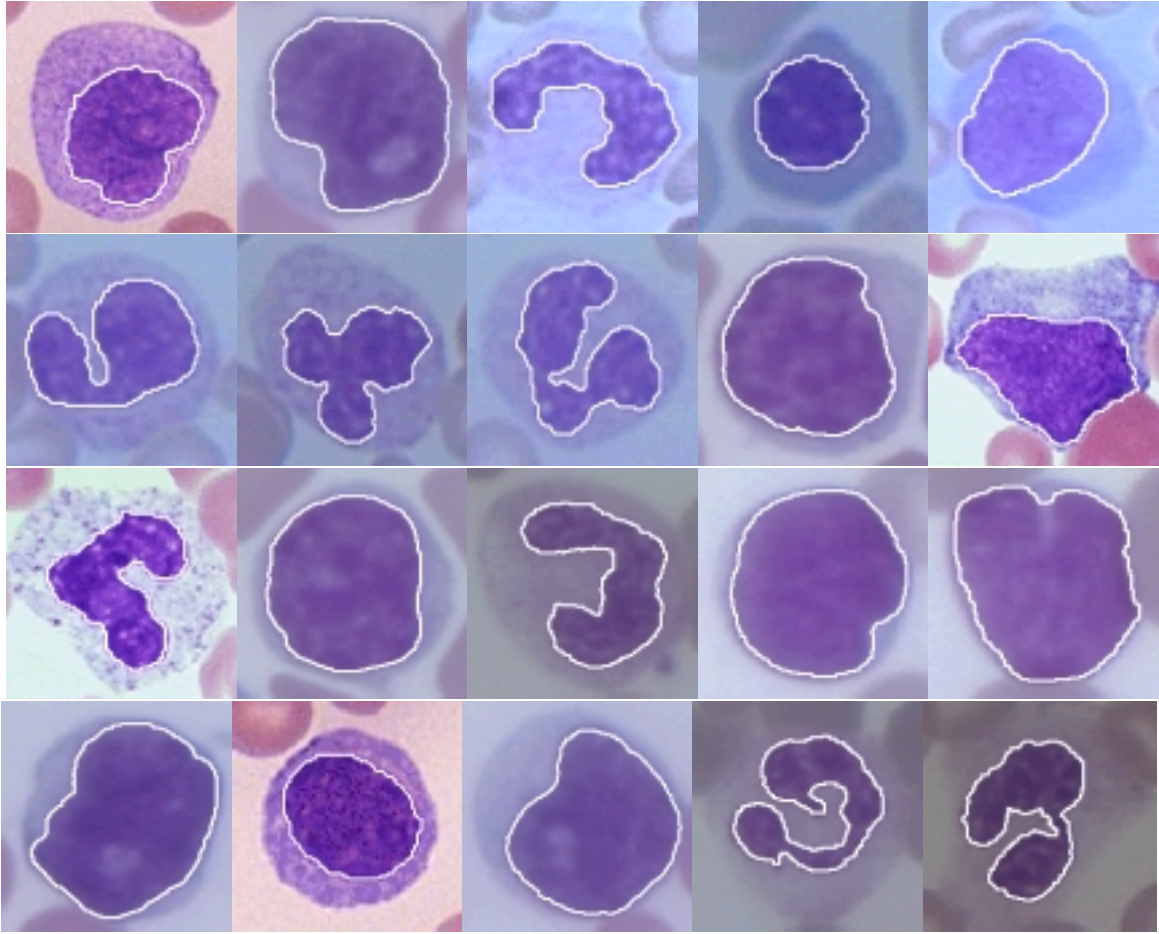
Figure 6: Nucleus segmentation for various cell categories. The nucleus border is marked with a white contour.

3   Decision Support System for Pathology

This section reviews a prototype system for clinical pathology that assists the user in the diagnosis process. The IGDS system [14, 13] segments and analyzes the elemental structures of the input image, searches remote databases and retrieves relevant images based on their content.

After a short description of the problem domain, we present an overview of the system

and the currently used database of ground truth cases. The analysis of the visual attributes of the query is explained next. The optimization of the overall dissimilarity measure is then presented, together with the retrieval performance assessment by cross-validation. Finally, comparisons to the human experts performance on the same database are given.

For more information about the IGSD system, including demonstration videos, the reader is referred to

$$http://www.caip.rutgers.edu/\sim comanici/jretrieval.html$$

3.1 Problem Domain

The subjective, visual analysis of malignant lymphomas and chronic lymphocytic leukemia gives rise in practice to a significant number of false negatives (malignant cells classified as benign). If suspicious cells are detected, subsequent morphological evaluation of specimens by even experienced pathologists is often inconclusive. In these cases differential diagnosis can only be made after supporting tests such as immunophenotyping by flow cytometry.

Mantle Cell Lymphoma (MCL) [2, 4] is of particular interest among the indolent lymphomas since it is often misdiagnosed as Chronic Lymphocytic Leukemia (CLL) or Follicular Center Cell Lymphoma (FCC) [5]. In addition, the survival of patients with MCL is much shorter than that of patients with other low-grade lymphomas, and standard therapy for CLL and FCC is ineffective with MCL. Timely and accurate diagnosis of MCL has therefore significant therapeutic and prognostic implications.

The literature in diagnostic hematopathology ascribes much of the difficulty in rendering consistent diagnoses to subjective impressions of observers and shows that when morphologic

cell classification is based upon computer aided analysis, the level of objectivity and reproducibility improves [3]. However, only recently the potential of diagnosis support through the presentation of relevant cases (as opposed to automatic diagnosis) has been recognized [40].

The technologies that capture, describe, and index the content of multimedia objects rely on methods from image analysis, pattern recognition, and database theory. A new family of information retrieval systems emerged in the recent years, exploiting the richness of visual information and covering a large spectrum of applications [16, 18, 26, 30, 31]. These systems differ according to their degree of generality (general purpose versus domain specific), level of feature abstraction (primitive features versus logical features), overall dissimilarity measure used in retrieval ranking, database indexing procedure, level of user intervention (with or without relevance feedback), and evaluation methodology.

The problem domain of the IGDS system, discriminating among lymphoproliferative disorders, contrasts to that of general retrieval systems engines. It is well defined and allows the quantitative evaluation of the system's performance and comparison to the human expert results. The reason of this comparison however is only to assess the usefulness of the system. In a real analysis scenario, a lot of context information difficult to quantize is taken into account for the diagnosis and no technique can ever replace the pathologist and light microscopy. Our system is designed as a tool to help the physician during its own analysis, by presenting cases consistent to the case in question, and not as an automatic cell classifier.

3.2   System Overview

The IGDS system has a platform-independent implementation in Java and consists mainly of two software components that are described below.

3.2.1   Micro-Controller Component

The Micro-Controller allows one primary user and multiple secondary users to connect to the image server located at the microscope site. The primary user can control the remote microscope (AX70 Olympus equipped with motorized stage), receive and visualize diagnostic-quality images of tissue samples. The transfered images can simultaneously be observed and analyzed by the secondary users. Thus, the IGDS system provides support for consultation, when a fellow pathologist is logged in as secondary user, or teaching, when a group of students is connected to the image server.

A display capture of the Micro-Controller is shown in Figure 7. The image 1 is obtained during the initialization and represents the low resolution panoramic view of the specimen on the robotic stage. The image 2 is the current view using a lens of $100X$ and corresponds to the small rectangular region marked on the panoramic image.

The primary user can adjust the light path or focus of the microscope, change the objective lens, move the specimen on the robotic stage, or copy the current image to the Decision-Support part for further analysis. These actions are possible by mouse input or by speech.

A distinct feature of the IGDS system is its bimodal human-computer interaction [39]. A fusion agent capable of multimodal inputs interprets the commands, calls the appropriate method, and gives voice feedback. Currently the system employs a speech recognizer engine
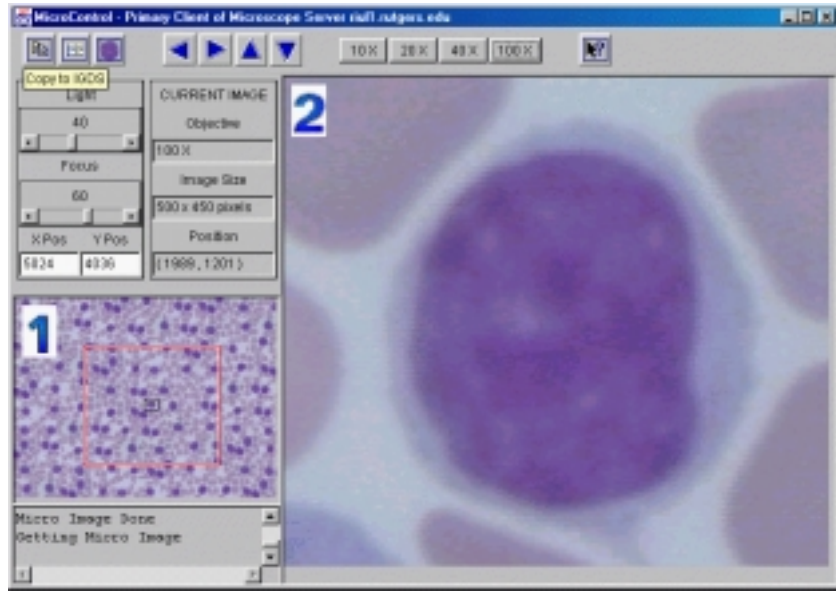
Figure 7: The Micro-Controller component. 1 - Panoramic image. 2 - Current view of the selected region in the panoramic image using the $100X$ lens.

with finite-state grammar. The use of a small, task-specific vocabulary results in very high recognition rate. The recognition is speaker-independent. Examples of voice commands for the Micro-Controller are: Set Light ##, Set Focus ##, Change #, Transfer, Move Right (Left, Up, Down), Update the System.

3.2.2 Decision-Support Component

This component allows the user to load a (remote) query image and select a rectangular region which contains the cells of interest. The elemental structures from the selected region (eg., leucocyte nuclei and cytoplasm areas) are then delineated through the segmentation described in Section 2. By choosing a cell nucleus, the user initiates first the analysis of the nucleus attributes (shape, texture, area, and color), then the search in a remote database of digitized specimens. As a response, the system retrieves and displays the images from the

database that are the closest to the query.

A display capture of the Decision-Support component is shown in Figure 8, containing the query image with the region of interest, the selected nucleus during texture analysis, the normalized shape of the nucleus, eight retrieved images, and the control panel.
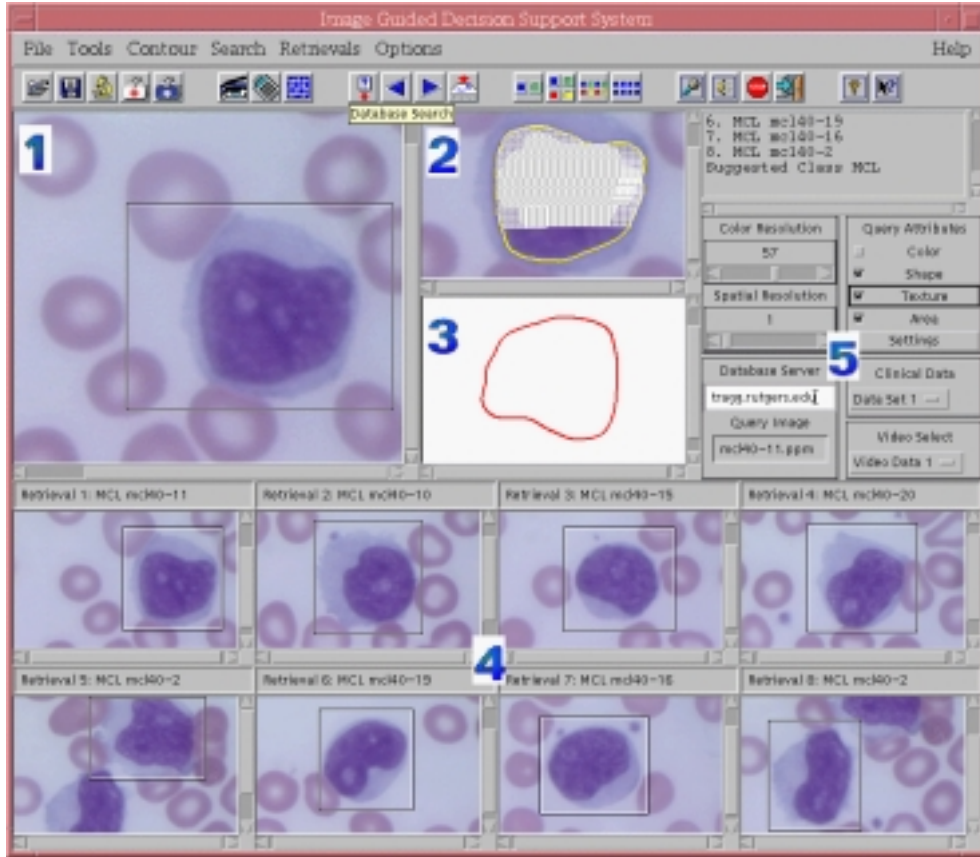

Figure 8: The Decision-Support component. 1- Query image. 2 - Delineated nucleus. 3 - Normalized shape of the nucleus. 4 - Retrieved images. 5 - Control panel.

The user can modify the color resolution and spatial resolution of the segmentation, which are defined as the inverses of the segmentation parameters h and T_2, respectively (see Section 2.4). Access to the resolution parameters is only for experiments and maintenance, in normal operations of the system they are set by default. While the segmentation produces reliable results for almost all the images in the database, the system provides a user-handled

contour correction tool based on cubic splines. It is also possible to select different query attributes, browse the retrievals, select a different scale for visualization, and display specific clinical data and video clips.

As in the Micro-Controller case, the commands can be initiated using speech recognition or graphical input. Typical voice commands for the Decision-Support are: Open Image ##, Save Image ##, Segment the Image, Search the Database, Show 2 (4, 8) Retrievals, Show First (Next, Previous) Retrievals, Show Video, Clinical Data #. Examples of voice feedback are: Image ## Opened, Segmentation Completed, Analyzing Texture, Database Search Completed, Suggested Class: CLL (FCC, MCL, Normal).

The Decision-Support component has the client-server architecture shown in Figure 9. The client part is intended to be used in small hospitals and laboratories to access through the Internet the database at the image server site.
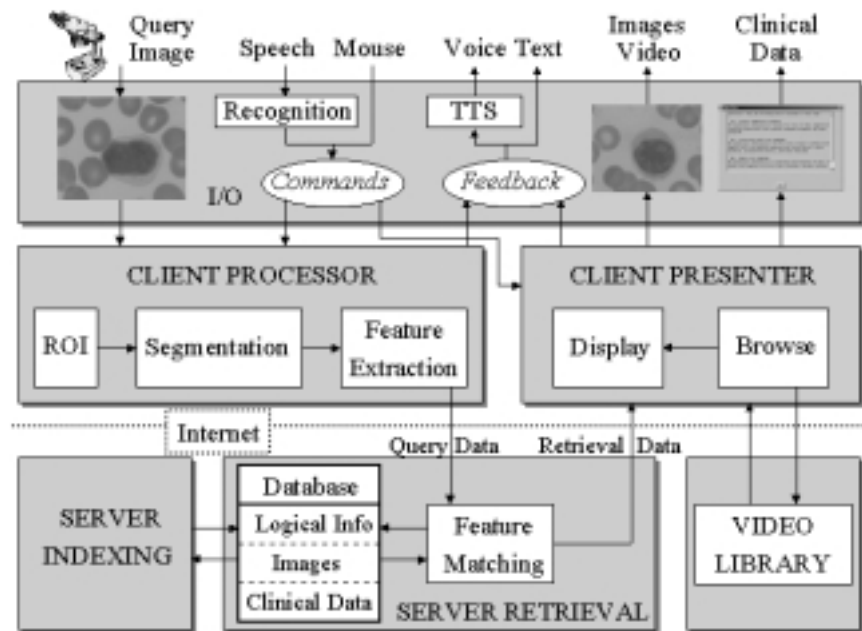

Figure 9: Decision-Support architecture (from [14]).

The I/O module loads the query image from a local or remote microscope and saves the retrieved information. It also includes the Speech Recognizer and TTS (Text to Speech) engines. The Client Processor contains the query formation stages, performing region of interest selection, color segmentation and feature extraction. Based on the retrieval data, the Client Presenter communicates the suggested classification to the user, and allows the browsing of cases of interest including their associated clinical data and video clips.

The server is formed of two parts: the Retrieval and Indexing modules. The retrieval process is multithreaded, simultaneous access to the database being authorized. During feature matching, the query data and logical information in the database are compared to derive a ranking of the retrievals. For indexing the server uses the same client processing stages plus an optimization stage.

3.3   Current Database

The IGDS system is currently using a database containing 98 CLL, 38 FCC, 66 MCL, and 59 Normal cells, a total of 261 images. The ground truth of the recorded cases was obtained off line through immunophenotyping and used to maximize the probability of correct classification.

Immunophenotyping is the characterization of white blood cells by determining the cell surface antigens they bear. The cells are isolated and incubated with fluorescently-tagged antibodies directed against specific cell-surface antigens. Then, they pass through the flow cytometer past a laser beam. When the cells meet the laser beam, they emit fluorescent signals in proportion to the amount of the specific cell surface antigen they have and a

16

computer calculates the percentage of cells expressing each antigen.

3.4 Analysis of Visual Attributes

The query data is determined by four visual attributes of the delineated cell nucleus: shape, texture, area, and color. Medical literature employs frequently the first three of the above attributes to morphologically describe the appearance of malignant cells [2]. The Fourier coefficients describing the nuclear shape are made invariant to changes in location, orientation, and scale (i.e., similarity invariant), by following the approach of Kuhl and Giardina [23]. The texture analysis is based on a multiresolution simultaneous autoregressive model (MRSAR) [27].

3.4.1 Similarity Invariant Shape Descriptors

The number of Fourier harmonics that reliably represent a nuclear shape is dependent on the uncertainty introduced by prior processing stages (e.g., region of interest delineation, segmentation). Since the segmentation process is global, any change in the region of interest selected by the user may have effect on the nucleus delineation. In addition, due to its probabilistic nature the segmentation produces slightly different results when repeatedly applied to the same image.

For example, Figure 10b shows the result of superimposing the contours obtained by segmenting 25 times the image from Figure 10a. The darker a pixel in the contour image, the more stable is the contour passing through that pixel.
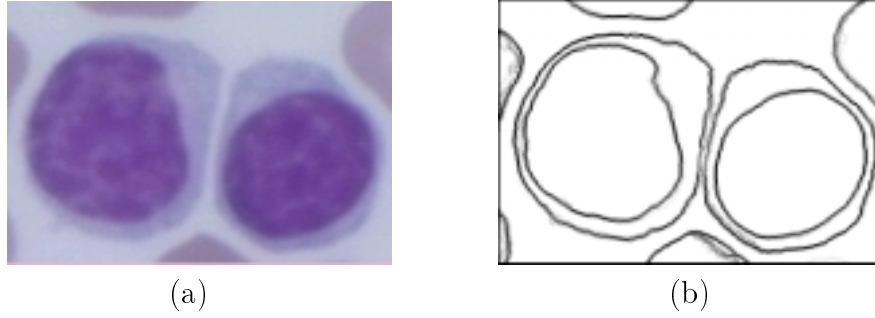
(a)   (b)

Figure 10: (a) Input image. (b) Contours showing the stability of the segmentation. The input image has been segmented 25 times and the resulting contours were superimposed. The regions between two cells are the least stable (from [14]).

To estimate the influence of this uncertainty on the Fourier coefficients, experiments with several images were conducted and the normalized variance (variance over the squared mean) of each coefficient was computed. For a given image, a user delineated 25 times the region of interest (a leukocyte). The region was then segmented and the first 64 Fourier coefficients were determined for the nucleus. Since the normalized variances of the coefficients were typically rapidly increasing for the coefficients with index larger than 40, we concluded that the segmentation is sufficiently stable for the use of only the first 40 coefficients (10 harmonics). Consequently, we compare a query contour with a reference contour in the database by computing the Euclidean distance between the corresponding 40-dimensional vectors of Fourier invariants

$$D_1 = \sqrt{\left(\mathbf{f}_{query} - \mathbf{f}_{reference}\right)^{\top} \left(\mathbf{f}_{query} - \mathbf{f}_{reference}\right)} \ . \tag{2}$$

3.4.2   Texture, Area, and Color Metrics

The nuclear texture is a representation of the chromatin density, being relatively unstructured. It is characterized by random patterns, and shows no presence of periodicity or

18

directionality. We therefore describe the gray level texture data with the multiscale simultaneous autoregressive (MRSAR) model [27]. This is a second-order noncausal model that contains five parameters at each resolution level. For a given resolution, the center pixel value in a neighborhood is expressed as a linear combination of the neighboring pixel values and an additive white Gaussian noise term. Four model parameters (the weights associated with selected pixel values) are estimated through least squares. The model parameters and the estimation error define a five-dimensional feature vector for the given neighborhood. The multiresolution feature vector is obtained by varying the neighborhood size and concatenating the obtained features.

In [34] it was shown that the MRSAR features computed with 5×5, 7×7, and 9×9 neighborhoods provide the best overall retrieval performance for the entire Brodatz database. The same neighborhoods were used here to form fifteen-dimensional multiresolution feature vectors whose mean and covariance we computed for each database entry.

Thus, the texture dissimilarity has to be measured by the distance between two multivariate distributions with known mean vectors and covariance matrices. We use the Mahalanobis distance between the MRSAR feature vectors to express this dissimilarity

$$D_2 = \sqrt{(\mathbf{t}_{query} - \mathbf{t}_{reference})^{\top} \mathbf{\Sigma}_{reference}^{-1} (\mathbf{t}_{query} - \mathbf{t}_{reference})} \ , \tag{3}$$

where $\mathbf{\Sigma}_{reference}^{-1}$ represents the inverse of the covariance matrix of $\mathbf{t}_{reference}$. For each entry in the database $\mathbf{\Sigma}_{reference}^{-1}$ is obtained and stored off-line for each indexed nucleus.

Note that for the current database, the use of $\mathbf{\Sigma}_{reference}$ for Mahalanobis computation

resulted in better classification than that obtained with Σ_{query}. By using the Mahalanobis distance, the assumption we make is that the covariance of the query Σ_{query} and the covariances of the references from the database $\Sigma_{reference}$ are similar.

However, our own research [12] showed that the retrieval performance can be improved by using the Bhattacharyya distance [20, p. 99] as a dissimilarity measure. This distance takes into account not only the separation induced by different mean vectors, but also the separation due to the difference in covariance matrices. In addition, we showed that efficient computation of the Bhattacharyya distance is possible when most of the energy in the feature space is restricted to a low dimensional subspace. The improved representation will be implemented into the IGDS system.

The digitized specimens in the database have all the same magnification, therefore, the nuclear area is computed as the number of pixels inside the delineated nucleus. The dissimilarity between two nuclei in terms of their areas is expressed as

$$D_3 = \sqrt{\left(a_{query} - a_{reference}\right)^2} \ . \tag{4}$$

The nuclear color is expressed as a 3-D vector in the $L^*u^*v^*$ space and is determined during the segmentation as the center of the associated color cluster. However, since the colors of the nuclei in the database cannot discriminate among the digitized specimens, the current implementation of the system uses the color attribute only for nucleus separation from the background.

3.5 Overall Dissimilarity Metric

Recall that the cases in the reference database fall into one of the four categories: CLL, FCC, MCL, or normal (benign). The suggested classification of the query image is based on the voting kNN rule [20, p. 305] among the classes of the closest k matches. That is

$$k_i = max\{k_1, \ldots, k_4\} \;\; \rightarrow \;\; X \in \omega_i \tag{5}$$

where k_i is the number of neighbors from the class ω_i $(i = 1, \ldots, 4)$ among the kNN's, and $k_1 + \ldots + k_4 = k$. In addition to the four original cell classes, the kNN rule may also produce a NO DECISION class, in the case when the value of i verifying (5) is not unique.

We measure the system performance through the *confusion matrix* $\mathbf{R}$ defined as having as element $r_{j,i}$ the empirical probability of classification in class j when the query image belonged to class i, $P(j|i)$. The criterion that should be maximized is the sum of conditional probabilities of correct decision

$$J = \sum_{j=1}^{4} P(j|j) \; . \tag{6}$$

The dissimilarity between two cell nuclei is expressed as a linear combination of the distances corresponding to each query attribute. Thus, for three attributes (e.g., shape, texture, and area) we have the overall distance

$$D = \sum_{i=1}^{3} w_i D_i \; , \tag{7}$$

where w_i represents the relevance of the i-th attribute and $\sum_{i=1}^{3} w_i = 1$.

The best weights w_i were derived off-line by employing the downhill simplex method [36, p. 408] with the objective function J (6). A simplex in N dimensions consists of $N+1$ totally connected vertices. The optimization is based on a series of steps which reflect, expand, and contract the simplex such that it converges to a maximum of the objective function. As an advantage, the downhill simplex requires only function evaluations and no computation of derivatives.

Table 1: *Best weights and the value of optimization criterion corresponding to the global maximum.*

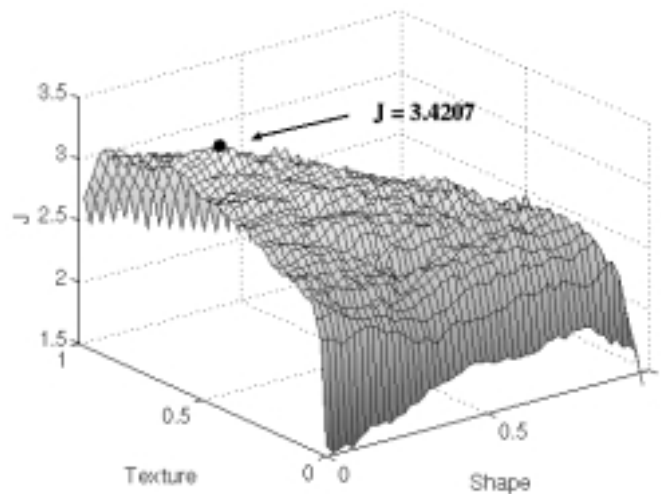| Shape	Texture	Area	J
0.1140	0.5771	0.3089	3.4207

Figure 11: Plot of the objective surface (resolution is 0.02 on each dimension). The downhill simplex converged in this case to the global maximum (from [14]).

Table 1 presents the best set of weights, obtained by running the optimization with seven retrievals over the entire database. It corresponds to the highest obtained value ($J = 3.4207$) of the objective function (6). Figure 11 shows the objective surface as a function of the two

independent weights.

3.6 Performance Evaluation and Comparisons

At present, the IGDS system is being evaluated in real retrieval scenarios at the Department of Pathology, UMDNJ-RWJ Medical School. Since the sequential searching and ranking of the logical database take about 50 ms on a Pentium II at 266 MHz, the extension of the current database to thousands of images is possible with no noticeable increase in the delay at the end user. The retrieval delay depends mostly on the bandwidth available for the client-server communication.

To obtain a more realistic estimation of the retrieval performance, we performed the ten-fold cross-validated classification [17, p. 238] of the entire database. The data set was randomized and split into 10 approximately equal test sets, each containing about 9 CLL, 3 FCC, 6 MCL, and 5 Normal cases. For the q-th test set its complement was used to obtain the best weights through the downhill simplex method described above. The confusion matrix $\mathbf{R}_q$ of the resulting classifier was then computed over the q-th test set for seven retrievals. The elements of the cross-validated confusion matrix were defined as

$$P_{cv}(j|i) = \frac{1}{10} \sum_{q=1}^{10} P_q(j|i) \ , \tag{8}$$

for $i = 1 \ldots 4$, and $j = 1 \ldots 5$.

According to the data in Table 2, the system performance is satisfactory, especially when related to the current difficulties in differentiating among lymphoproliferative disorders based

solely on morphological criteria [4].

Table 2: *Ten-fold cross-validated confusion matrix (7 retrievals): IGDS system.*

	CLL	FCC	MCL	NRML	NO DEC
CLL	**.8389**	.0200	.0711	**.0700**	.0000
FCC	.0250	**.9000**	.0000	**.0500**	.0250
MCL	.1357	.0143	**.8333**	**.0000**	.0167
NRML	.1333	.1200	.0000	**.7300**	.0167

Table 3: *Confusion matrix: Human experts.*

	CLL	FCC	MCL	NRML	NO DEC
CLL	**.5647**	.0352	.2117	**.1764**	.0117
FCC	.0285	**.9428**	.0000	**.0285**	.0000
MCL	.1538	.0769	**.5538**	**.1692**	.0461
NRML	.1228	.0000	.1053	**.7543**	.0175

	CLL	FCC	MCL	NRML	NO DEC
CLL	**.4000**	.0588	.1647	**.3765**	.0000
FCC	.0000	**1.000**	.0000	**.0000**	.0000
MCL	.0769	.0923	**.5538**	**.1692**	.0923
NRML	.0000	.0877	.1053	**.7719**	.0351

	CLL	FCC	MCL	NRML	NO DEC
CLL	**.4941**	.0235	.2118	**.2000**	.0471
FCC	.0000	**.8857**	.0857	**.0286**	.0000
MCL	.4308	.0154	**.3077**	**.0308**	.2154
NRML	.2000	.0364	.1455	**.3455**	.2727

The confusion matrices representing the results of three human experts classifying the digitized specimens from the same database are presented in Table 3. The human experts were shown one digitized specimen at a time on a high resolution screen with no other distractor displayed.

By comparing the Table 2 and Table 3 we observe that the human performance is slightly better for FCC and Normal cases, but it is worse for the CLL and MCL cases, both in terms of probabilities of correct decision (the marked diagonals) and probabilities of false negatives (the NRML column). The correlation between the human and machine results is also noteworthy. The classification of the FCC cells proved to be the easiest task while the CLL and MCL cells resulted in similar levels of difficulty.

We note here that in a real classification scenario, the human expert uses a lot of context information including both patient data and additional data inferred from the digitized specimens. We therefore stress the Decision Support function of the IGDS system. The system is not intended to provide automatic identification of the disorder, but to assist the pathologist to improve its own analysis. The pathologist combines the objective classification suggested by the system with the context information to obtain a robust diagnostic decision.

4   Conclusion

This chapter discussed an effective algorithm for cell segmentation and showed its integration in a real-time system that supports decision making in clinical pathology. The nonparametric nature of the segmentation and its robustness to noise allowed the use of a fixed resolution for the processing of hundreds of digital specimens captured under different conditions.

The segmentation has been indirectly evaluated through the IGDS system which demonstrated satisfactory overall performance. As a broader conclusion, however, this research proved that the segmentation, although a very difficult task in its general form, can become

a successful processing step when the goal of the vision application is well defined.

Acknowledgment

References

[1] S. Antani, R. Kasturi, R. Jain, Pattern Recognition Methods in Image and Video Databases: Past, Present and Future, In: Amin A et al.(ed), Advances in Pattern Recognition, Lecture Notes in Comp. Science, Springer, Vol. 1451, pp. 31-53, 1998.

[2] P.M. Banks et al., Mantle Cell Lymphoma: A proposal for Unification of Morphologic, Immunologic, and Molecular Data, Am. J. Surg. Pathology, Vol. 16, No. 7, pp. 637-640, 1992.

[3] I. Bauman, R. Nenninger, H. Harms, H. Zwierzina, K. Wilms, A.C. Feller, V.T. Meulen, H.K. Muller-Hermelink, Image Analysis Detects Lineage-Specific Morphologic Markers in Leukemic Blast Cells, Am. J. Clin. Pathol., Vol. 105, No. 1, pp. 23-30, 1995.

[4] E. Campo, E. Jaffe E, Mantle Cell Lymphoma, Arch. Pathol. Lab. Med., Vol. 120, No. 1, pp. 12-14, 1996.

[5] J. Chan, B.M. Banks, et al., A Revised European-American Classification of Lymphoid Neoplasms Proposed by the International Lymphoma Study Group, Am. J. Clin. Pathol., Vol. 103, No. 5, pp. 543-560, 1995.

[6] Y. Cheng, Mean shift, mode seeking, and clustering, IEEE Trans. Pattern Anal. Machine Intell., Vol. 17, No. 8, pp. 790-799, 1995.

[7] K. Cho, P. Meer, Image Segmentation from Consensus Information, Comp. Vis. and Image Understanding, Vol. 68, No. 1, pp. 72-89, 1997.

[8] D. Comaniciu, D. Foran, P. Meer, Shape-Based Image Indexing and Retrieval for Diagnostic Pathology, Proc. Int'l Conf. on Pattern Recognition, Brisbane, Australia, pp. 902-904, 1998.

[9] D. Comaniciu, P. Meer, D. Foran, A. Medl, Bimodal System for Interactive Indexing and Retrieval of Pathology Images, IEEE Workshop on Applications of Comp. Vis., Princeton, New Jersey, pp. 76-81, 1998.

[10] D. Comaniciu, P. Meer, Distribution Free Decomposition of Multivariate Data, Pattern Analysis and Applications, Vol. 2, No. 1, pp. 22-30, 1999.

[11] D. Comaniciu, P. Meer, Mean Shift Analysis and Applications, IEEE Int'l Conf. Comp. Vis., Kerkyra, Greece, pp. 1197-1203, 1999.

[12] D. Comaniciu, P. Meer, K. Xu, D. Tyler, Retrieval Performance Improvement through Low Rank Corrections, IEEE Workshop on Content-Based Access of Image and Video Lib., Fort Collins, Colorado, pp. 50-54, 1999.

[13] D. Comaniciu, B. Georgescu, P. Meer, W. Chen, D. Foran, Decision Support System for Multiuser Remote Microscopy in Telepathology, IEEE Symposium on Computer-Based Medical Systems, Stamford, Connecticut, pp. 150-155, 1999.

[14] D. Comaniciu, P. Meer, D. Foran, Image Guided Decision Support System for Pathology, Machine Vision and Applications, In press, 1999.

[15] I.J. Cox, M.L. Miller, T.P. Minka, P.N. Yianilos, An Optimized Interaction Strategy for Bayesian Relevance Feedback, IEEE Conf. on Comp. Vis. and Pattern Recognition, Santa Barbara, California, pp. 553-558, 1998.

[16] M. Das, E.M. Riseman, B.A. Draper, FOCUS: Searching for Multi-colored Objects in a Diverse Image Database, IEEE Conf. on Comp. Vis. and Pattern Recognition, San Juan, Puerto Rico, pp. 756-761, 1997.

[17] B. Efron, R. Tibshirani, An Introduction to the Bootstrap, Chapman & Hall, New York, 1993.

[18] M. Flickner et al., Query by Image and Video Content: The QBIC System, Computer, Vol. 9, No. 9, pp. 23-31, 1995.

[19] K. Fukunaga, L.D. Hostetler, The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition, IEEE Trans. Info. Theory, Vol. IT-21, No. 1, pp. 32-40, 1975.

[20] K. Fukunaga, Introduction to Statistical Pattern Recognition, Second Ed., Academic Press, Boston, 1990.

[21] H. Kauppinen, T. Seppanen, M. Pietikainen, An Experimental Comparison of Autoregressive and Fourier-Based Descriptors in 2D Shape Classification, IEEE Trans. Pattern Analysis Machine Intell., Vol. 17, No. 2, pp. 201-207, 1995.

[22] J. Kittler, M. Hatef, R.P.W. Duin, J. Matas, On Combining Classifiers, IEEE Trans. Pattern Analysis Machine Intell., Vol. 20, No. 3, pp. 226-238, 1998.

[23] F.P. Kuhl, C.R. Giardina, Elliptic Fourier Features of a Closed Contour, Comp. Graphics Image Process., Vol. 18, pp. 236-258, 1982.

[24] F. Liu, R.W. Picard, Periodicity, Directionality, and Randomness: Wold Features for Image Modeling and Retrieval, IEEE Trans. Pattern Analysis Machine Intell., Vol. 18, No. 7, 722-733, 1996.

[25] Y. Liu, F. Dellaert, A Classification Based Similarity Metric for 3D Image Retrieval,, IEEE Conf. on Comp. Vis. and Pattern Recognition, Santa Barbara, California, pp. 800-805, 1998.

[26] W.Y. Ma, B.S. Manjunath, NETRA: A Toolbox for Navigating Large Image Databases, IEEE Int'l Conf. Image Process., Santa Barbara, California, Vol. 1, pp. 568-571, 1997.

[27] J. Mao, A.K. Jain, Texture Classification and Segmentation using Multiresolution Simultaneous Autoregressive Models, Pattern Recognition, Vol. 25, No. 2, pp. 173-188, 1992.

[28] B. Moghaddam, A. Pentland, Probabilistic Visual Learning for Object Representation, IEEE Trans. Pattern Analysis Machine Intell., Vol. 19, No. 7, pp. 696-710, 1997.

[29] C. Nastar, M. Mitschke, C. Meihac, Efficient Query Refinement for Image Retrieval, IEEE Conf. on Comp. Vis. and Pattern Recognition, Santa Barbara, California, pp. 547-552, 1998.

[30] M. Ortega, Y. Rui, K. Chakrabarti, S. Mehrotra, T.S. Huang, Supporting Similarity Queries in MARS, ACM Multimedia'97, Seattle, Washington, pp. 403-413, 1997.

[31] A. Pentland, R.W. Picard, S. Sclaroff, Photobook: Content-based manipulation of image databases,Int'l. J. of Comp. Vis., Vol. 18, No. 3, pp. 233-254, 1996.

[32] E. Persoon, K.S. Fu, Shape Discrimination using Fourier Descriptors,IEEE Trans. on Systems, Man, and Cybern., Vol. 7, No. 3, pp. 170-179, 1977.

[33] E.G.M. Petrakis, C. Faloutsos, Similarity Searching in Large Databases,IEEE Trans. on Knowledge and Data Engn., Vol. 9, No. 3, pp. 435-447, 1997.

[34] R.W. Picard, T. Kabir, F. Liu, Real-time Recognition with the Entire Brodatz Texture Database, IEEE Conf. on Comp. Vis. and Pattern Recognition, New York, pp. 638-639, 1993.

[35] K. Popat, R.W. Picard, Cluster-Based Probability Model and Its Application to Image and Texture Processing, IEEE Trans. Image Process., Vol. 6, No. 2, pp. 268-284, 1997.

[36] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in C, Second Ed., University Press, Cambridge, 1992.

[37] Y. Rui, A.C. She, T.S. Huang, A Modified Fourier Descriptor for Shape Matching in MARS, In: Chang SK (ed) Image Databases and Multimedia Search, Series of Software Engn. and Knowledge Engn., World Scientific Publishing, Singapore, Vol. 8, pp. 165-180, 1998.

[38] R. Sedgewick, Algorithms in C++, New York: Addison-Wesley, 1992.

[39] R. Sharma, V.I. Pavlovic, T.S. Huang, Toward Multimodal Human-Computer Interface,Proceedings of the IEEE, Vol. 86, No. 5, pp. 853-869, 1998.

[40] H.D. Tagare, C.C. Jaffe, J. Duncan, Medical Image Databases: A Content-based Retrieval Approach, J. Am. Medical Inform. Assoc., Vol. 4, No. 3, pp. 184-198, 1997.

[41] E. Truco, A. Verri,Introductory Techniques for 3-D Computer Vision, NJ: Prentice Hall, 1998.

[42] G. Wyszecki, W.S. Stiles, Color Science: Concepts and Methods, Quantitative Data and Formulae, Second Ed. New York: Wiley, 1982.

[43] S.C. Zhu, A. Yuille, Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation, IEEE Trans. Pattern Anal. Machine Intell., Vol. 18, No. 9, pp. 884-900, 1996.