Detecting Diabetes Mellitus and Nonproliferative Diabetic Retinopathy Using Tongue Color, Texture, and Geometry Features

Bob Zhang*, *Member, IEEE*, B. V. K. Vijaya Kumar, *Fellow, IEEE*, and David Zhang, *Fellow, IEEE*

***Abstract*—Diabetes mellitus (DM) and its complications leading to diabetic retinopathy (DR) are soon to become one of the 21st century's major health problems. This represents a huge financial burden to healthcare officials and governments. To combat this approaching epidemic, this paper proposes a noninvasive method to detect DM and nonproliferative diabetic retinopathy (NPDR), the initial stage of DR based on three groups of features extracted from tongue images. They include color, texture, and geometry. A noninvasive capture device with image correction first captures the tongue images. A tongue color gamut is established with 12 colors representing the tongue color features. The texture values of eight blocks strategically located on the tongue surface, with the additional mean of all eight blocks are used to characterize the nine tongue texture features. Finally, 13 features extracted from tongue images based on measurements, distances, areas, and their ratios represent the geometry features. Applying a combination of the 34 features, the proposed method can separate Healthy/DM tongues as well as NPDR/DM-sans NPDR (DM samples without NPDR) tongues using features from each of the three groups with average accuracies of 80.52% and 80.33%, respectively. This is on a database consisting of 130 Healthy and 296 DM samples, where 29 of those in DM are NPDR.**

***Index Terms*—Diabetes mellitus (DM) detection, nonproliferative diabetic retinopathy (NPDR) detection, tongue color features, tongue geometry features, tongue texture features.**

I. Introduction

WORLD Health Organization (WHO) has estimated that in 2000 there were 171 million people worldwide with diabetes mellitus (DM), and the number will increase to 366 million by 2030 [1] making the disease among the leading causes of death, disabilities, and economic hardship in the world. Two main types of DM exist, Type 1 DM and Type 2 DM. People with Type 1 DM fail to produce insulin, and therefore require injections of it. Type 2 DM is the most common type and can be categorized by insulin resistance. Currently, there is no cure for Type 1 DM or Type 2 DM. However, Type 2 DM can be managed by eating well, exercising, and maintaining a healthy lifestyle.

A fasting plasma glucose (FPG) test is the standard method practiced by many medical professionals to diagnose DM. FPG test is performed after the patient has gone at least 12 h without food, and requires taking a sample of the patient's blood (by piercing their finger) in order to analyze its blood glucose levels. Even though this method is accurate, it can be considered invasive, and slightly painful (piercing process). Diabetic retinopathy (DR) is a microvascular complication of DM that is responsible for 4.8% of the 37 million cases of blindness in the world, estimated by WHO [1]. In its earliest stage known as nonproliferative diabetic retinopathy (NPDR), the disease if detected can be treated to prevent further progression and sight loss. Various imaging modalities such as red-free [2], [3], angiography [3], [4], and color [5]–[10] fundus imaging are used to examine the human retina in order to detect DR and subsequently NPDR. These methods are based on the detection of relevant features related to DR, including but not limited to hemorrhages, microaneurysms, various exudates, and retinal blood vessels. These imaging modalities themselves can be regarded as invasive, exposing the eye to bright flashes or having fluorescein injected into a vein in the case of angiography. Therefore, there is a need to develop a noninvasive yet accurate DM and NPDR detection method.

As a result, this paper deals with the aforementioned problems and proposes a noninvasive automated method to detect DM and NPDR by distinguishing Healthy/DM and NPDR/DM-sans NPDR (DM without NPDR) samples using an array of tongue features consisting of color, texture, and geometry. The human tongue contains numerous features that can used to diagnose disease [11]–[25], with color, texture, and geometry features being the most prominent [11]–[25]. Traditionally, medical practitioners would examine these features based on years of experience [11]–[25]. However, ambiguity and subjectivity are always associated with their diagnostic results. To remove these qualitative aspects, quantitative feature extraction and analysis from tongue images can be established. To the best of our knowledge, there is no other published work to detect DM or NPDR using tongue color, texture, and geometry features.

Manuscript received December 8, 2012; revised April 26, 2013 and July 31, 2013; accepted September 3, 2013. Date of publication September 18, 2013; date of current version January 16, 2014. This work was supported in part by the GRF fund from the HKSAR Government, the central fund from the Hong Kong Polytechnic University, and the NSFC Oversea fund (61020106004), China. *Asterisk indicates corresponding author.*

*B. Zhang is with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA and also with the Department of Computing, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong (e-mail: yibo1@ece.cmu.edu).

B. V. K. Vijay Kumar is with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: kumar@ece.cmu.edu).

D. Zhang is with the Department of Computing, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong (e-mail: csdzhang@comp.polyu.edu.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2013.2282625

Fig. 1. Tongue capture device.

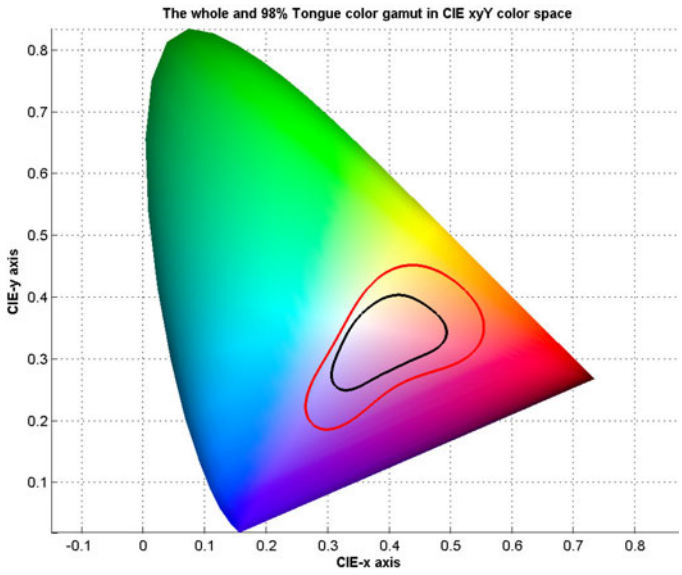

Fig. 2. Color gamut in the CIExyY color space depicting the tongue color gamut inside the red boundary. Furthermore, 98% of the tongue color gamut can be located within the black boundary.

Tongue images were captured using a especially designed in-house device taking into consideration color correction [26]. Each image was segmented [27] in order to locate its foreground pixels. With the relevant pixels located, three groups of features namely color, texture, and geometry were extracted from the tongue foreground. To conduct experimental results, a dataset consisting of 130 Healthy samples taken from Guangdong Provincial Hospital of Traditional Chinese Medicine, Guangdong, China, and 296 DM samples consisting of 267 DM-sans NPDR, and 29 NPDR processed from Hong Kong Foundation for Research and Development in Diabetes, Prince of Wales Hospital, Hong Kong SAR were used. Classification was performed between Healthy versus DM in addition to NPDR versus DM-sans NPDR initially using every feature individually (from the three groups), followed by an optimal combination of all the features.

The rest of this paper is organized as follows. Section II describes the tongue image capture device, color correction, and tongue segmentation, while Section III discusses tongue color feature extraction. In Section IV, tongue texture feature extraction is given in detail, with tongue geometry feature extraction presented in Section V. Section VI describes the experimental results and discussion, followed by concluding remarks in Section VII.

II. Capture Device and Tongue Image Preprocessing

The capture device, color correction of the tongue images, and tongue segmentation are given in this section. Fig. 1 shows the in-house designed device consisting of a three-chip CCD camera with 8 bit resolution, and two D65 fluorescent tubes placed symmetrically around the camera in order to produce a uniform illumination. The angle between the incident light and emergent light is 45°, recommended by Commission Internationale de l'Eclairage (CIE). During image capture, patients placed their chin on a chinrest while showing their tongue to the camera. The images captured in JPEG format that ranged from 257 × 189 pixels to 443 × 355 pixels were color corrected [26] to eliminate any variability in color images caused by changes of illumination and device dependence. This allows for consistent feature extraction and classification in the following steps. The idea of Wang and Zhang [26] (based on the Munsell ColorChecker) is to map a matrix generated from the input RGB vector to an objective sRGB vector, thereby obtaining a transformation model. Compared with the retinal imaging modalities mentioned previously, this capture device is noninvasive, neither requiring a bright flash nor injection of dye into a patient's blood stream.

Once the tongue images are captured, automatic segmentation [27] is next applied to each image in order to separate its foreground pixels from its background. This is accomplished by combining a bielliptical deformable template (BEDT), and an active contour model known as bielliptical deformable contour (BEDC). In [27], the segmented tongue is obtained by first minimizing the energy function of BEDT, followed by the energy function of BEDC. BEDT captures the overall tongue shape features, while BEDC can be deformed to match the local tongue details. The result is a binary tongue image clearly defining foreground pixels (tongue surface area and its edges) from its background pixels (area outside the tongue edges). This allows for three groups of features, color, texture, and geometry to be extracted from a tongue foreground image in the proceeding steps.

III. Tongue Color Features

The following section describes how color features are extracted from tongue images. The tongue color gamut is first summarized in Section III-A. In Section III-B, every foreground tongue pixel is compared to 12 colors representing the tongue color gamut and assigned its nearest color. This forms the color features.

A. Tongue Color Gamut

The tongue color gamut [28] represents all possible colors that appear on the tongue surface, and exists within the red

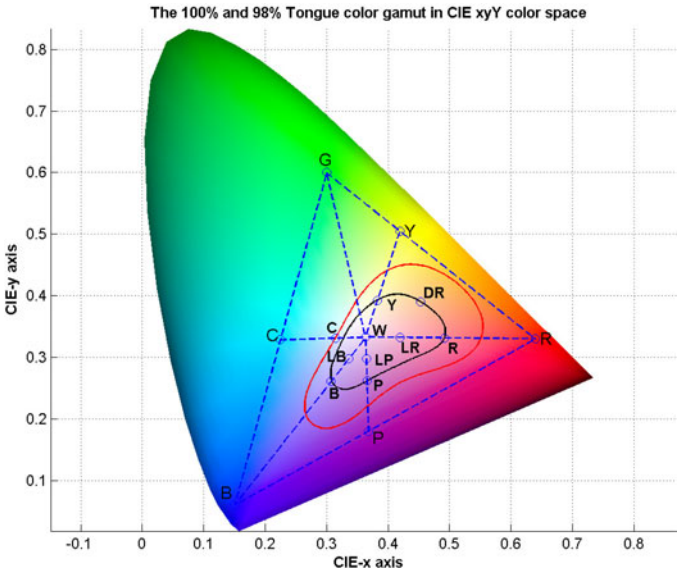

Fig. 3. Tongue color gamut can be represented using several points by drawing lines from the RGB color space.

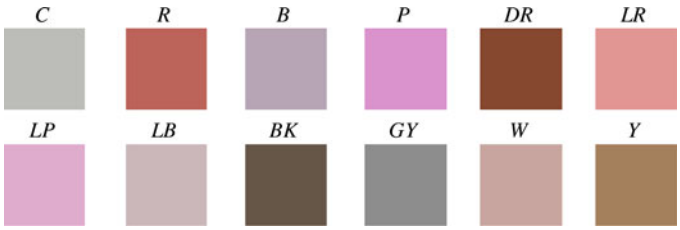

Fig. 4. Twelve colors representing the tongue color gamut with its label on top.

boundary shown in Fig. 2. It was created by plotting each tongue foreground pixel in our dataset onto the CIE 1931 chromaticity diagram (refer to Fig. 2), which shows all possible colors in the visible spectrum. Further investigation revealed that 98% of the tongue pixels lie inside the black boundary. To better represent the tongue color gamut, 12 colors plotted in Fig. 3 were selected with the help of the RGB color space. On the RG line, a point Y (Yellow) is marked. Between RB, a point P (Purple) is marked, and C (Cyan) is marked between GB. The center of the RGB color space is calculated and designated as W (White), the first of the 12 colors (see Fig. 3). Then, for each R (Red), B (Blue), Y, P, and C point, a straight line is drawn to W. Each time these lines intersect the tongue color gamut, a new color is added to represent the 12 colors. This accounts for R, Y, C, B, and P. *LR* (Light red), *LP* (Light purple), and *LB* (Light blue) are midpoints between lines from the black boundary to W, while *DR* (Deep red) is selected as no previous point occupies that area. More details about the tongue color gamut can be found in [28]. *GY* (Gray) and *BK* (Black) are not shown in Fig. 3 since both belong to grayscale.

The 12 colors representing the tongue color gamut are extracted from Fig. 3 and shown in Fig. 4 as a color square with its label on top. Correspondingly, its RGB and CIELAB values are given in Table I.

TABLE I
RGB AND CIELAB VALUES OF THE 12 COLORS

Color	[R G B]	[L A B]
C (Cyan)	[188 188 185]	[76.0693 -0.5580 1.3615]
R (Red)	[189 99 91]	[52.2540 34.8412 21.3002]
B (Blue)	[183 165 180]	[69.4695 9.5423 -5.4951]
P (Purple)	[226 142 214]	[69.4695 42.4732 -23.8880]
DR (Deep red)	[136 72 49]	[37.8424 24.5503 25.9396]
LR (Light red)	[227 150 147]	[69.4695 28.4947 13.3940]
LP (Light purple)	[225 173 207]	[76.0693 24.3246 -9.7749]
LB (Light blue)	[204 183 186]	[76.0693 7.8917 0.9885]
BK (Black)	[107 86 56]	[37.8424 3.9632 20.5874]
GY (Gray)	[163 146 143]	[61.6542 5.7160 3.7317]
W (White)	[200 167 160]	[70.9763 10.9843 8.2952]
Y (Yellow)	[166 129 93]	[56.3164 9.5539 24.4546]

B. Color Feature Extraction

For the foreground pixels of a tongue image, corresponding RGB values are first extracted, and converted to CIELAB [29] by transferring RBG to CIEXYZ using

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

followed by CIEXYZ to CIELAB via

$$\begin{cases} L^* = 166 \cdot f\left(Y/Y_0\right) - 16 \\ a^* = 500 \cdot \left[f\left(X/X_0\right) - f\left(Y/Y_0\right)\right] \\ b^* = 200 \cdot \left[f\left(Y/Y_0\right) - f\left(Z/Z_0\right)\right] \end{cases} \tag{2}$$

where $f(x) = x^{1/3}$ if $x > 0.008856$ or $f(x) = 7.787x + 16/116$ if $x \leq 0.008856$.

X_0, Y_0, and Z_0 in (2) are the CIEXYZ tristimulus values of the reference white point. The LAB values are then compared to 12 colors from the tongue color gamut (see Table I) and assigned the color which is closest to it (measured using Euclidean distance). After evaluating all tongue foreground pixels, the total of each color is summed and divided by the total number of pixels. This ratio of the 12 colors forms the tongue color feature vector v, where $v = [c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, c_9, c_{10}, c_{11}, c_{12}]$ and c_i represents the sequence of colors in Table I. As an example, the color features of three tongues are shown in visual form (refer to Figs. 5–7) along with its extracted color feature vector, where the original image is decomposed into one of the 12 colors. Fig. 5 shows a Healthy sample, Fig. 6 shows a DM sample, while an NPDR sample is given in Fig. 7. In these three samples, the majority of pixels are R.

The mean colors of Healthy, DM, and NPDR are displayed in Table II along with their standard deviation (std). DM tongues which have a higher ratio of *DR*, *LR*, and Y are greater in Healthy samples, and GY is higher in NPDR. The rest of the mean color features are similar. Only seven colors are listed out of the 12 as the remaining five have ratios less than 1%.

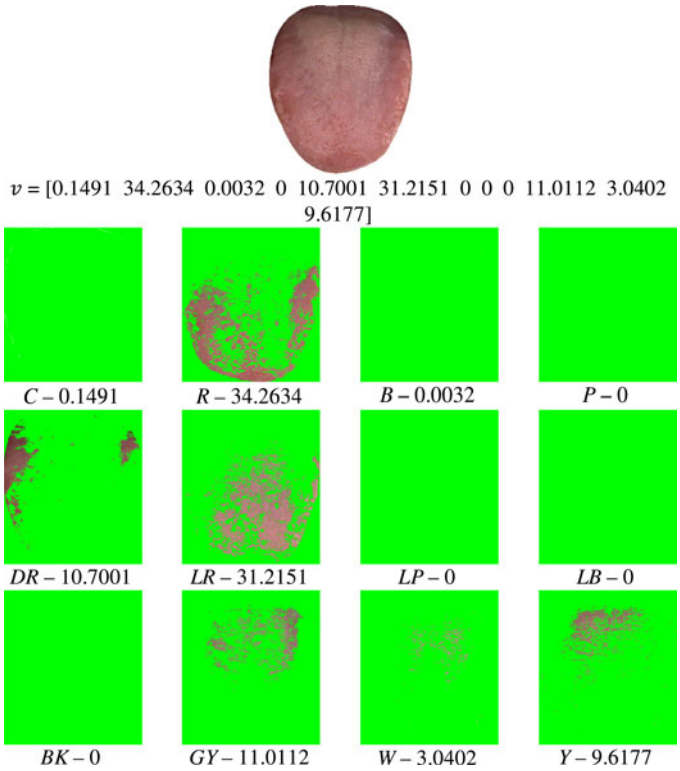

Fig. 5. Healthy tongue sample, its tongue color feature vector and corresponding 12 color makeup with most of the pixels classified as R.

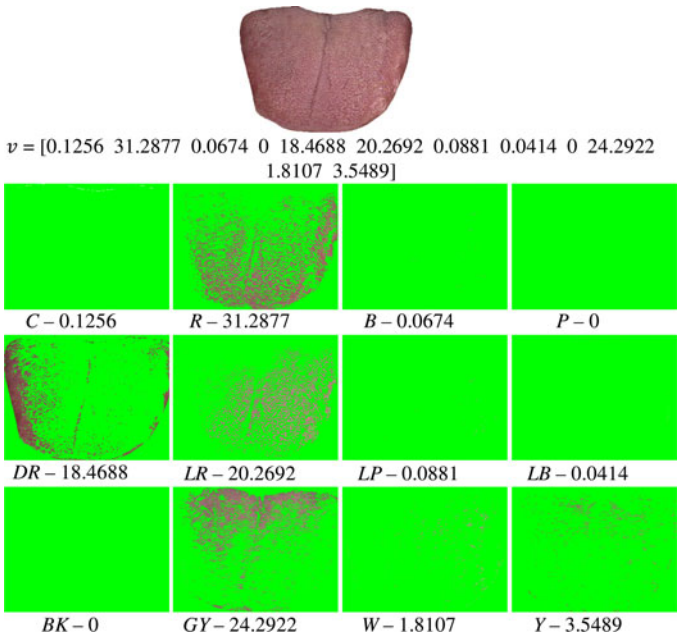

Fig. 6. DM tongue sample, its tongue color feature vector and corresponding 12 color makeup with most of the pixels classified as R.

IV. Tongue Texture Features

Texture feature extraction from tongue images is presented in this section. To better represent the texture of tongue images, eight blocks (see Fig. 8) of size 64 × 64 strategically located on the tongue surface are used. A block size of 64 × 64 was chosen

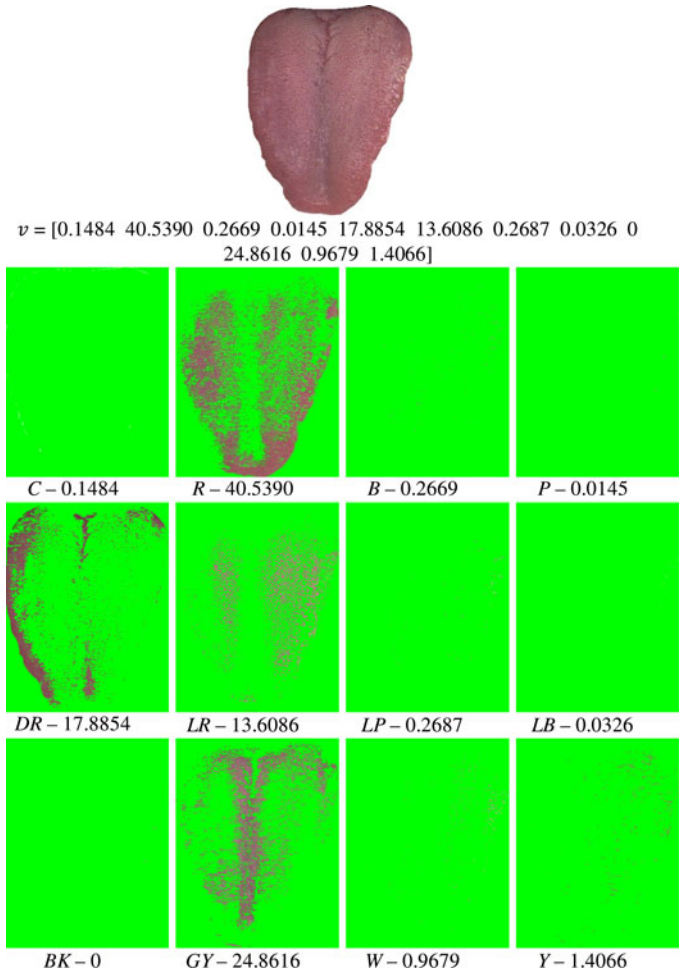

Fig. 7. NPDR tongue sample, its tongue color feature vector and corresponding 12 color makeup with most of the pixels classified as R.

TABLE II
Mean (std) of the Color Features for Healthy ($std_{\mathrm{Healthy}} = 11.71$), DM ($std_{\mathrm{DM}} = 12.50$), and NPDR ($std_{\mathrm{NPDR}} = 12.07$)

	C	*R*	*DR*	*LR*	*GY*	*W*	*Y*
Healthy	23.77 (3.76)	30.94 (10.73)	12.71 (8.11)	14.12 (9.36)	10.54 (10.80)	1.3 (1.84)	6.53 (6.35)
DM	24.80 (4.84)	30.70 (11.47)	18.40 (12.68)	10.53 (8.73)	10.80 (12.03)	1.07 (1.79)	3.55 (5.62)
NPDR	24.14 (4.86)	28.54 (13.13)	14.31 (10.38)	11.12 (7.74)	15.50 (13.92)	1.73 (2.16)	4.48 (6.82)

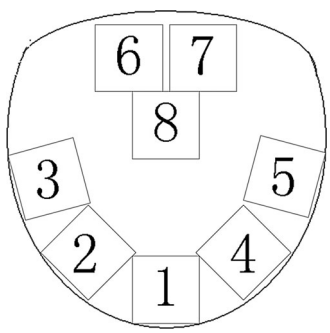

Fig. 8. Location of the eight texture blocks on the tongue.

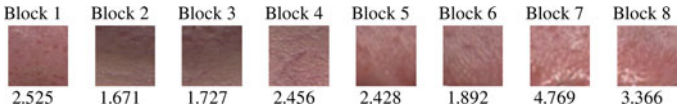

Fig. 9. Healthy texture blocks with its texture value below.

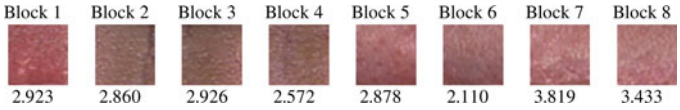

Fig. 10. DM texture blocks with its texture value below.

due to the fact that it covers all eight surface areas very well, while achieving minimum overlap. Larger blocks would cover areas outside the tongue boundary, and overlap more with other blocks. Smaller block sizes would prevent overlapping, but not cover the eight areas as efficiently. The blocks are calculated automatically by first locating the center of the tongue using a segmented binary tongue foreground image. Following this, the edges of the tongue are established and equal parts are measured from its center to position the eight blocks. Block 1 is located at the tip; Blocks 2 and 3, and Blocks 4 and 5 are on either side; Blocks 6 and 7 are at the root, and Block 8 is at the center.

The Gabor filter is a linear filter used in image processing, and is commonly used in texture representation. To compute the texture value of each block, the 2-D Gabor filter is applied and defined as

$$G_k(x,y) = \exp\left(\frac{x'^2 + \gamma^2 \cdot y'^2}{-2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}\right) \tag{3}$$

where $x' = x \cdot \cos\theta + y \cdot \sin\theta, y' = -x \cdot \sin\theta + y \cdot \cos\theta, \sigma$ is the variance, λ is the wavelength, γ is the aspect ratio of the sinusoidal function, and θ is the orientation. A total of three σ (1, 2, and 3) and four θ (0°, 45°, 90°, and 135°) choices were investigated to achieve the best result. Each filter is convolved with a texture block to produce a response $R_k(x,y)$:

$$R_k(x,y) = G_k(x,y) * im(x,y) \tag{4}$$

where $im(x,y)$ is the texture block and $*$ represents 2-D convolution. Responses of a block are combined to form FR_i, and its final response evaluated as follows:

$$FR_i(x,y) = \max\left(R_1(x,y), R_2(x,y), \ldots, R_n(x,y)\right) \tag{5}$$

which selects the maximum pixel intensities, and represents the texture of a block by averaging the pixel values of FR_i.

In the end, σ equal to 1 and 2 with three orientations (45°, 90°, and 135°) was chosen. This is due to the fact that the sum of all texture blocks between Healthy and DM had the largest absolute difference. Figs. 9–11 illustrate the texture blocks for Healthy, DM, and NPDR samples, respectively. Below each block, its corresponding texture value is provided.

Table III depicts the texture value mean for Healthy, DM, and NPDR together with their standard deviation. Healthy samples have a higher texture value in Block 7, whereas NPDR texture

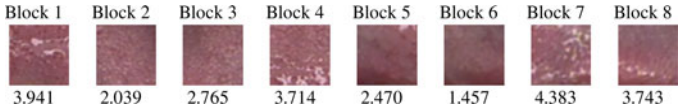

Fig. 11. NPDR texture blocks with its texture value below.

values are greater for the remaining blocks. The mean of all eight blocks is also included as an additional texture value. This brings the total number of texture features extracted from tongue images to be 9.

V. Tongue Geometry Features

In the following section, we describe 13 geometry features extracted from tongue images. These features are based on measurements, distances, areas, and their ratios.

Width: The width w feature (see Fig. 12) is measured as the horizontal distance along the x-axis from a tongue's furthest right edge point $(x_{\max})$ to its furthest left edge point $(x_{\min})$:

$$w = x_{\max} - x_{\min}. \tag{6}$$

Length: The length l feature (see Fig. 12) is measured as the vertical distance along the y-axis from a tongue's furthest bottom edge $(y_{\max})$ point to its furthest top edge point $(y_{\min})$:

$$l = y_{\max} - y_{\min}. \tag{7}$$

Length–width ratio: The length–width ratio lw is the ratio of a tongue's length to its width

$$lw = l/w. \tag{8}$$

Smaller half-distance: Smaller half-distance z is the half distance of l or w depending on which segment is shorter (see Fig. 12)

$$z = \min(l,w)/2. \tag{9}$$

Center distance: The center distance (cd) (refer to Fig. 13) is distance from $w's$ y-axis center point to the center point of l (y_{cp})

$$cd = \frac{\left(\max\left(y_{x_{\max}}\right) + \max\left(y_{x_{\min}}\right)\right)}{2} - y_{cp} \tag{10}$$

where $y_{cp} = (y_{\max} + y_{\min})/2$.

Center distance ratio: Center distance ratio (cdr) is ratio of cd to l:

$$cdr = \frac{cd}{l}. \tag{11}$$

Area: The area (a) of a tongue is defined as the number of tongue foreground pixels.

Circle area: Circle area (ca) is the area of a circle within the tongue foreground using smaller half-distance z, where $r = z$ (refer to Fig. 14):

$$ca = \pi r^2. \tag{12}$$

TABLE III
MEAN (std) OF THE TEXTURE FEATURES FOR HEALTHY ($std_{\mathrm{Healthy}} = 1.160$), DM ($std_{\mathrm{DM}} = 1.238$), AND NPDR ($std_{\mathrm{NPDR}} = 1.196$)

	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Mean of Blocks 1-8
Healthy	3.111 (0.978)	1.660 (0.632)	1.861 (0.677)	2.423 (0.875)	2.733 (0.989)	1.870 (0.650)	3.893 (1.167)	3.538 (1.326)	2.636 (0.486)
DM	2.952 (1.039)	2.142 (0.790)	2.359 (1.051)	2.711 (1.015)	2.522 (1.111)	1.859 (0.807)	3.710 (1.371)	3.887 (1.509)	2.768 (0.562)
NPDR	3.221 (1.118)	2.341 (0.777)	2.630 (0.892)	3.392 (1.333)	2.977 (1.130)	2.255 (1.235)	3.686 (1.343)	3.902 (1.133)	3.050 (0.521)

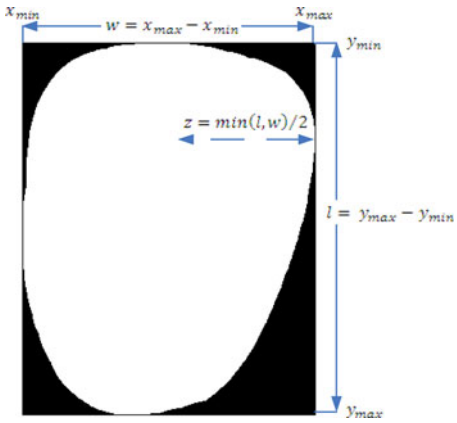

Fig. 12. Illustration of features 1, 2, and 4.

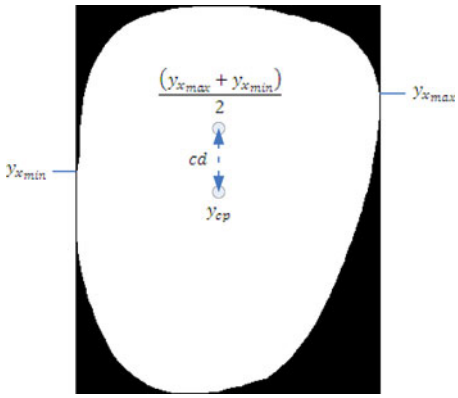

Fig. 13. Illustration of feature 5.

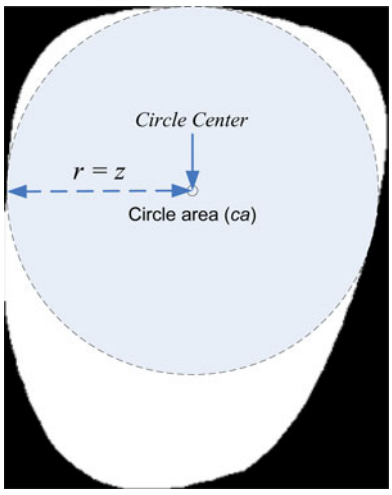

Fig. 14. Illustration of feature 8.

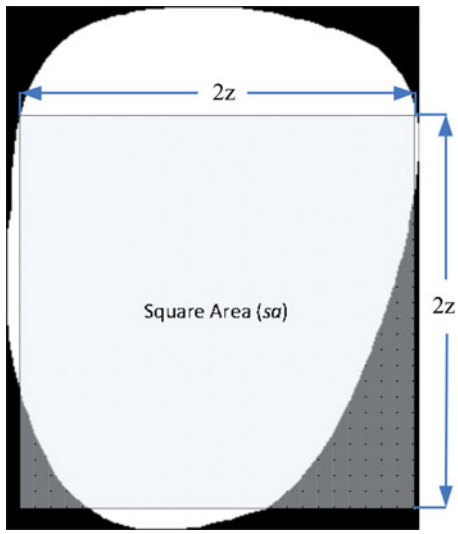

Fig. 15. Illustration of feature 10.

Circle area ratio: Circle area ratio (car) is the ratio of ca to a:

$$car = \frac{ca}{a}. \tag{13}$$

Square area: Square area (sa) is the area of a square defined within the tongue foreground using smaller half-distance z (refer to Fig. 15):

$$sa = 4z^2. \tag{14}$$

Square area ratio: Square area ratio (sar) is the ratio of sa to a:

$$sar = \frac{sa}{a}. \tag{15}$$

Triangle area: Triangle area (ta) is the area of a triangle defined within the tongue foreground (see Fig. 16). The right point of the triangle is $x_{\max}$, the left point is $x_{\min}$, and the bottom is $y_{\max}$.

Triangle area ratio: Triangle area ratio (tar) is the ratio of ta to a:

$$tar = \frac{ta}{a}. \tag{16}$$

The mean geometry features of Healthy, DM, and NPDR are shown in Table IV along with their standard deviation.

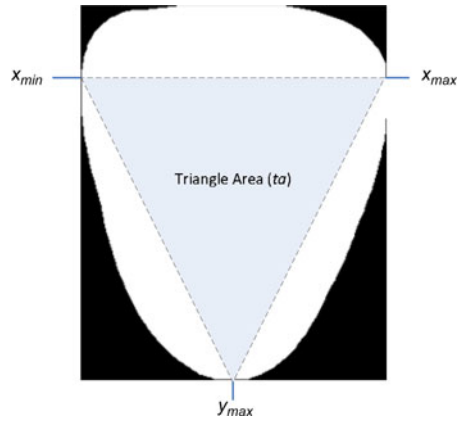

Fig. 16. Illustration of feature 12.

TABLE IV
MEAN (std) OF THE GEOMETRY FEATURES FOR HEALTHY ($std_{\mathrm{Healthy}} = 33\,013.78$, DM ($std_{\mathrm{DM}} = 33\,723.85$), AND NPDR ($std_{\mathrm{NPDR}} = 35\,673.93$)

	Healthy	DM	NPDR
w	320.8077 (36.5289)	335.9189 (42.7073)	344.3103 (39.4788)
l	302.6231 (43.4015)	295.2703 (62.2306)	309.8621 (64.9713)
lw	0.9527 (0.1638)	0.8897 (0.2067)	0.9117 (0.2202)
z	144.5385 (16.6250)	141.1875 (25.1676)	146.069 (23.6574)
cd	-49.6231 (29.2308)	-66.6926 (30.9031)	-64.6552 (34.3067)
cdr	-0.1631 (0.0871)	-0.2249 (0.0900)	-0.2097 (0.1029)
a	76709.14 (15525.3172)	76961.31 (21599.4127)	83286.67 (22629.9217)
ca	66493.77 (15079.8031)	64607.43 (21983.2771)	68727.14 (20900.1437)
car	0.8635 (0.0873)	0.8232 (0.1232)	0.8155 (0.1153)
sa	84662.52 (19200.1304)	82260.71 (27989.9621)	87506.07 (26610.9335)
sar	0.8908 (0.0703)	0.871689 (0.0848)	0.886897 (0.0920)
ta	32092.11 (7336.0657)	36077.43 (10624.3571)	37959.16 (9973.8946)
tar	0.4212 (0.0631)	0.4722 (0.0745)	0.4624 (0.0807)

TABLE V
CLASSIFICATION RESULT OF k-NN AND SVM USING EACH COLOR INDIVIDUALLY TO DISCRIMINATE HEALTHY VERSUS DM

Feature Number	Feature Name	Average Accuracy (%)
1	*C*	52.17
2	*R*	48.38
3	*B*	55.24
4	*P*	43.90
5	*DR*	60.16
6	*LR*	57.44
7	*LP*	54.30
8	*LB*	51.98
9	*BK*	50.00
10	*GY*	48.55
11	*W*	51.80
12	*Y*	64.78

TABLE VI
CLASSIFICATION RESULT OF k-NN AND SVM USING EACH TEXTURE BLOCK INDIVIDUALLY TO DISCRIMINATE HEALTHY VERSUS DM

Feature Number	Feature Name	Average Accuracy (%)
13	Block 1	53.71
14	Block 2	62.66
15	Block 3	61.30
16	Block 4	56.63
17	Block 5	56.50
18	Block 6	48.58
19	Block 7	54.59
20	Block 8	54.17
21	Mean of Blocks 1-8	51.15

VI. NUMERICAL RESULTS AND DISCUSSION

The ensuing section presents the numerical results. Healthy versus DM classification is first provided in Section VI-A. This is followed by NPDR versus DM-sans NPDR classification in Section VI-B.

A. Healthy Versus DM Classification

The numerical results were obtained on the tongue image database comprised of 426 images divided into 130 Healthy, and 296 DM (refer to Section I). Healthy samples were verified through a blood test and other examinations. If indicators from these tests fall within a certain range, they were deemed healthy. In the DM class, FPG test was used to diagnose diabetes.

Half of the images were randomly selected for training, while the other half was used as testing. This process was repeated five times. Classification was performed using k-nearest neighbor (k-NN) [30] (with $k = 1$) and a support vector machine (SVM) [31], where the kernel function (linear) mapped the training data into kernel space. To measure the performance, average accuracy was employed

$$\text{AverageAccuracy} = \frac{\text{sensitivity} + \text{specificity}}{2} \tag{17}$$

with the average of all five repetitions recorded as the final classification rate. In the first step, each individual feature (from the three groups) was applied to discriminate Healthy versus DM. This result can be seen in Tables V–VII. It should be noted that both k-NN and SVM achieved the same average accuracy for all 34 features. The highest average accuracy of 66.26% from this step was obtained using geometry feature ta (refer to Table VII).

In the next step, optimization by feature selection using sequential forward selection (SFS) was performed. SFS is a feature selection method that begins with an empty set of features. It adds additional features based on maximizing some criterion J, and terminates when all features have been added. In our case, J is the average accuracy of the classifier (k-NN and SVM). Tables VIII (k-NN) and IX (SVM) illustrate this result applied to each of the three main feature groups. From color features, the best combination is 3 and 12, which obtained an average accuracy of 68.76% using an SVM (see Table IX). In texture

TABLE VII
CLASSIFICATION RESULT OF k-NN AND SVM USING EACH GEOMETRY FEATURE INDIVIDUALLY TO DISCRIMINATE HEALTHY VERSUS DM

Feature Number	Feature Name	Average Accuracy (%)
22	w	60.81
23	l	50.25
24	lw	58.44
25	z	53.33
26	cd	60.29
27	cdr	61.61
28	a	43.76
29	ca	55.02
30	car	59.15
31	sa	55.02
32	sar	54.76
33	ta	66.26
34	tar	64.68

TABLE VIII
OPTIMIZATION OF HEALTHY VERSUS DM CLASSIFICATION USING FEATURE SELECTION WITH k-NN

Grouping	Feature(s) Number(s)	Feature(s) Name(s)	Average Accuracy (%)
Color	12	Y	64.78
Texture	14, 17, 16, 19	Blocks 2, 4, 5, 9	67.48
Geometry	22-30, 32-34	w-car, sar-tar	67.87
Best of Color, Texture, and Geometry	12, 14, 16, 17, 19, 22-30, 32-34	Y, Blocks 2, 4, 5, 9, w-car, sar-tar	67.87
All Features	1-30, 32-34	C-car, sar-tar	67.87

TABLE IX
OPTIMIZATION OF HEALTHY VERSUS DM CLASSIFICATION USING FEATURE SELECTION WITH THE SVM

Grouping	Feature Numbers	Feature Names	Average Accuracy (%)
Color	3, 12	B, Y	68.76
Texture	14-17	Blocks 2-5	67.67
Geometry	22, 30, 32-34	w, car, sar-tar	69.09
Best of Color, Texture, and Geometry	3, 12, 14, 15, 17, 22, 30, 32-34	B, Y, Blocks 2, 3, 5, w, car, sar-tar	77.39
All Features	3, 5, 12, 15, 22, 27, 30, 33, 34	B, DR, Y, Block 3, w, cdr, car, ta, tar	80.52

features, 14, 15, 16, and 17 attained an average accuracy of 67.67%, again using the SVM. With geometry features, 22, 30, 32, 33, and 34 distinguished Healthy versus DM with an average accuracy of 69.09% (in Table IX). Combining the features in these three groups by applying SFS, an average accuracy of 77.39% was achieved in Table IX using 3, 12, 14, 15, 17, 22, 30, 32, 33, and 34. Finally, by examining the best combination from all features (SFS), the highest average accuracy of 80.52% can be accomplished (via SVM), with a sensitivity of 90.77% and a specificity of 70.27%. Receiver operating characteristic (ROC) analysis was also performed on this classification as shown by the red ROC curve in Fig. 17. The average accuracy of this result is higher than the optimal combination from the three feature groups (77.39%), and contains fewer features. At the same time, it significantly improves upon the use of all features without feature selection, which obtained an average accuracy of 58.06% (k-NN) and 44.68% (SVM). The mean of features, 3, 5, 12, 15, 22, 27, 30, 33, and 34 from the best overall grouping for Healthy and DM, is shown in Table X, while Fig. 18 depicts three typical samples from Healthy and DM.

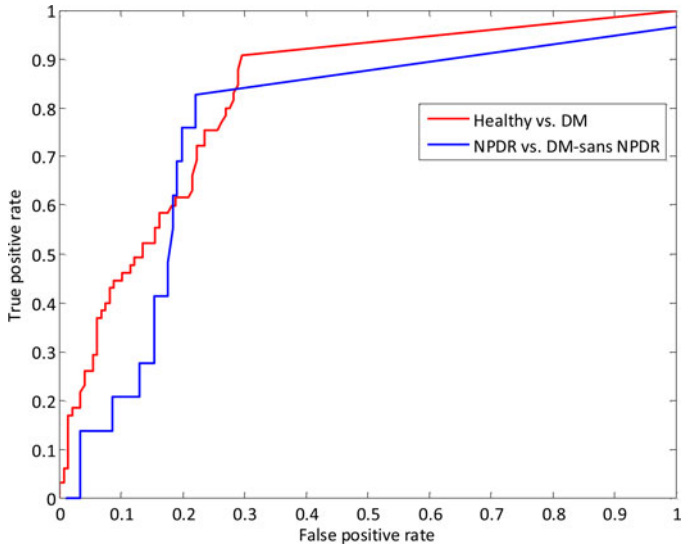

Fig. 17. ROC curves for Healthy versus DM (red), and NPDR versus DM-sans NPDR (blue).

B. NPDR Versus DM-Sans NPDR Classification

Of the 296 DM samples, 29 were marked as NPDR (refer to Section I). The NPDR samples were verified by medical professionals after examining the retina of patients. Using the same experimental setup as in Section VI-A, the results of NPDR versus DM-sans NPDR (267 samples) classification are illustrated in Tables XI–Tables XIV. Since it was established in the previous section that the SVM outperforms k-NN, only the former classifier was used. The average accuracies of applying each feature individually from the three groups are shown in Tables XI–XIII, and Table XIV displays the optimized result using SFS. As can be seen in the last row of Table XIV, the best result of 80.33% was achieved with features 7, 10, 11, 14, and 25 (sensitivity—82.76% and specificity—77.90% based on its blue ROC curve in Fig. 17). This compares to 59.05% and 53.35% average accuracies for k-NN and SVM, respectively, using all features without feature selection. Mean of the five optimal features for DM-sans NPDR and NPDR can be found in Table XV along with its three typical samples (refer to Fig. 19).

For completeness, NPDR versus Healthy classification was also conducted. An average accuracy of 87.14% was

TABLE X
MEAN OF THE OPTIMAL TONGUE FEATURES FOR HEALTHY AND DM

	3	5	12	15	22	27	30	33	34
Healthy	29.053	0.000831	1.519	1.613	2.628	-49.108	68184.26	0.892	32962.16
DM	29.600	0.000804	1.256	2.168	2.770	-70.392	66629.30	0.866	37344.22

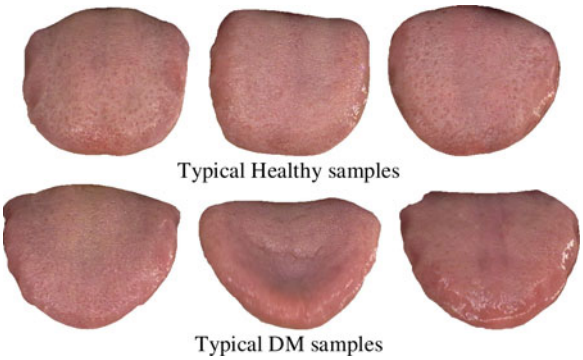

Fig. 18. Typical Healthy (top) and DM (bottom) tongue samples.

TABLE XI
CLASSIFICATION RESULTS OF USING EACH COLOR INDIVIDUALLY TO DISCRIMINATE NPDR VERSUS DM-SANS NPDR USING THE SVM

Feature Number	Feature Name	Average Accuracy (%)
1	*C*	54.26
2	*R*	40.09
3	*B*	57.24
4	*P*	50.29
5	*DR*	61.07
6	*LR*	52.98
7	*LP*	45.19
8	*LB*	55.07
9	*BK*	49.44
10	*GY*	70.81
11	*W*	64.14
12	*Y*	53.12

TABLE XII
CLASSIFICATION RESULTS OF USING EACH TEXTURE BLOCK INDIVIDUALLY TO DISCRIMINATE NPDR VERSUS DM-SANS NPDR USING THE SVM

Feature Number	Feature Name	Average Accuracy (%)
13	Block 1	50.36
14	Block 2	58.22
15	Block 3	55.22
16	Block 4	67.07
17	Block 5	61.45
18	Block 6	52.82
19	Block 7	48.41
20	Block 8	49.15
21	Mean of Blocks 1-8	61.00

accomplished using SFS with the SVM, achieving a sensitivity of 89.66% and a specificity of 84.62% via features 3, 9,15, 16, and 33.

TABLE XIII
CLASSIFICATION RESULTS OF USING EACH GEOMETRY FEATURE INDIVIDUALLY TO DISCRIMINATE NPDR VERSUS DM-SANS NPDR USING THE SVM

Feature Number	Feature Name	Average Accuracy (%)
22	*w*	30.12
23	*l*	52.98
24	*lw*	46.08
25	*z*	56.43
26	*cd*	43.91
27	*cdr*	48.48
28	*a*	62.57
29	*ca*	58.30
30	*car*	46.45
31	*sa*	58.30
32	*sar*	58.08
33	*ta*	32.07
34	*tar*	45.49

TABLE XIV
OPTIMIZATION OF NPDR VERSUS DM-SANS NPDR CLASSIFICATION USING FEATURE SELECTION WITH THE SVM

Grouping	Feature Numbers	Feature Names	Average Accuracy (%)
Color	1, 3, 5, 7-11	*C*, *B*, *DR*, *LP-W*	75.84
Texture	13-17, 21	Blocks 1-5, Mean of Blocks 1-8	72.09
Geometry	23, 28, 32	*l*, *a*, *sar*	65.27
Best of Color, Texture, and Geometry	3, 7, 10, 28	*B*, *LP*, *GY*, *a*	79.21
All Features	7, 10, 11, 14, 25	*LP*, *GY*, *W*, Block 2, *z*	80.33

TABLE XV
MEAN OF THE OPTIMAL TONGUE FEATURES FOR DM-SANS NPDR AND NPDR

	10	25	7	14	11
DM-sans NPDR	10.2884	140.6573	0.0647	2.1205	1.0025
NPDR	15.5041	146.0690	0.0565	2.3406	1.7322

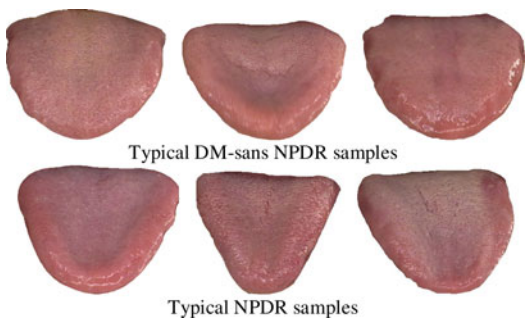

Fig. 19. Typical DM-sans NPDR (top) and NPDR (bottom) tongue samples.

VII. Conclusion

In this paper, a noninvasive approach to classify Healthy/DM and NPDR/DM-sans NPDR samples using three groups of features extracted from tongue images was proposed. These three groups include color, texture, and geometry. A tongue color gamut was first applied such that each tongue image can be represented by 12 colors. Afterward, eight blocks strategically located on the tongue were extracted and its texture value calculated. Finally, 13 geometry features from tongue images based on measurements, distances, areas, and their ratios were extracted. Numerical experiments were carried out using 130 Healthy and 296 DM tongue images. By applying each feature individually to separate Healthy/DM, the highest average accuracy achieved (via SVM) was only 66.26%. However, employing SFS with the SVM, nine features (with elements from all the three groups) were shown to produce the optimal result, obtaining an average accuracy of 80.52%. As for NPDR/DM-sans NPDR classification, the best result of 80.33% was attained using five features: three from color, one from texture, and one from geometry. This lays the groundwork for a potentially new way to detect DM, while providing a novel means to detect NPDR without retinal imaging or analysis.

Acknowledgment

The authors would like to thank Xingzheng Wang for his help in the data collection.

References

[1] *Prevention of Blindness From Diabetes Mellitus*, World Health Organization, Geneva, Switzerland, 2006.

[2] 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. 7, pp. 588–594, 2000.

[3] M. E. Martínez-Pérez, A. Hughes, S. Thom, A. Bharath, and K. Parker, "Segmentation of blood vessels from red-free and fluorescein retinal images," *Med. Image Anal.*, vol. 11, pp. 47–61, 2007.

[4] T. Spencer, J. A. Olson, K. C. McHardy, P. F. Sharp, and J. V. Forrester, "An image-processing strategy for the segmentation of microaneurysms in fluorescein angiograms of the ocular fundus," *Comput. Biomed. Res.*, vol. 29, pp. 284–302, 1996.

[5] M. Niemeijer, B. van Ginneken, J. Staal, M. S. A. Suttorp-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] M. Niemeijer, B. van Ginneken, M. J. Cree, A. Mizutani, G. Quellec, C. I. Sanchez, B. Zhang, R. Hornero, M. Lamard, C. Muramatsu, X. Wu, G. Cazuguel, J. You, A. Mayo, Q. Li, Y. Hatanaka, B. Cochener, C. Roux, F. Karray, M. Garcia, H. Fujita, and M. D. Abramoff, "Retinopathy online challenge: Automatic detection of microaneurysms in digital color fundus photographs," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 185–195, Jan. 2010.

[7] J. J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, Apr. 2004.

[8] B. Zhang, F. Karray, Q. Li, and L. Zhang, "Sparse representation classifier for microaneurysm detection and retinal blood vessel extraction," *Inf. Sci.*, vol. 200, pp. 78–90, 2012.

[9] B. Zhang, X. Wu, J. You, Q. Li, and F. Karray, "Detection of microaneurysms using multi-scale correlation coefficients," *Pattern Recog.*, vol. 43, pp. 2237–2248, 2010.

[10] B. Zhang, L. Zhang, L. Zhang, and F. Karray, "Retinal vessel extraction by matched filter with first-order derivative of Gaussian," *Comput. Biol. Med.*, vol. 40, pp. 438–445, 2010.

[11] C. C. Chiu, "The development of a computerized tongue diagnosis system," *Biomed. Eng.: Appl., Basis Commun.*, vol. 8, pp. 342–350, 1996.

[12] C. C. Chiu, "A novel approach based on computerized image analysis for traditional Chinese medical diagnosis of the tongue," *Comput. Methods Programs Biomed.*, vol. 61, pp. 77–89, 2000.

[13] B. Kirschbaum, *Atlas of Chinese Tongue Diagnosis*. Seattle, WA, USA: Eastland Press, 2000.

[14] D. Zhang, B. Pang, N. Li, K. Wang, and H. Zhang, "Computerized diagnosis from tongue appearance using quantitative feature classification," *Amer. J. Chin. Med.*, vol. 33, pp. 859–866, 2005.

[15] Y. Zhang, R. Liang, Z. Wang, Y. Fan, and F. Li, "Analysis of the color characteristics of tongue digital images from 884 physical examination cases," *J. Beijing Univ. Traditional Chin. Med.*, vol. 28, pp. 73–75, 2005.

[16] D. Zhang, *Automated Biometrics: Technologies and Systems*. Seattle, WA, USA: Kluwer, 2000.

[17] W. Su, Z. Y. Xu, Z. Q. Wang, and J. T. Xu, "Objectified study on tongue images of patients with lung cancer of different syndromes," *Chin. J. Integrative Med.*, vol. 17, pp. 272–276, 2011.

[18] K. Q. Wang, D. Zhang, N. M. Li, and B. Pang, "Tongue diagnosis based on biometric pattern recognition technology," in *Pattern Recognition From Classical To Modern Approaches*, 1st ed. S. K. Pal and A. Pal,, Eds. Singapore: World Scientific, 2001, pp. 575–598.

[19] B. Huang, J. Wu, D. Zhang, and N. Li, "Tongue shape classification by geometric features," *Inf. Sci.*, vol. 180, pp. 312–324, 2010.

[20] B. Li, Q. Huang, Y. Lu, S. Chen, R. Liang, and Z. Wang, "A method of classifying tongue colors for traditional Chinese medicine diagnosis based on the CIELAB color space," in *Proc. Int. Conf. Med. Biometrics*, 2007, pp. 153–159.

[21] C. H. Li and P. C. Yuen, "Tongue image matching using color content," *Pattern Recog.*, vol. 35, pp. 407–419, 2002.

[22] N. M. Li, "The contemporary investigations of computerized tongue diagnosis," in *The Handbook of Chinese Tongue Diagnosis.* Peking, China: Shed-Yuan Publishing, 1994.

[23] G. Maciocia, *Tongue Diagnosis in Chinese Medicine*. Seattle, WA, USA: Eastland Press, 1995.

[24] N. Ohta and A. R. Robertson, *Colorimetry: Fundamentals and Applications*. Hoboken, NJ, USA: Wiley, 2006.

[25] B. Pang, D. Zhang, and K. Wang, "Tongue image analysis for appendicitis diagnosis," *Inf. Sci.*, vol. 17, pp. 160–176, 2005.

[26] X. Wang and D. Zhang, "An optimized tongue image color correction scheme," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 6, pp. 1355–1364, Nov. 2010.

[27] B. Pang, D. Zhang, and K. Q. Wang, "The bi-elliptical deformable contour and its application to automated tongue segmentation in Chinese medicine," *IEEE Trans. Med. Imag.*, vol. 24, no. 8, pp. 946–956, Aug. 2005.

[28] X. Wang and D. Zhang, "Statistical tongue color distribution and its application," in *Proc. Int. Conf. Comput. Comput. Intell.*, 2011, pp. 281–292.

[29] H. Z. Zhang, K. Q. Wang, X. S. Jin, and D. Zhang, "SVR based color calibration for tongue image," in *Proc. Int. Conf. Mach. Learning Cybern.*, 2005, pp. 5065–5070.

[30] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd ed. Hoboken, NJ, USA: Wiley, 2000.

[31] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learning*, vol. 20, pp. 273–297, 1995.

Bob Zhang (M'12) received the M.A.Sc. degree in information systems security from Concordia University, Montreal, QC, Canada, in 2007, and subsequently the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2011.

He is currently a Postdoctoral Researcher in the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA. His research interests include pattern recognition, machine learning, and medical biometrics.

B. V. K. Vijay Kumar (F'10) received the B.Tech. and M.Tech. degrees in electrical engineering from the Indian Institute of Technology, Kanpur, Kanpur, India, and the Ph.D. degree in electrical engineering from Carnegie Mellon University (CMU), Pittsburgh, PA, USA.

Since 1982, he has been a faculty member in the Department of Electrical and Computer Engineering (ECE), CMU, where he is currently a Professor and the Associate Dean for Graduate and Faculty Affairs in the College of Engineering. He served as the Associate Department Head in the ECE Department from 1994 to 1996 and as the Acting Head in the Department during 2004–2005. His publications include the book entitled *Correlation Pattern Recognition* (coauthored with Dr. Abhijit Mahalanobis and Dr. Richard Juday, Cambridge, U.K.: Cambridge University Press, November 2005), twenty book chapters, 370 conference papers, and 180 journal papers. He is also a coinventor of eight patents. His research interests include computer vision and pattern recognition algorithms and applications and coding and signal processing for data storage systems.

Dr. Bhagavatula served as a Pattern Recognition Topical Editor for the Information Processing division of *Applied Optics* and as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY. He has served on many conference program committees and was a Co-General Chair of the 2005 IEEE AutoID Workshop, a Co-Chair of the 2008–2010 SPIE conferences on Biometric Technology for Human Identification and a Co-Program Chair of the 2012 IEEE Biometrics: Theory, Applications and Systems (BTAS) conference. He is a Fellow of SPIE, a Fellow of Optical Society of America, and a Fellow of the International Association of Pattern Recognition. In 2003, he received the Eta Kappa Nu Award for *Excellence in Teaching* in the ECE Department at CMU and the Carnegie Institute of Technology's *Dowd Fellowship* for educational contributions and he was also a corecipient of the 2008 Outstanding Faculty Research Award in CMU's College of Engineering.

David Zhang (M'89–SM'95–F'08) received the Graduation degree in computer science from Peking University, Beijing, China, the M.Sc. degree in computer science in 1982, and the Ph.D. degree in 1985 from the Harbin Institute of Technology (HIT), Harbin, China. In 1994, he received the second Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, Ontario, Canada.

From 1986 to 1988, he was a Postdoctoral Fellow with Tsinghua University and then an Associate Professor at the Academia Sinica, Beijing. He is currently a Chair Professor at the Hong Kong Polytechnic University, Kowloon, Hong Kong, where he is the Founding Director of the Biometrics Technology Centre (UGC/CRC) supported by the Hong Kong SAR Government in 1998. He also serves as a Visiting Chair Professor at Tsinghua University, Beijing, and an Adjunct Professor at Peking University, Shanghai Jiao Tong University, Shanghai, China, HIT, and the University of Waterloo. He is the author of more than 10 books and 250 journal papers.

Dr. Zhang is the Founder and Editor-in-Chief, *International Journal of Image and Graphics* (IJIG); Book Editor, Springer International Series on Biometrics (KISB); Organizer, the International Conference on Biometrics Authentication (ICBA); and Associate Editor of more than ten international journals including IEEE TRANSACTIONS AND PATTERN RECOGNITION. He is a Croucher Senior Research Fellow, Distinguished Speaker of the IEEE Computer Society, and a Fellow of International Association of Pattern Recognition.