

Automated Microaneurysm Detection Using Local Contrast Normalization and Local Vessel Detection

Alan D. Fleming*, Sam Philip, Keith A. Goatman, John A. Olson, and Peter F. Sharp

Abstract—Screening programs using retinal photography for the detection of diabetic eye disease are being introduced in the U.K. and elsewhere. Automatic grading of the images is being considered by health boards so that the human grading task is reduced. Microaneurysms (MAs) are the earliest sign of this disease and so are very important for classifying whether images show signs of retinopathy. This paper describes automatic methods for MA detection and shows how image contrast normalization can improve the ability to distinguish between MAs and other dots that occur on the retina. Various methods for contrast normalization are compared. Best results were obtained with a method that uses the watershed transform to derive a region that contains no vessels or other lesions. Dots within vessels are handled successfully using a local vessel detection technique. Results are presented for detection of individual MAs and for detection of images containing MAs. Images containing MAs are detected with sensitivity 85.4% and specificity 83.1%.

Index Terms—Automated detection, contrast normalization, microaneurysm, retinal imaging, screening.

I. INTRODUCTION

SCREENING programs for diabetic retinopathy are being set up in the U.K. with the aim of detecting and treating the disease before it causes blindness. These are based on digital photography of the retina with or without mydriasis (pupil dilation). Automated analysis of retinal images has the potential to reduce the screening program costs compared to manual image grading. Diabetic retinopathy produces a variety of lesions on the retina. Of these, microaneurysms (MAs) are likely to be the only lesion present at the earliest stage of the disease and continue to be present as the disease develops. Thus MAs are of major importance for automated retinopathy detection.

Since retinopathy is present in only 30% of the diabetic population, automated detection of MAs has the potential for a large reduction of the manual grading workload. Hipwell *et al.* [1] have shown that an automated MA detection system can reduce the workload by 50%. According to the severity of retinopathy, manual graders would decide whether a patient should be recalled at the normal or reduced screening interval or whether referral to an ophthalmologist is required.

Manuscript received January 16, 2006; revised May 29, 2006. Asterisk indicates corresponding author.

*A. D. Fleming is with the Biomedical Physics, University of Aberdeen, Aberdeen, AB25 2ZD, U.K. (e-mail: a.fleming@biomed.abdn.ac.uk).

K. A. Goatman, and P. F. Sharp are with the Biomedical Physics, University of Aberdeen, Aberdeen, AB25 2ZD, U.K.

J. A. Olson and S. Philip are with the Grampian Diabetes Retinal Screening Programme, Woolmanhill Hospital, Aberdeen, AB25 1LD, U.K.

Digital Object Identifier 10.1109/TMI.2006.879953

MAs are swellings of the capillaries caused by a weakening of the vessel wall. In retinal photographs, although the capillaries are not visible, MAs appear as dark red isolated dots. In common with vessels, MAs appear with highest contrast in the green plane of the color image. In normal retinas, dots are also present though with lesser contrast compared to MAs. Although the human eye has no difficulty identifying most MAs, there will inevitably be disagreements between human experts, especially regarding the fainter dots. Factors which make identification difficult for an automated system include variation in image lighting, the variability of image clarity, and the highly variable image background texture. Dark spots commonly occur on vessels where their contrast relative to the background reaches a peak and also at junctions even where the vessels themselves are indistinct. Bright reflections are common in retinal images; small gaps may occur between bright areas and these, when interpreted too locally, are indistinguishable from MAs. Another difficulty is that MAs can appear in clusters so that individual MAs no longer have the characteristics of an isolated dot.

Thus, any reliable system for MA detection must be complex and multifaceted in order to cope with these individual effects. Nonetheless, the basic characteristics of MAs, being of a fairly uniform shape and having a limited range of size, has led to the use of basic morphology for the identification of candidate MAs [2]–[4]. For example, [2] and [3] use the maximum of top-hat operations based on linear structuring elements at a range of different inclinations. Any dark object which can contain one of the structuring elements is removed. Therefore, linear objects, such as vessels, tend to be removed but MAs of smaller diameter than the length of the structuring elements are retained. An alternative morphological top-hat operation is used by [4] and is based on a “bounding box closing.” This retains any dark object whose maximum diameter is less than a given size. In practice, most of the candidate points resulting from these methods are not MAs and many lie on vessels.

Hence, further vessel exclusion is performed as a separate step. In [2] and [3], this is achieved by intensity based region growing from the candidate center. On vessels, this is likely to produce an object whose shape is too large, too elongate, or too complex in form to be an MA. A similar approach is also employed in [5] for detection of MAs and larger red lesions known as haemorrhages. In [4] and [6], an independent vessel detection method is first employed so that MA candidates can be rejected where vessel detection was successful. Detection of haemorrhages in [7] uses a similar approach in which vessels are detected by an artificial neural network. An artificial neural network is also used by [8] to classify image regions as vessel, haemorrhage, or other lesion type.

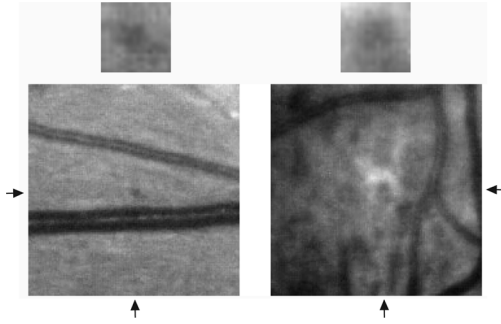


Fig. 1. Low contrast MA compared to a dot in a region of retina with high contrast. Upper images show the central dot extracted from each of the lower images and enlarged.

Another requirement for MA detection is to classify true MAs from dots which are caused by noise in the background. This is achieved in [2] and [3] by measuring features of the candidate MAs. These include, size, perimeter, complexity, and various intensity measures including intensity relative to the background. In [4], a threshold is derived from a measure of salt-and-pepper noise in the image; a higher threshold is used at higher noise levels. In [6], the energy of the candidate is compared with the energy of the background prior to vessel exclusion.

This paper describes methods for MA detection with emphasis on the role of local contrast normalization. Fig. 1 shows why local contrast evaluation is important. The upper small images each contain a dark spot of which only the one on the right is an MA. In the lower images showing a larger area the left image is easily observed to contain an MA at its center. The central spot in the right image, however, is not perceived as an isolated dot since similar retinal intensity variations occur throughout the image. From this we can deduce that the recognition of MAs is dependent on the contrast in the surrounding retina.

Other authors seem to have recognized the contrast sensitivity of retinal lesion recognition and have applied a contrast normalizing preprocessing technique [7], [9], [4]. They have used a constant sized running window for calculation of the original image contrast. The image is then normalized by dividing throughout by this contrast estimate. Another contrast normalization method is described by Foracchia *et al.* [10]. It uses an adaptive calculation of contrast by first identifying pixels which are likely to belong to the background retina. The standard deviation over these pixels is used as an estimate of local contrast. For reasons of efficiency, the calculation is performed only at points on a widely spaced grid. The contrast estimate is then extended across the entire image by interpolation.

This paper describes a new adaptive method for contrast normalization which is compared to alternative contrast normalization methods. We also describe a novel method for vessel exclusion and report its effectiveness. The ability of the system to identify individual MAs as well as images containing MAs is presented.

II. MATERIAL

The images used in this study cover a 45° retinal field of view containing disc and macula. They were acquired from diabetic patients attending the Grampian Diabetes

Retinal Screening Programme using a Canon EOS-D30 three mega-pixel (2160×1440) color digital camera attached to a Canon CR-6NM nonmydriatic fundus camera. Images were compressed by the camera, using its highest quality JPEG option. 1441 images were graded by a clinician for the presence of MAs. In 71 additional images, the clinician marked the locations of 422 individual MAs with between 1 and 37 MAs in individual images. The mean and standard deviation of area of regions filled during hand delineation was 69 ± 40 pixels.

III. METHODS

A. Notation

The following notation will be used.

$*$	Convolution operator.
$\text{gauss}(\sigma)$	Gaussian curve or surface with variance σ^2 .
$\text{boundary}(A)$	Boundary of region A defined as the set of pixels in region A that are 4-connected to the complement of A .
$\text{sd}(A)$	The standard deviation of values in set A or in image A .
$\text{grad}(A)$	Gradient magnitude of image A calculated by $\sqrt{(A * s_x)^2 + (A * s_y)^2}$ where s_x and s_y are Sobel gradient operators in the x and y directions.
$\text{area}(A)$	Number of pixels in region or image A .
A^c	Complement of region A .
$A \setminus B$	Set difference: $A \cap B^c$.

B. Preprocessing

The green plane of the image is filtered using a 3×3 median filter to remove occasional salt-and-pepper noise. Further noise removal is performed by convolving the image with $\text{gauss}(2)$ and the result is denoted I . Region growing and feature calculation is made more reliable by this step. The background intensity, denoted B , is estimated by applying a 68×68 median filter to I . A shade corrected image is generated with

$$S' = I/B - 1 \quad (1)$$

and is normalized for global image contrast by dividing by its standard deviation

$$S = S' / \text{sd}(S'). \quad (2)$$

The value of S at an MA or a vessel is typically -6 to -1.5 while retina is typically -0.5 to $+0.5$. Since the optic disc would tend to generate frequent false positive MA detections, the optic disc was detected automatically [11] and this region was omitted from further analysis.

C. Candidate Detection

Candidate MA detection has been previously reported by our group [2]. This section summarizes the method, parameters

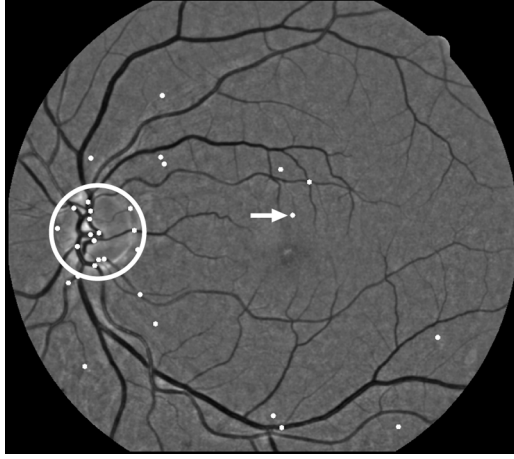


Fig. 2. Shade corrected image showing candidate MAs as white dots. Arrow shows the only true MA. Circle shows the automatically determined optic disc region in which candidate MAs are ignored.

being scaled appropriately for the images in this study. Let λ_θ be a linear structuring element which approximates a straight line of length 27 pixels at orientation θ . A top-hat operation [12] is performed by taking the minimum of morphological closing operations performed with the linear structuring element at eight different orientations. The original image is then subtracted

$$T = \min_{n=0 \dots \tau} [\phi(S, \lambda_{n\pi/8})] - S \quad (3)$$

where $\phi(A, E)$ represents morphological closing of image A by structuring element E . To enhance MAs with respect to noise, T is convolved with a Gaussian that approximates the expected profile of an MA using

$$T_f = T * \text{gauss}(2). \quad (4)$$

The intensity of the noise floor is estimated by taking the lowest threshold τ for which at most 5% of the pixels in T_f are greater than τ . Objects well above this intensity are segmented by applying a threshold 5 times greater than τ

$$T_c = \begin{cases} 1, & \text{if } T_f > 5\tau \\ 0, & \text{otherwise} \end{cases}. \quad (5)$$

Within each connected region D for which $T_c = 1$, a single point $q \in D$ is selected such that $S(q) = \min_{p \in D} (S(p))$. Each point q is known as a candidate MA. Fig. 2 shows a shade-corrected image and its candidate MAs.

D. Candidate Region Growing

Region growing was performed at each candidate MA using pixel q as a seed. Given a threshold value $t > 0$, $C(t)$ is defined as the largest 8-connected region which contains q and in which $S(p) \leq S(q) + t$ for all $p \in C(t)$. These regions were evaluated for $t = 0.1, 0.2, \dots, t_{\max}$ where t_{\max} is the largest value in this sequence for which $\text{area}(C(t)) < 3000$ pixels. Now an

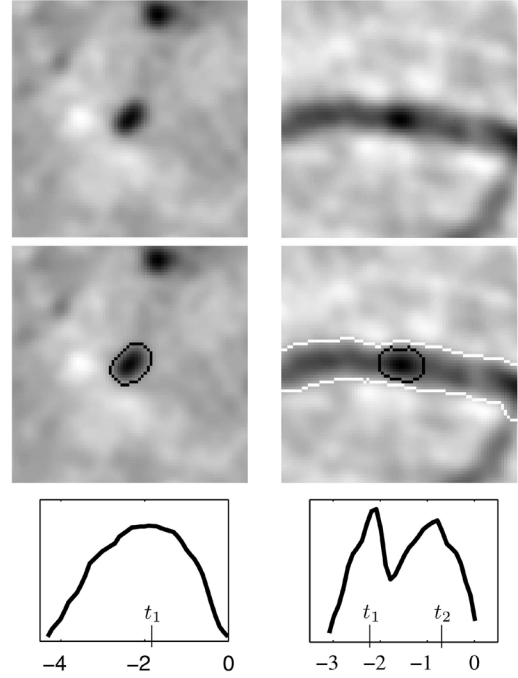


Fig. 3. Region growing at an MA (left) and at a dot within a vessel (right). Plot of energy, $E(t) * \text{gauss}(0.2)$, against threshold t shows one peak in the case of an MA and two peaks for a vessel. Black borders (in both images) are $C (= C(t_1))$ and the white border (in the right hand image) is $C(t_2)$.

“energy” function is defined as the mean squared gradient magnitude of S around the boundary of $C(t)$ using

$$G = \text{grad}(S)^2$$

$$E(t) = \text{mean}_{p \in \text{boundary}(C(t))} G(p). \quad (6)$$

Values of t at which $E(t)$ attains a peak are likely to be where the boundary of $C(t)$ separates an MA or a vessel from its background. Experimentation showed that smoothing of $E(t)$ with a Gaussian curve ($\sigma = 0.2$) removes many peaks occurring as a result of noise while leaving peaks that correspond to the boundary of an MA or of a vessel. Values of t at which $E(t) * \text{gauss}(0.2)$ attains a peak, as in Fig. 3, are denoted by $t_1, \dots, t_{N_{\text{peak}}}$, ($t_1 < \dots < t_{N_{\text{peak}}}$) where N_{peak} is the number of peaks. The region $C(t_1)$ will be abbreviated to C and is taken as the result of candidate region growing. C is used to define the extent of the MA candidate. A typical MA with a profile similar to a two-dimensional Gaussian function gives rise to a single peak in the $E(t)$ curve ($N_{\text{peak}} = 1$) as in the image on the left of Fig. 3. The right hand side of Fig. 3 shows a situation where the candidate is a dot within a vessel. In this case, $E(t)$ has two peaks since the gradient attains a peak at the boundary of the dot and again on the boundary of the vessel.

E. Candidate Evaluation

Least-squares fitting of Gaussian functions to vessels and to small round lesions in retinal images is described in [13]. From the parameters of the fitted function, width and orientation of the vessel or lesion are estimated. However, our experience in applying this technique to MAs has been that the function is affected by irregularities in the background. These can distort the function so that it does not fit well with the MA. We have

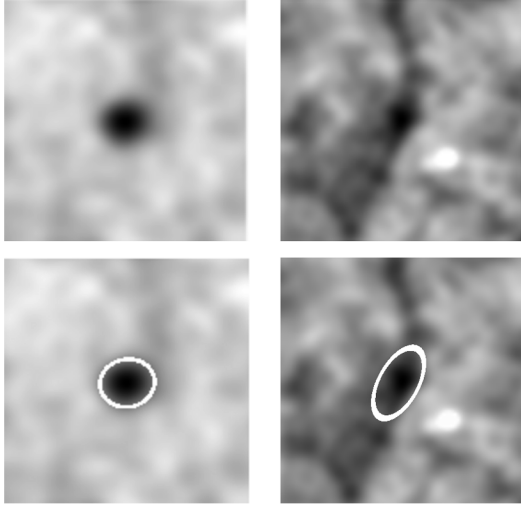


Fig. 4. White ellipses show the intersection of the fitted paraboloid with the plane $z = \bar{z}_b$ for an MA (left) and for a candidate MA on a vessel (right).

avoided this problem by using the Levenberg–Marquardt algorithm to minimize the sum of squares error between a paraboloid and S only for pixels, (x, y) , in C . The paraboloid is defined by

$$z = A_3(x - A_1)^2 + 2A_4(x - A_1)(y - A_2) + A_5(y - A_2)^2 + A_6 \quad (7)$$

where z corresponds to image intensity (of S) and A_1, \dots, A_6 are parameters to be varied. Let \bar{z}_b be the mean image intensity along the boundary of C . Then intersection of the paraboloid with the plane $z = \bar{z}_b$ is an ellipse which makes a good visual fit with the boundary of the MA candidate, as in Fig. 4. The axis lengths of this ellipse can be calculated by

$$u^\pm = \frac{1}{2} \left[A_3 + A_5 \pm \sqrt{(A_3 - A_5)^2 + 4A_4^2} \right] \quad (8)$$

$$a_{\text{minor}} = 2\sqrt{(\bar{z}_b - A_6)/u^+} \quad (9)$$

$$a_{\text{major}} = 2\sqrt{(\bar{z}_b - A_6)/u^-}. \quad (10)$$

These are used to calculate the candidate MA features in Table I grouped in two sets $\mathcal{F}_{\text{size}}$ and \mathcal{F}_{int} . Similar features can be calculated from the moments of the region C and from image statistics within region C [2]. $\mathcal{F}_{\text{size}}$ includes size measures, as well as d_I which is a measure of intensity of the candidate MA in the original image I relative to its background. This is included so that dots caused by digitization noise, which are particularly prevalent in low intensity images, can be correctly classified as non-MAs.

F. Watershed Retinal Region Growing

This section describes how a region of retina to be used for contrast normalization is determined around each candidate MA. Referring to the left image of Fig. 1 it can be seen that such a region must not intersect the nearby vessels otherwise local contrast would be overestimated. For reasons of efficiency, the processing is applied only to a subimage \check{S} centered on q

TABLE I
FEATURE SETS $\mathcal{F}_{\text{size}}$ AND \mathcal{F}_{int}

Symbol	Explanation
Feature set $\mathcal{F}_{\text{size}}$	
N_{peak}	The number of peaks in the smoothed energy function, $\text{gauss}(0.2) * E(t)$.
a_{major}	Major-axis length calculated with (10).
a_{mean}^{+1}	The geometric mean of minor and major axes of the cross-section of the paraboloid at $z = A_6 + 1$, (i.e. one unit above its minimum) calculated as $a_{\text{mean}}^{+1} = 2/(u^+u^-)^{1/4}$, see (8).
e	Eccentricity of the elliptical cross-section of the paraboloid calculated with $e = \sqrt{1 - u^-/u^+}$.
d_I	Depth of the candidate measured in the original image I , calculated with $d_I = B(q) - I(q)$ where B and I are defined in section III-B.
Feature set \mathcal{F}_{int}	
d_S	Depth of the MA candidate measured in S using the depth of the paraboloid, calculated with $d_S = \bar{z}_b - A_6$.
E	‘Energy’ in image S calculated at the boundary of C . Using the notation of sub-section D, this is $E(t_1)$.
d'_S	The depth of the MA candidate divided by its mean diameter, calculated with $d'_S = d_S/\sqrt{a_{\text{minor}}a_{\text{major}}}$, see (9) and (10).
E'	The energy at the boundary of the MA candidate with a depth correction, calculated with $E' = E/\sqrt{d_S}$.

extracted from S . Experiments showed that 121×121 pixels is sufficient for \check{S} . Region growing, based on the watershed transform of the gradient of \check{S} , is performed. To reduce the number of regions generated by the watershed transform, the h -minima transform [12] (denoted $\text{HMIN}_h(\cdot)$) was applied to $\text{grad}(\check{S})$. This suppresses all minima whose depth is less than or equal to h . The result is used as a marker function for watershed generation [12]. By successive approximation, a value of h was determined such that $\text{HMIN}_h(\text{grad}(\check{S}))$ contains between 20 and 60 minima. These limits are somewhat arbitrary but were found to be suitable for ensuring regions large enough for the subsequent algorithm. The watershed transform is then applied to $\text{HMIN}_h(\text{grad}(\check{S}))$ and divides \check{S} into regions which will be denoted W'_1, \dots, W'_m . Pixels on watersheds are assigned to an adjacent watershed region so that every pixel in \check{S} is assigned to one of the W'_i . The dash indicates that the watershed transform was applied to the gradient of \check{S} .

Region growing starts with region R_1 which is the union of the watershed-regions which intersect C

$$R_1 = \bigcup \{W'_i : W'_i \cap C \neq \emptyset\}. \quad (11)$$

Then at each iteration n a single member of $\{W'_1, \dots, W'_m\}$ is added to R_n to generate R_{n+1} . This watershed-region, denoted W'_{j_n} , is chosen from those adjacent to R_n , to minimize

$$|\text{mean}(\check{S}(W'_{j_n})) - M_n| \quad (12)$$

where

$$M_n = \begin{cases} \text{mean}(\check{S}(R_n \setminus R_1)), & \text{if } n > 1 \\ 0, & \text{if } n = 1 \end{cases} \quad (13)$$

is the mean of regions included so far, excluding the seed region. If possible, W'_{j_n} is chosen from regions whose shared border with R_n is greater than 10 pixels. This is much smaller than the

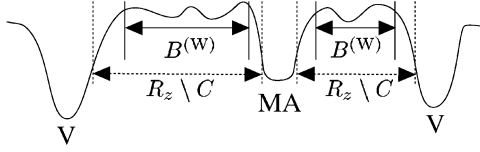


Fig. 5. Intensity profile through a microaneurysm (MA) and two nearby vessels (V). Maxima of the first derivative are indicated by dotted vertical lines. $R_z \setminus C$ includes borders of the MA and the vessels. These borders are removed to obtain $B^{(W)}$ by clipping at maxima of the second derivative, indicated by solid vertical lines.

average dimension of watershed regions. The result of the next iteration of region growing is

$$R_{n+1} = R_n \cup W'_{j_n}. \quad (14)$$

Two stopping conditions are used to terminate region growing, when $n \geq 3$

$$\text{area}(R_n) \geq \text{area}(\check{S})/2 \quad (15)$$

and

$$\frac{|\text{mean}(\check{S}(W'_{j_n})) - M_n|}{\text{sd}(\check{S}(R_n \setminus R_1))} \sqrt{\frac{\text{area}(R_n \setminus R_1)}{\text{area}(\check{S})}} > 1. \quad (16)$$

The first condition (15) causes growing to stop if R_n is large enough to evaluate the retinal contrast, as described later. The second condition (16) causes growing to stop if the mean intensity of the watershed region to be added is very different from previously added regions. This is likely when further growing would incorporate part of a vessel. The square root term in (16) causes this to become more likely as R_n increases in area. The final result of region growing is denoted R_z .

R_z contains C the initial estimate of the candidate MA extent. This was removed since its inclusion may cause an overestimate of contrast. Also, R_z may overlap onto the borders of dark or light features, as in Fig. 5. This is another source of overestimation of contrast since W'_i are bounded by locations of locally maximum gradient which, for particularly dark or light features, may occur where the image is already becoming darker or lighter than the surrounding background. These borders were removed by eroding pixels from the edge of R_z back to where the absolute value of the second derivative of \check{S} attains a maximum. This was achieved by using the watershed transform applied to $\text{grad}(\text{grad}(\check{S}))$ to segment \check{S} into regions W''_i . Pixels on watersheds are assigned to one of the regions. Region C and all of W''_k , where $W''_k \cap \text{boundary}(R_z \setminus C) \neq \emptyset$, are removed from R_z (see Fig. 5) to obtain $B^{(W)}$, the region used for contrast normalization

$$B^{(W)} = (R_z \setminus C) \setminus \left[\bigcup_k W''_k \right]. \quad (17)$$

The superscript on $B^{(W)}$ is a notation generalized in Section IV and indicates that it was obtained with the watershed retinal re-

gion growing method. Examples of the results of this method are displayed in Fig. 6.

G. Contrast Normalized Feature Evaluation

A contrast normalized version of \check{S} is obtained as follows. First, intensity gradients at a scale larger than the extent of a typical MA are removed by applying a high pass filter masked by $B^{(W)}$

$$\check{S}_{\text{hpf}} = \check{S} - \frac{(\check{S} \times B^{(W)}) * \text{gauss}(3)}{B^{(W)} * \text{gauss}(3)} \quad (18)$$

where $B^{(W)}$ is treated as an image which is 1 inside the region and 0 outside and where multiplication and division are on a pixel-by-pixel basis. Then contrast normalization is performed by dividing \check{S} by the standard deviation of \check{S}_{hpf} over $B^{(W)}$ to obtain $\check{S}^{(W)}$

$$\check{S}^{(W)} = \check{S} / \text{sd}(\check{S}_{\text{hpf}}(B^{(W)})). \quad (19)$$

A set of features, $\mathcal{F}^{(W)} = \{d_S^{(W)}, E^{(W)}, d'_S^{(W)}, E'^{(W)}\}$, is evaluated in the same way as the features in \mathcal{F}_{int} , defined in Table I, but using $\check{S}^{(W)}$ in place of S .

H. Vessel Exclusion

Many MA candidates are dark spots within a retinal vessel and so false positive MA detection would be reduced by reliable vessel detection in the vicinity of each MA candidate. Some reported methods for MA and other red lesion detection [7], [4] have attempted to locate all retinal vessels which are then excluded from further analysis. The method described here takes a different approach by performing analysis only on subimages surrounding each candidate. This method requires less processing time, since it is applied to a small portion of the image, and can be tailored to the typical vessel segments which are observed around candidate MAs. Moreover, the method is flexible and allows an MA to occur on a vessel, for example, if it is much darker or wider than the vessel. This section describes how a Boolean valued feature, *isvessel*, is derived for each MA candidate such that *isvessel* is true if the MA candidate appears to be part of a vessel.

Let $\tilde{S}^{(W)}$ be a subimage of dimensions 81×81 pixels extracted from the contrast normalized subimage, $\check{S}^{(W)}$, defined in (19), and centered on the candidate MA at q . Consider a transformation $\tilde{S}^{(W)} \rightarrow \tilde{S}_\alpha^{(W)}$ defined in polar coordinates centered on q , such that $\tilde{S}_\alpha^{(W)}(r, \theta) = \tilde{S}^{(W)}(r, \theta')$ where $\theta' = \theta + \tan^{-1}(\alpha/r)$. This transformation shifts each point, by a distance α , circumferentially on circles centered on q . An image which is positive near the centre of approximately radial linear dark features (which are likely to be vessels converging on the candidate MA) can be defined by

$$V = \min(\tilde{S}_{+\alpha}^{(W)} - \tilde{S}^{(W)}, \tilde{S}_{-\alpha}^{(W)} - \tilde{S}^{(W)}) \quad (20)$$

where $\alpha = 8$ pixels is chosen to be half of a medium sized vessel width.

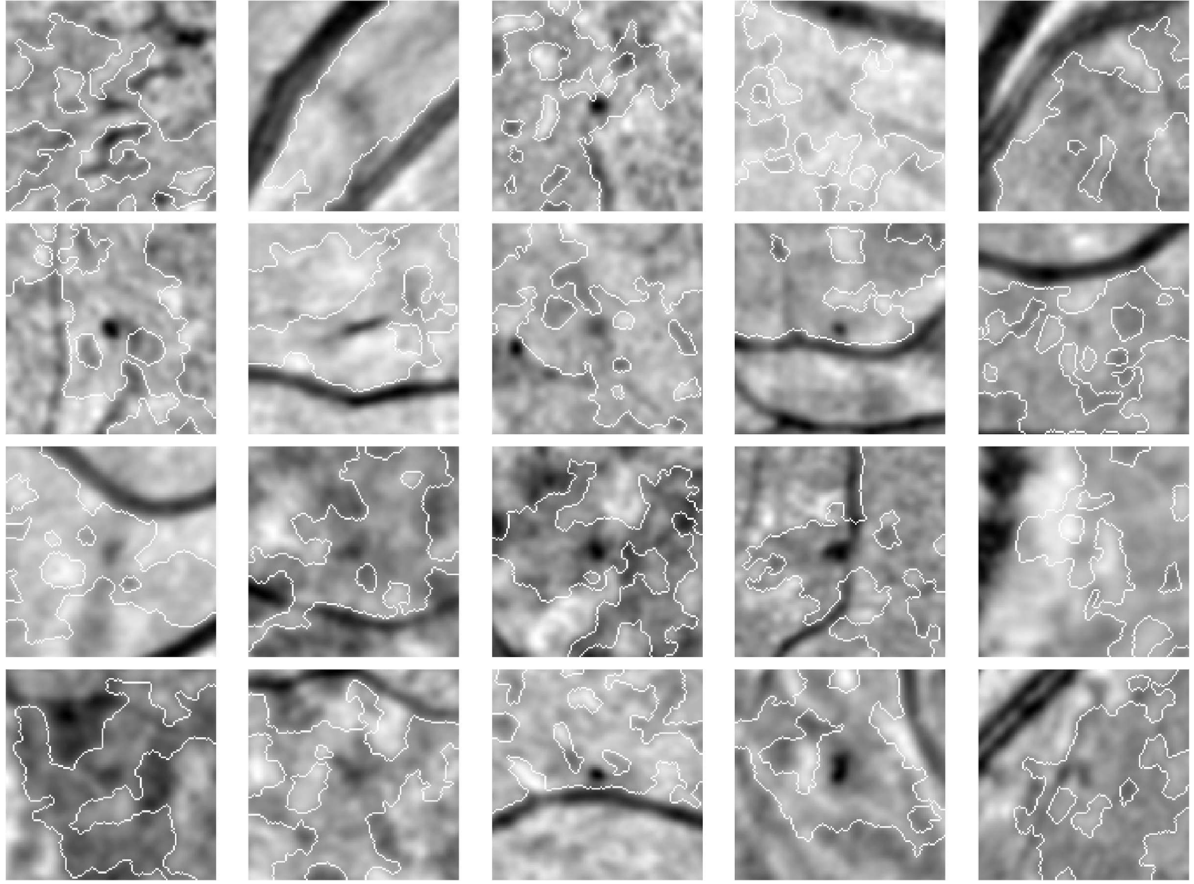


Fig. 6. Results of watershed retinal region growing around candidate MAs. Each panel shows a subimage \tilde{S} with an MA candidate q at its exact center. White lines are the boundary of $B^{(w)} \cup C$ (C has been included to clarify interpretation of the boundaries). Internal side of the boundaries can be identified since it contains the center of the image. Figure was created from candidate MAs which were randomly selected except that candidates falling on vessels have not been included.

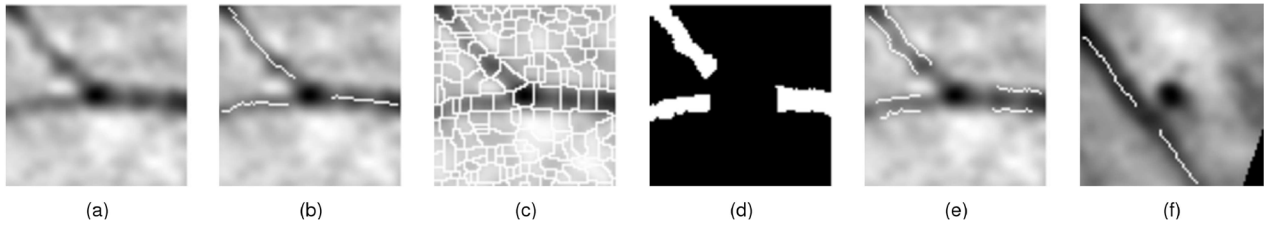


Fig. 7. (a) MA candidate at a vessel junction. (b) Candidate vessel segments. (c) Watershed regions. (d) Vessel segments grown to fill watershed regions. (e) Edges of grown regions used in calculation of E_v . (f) MA adjacent to a vessel; vessel segments, if extrapolated, do not pass through the MA.

Since a universal threshold that segments the vessel structures in V does not exist, multiple thresholds were applied. At each value of the threshold, objects with a shape suggestive of a vessel segment were extracted. The threshold takes three values, 0.5, 0.4, and 0.3, which are applied to V to generate three binary images which are skeletonised. Each skeleton is improved by spur removal and is then broken at its junctions. Examination of many training examples showed that a resulting two-ended skeletal segment has sufficient extent to be accepted as a candidate vessel segment if it has one end less than a distance of 13 pixels from q and if the other end is over 35 pixels from q . The candidate vessel segments are obtained from each binary image and duplicates are rejected, Fig. 7(b).

Each skeletal candidate vessel segment is then grown to fill the watershed-regions, W'_i , defined in Section III-F, which it intersects as in Fig. 7(c) and (d). W'_i are bounded by local gra-

dient maxima and so are bounded by the vessel walls. The mean squared gradient of $\tilde{S}^{(w)}$ along the edges of the vessel is calculated and denoted E_v . In training examples, candidate vessel segments with $E_v \leq 3$ were due to noise. Therefore, a candidate is classified as vessel if $E_v > 3$.

The vessel may however be running beside an actual MA, as in Fig. 7(f). A straight line is fitted to the pixels in the third of the vessel segment which is closest to the candidate MA. If this line passes further than a_{\min} , defined in (9), from q then the vessel segment is removed from further analysis because it does not pass through the candidate MA.

Having identified that the MA candidate lies on a vessel, a decision must be made as to whether the candidate is part of the vessel or whether it is an MA which is straddling the course of a vessel. A dark spot commonly occurs at a vessel junction or at a vessel crossover. These cases are dealt with by defining a vessel

junction to be a candidate MA at which more than two vessel segments meet or at which there are exactly two such vessel segments separated by an angle less than 135° . If the candidate is not on a junction, according to this definition, then it should be regarded as being an MA if it is much wider or much darker than the vessel. The mean width of the vessel segment, denoted w_v and the width of the MA candidate in a direction perpendicular to the vessel direction, denoted w , are measured. Comparison of vessel and MA candidate intensity are made using $E^{(W)} \in \mathcal{F}^{(W)}$, defined in Section III-G, and E_v . Training examples showed that $w > 1.5w_v$ and $E^{(W)} > 2E_v$ gave good results. These measurements are combined to define a Boolean feature, *isvessel*:

$$isvessel = \begin{cases} 1, & \text{if } \begin{cases} \text{the vessel segments form} \\ \text{a junction} \\ \text{or, } w > 1.5w_v \\ \text{or, } E^{(W)} > 2E_v \end{cases} \\ 0, & \text{otherwise.} \end{cases} \quad (21)$$

I. Classification

A training set was constructed from the 71 images in which individual MAs had been marked plus 7 images containing no MAs. The automated method found 1417 MA candidates in these images. All feature values were calculated for each candidate. A strictly increasing transformation was applied to each feature to make its distribution approximately uniform over the range 0 to 1. The transformed features were used to train a k -nearest neighbor (k NN) classifier. A value of $k = 15$ was derived from empirical evidence using training set images.

For any candidate to be identified, the same transformation was applied to each feature. Let N_{TRUE} be the count of true MAs in the k neighbors of the training set which are nearest to an MA candidate under test. The sensitivity of the classifier can be adjusted by choosing $P \in \{1, 2, \dots, k\}$, such that a candidate is classified as an MA if

$$N_{\text{TRUE}} \geq P. \quad (22)$$

The greatest number of candidates are classified as MAs when $P = 1$.

IV. EXPERIMENTS

Three additional methods for contrast normalization were tested. The notation $Y^{(X)}$ will be used to denote object Y calculated using contrast normalization method $X \in \{\text{"W"}, \text{"Wa"}, \text{"K"}, \text{"F"}\}$. The methods are as follows.

- “W” Watershed retinal region growing was performed to obtain $B^{(W)}$ as in (17). This region is used for contrast normalization as described in Section III-G.
- “Wa” Watershed retinal region growing was performed but omitted the last step defined by (17). Region $B^{(\text{Wa})} = R_z \setminus C$, as in Fig. 5, is used for contrast normalization instead of $B^{(W)}$ to obtain $\mathcal{F}^{(\text{Wa})}$.

- “K” A region $B^{(K)}$ is defined as a circle centered on q with constant diameter and with pixels in C omitted. A diameter of 81 pixels was used since this gives a region area similar to the average region area obtained by watershed retinal region growing. Contrast normalization was performed using $B^{(K)}$ in place of $B^{(W)}$ to obtain $\mathcal{F}_{\text{int}}^{(K)}$.
- “F” We implemented the contrast normalization method described by Foracchia *et al.*, [10], summarized here. Let $\hat{\mu}_N$ and $\hat{\sigma}_N$ be the local mean and standard deviation of an image I calculated over a square region with 200×200 pixels swept across the image. A set of “background” pixels \mathcal{B} is defined at which $|I(x, y) - \hat{\mu}_N(x, y)| / \hat{\sigma}_N(x, y) \leq 1$. Estimates of local luminosity \hat{L} and contrast \hat{C} are now evaluated at the centers of nonoverlapping 200×200 pixel blocks of the image as the mean and standard deviation of pixels in the intersection of \mathcal{B} with each block. \hat{L} and \hat{C} are extended across the entire image by bicubic interpolation. The image I is then normalized for luminosity and contrast with
$$\hat{I}^o(x, y) = (I(x, y) - \hat{L}(x, y)) / \hat{C}(x, y). \quad (23)$$

Define $\check{S}^{(F)}$ as a subimage of \hat{I}^o centered on q with the same dimensions as \check{S} . Features in $\mathcal{F}_{\text{int}}^{(F)}$ were calculated in the same way as features in \mathcal{F}_{int} but using $\check{S}^{(F)}$ in place of S .

Classification of MA candidates was performed using $\mathcal{F}_{\text{size}}$ and one of $\mathcal{F}_{\text{int}}^{(W)}$, $\mathcal{F}_{\text{int}}^{(\text{Wa})}$, $\mathcal{F}_{\text{int}}^{(K)}$, or $\mathcal{F}_{\text{int}}^{(F)}$.

V. RESULTS

A. Feature Assessment

To compare the contrast normalization methods, the effectiveness of individual features for classifying candidate MAs not lying on vessels was assessed by plotting receiver operator characteristic (ROC) curves. Only candidates for which *isvessel* is false were included in this analysis since contrast normalization was intended to assist the discrimination of candidates which are not on vessels. Sensitivity and specificity for correct classification by a feature f at a threshold t can be calculated as

$$\begin{aligned} \text{Sensitivity} &= N_{\text{ma}}^{f \geq t} / N_{\text{ma}} \\ \text{Specificity} &= N_{\text{non}}^{f < t} / N_{\text{non}} \end{aligned}$$

where N_{ma} and N_{non} are the number of candidates for which *isvessel* is false and which are true MAs or non-MAs, respectively. $N_{\text{ma}}^{\text{cond}}$ and $N_{\text{non}}^{\text{cond}}$ are the number of candidates which satisfy condition “cond” in N_{ma} and N_{non} . The ROC curve is then generated by plotting sensitivity against 1-specificity as t varies between the minimum and maximum values of f . The area under this ROC curve is especially useful since it is a scalar measure of the ability of each feature to correctly classify MAs and non-MAs. Such an analysis is displayed in Fig. 8 comparing d_S with $d_S^{(X)}$, and in Fig. 9 comparing E with $X \in \{\text{"W"}, \text{"Wa"}, \text{"K"}, \text{"F"}\}$. Similar plots could be made for d_S'

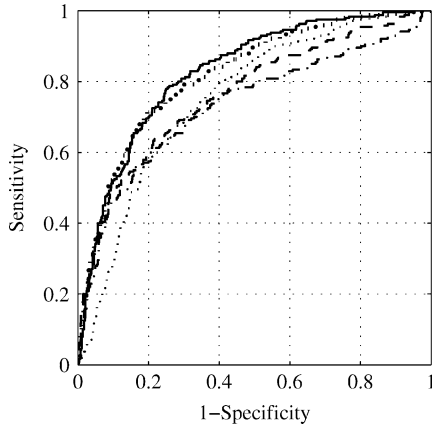


Fig. 8. Comparison of individual features based on d_S , depth of the MA candidate boundary, for discrimination between true MA and non-MA candidates for which $isvessel = \text{false}$.

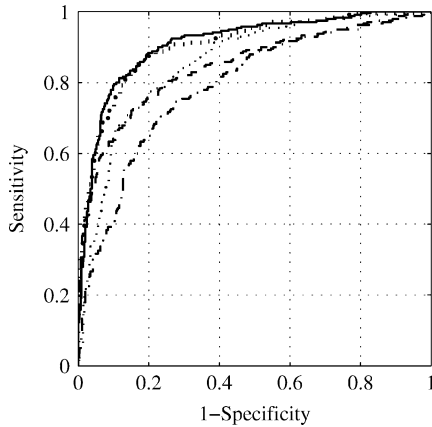


Fig. 9. Comparison of individual features based on E , “energy” at the candidate boundary, for discrimination between true MA and non-MA candidates for which $isvessel = \text{false}$.

TABLE II
AREAS UNDER ROC CURVES OBTAINED BY EACH LISTED FEATURE FOR CLASSIFYING CANDIDATES FOR WHICH $isvessel$ IS FALSE

Normalization method						
‘W’	‘Wa’	‘K’	‘F’	None		
$d_S^{(W)}$ 0.828	$d_S^{(Wa)}$ 0.819	$d_S^{(K)}$ 0.765	$d_S^{(F)}$ 0.733	d_S	0.753	
$E^{(W)}$ 0.911	$E^{(Wa)}$ 0.905	$E^{(K)}$ 0.847	$E^{(F)}$ 0.790	E	0.856	
$d'_S^{(W)}$ 0.907	$d'_S^{(Wa)}$ 0.900	$d'_S^{(K)}$ 0.848	$d'_S^{(F)}$ 0.792	d'_S	0.858	
$E'^{(W)}$ 0.882	$E'^{(Wa)}$ 0.884	$E'^{(K)}$ 0.852	$E'^{(F)}$ 0.793	E'	0.856	

and E' . However, for the sake of brevity and since the shapes of the curves are similar to those for d_S and E , these results are reported only as the areas under the ROC curves in Table II. In all cases, the features calculated with contrast normalization methods “W” and “Wa” show the best classification ability with “W” being slightly better than “Wa.” Method “K” shows almost no improvement over no contrast normalization and method “F” performed worse than no contrast normalization.

B. MA Detection

The recognition of individual MAs was assessed with the 71 images in which individual MAs had been identified plus an additional 165 images containing no MAs. This ensured preva-

lence of images with MAs similar to the local diabetic population. Since these images overlap the training set, a “leave-one-out” scheme was used in which each image was assessed using the training set with itself removed. Comparisons were made between the contrast normalization methods, Fig. 10. The improvement due to local vessel detection was also evaluated. Each curve was obtained by supplying the k NN-classifier with a set of features which, in all cases, included $\mathcal{F}_{\text{size}}$. The solid line was obtained with feature set $\mathcal{F}_{\text{size}} \cup \{isvessel\} \cup \mathcal{F}_{\text{int}}^{(W)}$. Other curves were obtained by replacing $\mathcal{F}_{\text{int}}^{(W)}$ with $\mathcal{F}_{\text{int}}^{(Wa)}$, $\mathcal{F}_{\text{int}}^{(K)}$, $\mathcal{F}_{\text{int}}^{(F)}$, or \mathcal{F}_{int} . The curve with triangular markers was obtained by removing $isvessel$. The points on each curve represent operating points given by $P = 1, \dots, 15$ in (22). The contrast normalizations methods have a similar ordering of effectiveness as suggested by the ROC analysis of the individual features.

Of the MAs identified by the clinician, 63.7% (267/422) were identified as candidate MAs. Feature $isvessel$ was true for 13.4% (154/1150) of candidates which were not MAs and for 1.9% (5/267) of candidates which were true MAs. Some failed cases of MA detection are shown in Fig. 11.

C. Image Classification

An automated image analysis scheme for diabetic retinopathy screening may be based on referring to manual graders all images with any MAs regardless of their number or location. To reflect this scenario, an image was regarded as positive if the automated system detected one or more MAs. The sensitivity for detection of images with one or more MAs gives the success rate for detection of this early sign of diabetic retinopathy and the specificity gives the reduction in workload for manual graders that can be achieved. 1441 images were assessed by a clinician, with 356 being classified as containing MAs and 1085 being classified as containing no MAs. ROC curves for the classification of images with any MAs are shown in Fig. 12. The feature sets supplied to the k NN-classifier are the same as those used in Fig. 10. Images containing MAs can be classified with sensitivity 85.4% and specificity 83.1%, (the point on the solid line indicated by the arrow in Fig. 12).

The average execution time per image of an implementation in Matlab and C was, for methods “W” and “Wa,” 58 s, for method “K,” 53 s, and for method “F,” 100 s.

VI. DISCUSSION

This paper has presented developments in automated methods for detection of microaneurysms in photographs of the retina. Comparative results for two parts of the automated method, local retinal contrast normalization and local vessel detection, have been reported.

Table II and Figs. 8 and 9 show that contrast normalization improves the discrimination ability of individual features. The best method for contrast normalization is watershed retinal region growing (method “W”) though the improvement over method “Wa” is very small. Method “K” is similar to no contrast normalization and method “F” performs worse than no contrast normalization. Fig. 10 shows that identification of individual MAs is affected in a similar way. At the bottom left end of Fig. 10, the number of false positives per image is too low for meaningful results.

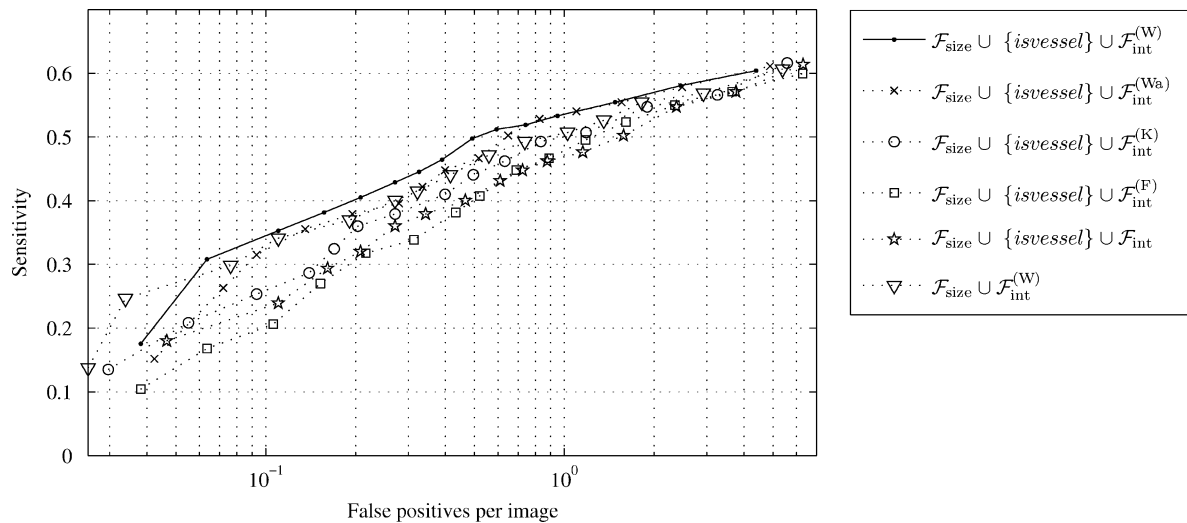


Fig. 10. Detection of individual MAs; comparison of contrast normalization methods and of no contrast normalization. Also shown is the effect of ignoring vessel detection. Legend lists the feature sets used in the k NN-classifier.

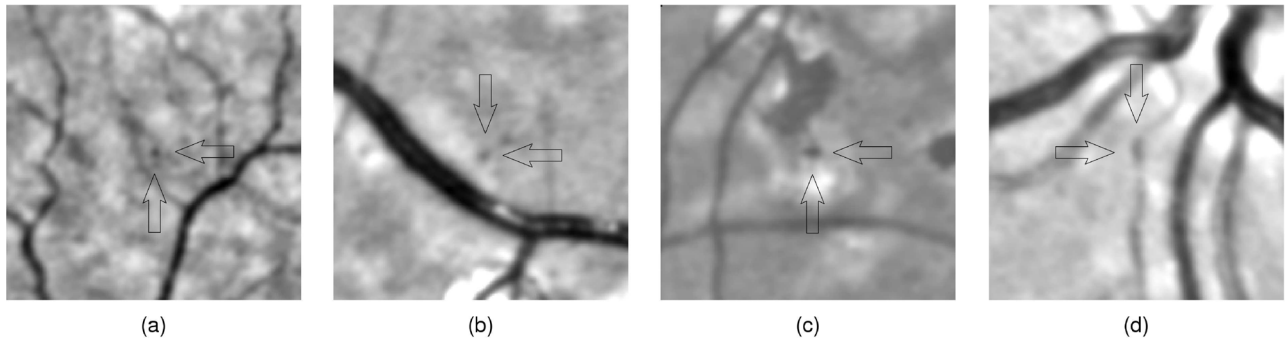


Fig. 11. (a) and (b) False negative detections due to small faint MA. (c) and (d) False positive detections. Dot in (c) is possibly due to pigmentation changes. Candidate in (d) was classified as an MA since the vessel was not recognized due to its uneven contrast.

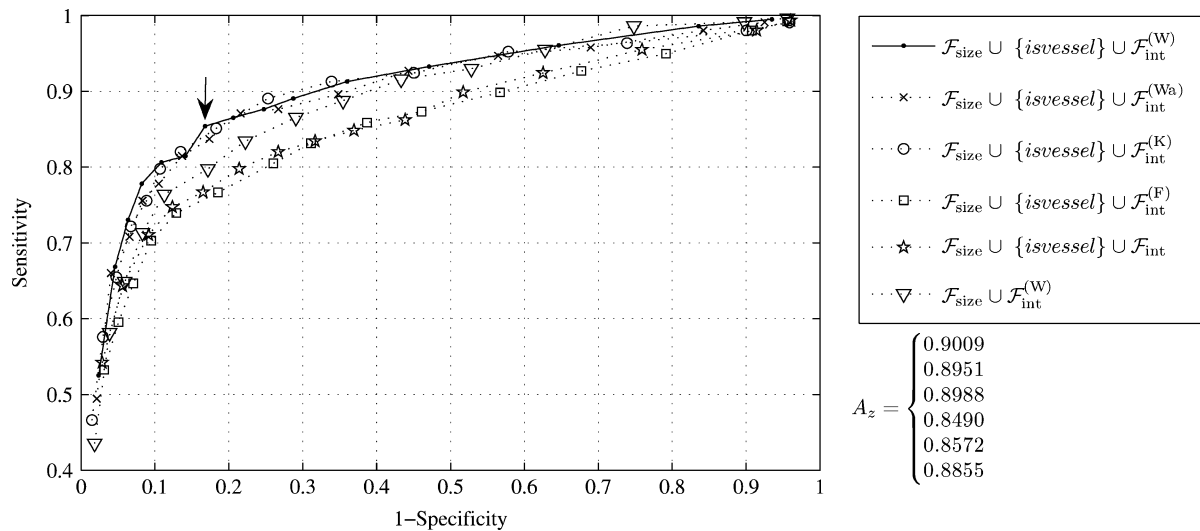


Fig. 12. Detection of images with MAs; comparison of contrast normalization methods and of no contrast normalization. Also shown is the effect of ignoring vessel detection. Legend lists the feature sets used in the k NN-classifier. Areas under the curves A_z are given in the order of the legend.

The results for image classification in Fig. 12 are less clear though a similar trend can be observed. This may be because of interdependencies between the occurrence of MAs in an image. For example, in images containing MAs at various stages of development and hence having a range of sizes, it is not important to detect the small faint MAs as long as the largest MA can be detected. Detection of large MAs is likely to be successful

even without contrast normalization. Fig. 12 shows that contrast normalization methods “W”, “Wa”, and “K” are better than no contrast normalization or contrast normalization method “F.” Method “K” performed similarly to methods “W” and “Wa” and this differs from results for individual MA detection.

The poor performance of method “F” requires some explanation. Image contrast estimates were made in constant sized

blocks (200×200 pixels). If a block overlaps dark vessels then the initial contrast estimate, $\hat{\sigma}_N$, (see paragraph "F" of IV) will be increased. It follows that the proportion of pixels truly belonging to background that are included in the intersection of \mathcal{B} with this block is larger than when the block does not contain dark vessels. Hence, the final contrast estimate is influenced by vessels. Also, method "F" uses interpolation to obtain contrast estimates at all points in the image since otherwise the computation time would be prohibitive. This means that overestimates of image contrast can be extended away from the dark vessels. An overestimate of image contrast is therefore made in a significant proportion of the image increasing the probability that MAs will be missed.

The results have shown that contrast normalization is beneficial to MA detection. Contrast normalization is most safely performed with a method which reliably avoids including vessels in the contrast estimate. Fig. 6 shows that method "W" achieves this. The success of this approach suggests that it could also produce benefits for identification of retinal vessels and of other types of retinal lesions. Adapting the method to larger lesions such as haemorrhages would require an estimate to be made of the boundary of each lesion candidate so that the lesion itself is not included in the contrast estimate.

The low number of MAs for which $isvessel$ is true indicate that a candidate is very unlikely to be an MA if $isvessel$ is true. Figs. 10 and 12 show that localized vessel detection improves MA detection. For example, sensitivity of individual MA detection is improved by approximately 3% when the false positive rate is between 0.2 and 0.4 per image (Fig. 10). Sensitivity for identification of images with MAs is improved by between 3% and 6% when the specificity is between 70% and 90% (Fig. 12).

Screening for diabetic retinopathy using digital photographs has already been widely introduced into the U.K. and elsewhere. Automation as a prescreening phase of image grading will make screening programs more cost effective by reducing the manual grading workload. Automated MA detection, as described in this paper, is able to perform this task. In addition, the techniques described are likely to find other applications particularly for identification of other features in retinal images.

REFERENCES

- [1] J. H. Hipwell, F. Strachan, J. A. Olson, K. C. McHardy, P. F. Sharp, and J. V. Forrester, "Automated detection of microaneurysms in digital red-free photographs: A diabetic retinopathy screening tool," *Diabetic Med.*, vol. 17, no. 8, pp. 588–594, 2000.
- [2] T. Spencer, J. A. Olson, K. C. McHardy, P. Sharp, and J. V. Forrester, "Image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus," *Comput. Biomed. Res.*, vol. 29, no. 4, pp. 284–302, 1996.
- [3] M. J. Cree, J. A. Olson, K. C. McHardy, P. F. Sharp, and J. V. Forrester, "A fully automated comparative microaneurysm digital detection system," *Eye*, vol. 11, no. 5, pp. 622–628, 1997.
- [4] T. Walter and J. C. Klein, "Automatic detection of microaneurysms in color fundus images of the human retina by means of the bounding box closing," in *Proceedings of Medical Data Analysis*. London, U.K.: Springer-Verlag, 2002, vol. 2526, Lecture Notes in Computer Science, pp. 210–220.
- [5] M. Niemeijer, B. van Ginneken, J. Staal, M. S. A. S. Schulten, and M. D. Abramoff, "Automatic detection of red lesions in digital color fundus photographs," *IEEE Trans. Med. Imag.*, vol. 24, no. 5, pp. 584–592, May 2005.
- [6] S. Abdelazeem, "Microaneurysm detection using vessels removal and circular Hough transform," in *IEEE Proc. 19th Natl. Radio Sci. Conf.*, Alexandria, Egypt, Mar. 2002, vol. 1–2, pp. 421–426.
- [7] C. Sinthanayothin, J. F. Boyce, T. H. Williamson, H. K. Cook, E. Mensah, S. Lal, and D. Usher, "Automated detection of diabetic retinopathy on digital fundus images," *Diabetic Med.*, vol. 19, no. 2, pp. 105–112, 2002.
- [8] G. G. Gardner, D. Keating, T. H. Williamson, and A. T. Elliott, "Automatic detection of diabetic retinopathy using an artificial neural network: A screening tool," *Br. J. Ophthalmol.*, vol. 80, no. 11, pp. 940–944, 1996.
- [9] A. Osareh, M. Mirmehdi, B. Thomas, and R. Markham, "Classification and localisation of diabetic-related eye disease," in *Computer Vision*. London, U.K.: Springer-Verlag, May 2002, Lecture Notes in Computer Science.
- [10] M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Med. Image Anal.*, vol. 9, no. 3, pp. 179–190, Jun. 2005.
- [11] A. D. Fleming, S. Philip, K. A. Goatman, J. A. Olson, and P. F. Sharp, "Automated assessment of retinal image field of view," in *Proc. Med. Image Understand. Anal.*, Sep. 2004, pp. 129–132.
- [12] P. Soille, *Morphological Image Analysis: Principles and Applications*, 2nd ed. Berlin, Germany: Springer-Verlag, 2003.
- [13] L. Wang, A. Bhalerao, and R. Wilson, J. Kittler, M. Petrou, and M. Nixon, Eds., "Robust modelling of local image structures and its application to medical imagery," in *Proc. 17th Int. Conf. Pattern Recogn.*, Aug. 2004, vol. 3, pp. 534–537.