31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Automatic Fundus Image Classification for Computer-Aided Diagonsis

Shijian Lu, Jiang Liu, Joo Hwee Lim, Zhuo Zhang, Tan Ngan Meng, Wing Kee Wong,
Huiqi Li, and Tian Yin Wong

Abstract— **With the advances of computer technology, more and more computer-aided diagnosis (CAD) systems have been developed to provide the "second opinion". This paper reports an automatic fundus image classification technique that is designed to screen out the severely degraded fundus images that cannot be processed by traditional CAD systems. The proposed technique classifies fundus images based on the image range property. In particular, it first calculates a number of range images from a fundus image at different resolutions. A feature vector is then constructed based on the histogram of the calculated range images. Finally, fundus images can be classified by a linear discriminant classifier that is built by learning from a large number of normal and abnormal training fundus images. Experiments over 644 fundus images of different qualities show that the classification accuracy of the proposed technique reaches above 96%.**

I. Introduction

With the advances of computer technology, various types of computer-aided diagnosis (CAD) systems [4], [5] have been developed in recent years. The developed CAD systems assist medical staff to interpret medical images and accordingly provide the "second opinion". They greatly help to relieve the limitations of the human eye/brain systems such as reader fatigue and distraction. Meanwhile a number of fundus image CAD systems have also been developed for the diagnosis of various types of ocular diseases such as glaucoma [6], [11], [12], macular diseases [1], and diabetic retinopathy [2], [3], [7]. These fundus image CAD systems have potentials to provide an alternative solution to mass screening programs that need to examine a vast number of fundus images as fast as possible.

So far, most fundus image CAD systems are still at the research stage and cannot be put in practical uses in clinics due to several reasons. One reason that prevents the practical uses of CAD systems lies with the diversity of fundus images across different people. In particular, fundus images may have far different qualities due to various types of lesion and artifacts shown in Fig. 1. As a result, it is very difficult to design an image processing algorithm that is capable of handling a large amount of fundus images with diverse characteristics. Take the optic disc location as an example. This simple task will not be so simple while dealing with fundus images with diverse characteristics as illustrated in Fig. 1 (where the "abnormal" fundus images shown in the last three subfigures are far different from the normal one shown in the first subfigure).

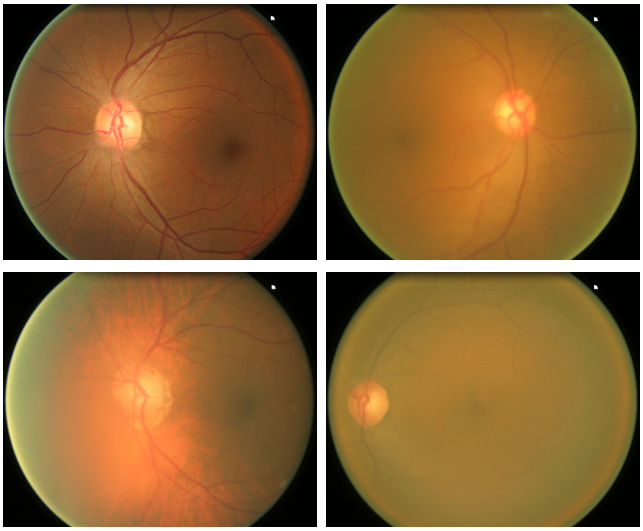

Fig. 1. Several sample normal and abnormal fundus images: The first on the top-left corner is a normal fundus image and the other three are abnormal fundus images.

One possible solution to the image diversity is to carry out a fundus image classification operation as a preprocessing procedure to screen out the "abnormal" fundus images (before feeding fundus images for further analysis). With such a preprocessing procedure a more robust CAD system can be designed to focus on the normal fundus images and those identified "abnormal" ones can be stored for later manual examination. In this paper we report a fundus image classification technique that efficiently identifies the abnormal fundus images illustrated in Fig. 1 automatically. The proposed technique is based on the observation that abnormal fundus images are usually smoother (compared with the normal ones) due to the blur of retina structures such as optic disc, blood vessels, and fovea. As a result, the histogram of the range image of abnormal fundus images will be much sharper (smaller variance) and centered more leftward compared with that of the normal fundus images. Experiments over 644 fundus images of different qualities show that the accuracy of the proposed classification technique reaches above 96%.

II. Proposed Method

This section presents the proposed fundus image classification technique. In the proposed technique, a fundus image is first converted into a feature vector based on histograms

Shijian Lu, Jiang Liu, Joo Hwee Lim, Zhuo Zhang, Tan Ngan Meng, Wing Kee Wong, and Huiqi Li are with Institute for Infocomm Research, A*STAR, Singapore {slu, jliu, joohwee, zzhang, nmtan, wkwong, huiqili}@i2r.a-star.edu.sg

Professor Tien Yin Wong is with Singapore National Eye Centre ophwty@nus.edu.sg

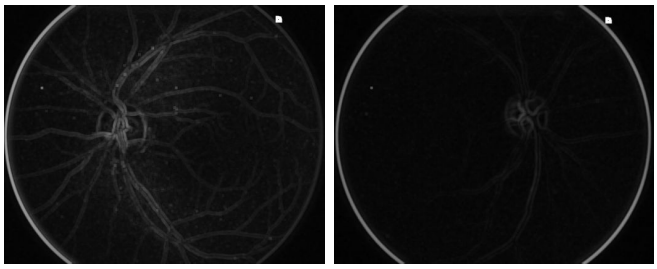

Fig. 2. The range images of the two fundus images in the first row in Fig. 1 that are calculated by using a neighborhood window of size 12.

of a number of range images that are converted from the fundus image at different resolutions. The fundus image can then be classified by a linear discriminant classifier that is trained by learning from feature vectors of a large number of normal and abnormal training fundus images.

A. Fundus Image Vectorization

We extract a feature vector from a fundus image by first converting the fundus image into a number of range images at different resolutions as follows:

$$R(x, y, w_i) = max(I(x, y, w_i)) - min(I(x, y, w_i)) \quad (1)$$

where x, y denote the position of the fundus image pixel under evaluation and w_i refers to the size of the i-th local neighborhood window that is used to calculate the range image. Therefore, $I(x, y, w_i)$ denotes the intensity of all fundus image pixels enclosed within the neighboring window that is centered at x, y with size w_i. For the two fundus images in the first row in Fig. 1, Fig 2 show the corresponding range images that are calculated by a neighborhood window of size 12 (i.e. 25×25).

We convert fundus images based on histograms of the corresponding range images, which have distinctive characteristics that can be used to differentiate the normal and abnormal fundus images. In particular, the histogram of the abnormal range images usually has a sharper peak (smaller variance) compared with that of normal range images. At the same time, it is usually centered more leftward compared with the histogram of normal range images. Such histogram characteristics can be explained by the fact that the range image of abnormal fundus images is usually much smoother as illustrated in Fig. 2 due to the blurrness of such retina structures as optic disc, blood vessels, and fovea. Fig. 3 illustrates such histogram characteristics where the two histograms (labeled by the solid graph and the dashed graph) are averaged over the histogram of the range image (with the window size set at 12) of 338 normal fundus images and 306 abnormal fundus images, respectively.

We capture the specific histogram characteristics (shown in Fig. 3) of normal and abnormal fundus images by using the histogram mean, histogram variance, and histogram energy

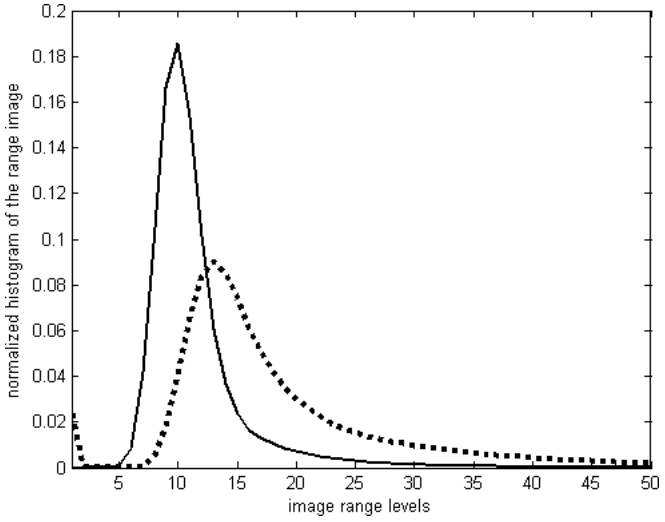

Fig. 3. The histograms of the range image of normal and abnormal fundus images: The dashed graph and the solid graph are determined by averaging the histograms of the range images of 338 normal fundus images and 306 abnormal fundus images, respectively.

that are defined as follows:

$$M = \sum_{b=0}^{L-1} bP(b)$$

$$V = \sum_{b=0}^{L-1} (b - M)^2 P(b) \quad (2)$$

$$E = \sum_{b=0}^{L-1} P(b)^2$$

where L denotes the number of gray levels of the fundus image under study, which is equal to 256. $P(b)$ refers to the normalized histogram that is determined through dividing the conventional histogram by the number of pixels within the fundus image under study.

A fundus image can thus be converted into a feature vector by combining the three histogram properties of multiple range images calculated at different resolutions as follows:

$$FV = [M_1, V_1, E_1, \cdots, M_i, V_i, E_i, \cdots, M_n, V_n, E_n] \quad (3)$$

where M_i, V_i, and E_i denote the histogram mean, histogram variance, and histogram energy of the range image, respectively, that is calculated by a local neighborhood window of size w_i defined in Eq. 1. Parameter n refers to the maximum size of the local neighborhood windows, which is set at 12 (i.e. the window size is 25×25) in our implemented system. Therefore, the size of the feature vector is $3 \times n$.

B. Fundus Image Classification

The feature vector described in the previous subsection has distinctive characteristics that can be used to classify normal and abnormal fundus images. This can be illustrated in Fig. 4 where the three figures show the histogram mean, histogram variance, and histogram energy that are averaged over the 338 normal fundus images (the graph labeled by squares) and the 306 abnormal fundus images (the graph labeled by circles), respectively, when the size of the neighborhood windows increases from 1 to 12. As Fig. 4 shows, the

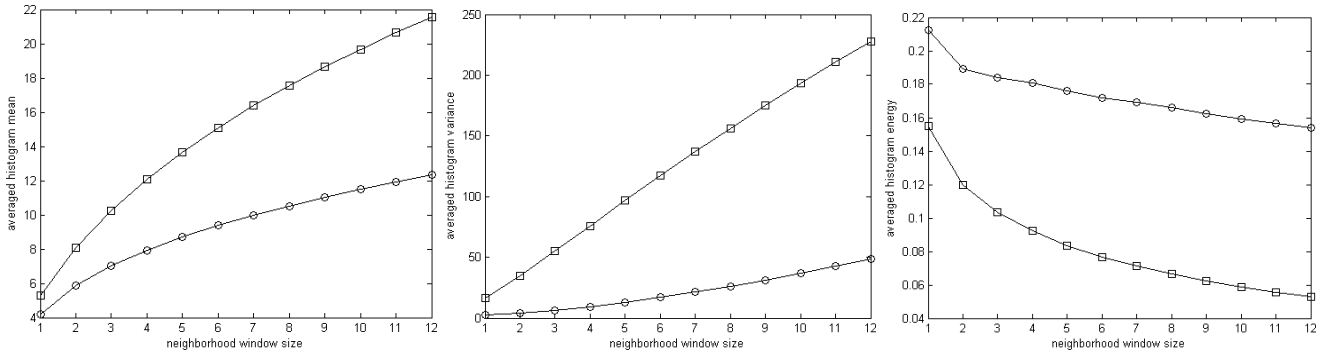

Fig. 4. The average mean, average variance, and average energy that are evaluated over the histograms of the range images of the 338 normal fundus (labeled by circles) images and 306 abnormal fundus images (labeled by squares), respectively, when the size of neighborhood windows (used to calculate the range images) increases from 1 to 12 gradually.

histogram mean, histogram variance, and histogram energy are consistently discriminative between the range image of normal and abnormal fundus images at different resolutions.

We classify fundus images by a classifier that is built based on Fisher's linear discriminant (FLD) [8]. In particular, the FLD attempts to find a linear vector w that projects a high-dimensional feature into a one-dimensional feature so that the ratio between the projected between-class and within-class variance given below can be maximized:

$$\underset{w}{\operatorname{argmax}} J(w) = \frac{w^T S_b w}{w^T S_w w} \tag{4}$$

where S_b and S_w correspond to the between-classes scatter matrix and within-classes scatter matrix as follows [9]:

$$\begin{aligned} S_b &= \sum_{c} N_c(\mu_c - \overline{FV})(\mu_c - \overline{FV})^T \\ S_w &= \sum_{c}\sum_{i \in c}(FV_i - \mu_c)(FV_i - \mu_c)^T, c = 1, 2 \end{aligned} \tag{5}$$

where N_c and μ_c denote the number and the mean of the training feature vectors in class c (c=1, 2 corresponds to normal and abnormal image classes), respectively. $\overline{FV}$ refers to the mean of all created training feature vectors. The linear projection vector w in Eq. 5 can be straightly estimated by $w = \Sigma^{-1}(\mu_1 - \mu_2)$ [9].

C. Design Considerations

It should be noted that the performance of the proposed technique has a close relation with the number of the range images calculated at different resolutions. Generally, the performance of the proposed technique will be improved when a larger number of range images are calculated and the corresponding histogram characteristics are incorporated into the feature vector. But the classification speed will be affected especially when the size of the local neighborhood window increases gradually (to be discussed in the next Section). The speed issue becomes even more crucial for mass screening programs where a vast number of fundus images of different qualities need to be classified as fast as possible before further processing.

III. Experiments

This section presents the experimentation of the proposed technique. In particular, we divide this section into three subsections, which deal with the data collection, the classification evaluation, and discussion, respectively.

A. Data Collection

We evaluate our proposed fundus image classification technique by using 644 fundus images that are used for the Singapore Malay Eye Study by the Singapore Eye Research Institute (SERI) [10]. The 644 fundus images include 338 normal fundus images and 306 abnormal images, which are all color images and taken with a fundus camera with resolution at 2048×3075. In particular, the 306 abnormal fundus images are manually labeled by professional fundus image graders.

B. Fundus Image Classification

We evaluate the proposed technique by using the fundus images described in the previous subsection. For each fundus image, twelve range images are first calculated by using twelve neighborhood windows whose size increases from 1 to 12 (i.e. from 3×3 to 25×25) gradually. A number of feature vectors are then determined based on the three histogram properties of the twelve range images, which will be used to evaluate the speed and accuracy of the proposed technique in relation to the number of range images used.

The proposed technique is evaluated by 10-fold cross-validation. In particular, the two types of feature vectors (i.e. normal and abnormal) are first partitioned into 10 groups each. One group of each type is then randomly selected (from the groups not selected before) as the validation data to test the classification model that is trained by using the remaining 9 groups. Such process is repeated ten times where each group is selected just once as the validation data. The classification results are then averaged to produce a single classification result. Fig. 5 shows the classification accuracy of the feature vectors described above.

In particular, the graph labeled by circles in Fig. 5 shows the accuracy when the feature vector is built based on the histogram of a single range image. As the figure shows, the

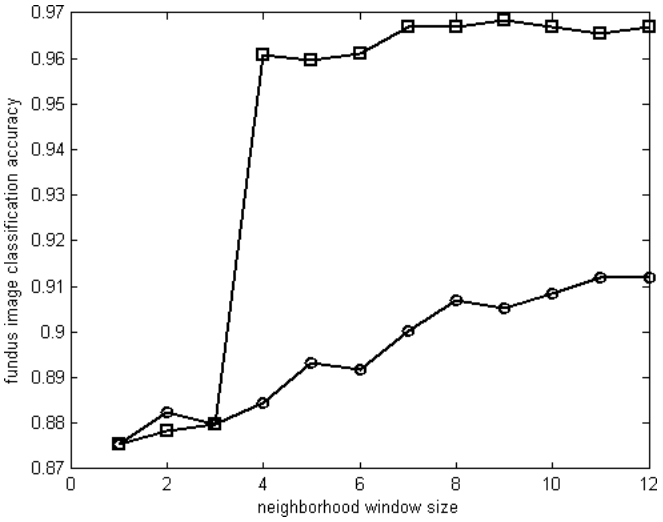

Fig. 5. Fundus image classification accuracy: The two graphs show the accuracy of the feature vectors built based on a single histogram (labeled by circles) and the ones built by combining multiple histograms (labeled by squares), respectively.

classification accuracy increases steadily when the size of the neighborhood window increases gradually. In particular, the accuracy tops at 91.17% when the size of the neighborhood window (used to calculate the range image) is at 12. At the same time, the accuracy increase becomes saturate when the size of the neighborhood window reaches around 8. The graph labeled by squares shows the accuracy when the feature vector is built by accumulating the histogram features of the range images when the size of the neighborhood window increases from the smallest to the largest step by step. As the figure shows, the accuracy of the accumulated feature vector is roughly the same as the feature vectors built based on a single histogram when the size of the neighborhood window is no larger than 3. However, the accuracy of the accumulated feature vector jumps and becomes much higher than that of the feature vector built based on a single histogram when the size of the neighborhood window becomes larger than 3. In addition, the accuracy reaches up to 96.34% when the size of the neighborhood window reaches up to 9.

In addition, experiments show that the execution time of the proposed technique increases rapidly with the increase of the number of the range images calculated. In particular, the execution time increases up to 26 seconds when twelve range images are calculated for each fundus image (the size of the neighborhood window from 1 to 12). Specifically, most of the execution time is spent on the calculation of the range image, which involves the determination of the maximum and minimum pixel value within the local neighborhood window. It should be noted that our system is developed by using Matlab and it can be speeded up greatly through the implementation by C++ and code optimization.

C. Discussion

As discussed in the previous subsection, there is a trade-off between the execution time and the accuracy of the proposed technique. In particular, the high classification accuracy requires more range images calculated with a larger number of neighborhood windows, which prolongs the fundus image classification process. One solution is to perform a feature selection procedure during the training stage to identify the most distinctive histogram features. Then the range images with less distinctive histogram features need not to be calculated. In addition, some other image features may also help to further improve the classification accuracy. We will study these two issues in our future work.

IV. Conclusion

This paper reports an automatic fundus image classification technique that will help the CAD of various types of eye diseases such as glaucoma, macular diseases, and diabetic retinopathy. The proposed technique classifies fundus images by converting a fundus image into a feature vector that is created based on the histogram of multiple range images calculated at different resolutions. Experiments show that the accuracy of the proposed technique reaches up to 96.34%.

References

[1] M. C. Luculescu and S. Lache, "Computer-aided diagnosis system for retinal diseases in medical imaging", WSEAS TRANSACTIONS on SYSTEMS, vol. 7, no. 3, pp. 264–276, 2008.

[2] M. D. Abrmoff, M. Niemeijer, M.S.A. Suttorp-Schulten, M. A. Viergever, S. R. Russell, and B. V. Ginneken, "Evaluation of a System for Automatic Detection of Diabetic Retinopathy From Color Fundus Photographs in a Large Population of Patients With Diabetes" Diabetes Care, vol. 31, no. pp. 193–198, 2008.

[3] G. G. Yen and W. F. Leong, "A Sorting System for Hierarchical Grading of Diabetic Fundus Images: A Preliminary Study", *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 1, pp. 118–130, 2008.

[4] H. Fujitaa, Y. Uchiyama, T. Nakagawa, D. Fukuoka, Y. Hatanaka, T. Hara, G.N. Lee, Y. Hayashi, Y. Ikedo, X. Gao, and X. Zhou, "Computer-aided diagnosis: The emerging of three CAD systems induced by Japanese health care needs", *Computer Methods and Programs in Biomedicine*, no. 92, pp. 238–248, 2008.

[5] K. Doi, "Computer-aided diagnosis in medical imaging: historical review, current status and future potential", *Computerized Medical Imaging and Graphics*, vol. 31, pp. 198-211, 2007.

[6] Y. Hayashi, T. Nakagawa, Y. Hatanaka, A. Aoyama, M. Kakogawa, T. Hara, H. Fujita, T. Yamamoto, "Detection of retinal nerve fiber layer defects in retinal fundus images using gabor filtering", *Proceedings of SPIE Medical Imaging 2007: Computer-Aided Diagnosis*, vol. 6514, pp. 65142Z-165142Z-8, 2007.

[7] Y. Hatanaka, T. Nakagawa, Y. Hayashi, A. Fujita, M. Kakogawa, K. Kawase, T. Hara, H. Fujita, "CAD scheme to detect hemorrhages and exudates in ocular fundus images", *Proceedings of SPIE Medical Imaging 2007: Computer-Aided Diagnosis*, vol. 6514, pp. 65142M-165142M-8, 2007,

[8] R. O. Duda, P. E. Hart and D. G. Stork, Pattern Classification, Second Edition Wiley, 2000.

[9] B. D. Ripley, Pattern Recognition and Neural Networks, Cambridge University Press, 1996.

[10] A. W. P. Foong, S. M. Saw, J. L. Loo, S. Shen, S. C. Loon, M. Rosman, T. Aung, D. T. H. Tan, E. S. Tai, and T. Y. Wong, "Rationale and methodology for a population-based study of eye diseases in Malay people: The Singapore Malay eye study (SiMES)", *Ophthal Epidemiol*, vol. 14, pp. 25–35, 2007.

[11] J. Liu., D.W.K. Wong, J.H. Lim., H. Li, F.Yin, X. Jia, K.L. Chan, N.M. Tan, T.Y. Wong, "ARGALI- an Automatic cup-to-disc Ratio measurement system for Glaucoma detection and AnaLysIs framework", SPIE Medical Imaging, vol. 7260, 2009.

[12] J. Liu, D.W.K. Wong, J.H. Lim, H. Li, N.M. Tan, Z. Zhang, T.Y. Wong, R. Lavanya, "ARGALI : An Automatic Cup-To-Disc Ratio Measurement System For Glaucoma Analysis Using Level-Set Image Processing", ICBME, 2008.