INTERNATIONAL JOURNAL OF ELECTRICAL ENGINEERING, VOL.21, NO.3 PP.75-83 (2014)

A CLASSIFICATION-BASED METHOD FOR RETINAL IMAGE QUALITY ASSESSMENT

Xiaowei Yan, Gongping Yang, Xianjing Meng and Zhe Han

School of Computer and Tec hnology, Shandong University

Yilong Yin

School of Computer and Tec hnology, Shandong University
School of Computer Science and Technology, Shandong University of Finance and Economics

ABSTRACT

Retinal fundus images are of great importance in terms of disease diagnosis and their quality affects the accuracy and reliability of the diagnosis result. Therefore, image quality assessment is essential for retinal image analysis. There are many aspects for quality assessment, such as blur, noise, poor illumination, etc. In this paper, a no-reference quality assessment method for retinal image based on supervised classification is introduced. It mainly focuses on the blur and noise of an image. First, the images are enhanced by curvelet transformation to reduce the influence of noise when choosing the anisotropic patches. Second, anisotropic patches are selected on the enhanced images and then extracted on the original images. Then, nine features are adopted on the original green channel at the extracted anisotropic patches, which are effective with regard to blur and noise. Subsequently, random forest is used to classify the selected patches in an image. Finally, the image quality is obtained by the voting of the classifying result. The method is tested on the public digital retinal images database (HRF) for quality assessment. The area under the receiver operating characteristic (ROC) curve (AUC) is 95.7%, which is superior compared to existing methods. When compared pair-wise, our method obtains 18 out of 18 pairs which agrees to the human observer 100%. The algorithm is demonstrated to be effective to measure quality of retina images.

Key words: no-reference image quality assessment, retinal image, supervised method, anisotropic patches

I. INTRODUCTION

Retinal fundus images are very important basis for diagnosis and patient treatment. They can be used in the diagnosis of diabetes [1], glaucoma [2], macular degeneration [3], vascular and neurological disorders and other diseases. Based on the importance of fundus photos, image quality has a great influence on the accuracy of disease diagnosis [4].

However, the image may suffer from various kinds of degradation during the process of acquisition which may interfere the diagnose procedure, such as patient's head or eye movement, blinking, poorly dilated, small pupils and media opacity. All these factors will lead to the presence of blur, noise and other issues in an image. As the number of retinal image is large, artificial judgment is time-consuming and hard sledding and cannot be easily performed in many scenarios. So an algorithm to automatically estimate the quality of fundus images is meaningful. Above all, it is indispensable to build a method which can measure the retinal image quality.

Some approaches to automatically determine the quality of retinal images have been developed. The existing methods can be mainly divided into two categories: (i) classification based approaches and (ii) quality metrics based methods. Quality metrics based methods can be further classified into a) segmentation based approaches and b) histogram based approaches.

Manuscript received Aug. 2, 2014; and revised Sep. 23, 2014; accepted Oct. 6, 2014.
DOI:10.6329/CIEE.2014.2.04

Classification based approaches generally extract features from retinal images and use a classifier to obtain the quality score. These kinds of methods usually take feature extraction on full images. A uniform quality across the whole image is assumed while the main parts affecting the fundus image quality are anatomical structures. In recent years, Niemeijer et al. [5] used a histogram of the image structure clustering (ISC) clustered pixels as well as the raw R, G and B histograms as features to assess quality of images. Giancardo et al. [6] used the elliptical local vessel density and color histograms as inputs to a linear SVM classifier to estimate quality. Davis et al. [7] calculated 17 simple features in the RGB and CIELab images for each channel to describe a retinal image quality assessment procedure. They made use of those features to evaluate the image along the dimensions of color, luminance and contrast. In 2010, an algorithm model was proposed by Paulus et al. [8] which combine clustering, sharpness metric and Haralick texture features to measure the image quality. Yu et al. [9] presented a supervised method which integrated global histogram features, textural features, and vessel density, as well as a local no-reference perceptual sharpness metric as inputs to a partial least square (PLS) classifier to distinguish low quality images from normal quality images. In 2012, four features (quantifying image color, focus, contrast and illumination) were computed using novel image processing techniques. These quality indicators were also combined and classified to evaluate the image suitability for diagnostic purposes [10]. In 2014, generic features consisting of local sharpness and textural features were extracted in [11].

One quality metrics based methods, segmentation based approaches are in common use. These methods utilize structures in retinal images such as vessel and other major anatomical structures to measure image quality. It is time-consuming and usually the segmentation result is inaccurate. In 2006, Fleming et al. [12] measured vessel definition in a region around the macula to assess quality of diabetic retinal images. In 2011, Hunter et al. [13] presented an approach which calculates the contrast and quantity of visible blood vessels within 1 Optic Disc diameter of the fovea, and measures the contrast between the fovea core region and the background retina. In 2012, the author compared methods of clarity assessment based on the degradation of visible structures and based on the deviation of image properties outside expected norms caused by clarity loss. They also demonstrated a new technique that makes a clarity assessment using only selected portions of the image [14]. In 2013, the authors of [15] presented a no-reference quality metric to quantify image noise and blur and its application to fundus image quality assessment. The proposed metric took the vessel tree visible on the retina as guidance to determine an image quality score.

Another quality metric based methods, histogram based approaches are also in general use. These approaches use information gained by image statistics. A shortcoming of these methods is that they use only a limited type of analysis and rely on one mean histogram for comparison, but it is unable to explain the natural variance in retinal images. In 1999, Lee and Wang [16] used a quality index of a retinal image to classify quality. The quality index was calculated by the convolution of a template intensity histogram obtained from a set of typically good retinal images and the intensity histogram of the retinal image. In 2001, Lalonde et al. [17] proposed a histogram matching function to calculate the proposed features from the histogram of the edge magnitude distribution in the image as well as the local histograms of pixel gray-scale values.

In this article, we focus on the supervised method. As supervised methods are designed based on pre-classified data, their performance is usually better than that of unsupervised ones. In spite of this, feature extraction in supervised methods is generally applied on the whole image, while the main parts affecting the fundus image quality are anatomical structures. So we do our task for image quality assessment to avoid the weakness mentioned above. We mainly focus on the blur and noise of an image.

The following describes the steps of the method. First, the images are enhanced by curvelet transformation to reduce the influence of noise when choosing the anisotropic patches, for image quality is mainly determined by anisotropic patches with dominant gradient direction. Second, anisotropic patches are selected on the enhanced images and then extracted on the original images. Then, nine features are adopted on the original green channel at the extracted anisotropic patches, which are effective with regard to blur and noise. Subsequently, random forest is used to classify the selected patches in an image. Finally, the image quality is obtained by voting the classifying result. The AUC is 95.7% while the value is 88.3% in original article [15]. The metric is also compared pair-wise between a good quality image and the corresponding image of poor quality: the ranking obtained for 18 out of 18 image pairs resulted in an agreement of 100% and it is also better than the result in the original article.

The novelty of this paper is as follows: (a) In this paper, a retinal image quality assessment method based on a supervised approach is proposed. (b) Curvelet transformation is applied on the green channel of images in order to effectively get anisotropic pixel patches. (c) In feature extraction, we combine 9 features which are demonstrated to be effective on the aspect of noise and blur. (d) It is the first time for Random Forest to be applied to retinal image quality assessment and the Random Forest classifier is suitable for the features we extracted. (e) Features are extracted on anisotropic patches instead of the whole image. Random forest is used to classify the selected patches and the score of image quality is based on the classification results.

The article is organized as follows. In Section 2, our method for fundus images is introduced. We analyze the experiments and results in Section 3. Finally, the paper ends with a conclusion in Section 4.

II. PROPOSED METHOD

2.1 Methodology overview

The steps of the proposed method are as follows: curvelet transformation, anisotropic patches selection, feature extraction, classification and classifying result voting. The green channel of the retinal image is extracted from the color image. In order to avoid the noise impact, curvelet transformation is performed at first and the anisotropic locations are masked. Then the original green channel image and anisotropic location mask are combined and a certain number of patches are selected randomly. Features on the aspect of blur and noise are extracted from these anisotropic locations on the green channel. Next, random forest is applied to classify the selected patches. Finally, retinal image quality is obtained by fusing the classification result. Fig.1 shows the flow chart of the proposed method.

2.2 Preprocessing

In this section, the green channel of the retinal image is extracted from the color image, where the contrast between vessel and background is higher than in the blue or red channel [8]. Undoubtedly, the existence of the image noise has a significant impact on the extraction of anisotropic block. So noise reduction is essential in the initial stage. As a preprocessing course, the curvelet transformation is applied to the downsampled version of the green channel.

Curvelet transformation [18] combines the anisotropic characteristics of Ridgelet transformation and multi-scale trait of wavelet transformation. Curvelet coefficients can be modified in order to do noise reduction in an image. Images can be reconstructed by the transformation of curvelet coefficients. The implementation of the second generation curvelet transformation has two ways. In this paper, we use the one which is processed based on the method of Wrapping, and get the discrete curvelet coefficients. Then, adjust the coefficients in descending order. A function must be defined which modifies the values of the curvelet coefficients. We modify curvelet coefficients as follows:

$$y_c(x) = \begin{cases} x & x \geq t \\ 0 & x < t \end{cases} \tag{1}$$

Here, $t = x_{\lceil \eta * n \rceil}$, where x denotes the curvelet coefficients and η is a parameter to adjust the threshold and n is the number of curvelet coefficients in each scale. We selected the $\lceil \eta * n \rceil_{th}$ in descending ordered x as the threshold to modify curvelet coefficients.

2.3 Anisotropic Patches Selection

The approach introduced by Zhu et al. [19] is adopted to calculate the value of the anisotropic block on the curvelet transformed image. First of all, the images are divided into non-overlapping blocks , and the size of each block P *is* $n \times n$. Next, they compute the local gradient matrix G of P.

$$G = \begin{pmatrix} P_x(1,1) & P_y(1,1) \\ \vdots & \vdots \\ P_x(n,n) & P_y(n,n) \end{pmatrix} \tag{2}$$

Where $P_x(x_i, y_i)$ and $P_y(x_i, y_i)$ denotes the derivative of P at pixel (x_i, y_i) in x- and y- direction, respectively. The singular value decomposition (SVD) of G is given by:

$$G = UDV^T = U \begin{pmatrix} s_1 & 0 \\ 0 & s_2 \end{pmatrix} V^T \tag{3}$$

For orthogonal matrices U, V and the singular values s_1, s_2. A local quality metric to quantify image noise and blur in an anisotropic patch P is given by:

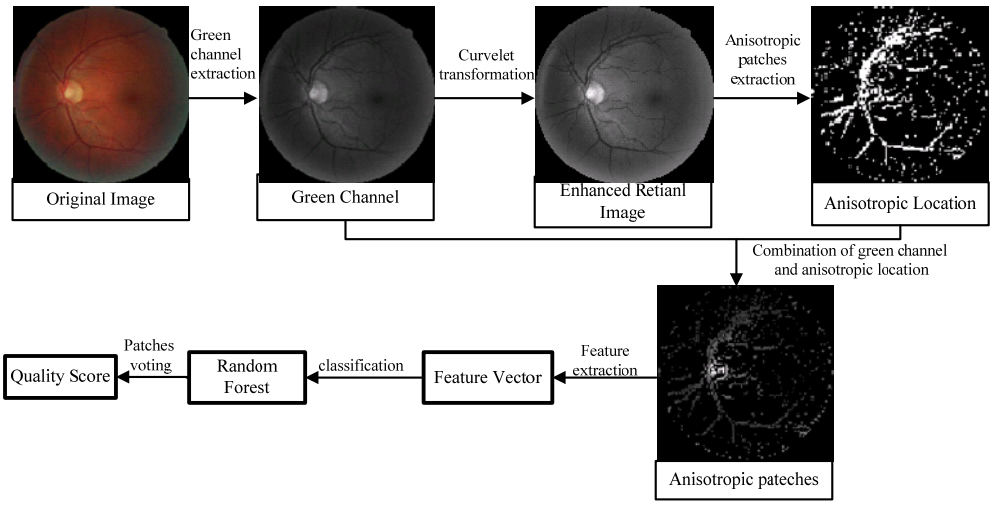


Fig. 1 The flow chart of the proposed method.

$$q(P) = s1 \cdot R \tag{4}$$

Where R denotes the coherence:

$$R = \frac{s1 - s2}{s1 + s2} \tag{5}$$

Patches are assumed to be anisotropic when $R > k * \tau_R$. Where, k is the coefficient for threshold value. Otherwise, it is isotropic, and isotropic is not meaningful. Threshold τ_R is calculated by

$$\tau_R = \sqrt{\frac{1\text{-}\alpha^{\frac{1}{n^2-1}}}{1+\alpha^{\frac{1}{n^2-1}}}} \tag{6}$$

It is important to know that the image quality is mainly determined by anisotropic patches with dominant gradient direction. Based on theory, the locations of anisotropic patches are masked and we obtain the feature vector from these positions.

2.4 Features Extraction

Features are extracted on selected patches from the original green plane of the retinal images. For the consideration of image blur and noise, the feature vector consists of 9 features [20, 21, 22] and is extracted from each anisotropic block. These features are testified to be effective on blurred image quality assessment and they are mainly divided into two parts. The first five features as first component are mainly for blur and noise. Variance (Var) is able to reflect the degree of dispersion between each pixel in a block. Energy of image gradient (EOG) is used to represent the change of gray value in the image. Energy of Laplacian (EOL) is used for analyzing high spatial frequencies associated with image border sharpness. Spatial frequency (SF) is a modified version of the energy of image gradient. The vesselness feature can denote blur and noise information of an image according to the local gradient variation. We utilize GLCM to compute the difference of texture feature with different blurred versions. Contrast, correlation, energy and homogeneity of GLCM are computed in this article. In all, we extract features mainly on the aspect of the fuzziness or noise in an image. They are all proven to be effective and the combination of these features also turn out to be valid in this article.

We describe each component of the feature vector with regard to a patch F with size $M \times N$ as follows.

1) The first component of features

(a) Var: The variance of gray scale patches is defined by:

$$Var = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} (F(x,y) - \mu)^2 \tag{7}$$

Where, μ is the mean value of the patch and is given by:

$$\mu = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} F(x,y) \tag{8}$$

(b) EOG: This measure is computed as follows:

$$EOG = \sum_{x=1}^{M\text{-}1} \sum_{y=1}^{N-1} (F_x{}^2 + F_y{}^2) \tag{9}$$

Where, $F_x = F(x+1, y) - F(x, y)$

and $F_y = F(x, y+1) - F(x, y)$

(c) EOL: The Laplacian operator is used for analyzing high spatial frequencies associated with patch border sharpness and is given by:

$$EOL = \sum_{x=2}^{M\text{-}1} \sum_{y=2}^{N-1} (F_{xx} + F_{yy})^2 \tag{10}$$

In the formula,

$$Fxx + Fyy = -F(x-1, y-1) - 4F(x-1, y) - F(x-1, y+1) - 4F(x, y-1) + 20F(x, y) - 4F(x, y+1) - F(x+1, y-1) - 4F(x+1, y) - F(x+1, y+1)$$

(d) SF: Spatial frequency is defined as:

$$SF = \sqrt{RF^2 + CF^2}$$

Where,

$$RF = \sqrt{\frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} (F(x,y) - F(x,y-1))^2}$$

$$CF = \sqrt{\frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} (F(x,y) - F(x-1,y))^2} \tag{11}$$

(e) Vesselness Feature

We employ the vesselness measure. The local Hessian matrix is computed for each pixel in green channel.

$$H = \begin{pmatrix} \frac{\partial d^2 Ig}{\partial x^2} & \frac{\partial d^2 Ig}{\partial x \partial y} \\ \frac{\partial d^2 Ig}{\partial x \partial y} & \frac{\partial d^2 Ig}{\partial y^2} \end{pmatrix} \tag{12}$$

We employ the vesselness measure:

$$V = \exp(-\frac{\lambda_1^2}{\lambda_2^2})(1 - \exp(-(\lambda_1^2 + \lambda_2^2))) \tag{13}$$

For the eigenvalues λ_1 and λ_2 of H where

$\lambda_2 \geq \lambda_1$.

2) The second component of features

Texture features are obtained by gray level co-occurrence matrix. We get four statics from GLCM including contrast, correlation, energy, and homogeneity. Let $p(i, j, d, \theta)$ stand for the GLCM. Then $p(i, j)$ denotes the value in GLCM at coordinate (i, j), d is the distance, and θ is the direction.

In our work, we set $d = 1$ and $\theta = 0^\circ$, 45°, 90° and 135°, respectively. The ultimate GLCM is obtained by computing the average of GLCM with four different directions. The formulas of calculating four statics are as follows:

$$Contrast(Gcon) = \sum_{n=1}^{L} n^2 \sum_{i=1}^{L} \sum_{j=1}^{L} P(i, j) \tag{14}$$

$$Correlation(Gcor) = \sum_{i=1}^{L} \sum_{j=1}^{L} \frac{ijP(i, j) - \mu_1 \mu_2}{\sigma_1^2 \sigma_2^2} \tag{15}$$

$$Energy(Gasm) = \sum_{i=1}^{L} \sum_{j=1}^{L} P(i, j)^2 \tag{16}$$

$$Homogeneity(Ghom) = \sum_{i=1}^{L} \sum_{j=1}^{L} | i - j | P(i, j) \tag{17}$$

Where, μ is the mean value and σ is the variance of each patch. Contrast reflects the clarity of the textures. Correlation represents the consistency of image texture. Energy measures the uniformity of distribution of gray level in the image. Homogeneity can be utilized for detecting similarities in the image.

2.5 Classification Using Random Forest

Random Forest [23] is a general term for ensemble methods which was first introduced by Leo Breiman in 2001. It uses a set of tree-type classifiers with $\{h(x, \beta_k), k = 1, ...\}$ for classification and regression. Where the $\{\beta_k\}$ is independent identically distributed random vectors and decides the process of single tree growth and x is an input pattern. $h(x, \beta_k)$ is the base classifier established by Classification And Regression Tree (CART) algorithm and each tree is grown to the largest extent possible and there is no pruning. In the process of the formation of every tree, bagging method is used to choose the training sets. Then it makes the choice of splitting attributes by random way.

Suppose there are X attributes. Define the number of attributes $H \leq X$. On each inside node, H properties were randomly extracted as splitting attribute sets. They get the best division method from H property and do split to the test node. Splitting attribute is selected based on *gini* index. Suppose that set T contains N categories of record, and then *gini* index is calculated by the equation as follows:

$$gi\ ni(T) = 1 - \sum_{j=1}^{n} p_j^2 \tag{18}$$

Where p_j represents the frequencies of j. If set T is divided into m parts, N_1, N_2, …, N_n, then the equation of *gini* index is:

$$gi\ ni_{split}(T) = \frac{N_1}{N} gi\ ni(T_1) + \cdots \frac{N_m}{N} gi\ ni(T_m) \tag{19}$$

Finally, we choose the minimum *gini* index as standard splitting. Do the classification based on the votes of the decision tree. We set the quality metric as follows:

$$Q_R = \frac{number(label = 1)}{number(label = 1) + number(label = 0)} \tag{20}$$

In our work, we select a RF classifier to do the classification. A RF classifier has many virtues. Firstly, it performs well when dealing with a lot of data sets, both discrete data and continuous data, and it has advantages over other classifiers. Secondly, RF is able to handle with high-dimensional features and is not necessary to do feature selection. In the process of training, it is capable of detecting the influence between features. After the training, it will show the importance of each feature. Thirdly, the RF classifier can balance the classification error when the data sets for training are unbalanced. The introduction of two random variables had a small possibility for random forest to generate over fitting. Fourthly, the training speed of the RF classifier is fast.

III. EXPERIMENTS AND RESULTS

3.1 Materials

We use all the retina images in High-resolution fundus (HRF) image database [15] to evaluate the proposed method and compared its performances with other methods. HRF database contains 18 image pairs of the same eye from 18 human subjects using a Canon CR-1 fundus camera with a field of view of 45° and different acquisition settings and a resolution of 5418 by 3456 pixels. For each pair, the first image has poor quality and thus the examination had to be repeated. Both images share approximately the same field of view, whereas small shifts were caused by eye movements between the acquisitions. Images of poor quality suffer from decreased sharpness locally or globally. Fig.2 shows the sample images in HRF database.

3.2 Experiments and Results

We use two measures to evaluate the performance

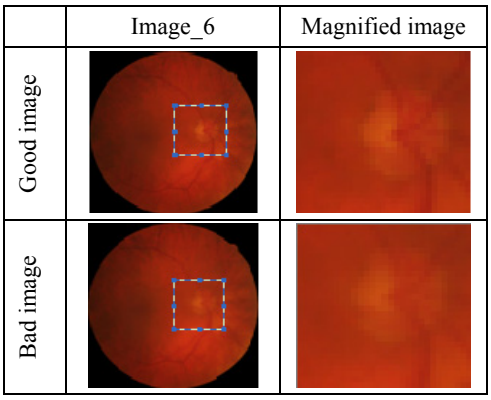


Fig.2 The instance of image in database.

of the method. A common measure to quantify the performance of method is to compute the area under the ROC curve. The area under the curve is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one. It is possible to calculate the AUC by using an average of a number of trapezoidal approximations. An area close to one means a good performance. Another measure is pair-wise comparison between a good acquisition and the corresponding image of poor quality.

Parameter Setting.

As the radius of blood vessel structure in retina image is usually to be 8 pixels. We set $n = 8$ for the patch size so as to find the edge of the vessel easily. The significance level $\alpha = 0.001$ in [19] and the parameter $\eta = 0.34$ by experiment. We draw the conclusion that if they are set to be much higher, important information of picture would be lost. On the contrary, if they are set to be lower, much noise information would remain on the image. The adjustment parameters are experimentally tuned based on intrinsic characteristics of the input image and the goal of work.

Performance Analysis.

The proposed method is used to do retina image quality assessment. We perform leave-one-out tests using all the 18 pairs of retinal images in the HRF dataset. Every image is measured in turn after training Random Forest with the other 17 pairs of retina images. As there are many patches in one photo, we randomly selected part of the blocks for training. In this section, 800 anisotropic patches are randomly selected in each training image and all anisotropic patches are chosen in the testing image. The proposed metric Q_R was compared to the Q_v[15] metric as well as to the Q[24] measurement and CPBD metric[25] and the anisotropy measure[26]. Fig.3 shows the ROC curve for quality classification based on different metrics and Fig.4 shows the quality score of every image in HRF database.

We got the area under ROC curve 95.7% compared to the above mentioned metrics: Qv, CPBD, Anisotropy and Q and their result is 88.9%, 50.9%, 75.3% and

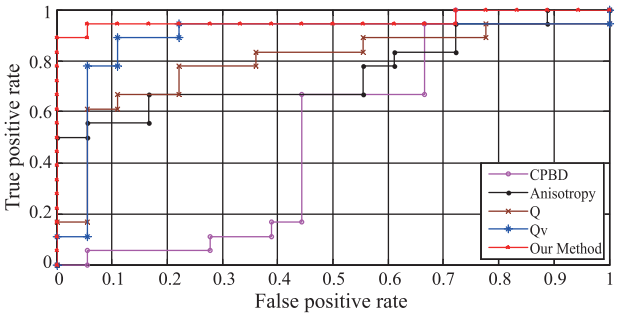


Fig.3 ROC curve for quality classification based on different metrics.

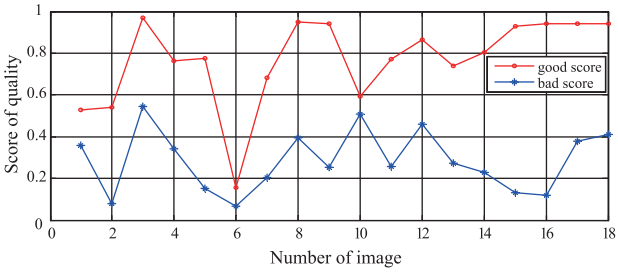


Fig.4 The quality score of each image.

79.6%, respectively. The result of pair-wise comparison is that 18 out of 18 pairs were meeting the human observer resulting in an agreement of 100%, while the result of the compared methods are 88.9%, 55.6%, 94.4% and 83.3%, respectively.

According to Fig.4, we draw the conclusion that the score of good image in HRF database is higher than the corresponding bad one. As each image contains different information, the behavior is not stable. As Fig.3 shows, some of good images have less information of blood vessels and other structures even though they have little blur and noise, they still get low scores.

Determination of Blocks.

In this experiment, the number of patches is selected and result of different classifiers are shown. Random forest classifier is used and compared with Support Vector Machine (SVM) and AdaBoost classifiers at the same time. The number of patches is decided from two aspects: area under the ROC curve and the number of errors when compared pair-wise. The trend of quality assessment results with different number of patches is shown in Fig.5 and Fig.6.

According to Fig.5 and Fig.6, the tendency is instable when the number of patches is less than 500. These two figures also demonstrate that Random Forest classifier is better than any other classifiers. As is shown in Fig.5, when the number of patches larger than 500, they all tend to be stable. In Fig.6, it is obvious that RF method is better than SVM and AdaBoost methods. Especially from 500 to 800, the error of RF classifier is 0 when compared pair-wise. To sum up, 800 anisotropic patches are chosen in the experiment. What's more, the

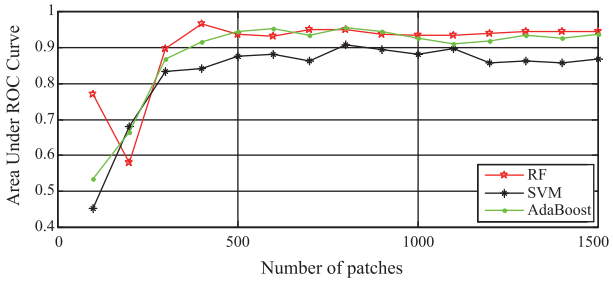


Fig.5 The AUC with different number of anisotropic patches.

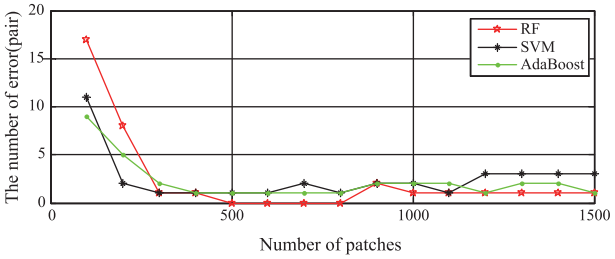


Fig.6 Number of error with different number of anisotropic patches.

running time of RF method is 0.25s for testing. It is less than the SVM and AdaBoost classifier with the time 0.27s and 3.17s, respectively.

Effects of Preprocessing

In this part, a number of anisotropic blocks in each picture are selected to be a training set and the left anisotropic blocks are to be the test set. As is mentioned above, R is the anisotropic value of patches in retina image. We set $k = 1.1$ and the number of the patches of training set is based on the experiment. We make an attempt to demonstrate the curvelet transformation before anisotropic patches extraction for retina image is effective.

As is described above, it is effective when the number of patches is no less than 500. We can see from Fig.7, curvelet transformation before anisotropic blocks extraction really works. There has been a significant increase in performance in some situations.

Effectivity of Feature Combination

In this section, parts of the nine features are combined to compare with the final combination of all features group. The result is shown in the following Fig.8 and it demonstrates that the features combination is effective.

IV. CONCLUSION

In this paper, we propose a method for no-reference image quality metric to measure the retinal fundus im

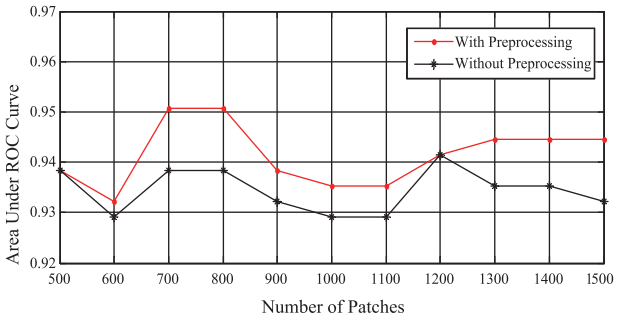


Fig.7 The comparison between preprocessing and without preprocessing.

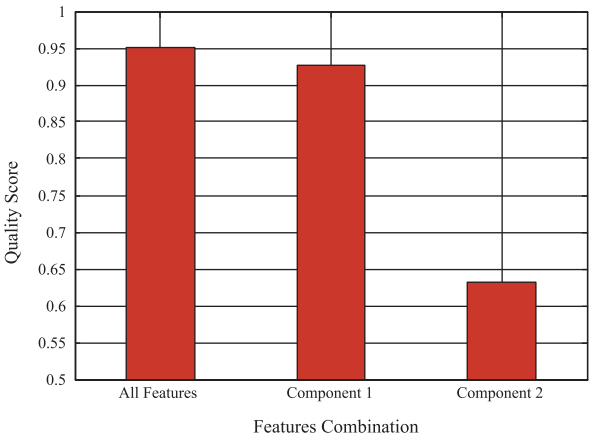


Fig.8 The result of different features combination.

ages. Our experiment adds curvelet transformation so as to obtain the anisotropic patches accurately and it proves to be effective by the result of quality measure. Anisotropic patches are chosen and features are extracted on the green channel at the position of selected blocks on the aspect of blur and noise. It demonstrates that 800 selected patches can help the performance, as is shown in Fig.5 and Fig.6. Finally, random forest is applied to this article, meanwhile, comparing with SVM and Adaboost classifiers, to measure retinal images quality.

The advantages of the proposed method can be seen in several aspects. Firstly, taking curvelet transformation before anisotropic patches extraction is in favor of noise suppression. Secondly, the 9 combined features demonstrate great discriminate ability for image quality assessment. Thirdly, random forest is used for quality classification for the first time and it has much superiority on solving this problem. Most important of all, a retinal image quality assessment method based on supervised approach is proposed, and the assessment result is much better than before from both aspects of evaluation.

The proposed method can be generalized to other kinds of medical image types. Though performance of the proposed method has demonstrated superiority over existing methods, there is still room for further improvement.

ACKNOWLEDGEMENTS

The work is supported by NSFC Joint Fund with Guangdong under Key Project U1201258, Shandong Natural Science Funds for Distinguished Young Scholar under Grant No. JQ201316 and the Research Found for the Doctoral Program of Higher Education under Grant No. 20110131130004.

REFERENCES

[1] A. Hutchinson, A. McIntosh, and J. O. Peters, "Effectiveness of screening and monitoring tests for diabetic retinopathy—a systematic review," *Diabetic Medicine*, vol. 17, no. 7, pp. 495-506, July. 2000.

[2] L. G. Nyúl, "Retinal image analysis for automated glaucoma risk evaluation," *Sixth International Symposium on Multispectral Image Processing and Pattern Recognition*, vol. 74971C-1, Oct. 2009.

[3] H. Bartlett and F. Eperjesi, "Use of fundus imaging in quantification of age-related macular change," *Surv Ophthalmol*, vol. 52, no. 6, pp. 655-671. Nov-Dec. 2007.

[4] N. Patton, T. M. Aslam, T. MacGillivray, I. J. Deary, B. Dhillon, R. H. Eikelboom, K. Yogesan, and I. J. Constable, "Retinal image analysis: concepts, applications and potential," *Prog Retin Eye Res*, vol. 25, no. 1, pp. 99-127, Jan. 2006.

[5] M. Niemeijer, M. D. Abramoff and van B. Ginneken, "Image structure clustering for image quality verification of color retina images in diabetic retinopathy screening," *Medical image analysis*, vol. 10, no. 6, pp. 888-898, Dec. 2006.

[6] L. Giancardo, M. D. Abramoff, E. Chaum, T. P. Karnowski, F. Meriaudeau and K. W. Tobin, "Elliptical local vessel density: a fast and robust quality metric for retinal images," *30th Annual International Conference on Engineering in Medicine and Biology Society*, Aug. 2008, pp. 3534-3537.

[7] H. Davis, S. Russell, E. Barriga, M. Abramoff and P. Soliz, "Vision-based, real-time retinal image quality assessment," *22nd IEEE International Symposium on Computer-Based Medical Systems*, Aug. 2009, pp. 1-6.

[8] M. D. Abramoff and M. S. Suttorp-Schulten, "Automated quality assessment of retinal fundus photos," *International journal of computer assisted radiology and surgery*, vol. 5, no. 6, pp. 557-564, May. 2010.

[9] H. Yu, C. Agurto, S. Barriga, S. C. Nemeth, P. Soliz and G. Zamora, "Automated image quality evaluation of retinal fundus photographs in diabetic retinopathy screening," *Southwest Symposium on Image Analysis and Interpretation*, Apr. 2012, pp.125-128.

[10] J. M. Pires Dias, C. M. Oliveira and L. A. da Silva Cruz, "Retinal image quality assessment using generic image quality indicators," *Information Fusion*, vol. 19, pp. 73-90, Sep. 2012.

[11] M. Fasih, J. M. P. Langlois, H. B. Tahar, et al., "Retinal image quality assessment using generic features." *SPIE Medical Imaging. International Society for Optics and Photonics*, vol. 9035, May. 2014.

[12] A. D. Fleming, S. Philip, K. A. Goatman, J. A. Olson and P. F. Sharp, "Automated assessment of diabetic retinal image quality based on clarity and field definition," *Investigative ophthalmology & visual science*, vol. 47, no. 3, pp. 1120-1125, Mar. 2006.

[13] A. Hunter, J. A. Lowell, M. Habib, B. Ryder, A. Basu and D. Steel, "An automated retinal image quality grading algorithm," *2011 Annual International Conference on Engineering in Medicine and Biology Society*, Sep. 2011, pp. 5955-5958.

[14] A. D. Fleming, S. Philip, K. A. Goatman, P. F. Sharp and J. A. Olson, "Automated clarity assessment of retinal images using regionally based structural and statistical measures," *Medical engineering & physics*, vol. 34, no. 7, pp. 849-859, Sep. 2012.

[15] T. Köhler, A. Budai, M. Kraus, J. Odstrcilik, G. Michelson and J. Hornegger, "Automatic No-Reference Quality Assessment for Retinal Fundus Images Using Vessel Segmentation," *26th IEEE Internatioal Symposium on Computer-Based Medical Systems*, Jun. 2013, pp. 95-100.

[16] S. C. Lee and Y. Wang, "Automatic retinal image quality assessment and enhancement," *International Society for Optics and Photonics*, pp.1581-1590, May. 1999.

[17] Ý. M. Lalonde, Ý. L. Gagnon and M. C. BoucherÞ, "Automatic visual quality assessment in optical fundus images," 2001.

[18] J. L. Starck, E. J. Candès and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Transactions on Image Processing*, vol. 11, no. 6, pp. 670-684, Jun. 2002.

[19] X. Zhu and P. Milanfar, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content." *IEEE Transactions on Image Processing*, vol. 19, no. 12, pp. 3116-3132, Dec. 2010.

[20] R. M. Bahy, G. I. Salama, and T. A. Mahmoud, "A no-reference blur metric guided fusion technique for multi-focus images," *Radio Science Conference (NRSC) 28th National*, Apr. 2011, pp. 1-9.

[21] A. F. Frangi, W. J. Niessen, K. L. Vincken and M. A. Viergever, "Multiscale vessel enhancement filtering," *Medical Image Computing and Comput-*

er-Assisted Intervention (MICCAI), vol. 1496, pp. 130–137, Oct. 1998.

[22] T. Zhang, D. Liang, X. Wang and X. Zhang, "Novel texture features based no-reference image blur metric," *Computer Engineering and Applications*, vol. 48, no. 26, pp. 185-191, 2012.

[23] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5-32, Oct. 2001.

[24] X. Zhu and P. Milanfar, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Transactions on Image Processing*, vol. 19, no. 12, pp. 3116-3132, Dec. 2010.

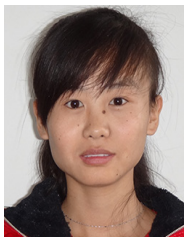
[25] N. D. Narvekar and L. J. Karam, "A no-reference image blur metric based on the cumulative probability of blur detection (CPBD)." *IEEE Transactions on Image Processing*, vol. 20, no. 9, pp. 2678-2683, Sep. 2011.

[26] S. Gabarda and G. Cristóbal, "Blind image quality assessment through anisotropy," *Journal of the Optical Society of America A*, vol. 24, no. 12, pp. 42-51, Dec. 2007.

Xiaowei Yan received the BSEE in 2012 from Qingdao Technological University. Now she is studying at Shandong University for a Masters degree in Computer Application Technology. Her main research interests are medical image processing and machine learning.

Gongping Yang received his Ph.D. degree in Computer Software and Theory from Shandong University, China in 2007. Now he is a professor in the Schoolof Computer Science and Technology, Shandong University. His research interests are biometrics, medical image processing, and so forth.

Yilong Yin is the director of MLA Lab and a professor of Shandong University. He received his PhD in 2000 from Jilin University. From 2000 to 2002, he worked as a postdoctoral fellow in the Department of Electronic Science and Engineering, Nanjing University. His research interests include machine learning, data mining, computational medicine, and biometrics.



Xianjing Meng received her BSEE in 2006 from Shandong University. Now she is studying at Shandong University for a Ph.D degree in Computer Science and Technology. Her main research interests are biometrics, medical image analysis and machine learning.

Zhe Han received her BS in 2012 from Shandong University of Finance and Ecnomics. Now she is studying at Shandong University for a Master degree in Software Engineering. Her main research interests are biomedical image processing and machine learning.