

Optimal algorithm for automatic detection of microaneurysms based on receiver operating characteristic curve

Lili Xu

Shuqian Luo

Capital Medical University
College of Biomedical Engineering
YouAnMneWai XiTouTiao 10
Beijing, China 100069

Abstract. Microaneurysms (MAs) are the first manifestations of the diabetic retinopathy (DR) as well as an indicator for its progression. Their automatic detection plays a key role for both mass screening and monitoring and is therefore in the core of any system for computer-assisted diagnosis of DR. The algorithm basically comprises the following stages: candidate detection aiming at extracting the patterns possibly corresponding to MAs based on mathematical morphological black top hat, feature extraction to characterize these candidates, and classification based on support vector machine (SVM), to validate MAs. Feature vector and kernel function of SVM selection is very important to the algorithm. We use the receiver operating characteristic (ROC) curve to evaluate the distinguishing performance of different feature vectors and different kernel functions of SVM. The ROC analysis indicates the quadratic polynomial SVM with a combination of features as the input shows the best discriminating performance. © 2010 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.3523367]

Keywords: diabetic retinopathy (DR); microaneurysms (MAs); receiver operation characteristic (ROC) curve; support vector machine (SVM).

Paper 10351R received Jun. 23, 2010; revised manuscript received Sep. 19, 2010; accepted for publication Oct. 12, 2010; published online Dec. 17, 2010.

1 Introduction

Diabetic retinopathy (DR) is one of the most serious and most frequent eye diseases in the world. It is a complication of diabetes and the most common cause of blindness in adults. In order to prevent the damage of DR, it is very important to diagnose DR early. The analysis of digital retinal images (Fig. 1), obtained by the fundus camera, is viewed as a feasible approach because the acquisition of the retinal image is noninvasive and low cost.

There are two categories of analysis based on retinal images. One is morphological analysis on arteries, veins, and optical cup.¹ The other is based on the detection of the pathology lesion, such as hemorrhages, microaneurysms (MAs), hard exudates, and cotton wool spots.^{2,3} MAs are the first unequivocal signs of DR, which appear as small reddish isolated patterns of circular shape in color fundus images. They are characterized by their diameters, which are always $< 125 \mu\text{m}$.⁴ Because they are situated on capillaries and capillaries are not visible in color retinal images, they appear as isolated patterns (i.e., disconnected from the vascular tree).

In this paper, we propose a new method to optimize the algorithm to identify MAs. The algorithm for automatic detection of MAs generally comprises three steps. The first step aims at detecting candidates i.e., all patterns possibly corresponding to MAs based on a mathematical morphological black top hat. Then, features are extracted to characterize these candidates. And finally, a support vector machine (SVM) is used to

distinguish MAs from all candidates. We use the receiver operating characteristic (ROC) curve to evaluate the distinguishing ability of different feature vectors and different kernel functions of SVM in this paper. Thus, the optimal feature vector and classifier can be determined by ROC analysis.

2 Automatic Detection of MA

2.1 Candidates Detection-Based Morphological Operation

Mathematical morphology operation is a way of nonlinear filtering used for image processing.⁵ The primary operations of mathematical morphology are dilation and erosion, denoted by \oplus and \ominus , on which other operations are based, such as opening and closing operations. In this paper, we use a black-top-hat transform. It can extract dark objects and structures in gray images. In the RGB images, the green channel exhibits the best contrast. We work on the gray image from the green channel, denoted by f . In the gray image f , MAs appear as dark patterns, small, isolated, and of circular shape.

Black-top-hat transform is the first closing operation on the image f and then subtracting the original image, defined as

$$f_b = (f \bullet e) - f, \quad (1)$$

where e is a structuring element and $f \bullet e$ means closing operation to f with e . The morphological closing operation dilates an image f and then erodes the dilated image using the same structuring element, which is defined as follows:

$$f \bullet e = (f \oplus e) \ominus e. \quad (2)$$

Address all correspondence to: Shuqian Luo, Capital Medical University, College of Biomedical Engineering, Beijing, China 100069; Tel: 86-8391-1000; E-mail: shuqian_luo@yahoo.com.cn

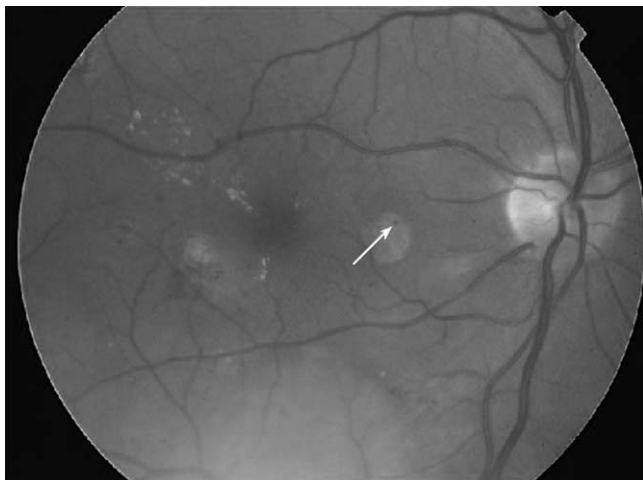


Fig. 1 Fundus retinal image (a MA marked with an arrow).

The blood vessels in the retinal images are usually considered as piecewise linear structures at different orientations. A total of 12 rotated linear structures are used with a radial resolution of 15 deg. The length of a linear structuring element should be such that it is larger than MAs in retinal images. MAs appear as dark patterns within s pixels, which is $<125 \mu\text{m}$. Thus, the length of the linear-structuring element must be longer than s pixels and s is variable for different sizes of original images. For each pixel, record the minimum response as the closed result applied with those 12 rotated linear-structuring elements. Then, we obtain the image f_b by subtracting the original gray image from the closed image. The image f_b contains MAs, which appear as the local bright regions in Fig. 2(a).

Only a certain number of candidates are reasonable (for example, several dozens for each image). A matched filter is used to extract regions of interest (ROI) from f_b . The matched filter is a 2-D Gaussian function with $\sigma = 1$ and has a size of $s \times s$ pixels. Gaussian difference d_G is an index image to evalu-

ate the difference between the local region in f_b and Gaussian function, calculated as

$$d_G(i, j) = \frac{\sum_{i', j' \in S} [f_b(i', j') - g(i', j')]^2}{s^2}, \quad (3)$$

where g is the normalized 2-D Gaussian function centered at (i, j) and (i', j') is the pixel position within the local region of f_b centered at (i, j) , with size of $s \times s$ pixels. Thus, $d_G(i, j)$ is summed over all the pixels (i', j') in the $s \times s$ region, and this local region is denoted by S in Eq. (3).

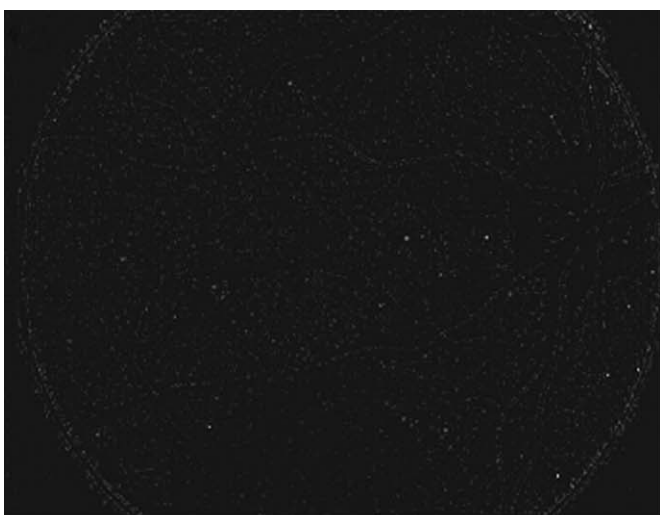
The ROI are the local bright regions in f_b with lower Gaussian difference value $[d_G(i, j)]$. They are extracted by a global threshold of Gaussian difference. Each ROI has a size of $s \times s$ pixels and is recorded by its center coordinates for facilitating to extract corresponding ROI from different source. The binary candidate region of each ROI is determined by thresholding, which minimizes the intraclass variance of candidate and surrounding [shown in Fig. 2(b)].

2.2 Feature Extraction for Candidates

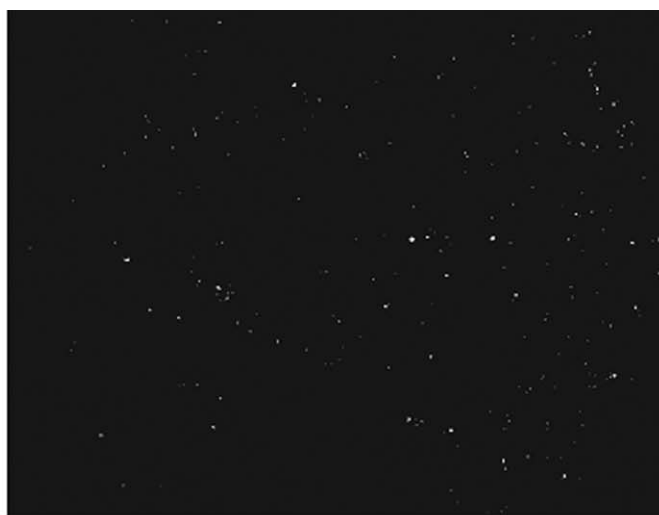
Because MAs are mainly characterized by their shape, size, and color, we use three types of features as follows:

1. shape features, such as relative size and compactness of candidate
2. texture features based on the gray-level co-occurrence matrix (GLCM) of ROI from the green channel
3. color features within the ROI from different color space

Shape is one of the essential characteristics of the object, which provides meaningful information. The Shape feature can be divided into two categories: one is based on the boundary of the candidate and the other is based on the region. We use the relative size and compactness to describe the shape of candidates. Each ROI is a region of $s \times s$ pixels, where s depends on the size of the original image to ensure the length of s pixels is not $<125 \mu\text{m}$. The binary candidate of each ROI is obtained by thresholding the ROI, making the surrounding black and the



(a)



(b)

Fig. 2 (a) Image after black-top-hat transform and (b) binary image after matched filter.

candidate white. In the binary ROI, area A and perimeter P of the candidate is calculated. Relative size, s_1 , is defined as the ratio of area of candidates and ROI. Compactness, s_2 , is used to describe circularity of the shape of candidate, which is defined as

$$\begin{aligned} s_1 &= A/s^2, \\ s_2 &= 1 - 4\pi A/P^2, \end{aligned} \quad (4)$$

where A is the number of pixels in a candidate and P is the total number of pixels around the boundary of the candidate.

Texture is a significant property in digital images. It has an important role in human visual perception and offers information for recognition and interpretation. The GLCM is a powerful statistical tool that has proved its usefulness in a variety of image-analysis applications. It captures second-order gray-level information, which is mostly related to human perception and the discrimination of textures. It is common practice to utilize the 14 well-known Haralick's coefficients as the GLCM-based features.⁶ The coefficients are usually calculated from the average GLCM obtained by averaging the matrices calculated for 0-, 45-, 90-, and 135-deg directions. In this work, we first convert the gray scale of each ROI to 11, then calculate the average GLCM, \bar{C} . Last, we select four coefficients based on GLCM. They are energy t_1 , contrast t_2 , local homogeneity t_3 , and entropy t_4 .

$$\begin{aligned} t_1 &= \sum_{i,j=1}^N \bar{C}^2(i, j), \\ t_2 &= \sum_{i,j=1}^N (i - j)^2 \bar{C}(i, j), \\ t_3 &= \sum_{i,j=1}^N \frac{1}{1 + (i - j)^2} \bar{C}(i, j), \\ t_4 &= \sum_{i,j=1}^N \bar{C}(i, j) \log \bar{C}(i, j). \end{aligned} \quad (5)$$

Intensity is the only available information in the gray image, but color images provide more abundant color information, except intensity. The color contrast is a useful feature. It is defined as the sum of squared differences between the candidate region and its surroundings.⁴ We extract color features in RGB color space first, defined as

$$\begin{aligned} c_1 &= [\mu_{\text{ext}}(R) - \mu_{\text{int}}(R)]^2, \\ c_2 &= [\mu_{\text{ext}}(G) - \mu_{\text{int}}(G)]^2, \\ c_3 &= [\mu_{\text{ext}}(B) - \mu_{\text{int}}(B)]^2, \end{aligned} \quad (6)$$

where μ_{int} is the mean value on the candidate region and μ_{ext} is the mean value on the surroundings in each ROI, and R, G, B in parentheses represent that the calculation is executed in R, G, and B channels, respectively.

The selection of these color features is a complicated task due to the variety of color models. The RGB space of the original image is transformed to hue, saturation, and value space (HSV) because HSV color space is more appropriate since it allows the value component to be separated from the other two color components. Then we extract two kinds of color feature from the hue and saturation components as

$$\begin{aligned} c_4 &= [\mu_{\text{ext}}(H) - \mu_{\text{int}}(H)]^2, \\ c_5 &= [\mu_{\text{ext}}(S) - \mu_{\text{int}}(S)]^2. \end{aligned} \quad (7)$$

In order to select the optimal feature vector, we construct several feature vectors, denoted by S , T , $C1$, $C2$, $A1$, and $A2$. Here, $S = [s_1, s_2]$, $T = [t_1, t_2, t_3, t_4]$, $C1 = [c_1, c_2, c_3]$, $C2 = [c_4, c_5]$, $A1 = \{S, T, C1\}$, and $A2 = \{S, T, C2\}$. We test these feature vectors by SVM in Section 2.3.

2.3 Validation of MA Based on SVM

The SVM, first introduced by Vapnik, is a learning algorithm for two-class classification. It is widely used in pattern recognition applications. It is based on strong foundations from the broad area of statistical learning theory according to structural risk minimization. A SVM is known for its good performance. The basic principle behind a binary SVM is to find the hyperplane that best separates vectors from both classes in feature space while maximizing the distance from each class to the hyperplane.⁷ There are both linear and nonlinear approaches to a SVM. If the two classes are linearly separable, then the SVM attempts to find the optimal separating hyperplane by maximizing the margin between both classes. The margin is $2/\|w\|^2$. For a linear SVM, to find the optimal hyperplane is equal to $\min \frac{1}{2} \|w\|^2$. When the two classes are nonlinearly separable, the SVM computes the optimal separating hyperplane by minimizing the following equation:

$$J(w) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i, \quad (8)$$

where the constant $C > 0$ is user defined and determines the trade-off between the maximization of the margin and minimization of the classification error and ξ_i are slack variables introduced for nonlinearly separable classes.

Different kernel functions determine the classification performance of the SVM. Thus, the selection of the kernel function is important to the SVM. We compare the performances of commonly used kernel functions based on the receiver operation characteristic (ROC) curve. The commonly used kernels are defined as

$$\begin{aligned} \text{Linear function: } K(x_1, x_2) &= \langle x_1, x_2 \rangle, \\ \text{Polynomial (poly): } K(x_1, x_2) &= (\langle x_1, x_2 \rangle + 1)^p, \\ \text{Gaussian radial basis function (rbf): } K(x_1, x_2) &= e^{-|x_1 - x_2|^2 / 2\sigma^2}, \end{aligned} \quad (9)$$

where p is order of polynomial and usually takes 2 or 3 and σ is width of the rbf and controls its functionary range.

3 Optimizing Features Vector and Classifier Based on ROC Curve

The ROC is a performance measure commonly used to compare different classifiers. The ROC curve can be drawn using sensitivity as the x coordinate and 1-specificity as the y coordinate.⁸ Sensitivity describes the probability of a positive test among all positive samples and indicates the probability of a negative test among all negative samples. Therefore, the diagonal in an ROC plot is the performance of random guessing. The ROC curves move toward the upper left corner, indicating rising accuracy of performance. The area under the ROC curve (AUC^{ROC}) is an appropriate performance measure. The AUC^{ROC} value ranges from

Table 1 Sensitivity and specificity of different feature vectors for SVM.

Feature	T		A1		A2	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
	1	0	1	0	1	0
	0.6121	0.7494	0.9720	0.0735	0.6916	0.7606
	0.5935	0.7806	0.8879	0.3062	0.6541	0.7873
	0.4860	0.8608	0.7991	0.5412	0.5981	0.8396
	0.4533	0.8719	0.6963	0.6158	0.4907	0.8998
Data	0.3972	0.9020	0.5981	0.7595	0.3972	0.9388
	0.3551	0.9287	0.4860	0.8686	0.3037	0.9610
	0.2991	0.9465	0.3925	0.9020	0.2897	0.9621
	0.1963	0.9666	0.2944	0.9399	0.1822	0.9788
	0.0935	0.9889	0.1963	0.9699	0.1449	0.9866
	0	1	0	1	0	1
AUC	0.7093		0.7359		0.7574	

0.5 to 1. The larger the AUC^{ROC} is and the closer it is to 1.0, the higher the validity of the classifiers is. Conversely, the nearer the AUC^{ROC} is to 0.5, the lower the validity of the classifiers is. If AUC^{ROC} is 1, then this means the classifier is perfect.

4 Results and Discussion

4.1 Sample Sets

MAs can easily be confounded with other dark patterns. One of the major problems in detection of a MA is to establish

Table 2 Sensitivity and specificity of different kernels of SVM.

Kernel	Linear		rbf		Poly 2		Poly 3	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
	1	0	1	0	1	0	1	0
	0.6963	0.6915	0.7103	0.4788	0.6916	0.7606	0.6355	0.7650
	0.5935	0.8107	0.6963	0.5078	0.6541	0.7873	0.5888	0.8274
	0.4860	0.8864	0.5981	0.6537	0.5981	0.8396	0.4953	0.8391
	0.4439	0.9065	0.4953	0.7706	0.4907	0.8998	0.4597	0.9198
Data	0.3925	0.9354	0.4533	0.8241	0.3972	0.9388	0.3972	0.9410
	0.3598	0.9399	0.3972	0.8664	0.3037	0.9610	0.3598	0.9488
	0.2897	0.9577	0.2991	0.9410	0.2897	0.9621	0.2944	0.9677
	0.1963	0.9800	0.1916	0.9577	0.1822	0.9788	0.1963	0.9800
	0.0981	0.9944	0.1262	0.9766	0.1449	0.9866	0.1729	0.9866
	0	1	0	1	0	1	0	1
AUC	0.7427		0.6677		0.7574		0.7364	

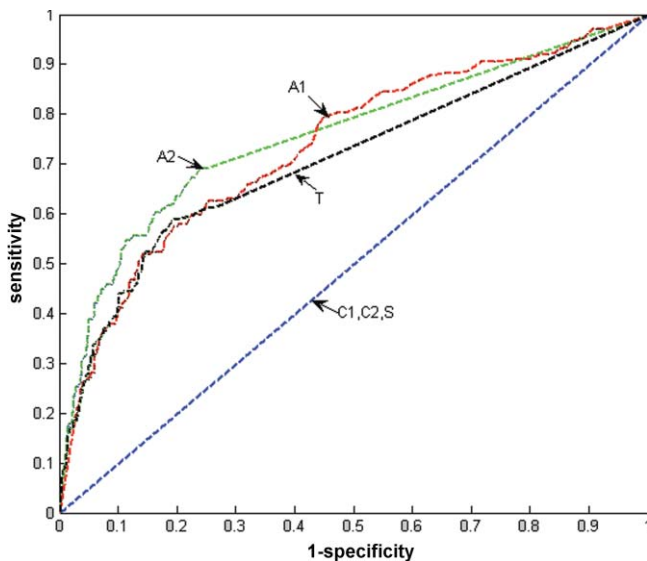


Fig. 3 ROC curves for different feature vectors.

a “gold standard,” namely, a set of annotated samples for learning and testing. In this paper, we used the 50 annotated retinal images from the Retinopathy Online Challenge database.⁹ The 50 images were from patients with diabetes without known diabetic retinopathy at the moment of photography; they represent a random sample of unique patients with “red lesions” from a large (>10,000 patients) diabetic retinopathy screening program. These images were taken with Topcon NW 100, NW 200, or Canon CR5-45NM “nonmydriatic” cameras at their native resolution and compression settings. The retinal specialist annotations were obtained from a combination of three ophthalmologists with retinal fellowship training. The first 20 images are for training, whereas the remaining 30 are for testing in this paper.

We extract all annotated MA regions as true positive samples and ~30 spurious objects as true negative samples from each training image and construct a training sample set, size of 727.

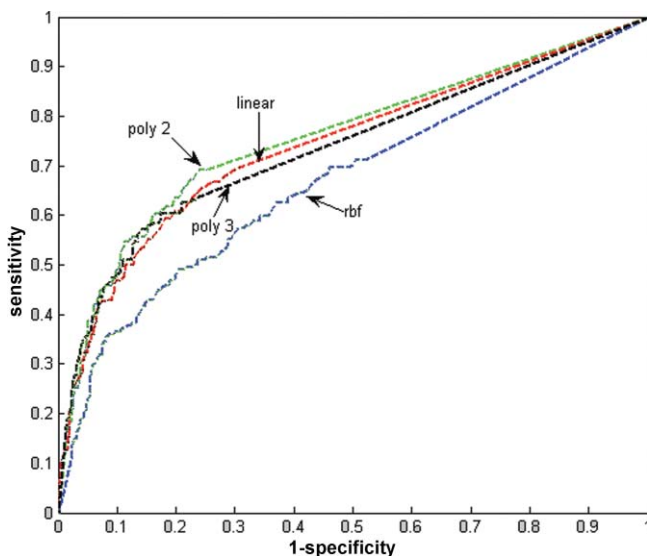


Fig. 4 ROC curves for different kernels of SVM.

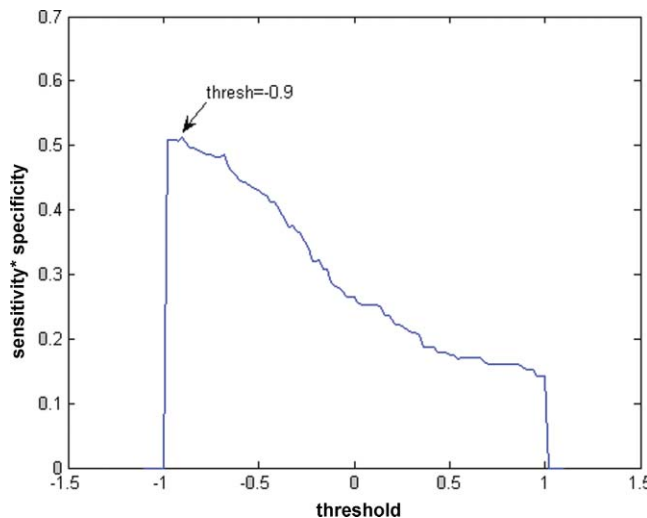


Fig. 5 Threshold for SVM.

And we obtain a test sample set, size of 1112, in the same manner.

4.2 Experimental Results

In order to identify the most favorable feature vector to distinguish real MAs from spurious objects, the feature vector shape features *S*, texture features *T*, and color features *C1*, *C2*, and the different combinations of them, *A1* and *A2*, are respectively used as the input of the SVM. ROC curves for different feature vectors are shown in Fig. 3.

The diagonal in Fig. 3 is the ROC of the SVM using *C1* or *C2* or *S* as feature vector, separately. They are laid over each other and have no diagnostic value. The ROC for *T*, *A1*, and *A2* have higher AUC^{ROC} . Table 1 shows that the ROC for *A2* has the highest AUC^{ROC} and the optimal classification performance among these feature vectors.

An SVM with different kernel functions has a significant difference of classification performance. In this paper, the commonly used kernel functions, such as linear function, Gaussian rbf, and polynomial kernel, are discussed. Figure 4 shows these ROC curves of SVM using these different kernel functions, and the best result is obtained by using the polynomial function kernel with $p = 2$. Table 2 shows that the ROCs for linear and polynomial functions have higher areas than the one for rbf, and the ROC for the polynomial function with $p = 2$ has the largest AUC^{ROC} value (0.7574), which is slightly larger than the one for the linear function. This indicates that the SVM using the polynomial function with $p = 2$ has the highest classification performance among the above-mentioned kernels.

Different thresholds can bring different classification results. Thus, the threshold is very important for classification performance of the classifier. Here, sensitivity × specificity is used as a performance measure to determine which threshold has the better diagnostic value.⁹ When threshold is set at -0.9 , sensitivity × specificity reached its maximum, as shown in Fig. 5. Because the threshold of classification is -0.9 , the sensitivity is 64% and the specificity is 80%.

5 Conclusion

In this paper, we present a new method to optimize the algorithm of automatic detection of MAs. The MA detection algorithm comprises three steps: candidate detection, feature extraction, and classification. We use a ROC curve to evaluate the distinguishing ability of different feature vectors and different kernel functions of an SVM. The feature vector A2 has the highest AUC^{ROC} in Fig. 3 and is used for the input of an SVM, and the polynomial function kernel with $p = 2$ shows the best discriminating performance according to those the ROC curves in Fig. 4. As mentioned above, the ROC curve is a useful technique for estimate classification performance of classifiers. It can be used to select the appropriate feature vector and to optimize the classifier.

References

1. W. L. Yun, U. R. Acharya, and Y. V. Venkatesh, "Identification of different stages of diabetic retinopathy using retinal optical images," *Inf. Sci.* **178**(1), 106–121 (2008).
2. Y. Hatanaka, T. Nakagawa, and Y. Hayashi, "CAD scheme to detect hemorrhages and exudates in ocular fundus images," *Proc. SPIE* **6514**, 65142M1 (2007).
3. B. Zhang, X. Wu, J. You, Q. Li, and F. Karray, "Detection of microaneurysms using multi-scale correlation coefficients," *Pattern Recogn.* **43**(6), 2237–2248 (2010).
4. T. Walter, P. Massin, and A. Erginay, "Automatic detection of microaneurysms in color fundus images," *Med. Image Anal.* **11**(6), 555–566 (2007).
5. A. Sopharak, B. Uyyanonvara, S. Barman, and T. H. Williamson, "Automatic detection of diabetic retinopathy exudates from non-dilated retinal images using mathematical morphology methods," *Comput. Med. Imaging Graphics* **32**(8), 720–727 (2008).
6. A. Gelzinis, A. Verikas, and M. Bacauskiene, "Increasing the discrimination power of the co-occurrence matrix-based features," *Pattern Recogn.* **40**(9), 2367–2372 (2007).
7. M. H. Horng, "Multi-class support vector machine for classification of the ultrasonic images of supraspinatus," *Expert Syst. Appl.* **36**(4), 8124–8133 (2009).
8. X. Yuan, G. Ao, C. Quan, J. Zhang, P. Wang, and Y. Tian, "ROC analysis of CT hemodynamic in the diagnosis of breast cancer," *Chin.-German J. Clin. Oncol.* **9**(3), 165–168 (2010).
9. Retinopathy Online Challenge, (<http://roc.healthcare.uiowa.edu>) (Mar. 3, 2010).