Automated Interpretation of Optic Nerve Images: A Data Mining Framework for Glaucoma Diagnostic Support

Syed SR. Abidi[a], Paul H. Artes[b], Sanjan Yun[a], Jin Yu[a]

[a] *NICHE Research Group, Faculty of Computer Science, Dalhousie University, Halifax, Canada*
[b] *Department of Ophthalmology and Visual Sciences, Dalhousie University, Halifax, Canada*

Abstract

Confocal Scanning Laser Tomography (CSLT) techniques capture high-quality images of the optic disc (the retinal region where the optic nerve exits the eye) that are used in the diagnosis and monitoring of glaucoma. We present a hybrid framework, combining image processing and data mining methods, to support the interpretation of CSLT optic nerve images. Our framework features (a) Zernike moment methods to derive shape information from optic disc images; (b) classification of optic disc images, based on shape information, to distinguish between healthy and glaucomatous optic discs. We apply Multi Layer Perceptrons, Support Vector Machines and Bayesian Networks for feature sub-set selection and image classification; and (c) clustering of optic disc images, based on shape information, using Self-Organizing Maps to visualize sub-types of glaucomatous optic disc damage. Our framework offers an automated and objective analysis of optic nerve images that can potentially support both diagnosis and monitoring of glaucoma.

Keywords:

Glaucoma, Data Mining, Feature Selection, Clustering, Confocal Scanning Laser Tomography, Support Vector Machines

Introduction

Glaucoma is an eye disease that is characterized by slow progressive damage to the optic disc and corresponding deterioration of the patient's vision [1]. At present, there is a gap in the understanding of the cause, the types and the natural course of glaucoma. The use of sophisticated imaging technologies, such as Confocal Scanning Laser Tomography (CSLT), capture 3-dimensional images of the optic disc that are used for diagnostic purposes [2]. However, the interpretation of CSLT images is a manual and subjective process—a trained professional has to manually define the margins of the optic nerve based on his/her training and expertise and then classify the optic nerve as normal or glaucomatous. The current process allows for misjudgments/errors in the interpretation of the CSLT image, inability to distinguish between actual and noisy images and variance in the diagnostic recommendations over a set of practitioners. The challenge, therefore, is to automate the analysis of CSLT images of the optic disc, in an objective and quantifiable manner, to support practitioners in the diagnosis and therapeutic management of glaucoma.

Researchers have analyzed optic nerve data and CSLT based images with varying results. Bowd et al [3], working with retinal tomography images applied forward and backward feature selection methods for training Multi Layer Perceptron (MLP), Support Vector Machine (SVM) and linear discriminant functions; Park et al [4] used correlation analysis and forward wrapper model to select features from optic disc data for training SVM classifiers; Swindale et al [5] used a wrapper model for feature selection to train SVM classifiers.

We have developed a data-driven Glaucoma Diagnostic Support (GDS) system that features the automatic interpretation of CSLT topography images of the optic nerve to support (a) the classification of the optic disc images to distinguish between healthy and diseased optic discs; (b) the identification of the sub-types of glaucomatous optic disc damage. This is to help further sub-classify the glaucoma patients in order to provide treatments in line with the specific morphological patterns of damage [6]; and (c) the visualization of the temporal progression of the disease for a patient over a period of time.

In this paper we present an automated approach to CSLT topography image analysis to support glaucoma diagnosis. Our multi-stage approach is a hybrid of image processing and data mining methods. In Stage 1, we apply image-processing techniques to CSLT images to derive image-defining features. In Stage 2, we apply data classification methods to the image's shape-defining features to develop classifiers that can discriminate between healthy and glaucomatous optic discs. An important task at this stage is feature selection whereby we select an optimal subset of image features that exhibit high image classification capabilities. In Stage 3, we apply data clustering techniques to the optimal subset of image-defining features in order to identify the different sub-types of glaucomatous images in the image data-set. The emergent image clusters are subsequently used to both visualize the progression of the disease and the identification of noisy optic nerve images. In Stage 4, we apply rule-induction techniques to the optimal subset of features to induce classification rules (not discussed here). These symbolic rules provide practitioners with a justification of the diagnostic recommendations by our

image classifiers. For our experiments, we worked with 1257 tomography images taken at different time intervals from 136 subjects (51 healthy subjects and 85 glaucoma patients).

Methods

Figure 1 illustrates the functional design of our GDS system. We explain the methods developed for each processing stage.

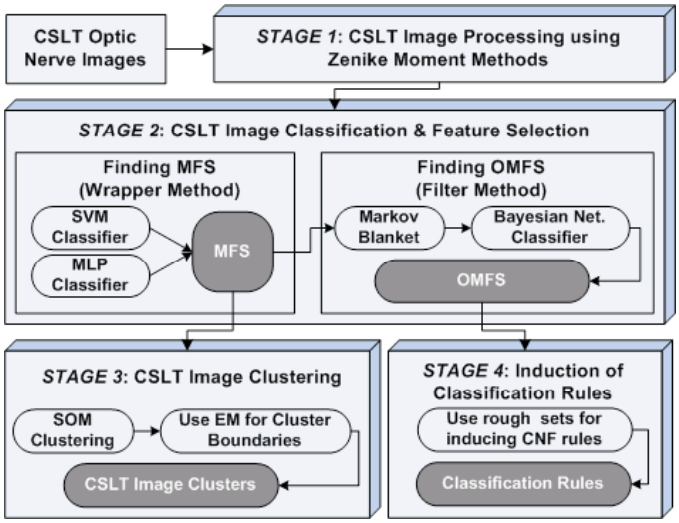

Figure 1 – Functional design of our GDS system

Stage 1: CSLT Image Processing

This stage involves the extraction of shape-defining features from CSLT images. These features are used to develop the image classification and clustering models. We use an image processing technique referred to as Moment Methods that describes the properties of connected regions in binary images as *Moment* features. We use Zernike moments [7] which use a set of complex polynomials to describe the image's properties by their order (n) and repetition (m) with respect to a digital image—the low order moments capture gross shape information and high order moments incrementally resolve high frequency information (representing detail) of the digital image. Two attractive features of Zernike moments for our purpose is that (a) moments can be made invariant to shifts, rotations and magnification changes; and (b) the optic nerve is centered in the image, thus avoiding the requirement for an independent segmentation stage in which the object is explicitly identified.

For each CSLT image we generated 254 Zernike moments, grouped in an incremental order ranging from 1 to 29—each group comprises a set of ordered moments. Low order moments capture fundamental geometric properties and high order moments represent detailed information of the image [7].

For efficient classification of CSLT images, it was important to select an optimal number of lower order moments. This is a non-trivial task because: (a) there is no objective measure to determine the exact number of (low order) moments needed to achieve high classification accuracy; and (b) there is no discernable relationship between the moments that can be utilized. Given these challenges, next we pursue feature subset selection in conjunction with image classification.

Stage 2: Classification of CSLT Images

In the previous stage, we derived a 254 moment based representation for each CSLT image. In this stage, we pursue the classification of CSLT images based on a sub-set of low order moments. This stage therefore involves two tasks—i.e. firstly feature (sub-set) selection and secondly image classification. We have developed a two-pass image classification method that incorporates feature sub-set selection as an integral element (see Figure 1). In the first pass, MLP and SVM based wrapper models are simultaneously used to generate a *Moment Feature Subset* (MFS) consisting of low order moment features. In the second pass, we apply a Markov blanket filter method [8] based on an inferred Bayesian network to select the highly relevant moments from the MFS—i.e. the *Optimal Moment Feature Subset* (OMFS)—that offers reasonably high image classification despite using a small number of moments.

Pass I: Using MLP and SVM

In the absence of any guiding principle to determine the size of the MFS, we devised an accumulative feature subset selection strategy as follows: (a) Generate training-set by incrementally adding the next higher order moments to an existing training set. We exploited the intrinsic partitioning of the 254 moments in terms of their order ranging from 1 to 29. Therefore, feature subset1 included moments with *order2*, feature subset2 includes moments with order 2 and 3, and so on. In total 29 different training sets were generated, where each training set covered all the images based on the moment orders chosen to represent it; (b) Train both a MLP and a SVM classifier separately using the 29 training sets. In total, 29 different MLP and SVM classifiers were trained. For training the MLP and SVM, we partitioned the images so that 75% images were used for training and 25% images were used for testing. For training the SVM, based on the training data a 5-fold cross validation was performed to find the optimal hyper parameters: C and λ; and (c) Determine the classification accuracy of both classifiers, using the test images that are represented by the same number of moments as used to train the classifier.

The next step was to determine the size of the MFS and based on it to select the most efficient MLP and SVM classifier. Our objective was to select the largest possible number of moments without compromising the classification accuracy. To do so, we plotted the classification accuracy of both classifiers and then identified the highest accuracy point on the plot (i.e. with respect to n moment groups) just prior to a downward trend in the classification accuracy as a result of the inclusion of the next higher moment group. The most low order moment groups that achieved the highest classification accuracy were selected as the MFS. And, the MLP and SVM classifiers trained using the MFS were deemed as the most efficient.

A comparison of the classification accuracy trends for both the MLP and the SVM classifiers (see figure 2) shows that both classifiers exhibited a similar classification accuracy trend—i.e. they both start with a relatively high accuracy with the first moment group and then the accuracy drops with the addition of the next few moment groups. But later the accuracy starts to pick up again such that for the MLP it peaks when the feature subset constitutes the first 8 moment groups, whereas for the

SVM the accuracy peaks for the first 11 moment groups. It is interesting to note that the classification accuracy with higher order moment groups is relatively low as compared to the peak achieved with just the lower order moments.

Based on classification accuracy trend for both classifiers (shown in figure 2), we determined the MFS to contain the first 11 moment groups—i.e. the first 47 moments. With 11 moment groups the SVM exhibited the highest accuracy and the MLP produced its second highest accuracy level.

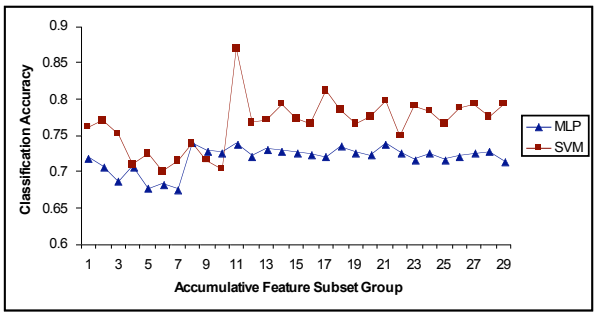

Figure 2 - Classification accuracy for both MLP and SVM

Pass II: Using Markov Blanket and Bayesian Network

In the second pass, we attempt to further reduce the size of the MFS in order to generate the OMFS that comprises only the highly salient moments. We use a filter model based on a Bayesian Network (BN) and the Markov blanket of the class label [8]. The choice of Markov blanket is guided by the observation that the correlation between most of the moments and their class label is weak, and the same is true for correlation between different moments. Hence, correlation based feature selection methods are not suitable here. We decided to use the Markov blanket approach as it considers every feature's probability dependence relationship during the learning procedure of the Bayesian network's structure.

In a BN where C_A is the set of children of node A, and Q_A is the set of parents of node A, the subset of nodes containing Q_A, C_A and the parents of C_A is called Markov blanket of A [8]. The Markov blanket of a specific feature is a subset of nodes in the BN; it comprises the feature's parent nodes, child nodes and all parent nodes of the child nodes. If we consider the class label node as the root node to learn a BN from data, then all nodes within the Markov blanket of the class node have probabilistic dependence relationship with it.

The steps to generate the Markov blanket were as follows: *Step 1*: We used the *K*2 algorithm to learn the BN. Initially, the 47 moments in the MFS were discretized using an entropy-based method, resulting in 29 moments to be discretized into a single value. These moments were removed from the MFS. Thus we were left with only 18 moments for training the BN. The following moments were retained: moments {1, 2, 5, 6, 7, 12, 16, 21, 23, 25, 27, 33, 36, 37, 43, 44, 45, 46}. *Step 2*: A BN was trained using the18 moments in their original order. Five-fold stratified cross validation was used to evaluate the classification accuracy (see table 2 for results). *Step 3*: The 18 moments were ordered based on the chi squared statistical test score x^2 between the moments and their class labels. The moments with the highest x^2 were: {1, 43, 16, 25, 21, 23, 6, 5, 36, 2, 27, 33, 37, 7, 46, 45, 44, 12}. A BN was learnt using the ordered moments (see table 2 for results). *Step 4*: From the BN learnt in step 3, we inferred the Markov blanket of the class label and found that only six (6) moments {1, 6, 16, 21, 37 46} were within the Markov blanket of the class label. These six moments were selected to form the OMFS. *Step 5*: In order to determine the classification capability of the selected OMFS we used them to train a BN. Next, 5-folds cross validation's classification accuracy was calculated (see table 2 for results) and it was noted that the OMFS offers quite high classification accuracy.

Stage 3: Clustering of CSLT Images

In this stage we pursued the clustering of the CSLT optic nerve images, represented using the 47 moments in the MFS, to differentiate between the different subtypes of healthy and glaucomatous optic nerves. It may be noted that an important theme in glaucoma research is to develop an understanding of the large variation in the appearance of the optic nerve, both within groups of healthy subjects and in patients with glaucoma. It is therefore important, from a clinical standpoint, to recognise and differentiate between such patterns. However, the problem with the sub-classification of patterns of optic nerve damage is that it is a subjective task, giving rise to considerable levels of disagreement between trained experts. In this context, our aim was to develop an objective and automated method to characterize optic nerve images.

Our two phase clustering strategy was to: (a) partition the images into distinct clusters using Self Organizing Maps (SOM); and (b) draw clear and distinct boundaries around the clusters using the Expectation Maximization (EM) algorithm [9].

Phase A: Data Clustering Using SOM

We used a SOM to cluster the CSLT images based on the similarities between image shapes, where each cluster may represent a different subtype of healthy and glaucomatous optic nerves. The training of the SOM was conducted as follows: (i) we determined the topology of the SOM to be hexagonal lattice comprising 192 units that were arranged as 16 rows and 12 columns; (ii) The units were linearly initialized along the two greatest eigenvectors of the covariance matrix of the training data—i.e. images represented using the 47 moments in the MFS; (iii) The SOM was trained using a sequential training algorithm by first running a rough training phase comprising 100 epochs starting with a large neighbourhood radius of 12 that was linearly reduced to 3 with a learning rate of 0.5. This was followed by a second fine-tuning phase comprising 1000 epochs with a small initial neighbourhood radius 3 that was reduced to 1 with learning rate of 0.1. In both cases a Gaussian neighbourhood function was used and the learning rate function was set to be inversely proportional to the training epochs; (iv) Finally, we achieved a trained SOM that placed similar images into close proximity, thus leading to the image clusters. We applied principal component projection to the learnt SOM to determine its projection. This was followed by developing a U-matrix representation of the learnt SOM by spreading a colour map on the projection. Based on the visualization offered by the SOM, it was noted that the data was partitioned into discernable clusters.

Phase B: Defining the Cluster Boundaries

After determining broad clusters of CSLT images, in this phase we objectively determine the cluster boundaries. The processing was guided by our assumption that the distribution of the clusters within the learnt SOM is Gaussian. Therefore, we used the EM algorithm [9] as it is suitable to find distinct components in the case of Gaussian mixtures. Functionally, the EM algorithm initiates with an estimate of the number of components and their parameters. Our strategy was to maximize the likelihood of the optic nerve images into distinct clusters given the parameters and a maximum likelihood measure that indicated how well the Gaussian mixtures fit the data into clusters. We used a Bayesian Information Criterion (BIC) [9], where the best estimate (e.g., number of clusters) was chosen based on the highest BIC value.

To achieve the cluster boundaries, using the EM method with BIC, we initialized the EM using 10 random re-starts method, and then selected a parameter setting to maximize the log-likelihood of our clusters from the SOM. EM clustering was performed for different number of clusters. Table 1 shows that the maximum BIC is achieved when K = 4. Hence, we determined that given the learnt SOM there are 4 clusters—one cluster represents health images and the three clusters for subtypes of glaucomatous images—in it that best fit the data (see Table 1). To finalize the cluster boundaries for the 4 clusters, we calculated the assignment probabilities of each CSLT image to all the cluster labels, the cluster label with the highest probability value was assigned to the CSLT image. Figure 2a shows the SOM with the emergent clusters, the clusters are coded using grey scale for visualization purposes.

Table 1 – Number of clusters vs. BIC values

K	2	3	4	5	6	7	8
BIC	29100	30409	31354	30516	29125	27456	25486

Evaluation and Discussion

In this section we present the evaluation results for the various methods developed for stages 2 and 3 of our GDS system.

Evaluation 1: Evaluating CSLT Image Classification

Table 2 presents the CSCLT image classification accuracy for the different classifiers trained in phase 2, using test images not previously seen by the classifiers. It is interesting to note that both the MLP and the SVM classifiers offered higher accuracy with the MFS as compared to the original 254 moments. This vindicates our hybrid feature sub-set selection strategy, and also confirmed the theoretical assumption that low order moments contain more shape information that is relevant for classification as compared to the information content of high order moments. In the second pass, we determined that the MFS could be further reduced to just 6 moments—i.e. the OMFS—without compromising the classification accuracy. The highest accuracy for MFS was offered by the SVM—i.e. 86.96%. The highest accuracy for the OMFS was 83.82% offered by a BN. Therefore, the compromise in classification accuracy is just 3 %, yet the gain in computational efficacy is quite significant. Note that the BN (with Markov Blanket) offers the most optimal classification results when compared with both MLP and SVM trained on the OMFS. We therefore selected the BN classifier trained with the OMFS to distinguish between healthy and glaucomatous optic nerves.

Table 2 - Classification accuracy for different classifiers

Feature Subset Size	Classifier	Accuracy
Pass I		
254 moments	MLP	72.88%
254 moments	SVM	77.50%
47 moments in MFS	MLP	74.00%
47 moments in MFS	SVM	86.96%
Pass II		
18 moments (original order)	BN	77.21%
18 moments (chi^2 order)	BN	80.88%
6 moments in OMFS	**BN**	**83.82%**
6 moments in OMFS	*SVM*	*80.26%*
6 moments in OMFS	*MLP*	*72.84%*

Evaluation 2: Examining the CSLT Image Clusters

Evaluation of the clustering stage involved mapping a series of optic nerve images for individual patients (i.e. test cases with explanations provided by experts) onto the SOM and noting the Compactness Factor (CF) between the activated units. The CF measures how close the images are with respect to each other in terms of the average distance between the centroid of all active units. The CF is an objective measure for evaluating the clustering goodness based on our initial observation that for a patient the series of optic nerve images are quite similar over a period of time; over time the differences are quite minute and should not lead to large variations between consecutive images. This implies that when visualizing the optic nerve images for a subject, the active units should be in close proximity and therefore yield a low CF.

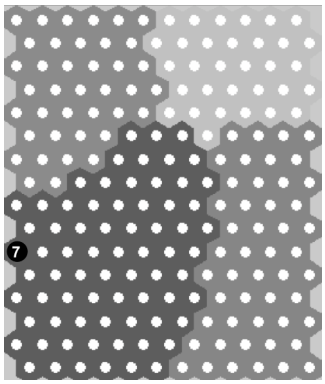

Figure 2a – SOM showing all images mapped to a unit

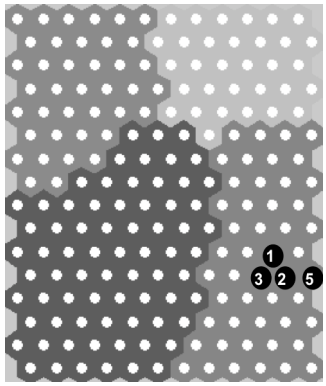

Figure 2b – SOM showing all images mapped in one cluster

Figure 2a show that the results for patient 4209643, and it maybe noted that the 7 optic nerve images, taken over a period of time, map on to a single SOM unit resulting in a compactness factor of to 0. The numeral within the active unit shows the number of images mapping on to that unit. This demonstrates the best possible clustering outcome as the learnt SOM recognizes the similarity between all the 'healthy' optic nerve images for this patient. Figures 2b shows the 11 optic nerve images of patient 112455 being mapping on to 4 neighboring SOM units within one cluster, with a compactness factor of 0.20808. This result again implies the close proximity of the

images for this patient. These results are in line with the visual observations of these images by experts, who also concurred that the images for these patients are quite similar in shape.

Evaluation 3: Visualizing Disease Progression Over Time

We use the learnt SOM to visualize the disease progression for a patient over a period of time. Images taken over time for a patient were mapped onto the SOM. The pattern of the active units indicated the potential progression of the disease from one cluster to another, where each cluster may represent images of a specific glaucoma sub-type. In Figure 3a the images fall into two adjoining clusters, and the path across the clusters suggests the progression of the disease from one sub-type to another. Figure 3b shows the progression over time.

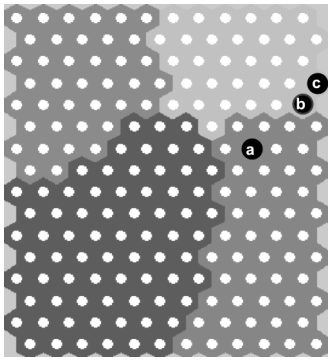

Figure 3a - SOM showing the dispersion of images over two adjoining clusters

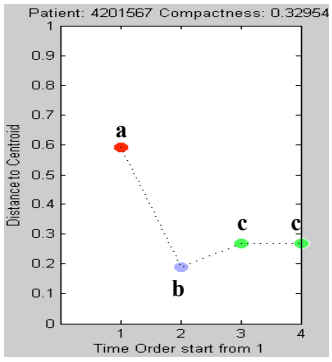

Figure 3b – The disease progression path. Note the high CF between the images

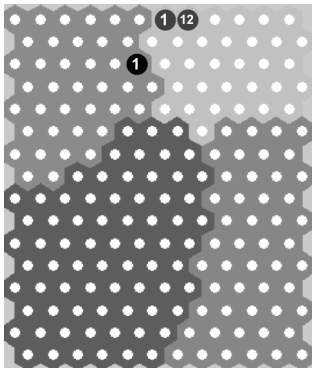

Figure 4a - SOM showing a noisy image that is distant from the other images.

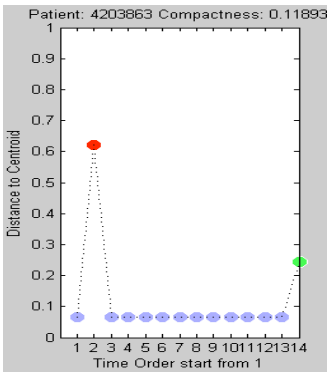

Figure 4b – The spike shows that the 2nd image is noisy, as it does not follow the pattern.

Evaluation 4: Identifying Noisy CSLT Images

We used the learnt SOM to identify noisy CSLT images that typically occur due to various factors related to the capture of the optic nerve image. With the knowledge that consecutive images do not manifest drastic changes, if an image is noted to be significantly dissimilar from its neighbors it can be regarded as a noisy image. At present there are no objective means to identify noisy CSLT images. Figure 4 (a-b) shows 14 images for a patient, where the 2nd image is identified as a single noisy image because it is in a different cluster, whereas the remaining images all map onto just two other units that are very close to each other.

Concluding Remarks

We presented a data mining framework to objectively analyze medical images, and applied it to investigate glaucoma. The novel features of our approach are that: (a) we process CSLT images to derive shape information by using image processing techniques. This is in contrast to the traditional approach of using morphological features to analyze CSLT images; (b) we have developed a feature selection strategy that identifies the most salient image-defining features without compromising the classification accuracy; and (c) we are able to visualize the CSLT images in terms of clusters of similar images. These clusters provide an opportunity to visualize the dispersion of multiple observations for a subject, and we show how this information can help to (i) determine a potential progression of the disease due to changes in the optic disc over time; and (ii) identify noisy CSLT images. We believe that our framework takes a step towards the automated and objective analysis of optic nerve images to support glaucoma diagnostics.

Acknowledgments

This project is supported by a research grant from Nova Scotia Health Research Foundation, Canada.

References

[1] Hoskins H, Kass M. Primary open-angle glaucoma, Becker-Shaffer's Diagnosis and Therapy of the Glaucomas. The C.V. Mosby Company, St.Louis, 1989: 277-307.

[2] Zinser G, Wijnaendts-van-Resand RW, Dreher AW. Confocal laser tomographic scanning of the eye. Proc. SPIE 1161, 1980: 337–344,

[3] Bowd C, Cban K, Zangwill LM, Goldbaum MH, Lee T, Sejnowski TJ. Comparing neural networks and linear discriminant functions for glaucoma detection using confocal scanning laser opthalmoscopy of the optic disc. Investigative Opthalmology & Visual Science, Vol. 43 (11), 2002.

[4] Park J, Reed J, Zhou Q. Active feature selection in optic disc nerve data using support vector machine. IEEE World Congress on Computational Intelligence, 2002.

[5] Swindale NV, Stjepanovic G, Cbin A, Mikelberg F. Automated analysis of normal and glaucomatous optic nerve head topography images. Investigative Opthalmology & Visual Science, Vol. 41 (7), 2000.

[6] Broadway DC, Nicolela MT, Drance SM. Optic disc morphology on presentation of chronic glaucoma. *Eye* 17 (6), 2003: 798-799.

[7] The CH, Chin RT. On Image Analysis by the methods of moments. IEEE Trans. Pattern Analysis by Machine Intelligence 10(4), July 1998: 96-513.

[8] Hruschka Jr. ER, Hruschka ER, Ebecken NEF. Feature selection by Bayesian network. Canadian AI Conference, LNAI 3060, Springer Verlag, 2004:370-279.

[9] McLachlan G, Krishnan T. The EM algorithm and extensions. New York: John Wiley & Sons, 1997.

Address for correspondence

Syed Sibte Raza Abidi, Faculty of Computer Science, Dalhousie University, Halifax, B3H 1W5, Canada. Email: sraza@cs.dal.ca