11

Quality Assessment of Retinal Fundus Images using Elliptical Local Vessel Density

Luca Giancardo[1,2], Fabrice Meriaudeau[1], Thomas P Karnowski[2], Dr Edward Chaum[3] and Kenneth Tobin[2]
[1]*Université de Bourgogne*
France
[2]*Oak Ridge National Laboratory*
USA
[3]*University of Tennessee - Hamilton Eye Institute*
USA

1. Introduction

Diabetic retinopathy is the leading cause of blindness in the Western world. The World Health Organisation estimates that 135 million people have diabetes mellitus worldwide and that the number of people with diabetes will increase to 300 million by the year 2025 (Amos et al., 1997). Timely detection and treatment for DR prevents severe visual loss in more than 50% of the patients (ETDRS, 1991). Through computer simulations is possible to demonstrate that prevention and treatment are relatively inexpensive if compared to the health care and rehabilitation costs incurred by visual loss or blindness (Javitt et al., 1994).
The shortage of ophthalmologists and the continuous increase of the diabetic population limits the screening capability for effective timing of sight-saving treatment of typical manual methods. Therefore, an automatic or semi-automatic system able to detect various type of retinopathy is a vital necessity to save many sight-years in the population. According to Luzio et al. (2004) the preferred way to detect diseases such as diabetic retinopathy is digital fundus camera imaging. This allows the image to be enhanced, stored and retrieved more easily than film. In addition, images may be transferred electronically to other sites where a retinal specialist or an automated system can detect or diagnose disease while the patient remains at a remote location.
Various systems for automatic or semi-automatic detection of retinopathy with fundus images have been developed. The results obtained are promising but the initial image quality is a limiting factor (Patton et al., 2006); this is especially true if the machine operator is not a trained photographer. Algorithms to correct the illumination or increase the vessel contrast exist (Chen & Tian, 2008; Foracchia et al., 2005; Grisan et al., 2006; Wang et al., 2001), however they cannot restore an image beyond a certain level of quality degradation. On the other hand, an accurate quality assessment algorithm can allow operators to avoid poor images by simply re-taking the fundus image, eliminating the need for correction algorithms. In addition, a quality metric would permit the automatic submission of only the best images if many are available.

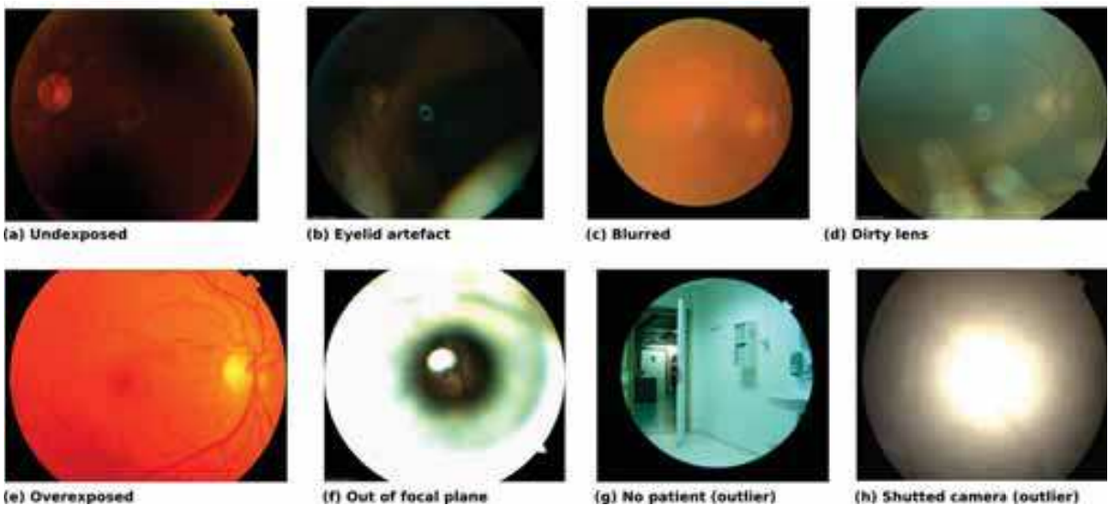

Fig. 1. Examples of Poor Quality Fundus Images (images extracted from datasets used in this study, see Section 4.1).

The measurement of a precise image quality index is not a straightforward task, mainly because quality is a subjective concept which varies even between experts, especially for images that are in the middle of the quality scale. In addition, image quality is dependent upon the type of diagnosis being made. For example, an image with dark regions might be considered of good quality for detecting glaucoma but of bad quality for detecting diabetic retinopathy. For this reason, we decided to define quality as the *"characteristics of an image that allow the retinopathy diagnosis by a human or software expert"*.

Fig. 1 shows some examples of macula centred fundus images whose quality is very likely to be judged as poor by many ophthalmologists. The reasons for this vary. They can be related to the camera settings like exposure or focal plane error (Fig. 1.(a,e,f)), the camera condition like a dirty or shuttered lens (Fig. 1.(d,h)), the movements of the patient which might blur the image (Fig. 1.(c)) or if the patient is not in the field of view of the camera (Fig. 1.(g)). We define an outlier as any image that is not a retina image which could be submitted to the screening system by mistake.

Existing algorithms to estimate the image quality are based on the length of visible vessels in the macula region (Fleming et al., 2006), or edges and luminosity with respect to a reference image (Lalonde et al., 2001; Lee & Wang, 1999). Another method uses an unsupervised classifier that employs multi-scale filterbanks responses (Niemeijer et al., 2006). The shortcomings of these methods are either the fact that they do not take into account the natural variance encountered in retinal images or that they require a considerable time to produce a result.

Additionally, none of the algorithms in the literature that we surveyed generate a "quality measure". Authors tend to split the quality levels into distinct classes and to classify images in particular ones. This approach is not really flexible and is error prone. In fact human experts are likely to disagree if many categories of image quality are used. Therefore, we think that a normalised "quality measure" from 0 to 1 is the ideal way to approach the classification problem.

Processing speed is another aspect to be taken into consideration. While algorithms to assess the disease state of the retina do not need to be particularly fast (within reason), the time

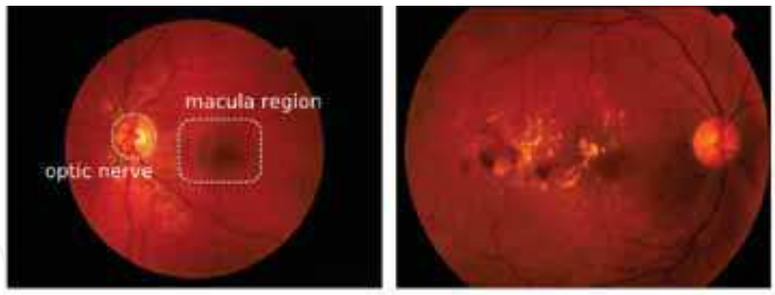

Fig. 2. Comparison of fundus images of an healthy and an unhealthy patient (images extracted from our datasets, see Section 4.1).

response of the quality evaluation method is key towards the development of an automatic retinopathy screening system.

This chapter is structured as follows. The rest of the introduction gives an brief overview of the anatomy of the retina and about diabetic retinopathy, which is useful to fully comprehend the algorithms that will be presented. Section 2 is a survey of existing techniques to evaluate the quality of retina fundus images. In Section 3 we introduce a quality assessment technique based on a new set of features called ELVD. Section 4 describes the tests and results obtained. Section 5 concludes the chapter.

1.1 Anatomy of the Retina

The retina is a multi-layered sensory tissue that lies on the back of the eye. It contains millions of photoreceptors that capture light rays and convert them into electrical impulses. These impulses travel along the optic nerve to the brain where they are converted into images. Many retinal blood vessels supply oxygen and nutrients to the inner and outer layers of the retina. The former are visible, the latter are not since they are situated in the choroid (the back layer of the retina) (Cassin & Solomon, 1990).

There are two types of photoreceptors in the retina: rods and cones, named after their shape. Rod cells are very sensitive to changes in contrast even at low light levels, hence able to detect movement, but they are imprecise and insensitive to colour. They are generally located in the periphery of the retina and used for scotopic vision (night vision). Cones, on the other hand, are high precision cells capable of detecting colours. They are mainly concentrated in the macula, the area responsible for photopic vision (day vision). The very central portion of the macula is called the fovea, which is where the human eye is able to distinguish visual details at its best. While loss of peripheral vision may go unnoticed for some time, damage to the macula will result in loss of central vision, which has serious effects on the visual perception of the external world (Wyszecki & Stiles, 2000).

All the photoreceptors are connected to the brain through a dense network of roughly 1.2 million of nerves (Jonas et al., 1992). These leave the eye in a unique bundle, the optic nerve. Fig. 2.(a) shows where the macula and fovea areas are located.

1.2 Diabetic Retinopathy

Diabetes mellitus (DM) is a chronic, systemic, life-threatening disease characterised by disordered metabolism and abnormally high blood sugar (hyperglycaemia) resulting from low levels of the hormone insulin with or without abnormal resistance to insulin's effects (Tierney

et al., 2002). DM has many complications that can affect the eyes and nervous system, as well as the heart, kidneys and other organs. Diabetic retinopathy (DR) is a vascular complication of DM which causes damages to the retina which leads to serious vision loss if not treated promptly. People with diabetes are 25 times more likely to develop blindness than individuals without diabetes. For any type of diabetes, the prevalence of diabetic retinopathy in people more than 40 years of age was reported to be 40.3% (Baker et al., 2008).
The National Eye Institute divides diabetic retinopathy in four subsequent stages:

- *Mild Nonproliferative Retinopathy:* At this earliest stage, microaneurysms occur. They are small areas of balloon-like swelling in the retina's tiny blood vessels.
- *Moderate Nonproliferative Retinopathy:* As the disease progresses, some blood vessels that nourish the retina are blocked. Lesions like exudates (fat deposits) and haemorrhages start to appear.
- *Severe Nonproliferative Retinopathy:* Many more blood vessels are blocked, depriving several areas of the retina with their blood supply. These areas of the retina send signals to the body to grow new blood vessels for nourishment.
- *Proliferative Retinopathy (PDR):* At this advanced stage, the signals sent by the retina for nourishment trigger the growth of new blood vessels. These new blood vessels are abnormal and fragile. They grow along the retina and along the surface of the clear, vitreous gel that fills the inside of the eye. By themselves, these blood vessels do not cause symptoms or vision loss. However, they have thin, fragile walls. If they leak blood, vision loss and even blindness can result.

2. State of the Art of Fundus Images Quality Assessment

Computerised evaluation of image quality is a problem not only in the field of medical imaging but in many other image processing systems, such as image acquisition, compression, restoration and enhancement. Over the years, a number of researchers have developed general purpose algorithms to objectively assess the image quality with a good consistency with human judgements, regardless the type, scale or distortion of the image (Sheikh et al., 2006). In this section we present only techniques that are designed specifically for retinal fundus images. These methods attempt to simulate the judgement of an expert ophthalmologist rather than a generic human vision system. The methods are grouped in three different categories depending on the technique used: Histogram Based Methods, Retina Morphology Methods and "Bag-of-Words" Methods.

2.1 Histogram Based Methods

Besides some reference to Quality Assessment (QA) in research reports in the OPHTEL EU project (Mann, 1997), the first authors that have explicitly addressed the problem of automatic detection of fundus image quality are Lee and Wang (Lee & Wang, 1999). Their approach starts from a pure signal processing perspective with the aim of providing a quantitative measure to compare and evaluate retinal images enhancement methods in further studies. They used 20 images with excellent quality extracted from a set of 360. These reference images are used to compute an ideal template intensity histogram discarding any information about the colour. The template histogram is adjusted in order to approximate a Gaussian distribution as follows:

$$f(i) = A \cdot \exp\left(\frac{-(i-M)^2}{2\sigma^2}\right) \tag{1}$$

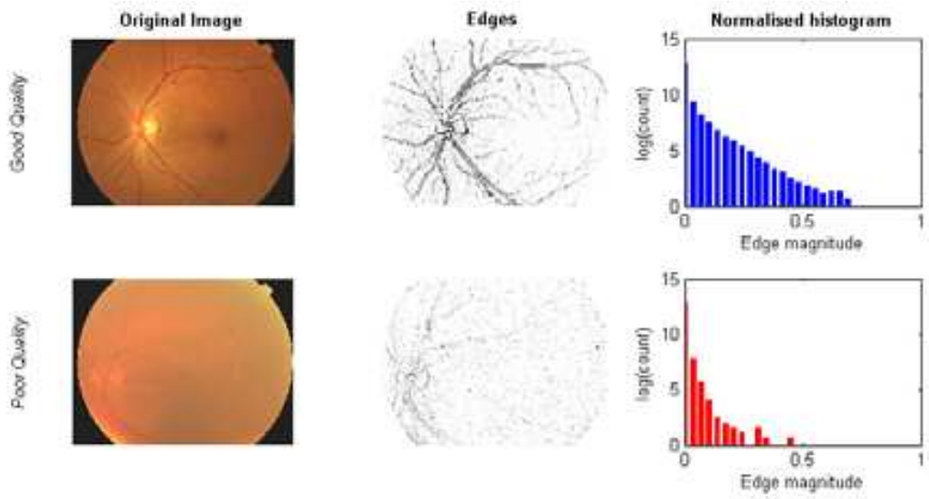

Fig. 3. Comparison of edges between a good and poor quality image

where i (from 0 to 255) is the pixel intensity, A is the peak value of the Gaussian curve, M and σ are respectively the mean and standard deviation of all the training histograms. In their tests, the authors estimated that $\sigma = R/6$ where R is the histogram spread. The quality of a target image is assessed by convolving its histogram with the template histogram and by computing a quality index Q. The index Q is normalised between 0 and 1 by employing the self correlation of the template histogram as the maximum value possible.

The key discriminating features in this method are the image contrast (i.e. the histogram spread), brightness and signal-to-noise ratio (SNR). New subsequent publications challenged the idea that pure histogram similarity is correlated with image quality. For example, Lalonde et al. (2001) found poor quality images whose histogram resembled the template histogram and also good quality images with markedly different histograms. Therefore, they tried to extend the approach of Lee and Wang maintaining the idea that a model of a good image is defined using a set of images of excellent quality but using two different sets of features: the distribution of the edge magnitudes in the image and the local distribution of the pixel intensity, as opposed to the global histogram of Lee and Wang. Their notion of quality differs from the one of Lee and Wang. Rather than viewing it from a pure signal processing perspective where quality is correlated with noise, they are closer to the medical field needs, whose concept of quality depends on the experts' ability to diagnose retinopathy.

Lalonde et al. notice that the edge magnitude histogram in a ophthalmic image has a shape that is similar to a Rayleigh distribution. In Fig. 3 the edge distributions are compared. The authors found that the edge distribution of poor images fall more rapidly than good images (notice that in the figure the histogram is plotted on a logarithmic scale). They evaluate the difference between two edges magnitude histogram using an equation similar to the χ^2 statistic:

$$d_{edge}(T,R) = \sum_i \frac{(R_i - T_i)^2}{R_i + T_i}, \quad \forall i | R_i + T_i \neq 0 \tag{2}$$

where R is the reference histogram and T is the edge histogram of the target image.

The second set of features used is a localised version of the global histogram of Lee and Wang. They retrieve the global histogram and segment it into uniform region by the standard

histogram-splitting algorithm from Ohlander et al. (1978). Regions below a certain threshold are discarded. The dissimilarity W between reference and target image is calculated as follows:

$$W(h_1, h_2) = \left[\frac{\mu_{h_1} - \mu_{h_2}}{min(\mu_{h_1} \mu_{h_2})} \right] \quad (3)$$

It should be noticed that only the mean of the histogram is used in the equation; all the other information is discarded.
Finally, they classified the target images into three classes by using on the similarity measures given from Eq. 2 and 3. Using a dataset of 40 images they obtained 77% images classified correctly.

2.2 Retina Morphology Methods

Usher et al. (2003) were the first authors to consider features unique to retina images for QA. They noticed a correlation between image blurring and visibility of the vessels. By running a vessel segmentation algorithm and measuring the area of detected vessels over the entire image, the authors estimated if the quality of the image was sufficient for screening since images that are out of focus or blurred will not have visible smaller vessels. The classification between good and poor is performed by means of a threshold value. The authors employed a dataset of 1746 images taken from a retinopathy screening program obtaining a sensitivity of 84.3% and a specificity of 95.0%.
Fleming et al. (2006) found a problem in the previous approach: even if some vessels are undetectable in cases of image distortions, large vessels can remain visible, especially in the main arcades coming out from the optic nerve. These vessels have a substantial area which can easily be greater than the classifier threshold.
Consequently, they developed a method based on the image grading system used in the Grampian Diabetes Retinal Screening Programme in Aberdeen, Scotland. The QA score is divided into two aspects: image clarity and field definition. Image clarity is graded as follows:

- *Excellent:* Small vessels are clearly visible and sharp within one optic disc diameter around the macula. The nerve fibre layer is visible.
- *Good:* Either small vessels are clearly visible but not sharp within one optic disc diameter around the macula or the nerve fibre layer is not visible.
- *Fair:* Small vessels are not clearly visible within one optic disc diameter around the macula but are of sufficient clarity to identify third-generation branches within one optic disc diameter around the macula.
- *Inadequate:* Third-generation branches within one optic disc diameter around the macula cannot be identified.

Field definition is graded as follows:

- *Excellent:* The entire macula and optic disc are visible. The macula is centred horizontally and vertically in the image.
- *Good:* The entire macula and optic disc are visible. The macula is not centred horizontally and vertically in the image, but both main temporal arcades are completely visible and the macula is complete.
- *Inadequate:* Either a small-pupil artefact is present, or at least one of the macula, optic disc, superior temporal arcade, or inferior temporal arcade is incomplete.

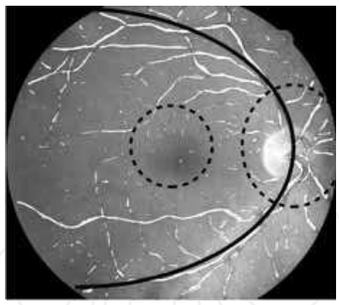

Fig. 4. Detected vessels (white) with the semiellipse fitted to the temporal arcades and the search regions for fovea and optic disk with the method described by Fleming et al. (2006).

First, a search is made for the arcade vessels. The generalised Hough transform (Ballard, 1981) is used to identify large-scale vessels between 10 and 30 pixels by employing semielliptical templates with different sizes, orientations and eccentricities. The process is quite computationally expensive. Hence, the image is subsampled by a factor of 32.
The authors estimate the average optic nerve diameter (OND) to be 246 pixels, based on a manual estimation of its the mean size in the dataset. The rightmost (or leftmost depending on the arcade template detected) point of the semiellipse fitted to the temporal arcades is used as a search centre for the optic disk. The search space is restricted to a region with height 2.4 x 2.0 times OND. Within this region a Hough transform is applied to detect the optic disk with a circular template.
The search area for the fovea is restricted to a circular region with diameter 1.6 OND centred on a point that is 2.4 OND from the optic disk and on a line between the detected optic disk and the centre of the temporal arcades. The fovea is actually found by identifying the maximum cross-correlation between the image and a predefined foveal model in the search area. Figure 4 shows the search region for the optic disk and fovea.
The image clarity was assessed taking into consideration the vessel area. However, instead of measuring it globally like Usher et al. (2003), only the area in the foveal region is used. The size of measured area is again relative to OND: a square of 3.5 OND if the foveal cross-correlation coefficient is large enough, otherwise a square sized 4.5 OND. The rationale for the choice of such area is the fact that in the foveal region there are the thinnest vessels, the ones that are more likely to disappear when the image is degraded.
The second aspect considered is the field definition. A fundus image with an adequate field definition has to satisfy the following constraints[1]:

- Distance between optic disk and the edge of the image < 0.5 OND
- Distance from the fovea to the edge of the image > 2 OND
- Angle between the fovea and the optic disk between 24.7° and -5.7°
- Length of the vessel arcades > 2.1 OND

The final classification of the overall quality is obtained by combining the two measures of image clarity and field definition. The authors reported a sensitivity and specificity respectively of 99.1% and 89.4% on a dataset of 1039 images. In this context, the sensitivity represents the "good quality" images correctly classified, while the specificity represents the correct classification on "poor quality" images.

[1] The measurement of all these constraints are possible thanks to the initial segmentation step.

2.3 "Bag of Words" Methods

Niemeijer et al. (2006) found various deficiencies in previous QA methods. They highlight that it is not possible to consider the natural variance encountered in retinal images by taking into account only a mean histogram of a limited set of features like Lalonde et al. (2001); Lee & Wang (1999). Niemeijer et al. acknowledge the good results of Fleming et al. (2006) but having to segment many retinal structures is seen as a shortcoming. In fact, detecting the segmentation failure in case of low quality is not trivial. Finally, they proposed a method that is comparable to the well known "Bag-of-Words" classification technique, used extensively in pattern recognition tasks in fields like image processing or text analysis (Fei-Fei & Perona, 2005; Sivic et al., 2005).

"Bag-of-Words" methods work as follows. First, a feature detector of some sort is employed to extract all the features from the complete training set. Because the raw features are too numerous to be used directly in the classification process, a clustering algorithm is run to express the features in a compact way. Each cluster is analogue to a "word" in a dictionary. In the dictionary, words do not have any relative information about the class they belong to or their relative location respect others. Instead, they are simply image characteristics that are often repeated throughout the classes, therefore they are likely to be good representatives in the classification process. Once the dictionary is built, the features of each sample are mapped to words and a histogram of word frequencies for each image is created. Then, these histograms are used to build a classifier and the learning phase ends. When a new image is presented to this type of system, its raw features are extracted and their word representation is searched in the dictionary. Then, the word frequency histogram is built and presented to the trained classifier which makes a decision on the nature of the image.

Niemeijer et al. employ two sets of feature to represent image quality: colour and second order image structure invariants (ISI). Colour is measured through the normalised histograms of the RGB planes, with 5 bins per plane. ISI are proposed by Romeny (ter Haar Romeny, 2003) who employed filterbanks to generate features invariant to rotation, position or scale. These filters are based on the gauge coordinate system, which is defined in each point of the image L by its derivative. Each pixel has a local coordinate system $(\vec{v}, \vec{w})$ where $\vec{w}$ points in the direction of the gradient vector $\left(\frac{\delta L}{\delta x}, \frac{\delta L}{\delta y}\right)$, and $\vec{v}$ is perpendicular to it. Because the gradient is independent of rotation, any derivative expressed in gauge coordinates is rotation independent too. Table 1 shows the equations to derive the gauge coordinates from the (x,y) coordinate system up to the second order. Notice that L is the luminosity of the image, L_x is the first derivative in the x direction, L_{xx} is the second derivative on the x direction, etc.

The ISI are made scale invariant by calculating the derivatives using Gaussian filters at 5 different scales, i.e. Gaussian with standard deviation $\sigma = 1, 2, 4, 8, 16$. Therefore the total number of filters employed is 5 x 5 = 25.

In Niemeijer et al. (2006), the authors derived the "visual words" from the feature by randomly sampling 150 response vector from the ISI features of 500 images. All vectors are scaled to zero mean and unit variance, and k-means clustering is applied. The frequency of the words is used to compute a histogram of the ISI "visual words" which, in conjunction with the RGB histogram is presented to the classifier.

Niemeijer et al. tested various classifiers on a dataset of 1000 images: Support Vector Machine with radial basis kernel (SVM), a Quadratic Discriminant Classifier (QDC), a Linear Discriminant Classifier (LDC) and a k-Nearest Neighbour Classifier (kNNC). The best accuracy is 0.974 obtained through SVM classifier.

Feature	Expression
L	L
L_w	$\sqrt{L_x^2 + L_y^2}$
L_{vv}	$\frac{-2L_xL_{xy}L_y+LxxL_y^2+L_x^2L_{yy}}{L_x^2+L_y^2}$
L_{vw}	$\frac{-L_x^2L_{xy}+L_y^2L_{xy}+L_xL_y(Lxx-L_{yy})}{L_x^2+L_y^2}$
L_{ww}	$\frac{L_x^2L_{xx}+2L_xL_{xy}L_y+L_y^2L_{yy}}{L_x^2+L_y^2}$

Table 1. Derivation of the irreducible set of second order image structure invariants (Niemeijer et al., 2006).

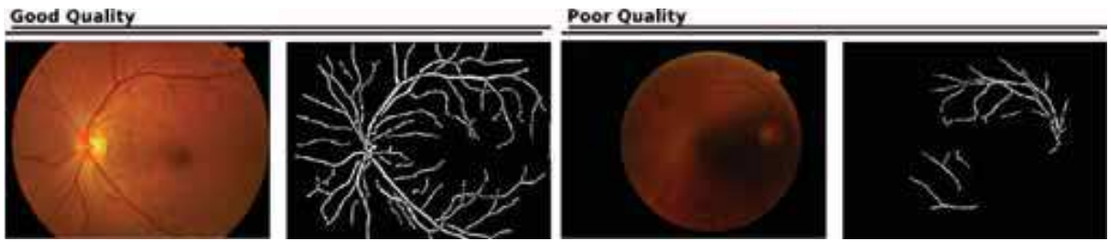

Fig. 5. Comparison of the vessel segmentation by our implementation of Zana & Klein (2001) in a good and a poor quality fundus image.

The whole QA process is called "image structure clustering" (ISC). They estimated a time of around 30 seconds to QA a new image[2].

3. Methodology

The QA proposed aims to be: *accurate* in its QA of patients of different ethnicities, *robust* enough to be able to deal with the vast majority of the images that a fundus camera can produce (outliers included), *independent* of the camera used, *computationally inexpensive* so that it can produce a QA in a reasonable time and, finally it should produce a *quality index* from 0 to 1 which can be used as input for further processing.

Our approach is based on the hypothesis that a vessel segmentation algorithm's ability to detect the eye vasculature correctly is partly related to the overall quality of an image. Fig. 5 shows the output of the vessel segmentation algorithm in images with different quality. It is immediately evident that the low vessel density in the bottom part of the right image is due to an uneven illumination and possibly to some blurring. However, a global measure of the vessel area (or vessel density) is not enough to discriminate good from bad quality images. One reason is that a considerable quantity of vessels area is taken by the two arcades which are likely to be detected even in a poor quality image as in Usher et al. (2003). Another problem is that the illumination or blurring might be uneven, making only part of the vessels undetectable. The visible vessels area can be enough to trick the QA into a wrong decision. Finally, this type of measure does not take into account outliers, artefacts caused by smudges on the lens or different Field of View (FOV) of the camera.

[2] Niemeijer et al. did not reported the hardware configuration for their tests, however in our implementation we obtained similar results (see Section 4.4)

The algorithm presented is divided in three stages: Preprocessing, Features Extraction and Classification. An in depth illustration of the full technique follows in the next sections.

3.1 Preprocessing

Mask Segmentation

The mask is defined as "a binary image of the same resolution of the fundus image whose positive pixels correspond to the foreground area". Depending on the settings, each fundus camera has a mask of different shape and size. Knowing which pixels belongs to the retina is a step that helps subsequent analysis as it gives various information about the effective size and shape of the image analysed.
Some fundus cameras (like the Zeiss Visucam PRO NM^{TM}) already provide the mask information. However, having the ability to automatically detect the mask has some benefits. It improves the compatibility across fundus cameras because it does not need to be interfaced with any sort of proprietary format to access the mask information. Also, if the QA is performed remotely, it reduces the quantity of information to be transmitted over the network. Finally, some image archives use a variety of fundus cameras and the mask is not known for each image.
The mask segmentation is based on region growing (Gonzales & Woods, 2002). It starts by extracting the green channel of the RGB fundus image, which contains the most contrast between the physiological features in the retina (Teng et al., 2002), hence this channel best describes the boundary between background and foreground. It is also the channel that is typically used for vessel segmentation. Then, the image is scaled down to 160x120, an empirically derived resolution which keeps the computational complexity as low as possible. Four seeds are placed on the four corners of the image with an offset equals to 4% of the width or height:

$$
\begin{aligned}
offset_w &\leftarrow round(imageWidth \cdot 0.04) \\
offset_h &\leftarrow round(imageHeight \cdot 0.04) \\
seed_{tl} &= [offset_w; offset_h] \\
seed_{tr} &= [imageWidth - offset_w; offset_h] \\
seed_{bl} &= [offset_w; imageHeight - offset_h] \\
seed_{br} &= [imageWidth - offset_w; imageHeight - offset_h]
\end{aligned}
$$

where $seed_{xy}$ is the location of a seed. The reason for the offsets is to avoid regions getting "trapped" by watermarks, ids, dates or other labels that generally appear on one of the corners of the image.
The region growing algorithm is started from the 4 seeds with the following criteria:

1. The absolute grey-level difference between any pixel to be connected and the mean value of the entire region must be lower than 10. This number is based on the results of various experiments.
2. To be included in one of the regions, the pixel must be 4-connected to at least one pixel in that region.
3. When no pixel satisfies the second criterion, the region growing process is stopped.

When four regions are segmented, the mask is filled with negative pixels when it belongs to a region and positive otherwise. The process is completed scaling back the image to its original size by using bilinear interpolation. Even if this final step leads to a slight quality loss, the

advantages in terms of computational time are worth the small imperfections at the edges of the mask.

"Virtual" FOV Identification

During the acquisition of a macula centred image, the patient is asked to look at fixed point visible at the back of the camera lens. In this way the macula is roughly located at the centre of the image Field of View (FOV). Even if the area viewed by different cameras is standardised, various vendors crop some part of the fundus images that do not contain useful information for diagnosis purposes.

In order to develop an algorithm that runs independently from the lost information, the "Virtual" FOV (VFOV) is extracted. The VFOV consists of an ellipse that represents the contour of the fundus image as if it was not cropped. This measure allows a simplification of the algorithm at further stages and it is the key component that makes the method independent of the camera FOV and resolution.

The classical technique to fit a geometric primitive such as an ellipse to a set of points is the use of iterative methods like the Hough transform (Leavers, 1992) or RANSAC (Rosin, 1993). Iterative methods, however, require an unpredictable amount of computational time because the size of the image mask could vary. Instead, we employ the non-iterative least squares based algorithm presented by Halir & Flusser (2000) which is extremely computationally efficient and predictable.

The points to be fitted by the ellipse are calculated using simple morphological operations on the mask. The complete procedure follows:

$\alpha \leftarrow$ erode($maskImage$)
$\gamma \leftarrow maskImage - \alpha$
fitEllipse(γ)

The erosion is computed with a square structuring element of 5 pixels. The binary nature of the image in this step (Fig. 6.b) makes the erosion very computationally efficient.

Vessel Segmentation

The ability to discern vessels from other structure is a preprocessing step of great importance in many medical imaging applications. For this reason many vessel segmentation algorithms

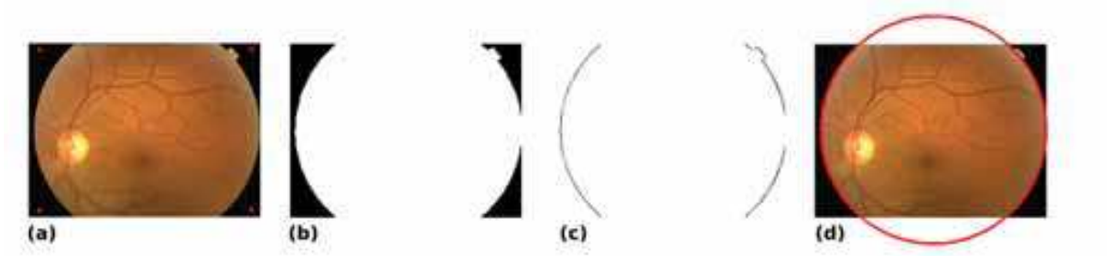

Fig. 6. (a) Original image with the 4 seeds (in red) placed. (b) Mask segmentation results. (c) Points used for VFOV detection. (d) VFOV detected.

have been presented in the literature (such as Lam & Hong, 2008; Patton et al., 2006; Ricci & Perfetti, 2007).
The technique chosen to segment veins and arteries visible in fundus images is based on the mathematical morphology method introduced by Zana and Klein (Zana & Klein, 2001). This algorithm proved to be effective in the telemedicine automatic retinopathy screening system currently developed in the Oak Ridge National Laboratory and the University of Tennessee at Memphis (Tobin et al., 2006). Having multiple modules that share the same vessel segmentation algorithm is a benefit for the system as a whole to prevent redundant processing.
Although there are more recently developed algorithms with somewhat improved performance relative to human observers, the Zana & Klein algorithm is useful because it does not require any training and its sensitivity to the quality of the image actually benefits the global QA.
This algorithm makes extensive use of morphological operations; for simplicity's sake the following abbreviations are used:
erosion: $\epsilon_B(S)$
dilation: $\delta_B(S)$
opening: $\gamma_B(S) = \delta_B(\epsilon_B(S))$
closing: $\phi_B(S) = \epsilon_B(\delta_B(S))$
geodesic reconstruction (or opening): $\gamma^{rec}_{S_{marker}}(S_{mask})$
geodesic closing: $\phi^{rec}_{S_{marker}}(S_{mask}) = N_{max} - \gamma^{rec}_{N_{max}-S_{marker}}(N_{max} - S_{mask})$
where B is the structuring element and S is the image to which it is applied, S_{marker} is the marker, S_{mask} is the mask and S_{max} is the maximum possible value of the pixel. A presentation of these morphological operators can be found in Vincent (1993).
The vessel segmentation starts using the inverted green channel image already extracted by the mask segmentation. In fact, the blue channel appears to be very weak without many information about vessels. On the other hand, the red band is usually too saturated since vessels and other retinal features emit most of their signal in the red wavelength.
The initial noise is removed while preserving most of the capillaries on the original image S_0 as follows:

$$S_{op} = \gamma^{rec}_{S_0}(Max_{i=1...12}\{\gamma_{L_i}(S_0)\}) \tag{4}$$

where L_i is a linear structuring element 13 pixels long and 1 wide for a fundus image. For each i, the element is rotated of 15°. The authors specify that the original method is not robust for changes of scale. However, since we have an estimation of the VFOV, we are in a position

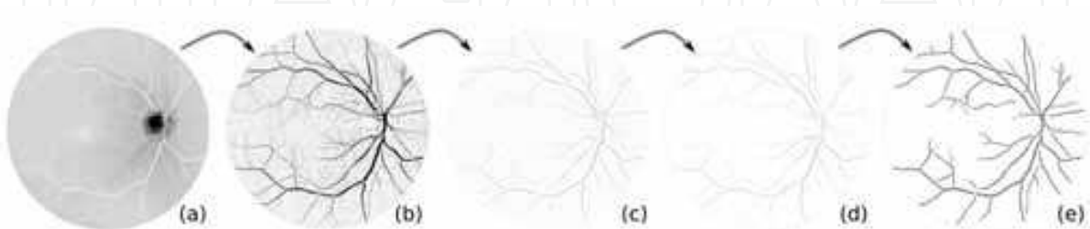

Fig. 7. Vessel segmentation summary. (a) Initial image (green channel). (b) Image after Eq. 5. (c) Image after Gaussian and Laplacian filter. (d) Image after Eq. 8. (e) Final segmentation after binarisation and removal of small connected components. All images, apart from the first one, have been inverted to improve the visualisation.

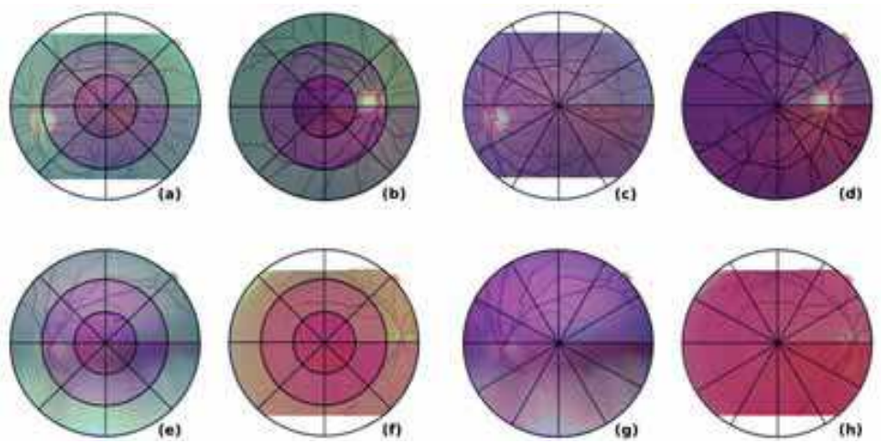

Fig. 8. Elliptical local vessel density examples. Even and odd columns respectively contain left and right retina images. In top row good quality images are shown, in the bottom row bad quality ones. The 4 images on the left use ELVD with $\theta = 8$ and $r = 3$; the 4 images on the right are the same ones but the parameters for ELVD are $\theta = 12$ and $r = 1$.

to improve it by dynamically changing the size elements depending on the length of the axes in the VFOV.
Vessels can be considered as linear bright shapes identifiable as follows:

$$S_{sum} = \sum_{i=1}^{1} 2(S_{op} - \gamma_{L_i}(S_0)) \tag{5}$$

The previous operation (a sum of top hats) improves the contrast of the vessels but at the same time various unwanted structures will be highlighted as well. The authors evaluate the vessel curvature with a Gaussian filter (width=7px; $\sigma = 7/4$) and a Laplacian (size=3x3) obtaining the image S_{lap}. Then alternating the following operation the final result is obtained and the remaining noise patterns eliminated:

$$S_1 = \gamma^{rec}_{S_lap}(Max_{i=1...12}\{\gamma_{L_i}(S_lap)\}) \tag{6}$$

$$S_2 = \phi^{rec}_{S_1}(Min_{i=1...12}\{\phi_{L_i}(S_1)\}) \tag{7}$$

$$S_{res} = (Max_{i=1...12}\{\gamma^2_{L_i}(S_2)\} \geq 1) \tag{8}$$

As the last step of our implementation, we binarise the image and remove all the connected components that have an area smaller than 250 pixels. Once again this value is scaled depending on the VFOV detected. Fig. 7 shows a visual summary of the whole algorithm.

3.2 Feature Extraction

Elliptical Local Vessel Density (ELVD)

By employing all information gathered in the preprocessing phase, we are able to extract a local measure of the vessel density which is camera independent and scale invariant. Other authors either measure a similar feature globally like Usher et al. (2003), or they use a computationally expensive method like Fleming et al. (2006) whose approach requires a vessel segmentation, a template cross correlation and two different Hough transforms. Instead, we

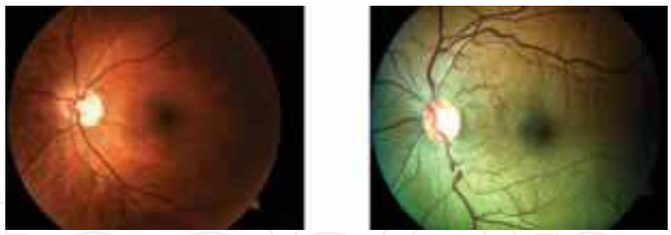

Fig. 9. Pigmentation difference between Caucasian (on the left) and African American (on the right) retinas. Images extracted from the datasets used in our tests (see section 4.1).

employ an "adaptable" polar coordinate system (θ, r) with the origin coincident with the origin of the VFOV. It is adaptable in the sense that its radius is not constant but it changes according to the shape of the ellipse. This allows to deal with changes of scale not proportional between height and width.
The Elliptical Local Vessel Density (ELVD) is calculated by measuring the vessel area under each local window, then normalised with zero mean and unit variance[3]. The local windows are obtained sampling r and θ. Different values of r and θ will tolerate or emphasize different problems with the image quality. In Fig. 8 for example, the 4 images on the left ($\theta = 8$ and $r = 3$) have 8 windows each on the centre of VFOV where the macula is located. In this fashion, ELVD features can detected a misaligned fundus image. On the other hand, the ELVD in the 4 images on the right ($\theta = 12$ and $r = 1$) will be more robust to macula misalignment, but more sensitive to vessel detection on both vascular arcades.
The idea behind ELVD is to create local windows that are roughly placed in consistent positions throughout different images. In the even or odd columns of Fig. 8, note that vessels close to the ON are in the same or nearby local windows, even if images have different FOVs. The power of this new style of windowing is its capability of capturing morphological information about fundus images without directly computing the position of ON, macula or arcade vessels, since these operations are computational expensive and prone to errors if the image has a very poor quality.

Luminosity/Colour Information

The analysis of the global colour information of the fundus image can contain useful information for the quality of the image. The method of Lee & Wang (1999) employed the histogram of the grey-level obtained from the RGB image as the only means to describe the image quality. The much more refined method of Niemeijer et al. (2006) uses 5 bins of each channel of the RGB histogram as additional features as input to the classifier. The authors presented results demonstrating that this piece of RGB information improved their classification respect to pure ISI features, even if ISI is representative of most of the retinal structures.
Inspired by Niemejer et al. we use colour information to represent aspects of quality that cannot be entirely measured with ELVD such as over/underexposed images in which the vasculature is visible or outliers with many features that are recognised as vessels.
All RGB channels are evaluated by computing the histogram for each plane. The histogram is normalised by the size of the mask in order to make this measure scale independent. It is noticed that people from different ethnic origin have a different pigmentation on the retina; this aspect is particularly noticeable in the blue and red channel. For example while Caucasians

[3] The zero mean and unit variance is calculated for each feature across all the training images.

have a fundus with a very strong red component people of African descent have a darker pigmentation with a much stronger blue component (see figure 9). In our case this is not an issue because we ensure we have adequate examples of different ethnic groups in our training library.
Also, the HSV colour space is employed as a feature. Only the saturation channel is used which seems to play an important role in the detection of the over/under exposition of the images. The reason is the channel relative independence from pigment and luminosity. Once again, the global histogram is extracted and normalised with the image mask.

Other Features

In addition to ELVD and colour information two other sets of features are considered as candidates to represent quality:

- *Vessel Luminosity:* Wang et al. (2001) noted that the grey level values of corresponding to the vessels can be used as a good approximation of the background luminosity. They proposed an algorithm that exploits this information to normalise the luminosity of the fundus images. If the vessel luminosity with the same elliptical windows used for the ELVD, we can measure the luminosity spread in the image. This can be particularly useful because poor quality images have often an uneven illumination.
- *Local Binary Patterns (LBP):* Texture descriptors are numerical measures of texture patterns in an image. LBP are capable of describing a texture in a compact manner independently from rotation and luminosity (Ojala & Pietikainen, 1996). The LBP processing creates binary codes depending on the relation between grey levels in a local neighbourhood. In the QA context this type of descriptor can be useful to check if the particular patterns found in a good quality retina are present in the image. This is accomplished by generating an histogram of the LBP structures found.

3.3 Classification

The majority of the authors who developed a QA metric for retinal images approached the classification in a similar way (Lalonde et al., 2001; Lee & Wang, 1999; Usher et al., 2003). The training phase consists of creating models of good and poor quality images (in some cases more intermediate models are employed) by calculating the mean of the features of the training sets. When a new retinal image is retrieved, its features are computed and the it is classified based on the shortest distance[4] to one of the models. This type of approach works reasonably well if the image to be classified is similar enough to one of the models. Also, it simplifies the calculation of a QA metric between 0 and 1 because distances can be easily normalised. However, this approach has a major drawback: the lack of generalisation on images with a large distance from the both models. This problem limits the method applicability in a real world environment.
Niemejer et al. (Niemeijer et al., 2006) are the only authors to our knowledge that approach the QA as a classic pattern classification problem. During the training phase they do not try to build a model or to make any assumption about the distribution of the data. Instead, they label each samples in one of the two classes and train one of the following classifiers: Support Vector Machines (SVM), Quadratic Discriminant Classifier (QDC), Linear Discriminant Classifier (LDC) and *k*-Nearest Neighbour Classifier (KNNC). Finally, they selected the classifier

[4] Distances calculations vary; some use Euclidean distance, others are based on correlation measures.

with the best performance (in their case a SVM with radial basis kernel) by testing it with a separate dataset.
Our classification technique is similar to the one of Niemeijer et al., but with two major differences. The first one is that the feature vector is created directly from the raw features without any need of pre-clustering, which can be computationally expensive, especially if a large number of features are used in a high dimensional space. The second difference is the fact that the classifier needs to output a posterior probability rather than a clear cut classification of a particular class. This probability will allow the correct classification of fair quality images even if the training is performed on two classes only.

4. Tests and Results

In this section, a summary of the most significant experiments performed during the development of the ELVD quality estimator are presented. The first section contains an overview of the datasets used. We then show the tests used for an initial evaluation of the QA proposed, the comparison with existing techniques and the choice of the classifier. Then, an analysis on possible optimisations of the feature set is performed. Finally the final QA system is tested on all the datasets and its computational performance is evaluated.

4.1 Data Sets

Various datasets are employed in the following tests. Each of them has peculiar characteristics that make it useful to test particular aspects of the QA classifier. The list follows:

- *"Abramoff":* dataset composed of 10862 retinal images compiled by M. Abramoff as part of a study in the Netherlands. They were obtained using with different settings (FOV, exposure, etc.) on healthy and ill patients (Niemeijer et al., 2007). Three different cameras were used: Topcon NW 100, Topcon NW 200 and the Canon CR5-45NM. Unfortunately their quality was not labelled by the physicians.
- *"Aspen":* dataset composed of 98 images that targets mainly patients with retinopathy conditions. This images were captured as part of a non-research teleophtalmology program to identify diabetic retinopathy in people leaving in the Aspen Health Region of Alberta, Canada (Rudnisky et al., 2007). Once again the quality was not labelled by the physicians.
- *"Chaum":* this set is composed of 42 images extracted from the Abramoff dataset labelled as good and poor quality. They are good representatives of various aspects of the quality aspects of fundus images. These images were labelled by an expert in the field (Dr. E. Chaum) in order to facilitate the development of the QA system.
- *"ORNL":* it is composed of 75 images extracted from the Abramoff dataset and labelled as good, fair and poor quality. These images were compiled at the Oak Ridge National Laboratory for the analysis of various aspects of the automatic diagnosis of diabetic retinopathy.
- *"African American":* it contains 18 retina images of African American patients. All these images were labelled as good quality by Dr. E. Chaum. This dataset is of particular importance because it is very likely that most of the patients in Netherlands are Caucasian[5], but our system deployment is targeted toward the deep-to-mid South region of the United States of America where there is a large population of African Americans.

- *"Outliers":* it is composed of 24 images containing various types of image outliers, all captured with a fundus camera.

4.2 Classifier Selection

In order to select the most appropriate classifier, a series of comparative tests is run on the "ORNL" and "Outliers" dataset. The results are compared with our implementation of the QA by Niemeijer et al. (2006), the most recent method found in the literature. The feature vector used by our classifiers is composed of ELVD with 3 slices and 8 wedges (ELVD 3x8) and the RGB colour histogram with 5 bins per channel. These tests were presented in the EMBC conference of 2008 and led to encouraging results (Giancardo et al., 2008).
The testing method used a randomised 2-fold validation, which works as follows. The samples are split in two sets A and B. In the first phase A is used for training and B for testing, then roles are inverted and B is used for training and A for testing. The performance of a classifier are evaluated using the Area Under the ROC curve (AUR) for $\frac{TruePositiveRate}{FalsePositiveRate}$ (TPR/FPR) and $\frac{TrueNegativeRate}{FalseNegativeRate}$ (TNR/FNR). See (Fawcett, 2004) for more details.

	ORNL set		ORNL + Outliers dataset	
Classifier	TPR/FPR	TNR/FNR	TPR/FPR	TNR/FNR
Nearest Neighbour	1	1	1	1
KNN (K=5)	1	1	0.99	0.98
SVM (Linear)	1	1	0.92	0.79
SVM (Radial)	1	1	1	1
ISC by Niemeijer et al.	1	0.88	1	0.88

Table 2. Good/Poor classifier test on "ORNL" and "Outliers" dataset. For the first four classifiers the feature vector used is ELVD 3x8 + RGB histogram with 5 bins.

In the two columns on left, table 2 shows the Good/Poor classification results for the "ORNL" dataset. All the classifiers using our feature vector have perfect or near-perfect performance in the selection between good and poor class, which is not the case for the Niemeijer et al. method (note that only the good and poor classes are used).
In the two columns on the right, all the Outliers dataset were added as test samples. An outlier image can have an enormous variability, therefore we feel that the training on this type of images might bias the classifier. Ideally, a classifier should be able to classify them as poor even if they are not fundus images as such. In this test, the classifiers performed differently, the best results are given by Nearest Neighbour classifier and SVM with a radial kernel.
Recall that the aim of this system is to generate a quality score from 0 to 1 to judge the image quality. In order to analyse this aspect, means and standard deviations of the scores obtained are displayed in Fig. 10. The classifiers are again trained on Good and Poor class (with 2-fold validation) but the Fair class is added to the testing samples without any explicit training on it. This allows to test the generalisation of the system. The most striking result of this test is the fact that the classifier with the poorest average AUR (SVM with a linear kernel) is also the one that achieves the best class separation, with an average score separation between Good and Poor classes of more than 0.8. The Fair class in this test has a mean score located at the middle of the scale.

[5] For privacy reasons the ethnicity of the subjects in the Abramoff dataset was not known.

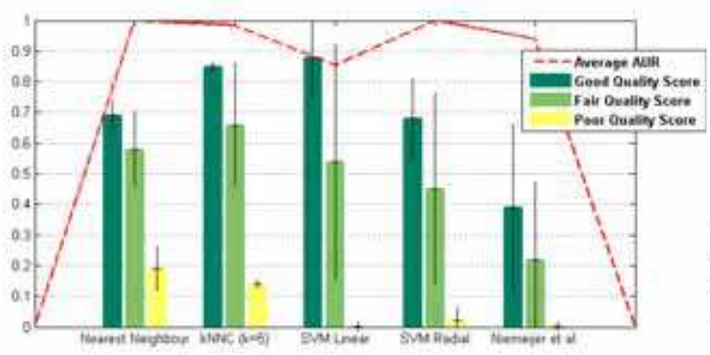

Fig. 10. Classifier scores test on "ORNL" dataset. For the first four classifiers the feature vector used is ELVD 3x8 + RGB histogram with 5 bins.

This apparent contradiction makes the selection of the classifier difficult. Therefore another series of tests was run on the more challenging "Chaum" dataset. In this case a leave-one-out strategy is used, i.e. the classifier is trained multiple times removing a different sample from the training set and using it as test target each time. This technique allows us to run complete tests using a relative small dataset.
Table 3 shows the results obtained employing the same classifiers and feature vector as before. While no classifier obtained ideal performance, the SVM with a linear kernel seems to have a good compromise between AUR and score separation. The small AUR advantages of KNN and Nearest Neighbour do not justify the computational performance issues that these type of classifiers have when many training samples in a high dimensional space are used, and also these classifiers have relatively low score difference between Good and Poor class.

Classifier	AUR TPR/FPR	AUR TNR/FNR	Average Good/Poor score difference
Nearest Neighbour	0.97	0.97	0.51
KNN (K=5)	0.97	0.97	0.51
SVM (Linear)	0.97	0.94	0.76
SVM (Radial)	0.94	0.91	0.54

Table 3. Good/Poor classifier test on "Chaum" dataset. The feature vector used is ELVD 3x8 + RGB histogram with 5 bins (the error bars show the average standard deviation).

The main problem of SVM with a linear kernel is its poor performance on the outliers, especially when compared with the results obtained by the other classifiers tested. For a better understanding of this behaviour part of the sample vectors are projected on a hyperplane which allows their representation in 2 dimensions. The hyperplane is calculated using the Linear Discriminant Analysis (Duda et al., 2001) on the Good and Poor samples, allowing a visualization of the space from the "point of view of the classifier".
Fig. 11 shows the result of the LDA calculation. While the distribution of Good and Poor class are well defined, the Outliers are spread throughout the LDA space. Nonlinear classifiers like Nearest Neighbour, KNN or SVM (Radial) can easily isolate the cluster of Good samples from

the rest, but this is not a problem solvable by a linear function like the one employed by SVM (Linear).

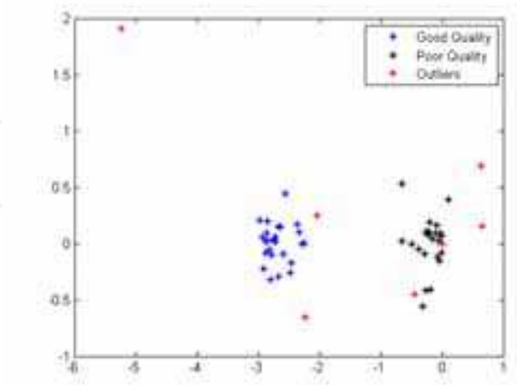

Fig. 11. 2D LDA Space Projection for ELVD features (Outliers not included in the LDA calculation).

4.3 Features Selection

It would be desirable to use the SVM (Linear) given its good score separation properties. One solution to this problem is the selection of new feature capable of linearising the space. However, the selection of adequate features allowing the SVM hyperplane to split the good quality samples from all the rest is not a straightforward task. Testing all the possible combination of the feature sets mentioned is impractical. Each feature set has many parameters: ELVD 36 (3 sets of radial section and 12 sets of wedges), Vessel Luminosity 36 (same as previously), RGB histogram 80 (all the channel combinations which can be normalised or not and 5 sets of histogram bins), HSV histogram 80 (same as previously) and LBP 4 (2 sets of radii length and 2 sets of LBP codes), for a total of 33 177 600 possible combinations.

Therefore an empirical approach was adopted. Firstly, it is assumed that all feature sets represent independent aspects of the fundus image quality. While this assumption is rather farfetched, it does allow us to run only 324 tests to check all the possible permutation in each feature set, and also gives a feeling for what features are worth testing. Table 4 shows which are the parameters that achieved the best results for each feature set on the "Chaum" dataset. This dataset was chosen because is the most authoritative representation of good and poor quality image in most of the different aspects.

If feature sets were actually independent, the ideal feature vector would be composed by all of them with the parameters shown in Table 4. However, because there is almost certainly some degree of correlation, various parameters of the feature sets are selected based on their relative AUR and Good/Poor score difference and they are combined together for a total of 16 800 tests.

Surprisingly optimal results (Avg AUR of 1) and excellent good/poor score separability (0.91) are obtained with a relatively simple feature vector composed of:

- ELVD with 6 wedges and a single radial section
- The mask normalised histogram of the saturation with 2 bins

Feature Set	Parameters	Avg AUR	Average Good/Poor score difference
ELVD	16 rad. sec. & 6 wedges	0.98	0.74
RGB Hist	4 bins per ch. & mask norm.	0.81	0.51
HSV Hist	5 bins of Sat. & mask norm.	0.85	0.59
Vessel Luminosity	16 rad. sec. & 6 wedges	0.98	0.74
LBP	8 px radius & 8 codes	0.85	0.59

Table 4. Best results of each independent feature set on the "Chaum" dataset. The test is a leave-one-out with a SVM Linear classifier.

As it was suspected, the parameters that lead to the best results in this test are not the combination of the parameters found in each independent feature set test (table 4). However, they allowed to reduce the parameters search space and obtain excellent results with a relative simple combination.

4.4 Computational Performance

The performance of the C++ implementation of the ELVD QA is evaluated with a standard benchmarking technique. The complete ELVD QA system is run on 25 images randomly chosen from the "Chaum" dataset, during each iteration the time required to run the total system and each separate algorithm is recorded and averaged. All the images are scaled to the common resolution of 756x576 in order to have fairly consistent measurements. All the test were run on a 3.4 GHz Intel Pentium 4 machine with 2 GB of RAM.

Stage	Time (in milliseconds)
Mask Detection	116
VFOV	16
Vessel Segmentation	1920
ELVD	15
Saturation Histogram	25
Classification + Memory Allocation	38
	Total 2130

Table 5. Relative performance of the different components in the ELVD QA C++ implementation.

The total time required to obtain a quality score for a single image is 2130 milliseconds. Table 5 shows how each system component contributes to the global computational time. The vessels segmentation is by far the main contributor having more the 10 times the computational cost of all the other algorithms summed together. The mask detection and the classification, two possibly expensive operations, are actually quite efficient considering the needs of this system. For comparison, a global benchmark was run on our implementation of the Niemeijer et al. QA classification (Niemeijer et al., 2006). The result obtained is well over 30 seconds, a time one order of magnitude greater than our approach. This is due to the many filterbanks that must be executed to calculate the raw features and the nearest neighbour operations to obtain

the "words". However, the comparison between the two techniques should be taken with a bit of perspective because of the different implementation platforms. In fact the Niemeijer et al. algorithm is implemented in Matlab, a slower language than C++ because of its interpreted language nature. Nevertheless, we should point out that in our tests Matlab uses fast native code thanks to the Intel IPP libraries (Intel, 2007) for all the filtering operations, and these are very computationally efficient regardless of programming language choice.

5. Conclusion

At the beginning of the chapter, the quality assessment for fundus images was defined as "the characteristics of an image that allow the retinopathy diagnosis by a human or software expert". The literature was surveyed to find techniques which could help to achieve this goal. General image QA does not seem well suited for our purposes, as they are mainly dedicated to the detection of artefacts due to compression and they often require the original non-degraded image, something that does not make much sense in the context of QA for retinal images.
Our survey found five publications which tackled a problem comparable to the one of this project. They were divided into 3 categories: "Histogram Based", "Retina Morphology" and "Bag-of-Words". The authors of the first category approached the problem by computing relatively simple features and comparing them to a model of a good quality image. Although this approach might have advantages like speed and ease of training, it does not generalise well on the natural variability of fundus images as highlighted by Niemeijer et al. (2006) and Fleming et al. (2006). "Retina Morphology" methods started to take into account features unique to the retina, such as vessels, optic nerve or temporal arcades. This type approach considerably increased the QA accuracy. Remarkably, Fleming et al. developed a very precise way to judge the quality of image clarity and field definition which closely resembled what an ophthalmologist would do. The main drawbacks are time required to locate the various structures and the fact that if the image quality is too poor, some of the processing steps might fail, giving unpredictable results. This is unlikely to happen in the problem domain of Fleming et al. because they worked with images taken by trained ophthalmologists, but this is not the case with systems that can be used by personnel with basic training. The only method of the "Bag-of-Words" category is the one developed by Niemeijer et al. Their technique is based on pattern recognition algorithms which gave high accuracy and specificity. The main drawback is again speed of execution.
The new approach described in this chapter was partially inspired by all these techniques: colour was used as features as in the "Histogram Based" technique, the vessels were segmented as a preprocessing step like in the "Retina Morphology" techniques and the QA was computed by a classifier similar to the one used in the "Bag-of-Words" techniques. New features were developed and used such as ELVD, VFOV and the use of the HSV colour space, which was not evaluated by any of the previous authors for QA of fundus images. This made possible the creation of a method capable of classifying the quality of an image with a score from 0 to 1 in a period of time much shorter than "Retina Morphology" and "Bag-of-Words" techniques.
Features, classifier types and other parameters were selected based on the results of empirical tests. Four different types of datasets were used. Although none are very large (none contained more than 100 images) they were fairly good representative of the variation of fundus images in terms of quality, camera used and patient's ethnicity. In the literature, the method which seemed to perform best and which had the best generalisation was the one of Niemeijer et al. It was implemented and compared to our algorithm. Our results are in favour of the

method presented in this chapter in terms of classification performance and speed. However, while our method has a clear advantage in terms of speed (it runs one order of magnitude faster because of the lower computational complexity), the comparison in terms of classification should be taken with care. In fact, Niemeijer et al. employed a dataset larger than ours to train the system.

The final algorithm was implemented in C++. Tests showed that it was able to produce a QA score in 2 seconds, also considering the vessel segmentation which can later be used by other modules of the global diabetic retinopathy diagnosis system.

In February 2009, the first clinic in a telemedicine network performing teleophthalmology went on-line in Memphis, Tennessee under the direction of Dr. E. Chaum. This network addresses an under served population and represents a valuable asset to broad-based screening of diabetic retinopathy and other diseases of the retina. A secure web-based protocol for submission of images and a database archiving system has been developed with a physician reviewing tool. All images are acquired from non-dilated retinal images obtained in primary care clinics and are manually reviewed by an ophthalmologist. As part of the submission process, all images undergo an automatic quality estimation using our C++ implementation of the ELVD QA.

6. References

Amos, A. F., McCarty, D. J. & Zimmet, P. (1997). The rising global burden of diabetes and its complications: estimates and projections to the year 2010., *Diabetic Medicine* **14 Suppl 5**: S1–85.

Baker, M. L., Hand, P. J., Wang, J. J. & Wong, T. Y. (2008). Retinal signs and stroke: revisiting the link between the eye and brain., *Stroke* **39**(4): 1371–1379.

Ballard, D. (1981). Generalizing the hough transform to detect arbitrary shapes, *Pattern Recognition* **13**: 111–122.

Cassin, B. & Solomon, S. (1990). *Dictionary of Eye Terminology*, Gainsville, Florida: Triad Publishing Company.

Chen, J. & Tian, J. (2008). Retinal vessel enhancement based on directional field, *Proceedings of SPIE*, Vol. 6914.

Duda, R. O., Hart, P. E. & Stork, D. G. (2001). *Pattern Classification*, Wiley-Interscience.

ETDRS (1991). Early photocoagulation fior diabetic retinopathy. early treatment diabetic retinopathy study report number 9, *Ophthalmology* **98**: 766–785.

Fawcett, T. (2004). Roc graphs : Notes and practical considerations for researchers, *Technical report*, HP Laboratories, 1501 Page Mill Road, Palo Alto, CA 94304, USA.

Fei-Fei, L. & Perona, P. (2005). A bayesian heirarcical model for learning natural scene categories, *Proceedings of CVPR*.

Fleming, A. D., Philip, S., Goatman, K. A., Olson, J. A. & Sharp, P. F. (2006). Automated assessment of diabetic retinal image quality based on clarity and field definition., *Investigative Ophthalmology and Visual Science* **47**(3): 1120–1125.

Foracchia, M., Grisan, E. & Ruggeri, A. (2005). Luminosity and contrast normalization in retinal images., *Medical Image Analysis* **9**(3): 179–190.

Giancardo, L., Abramoff, M. D., Chaum, E., Karnowski, T. P., Meriaudeau, F. & Tobin, K. W. (2008). Elliptical local vessel density: a fast and robust quality metric for retinal images, *Proceedings of IEEE EMBS*.

Gonzales, R. C. & Woods, R. E. (2002). *Digital Image Processing*, Prentice-Hall.

Grisan, E., Grisan, E., Giani, A., Ceseracciu, E. & Ruggeri, A. (2006). Model-based illumination correction in retinal images, *in* A. Giani (ed.), *Proceedings of 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro*, pp. 984–987.

Halir, R. & Flusser, J. (2000). Numerically stable direct least squares fitting of ellipses, *Department of Software Engineering, Charles University, Czech Republic* .

Intel (2007). *Intel Integrated Performance Primitives for the Windows OS on the IA-32 Architecture*, 318254-001us edn.
URL: *http://developer.intel.com*

Javitt, J., Aiello, L., Chiang, Y., Ferris, F., Canner, J. & Greenfield, S. (1994). Preventive eye care in people with diabetes is cost-saving to the federal government, *Diabetes Care* **17**: 909–917.

Jonas, J. B., Schneider, U. & Naumann, G. O. H. (1992). Count and density of human retinal photoreceptors, *Graefe's Archive for Clinical and Experimental Ophthalmology* **230**: 505–510.

Lalonde, M., Gagnon, L. & Boucher, M. C. (2001). Automatic visual quality assessment in optical fundus images, *Proceedings of Vision Interface*, pp. 259–264.

Lam, B. S. Y. & Hong, Y. (2008). A novel vessel segmentation algorithm for pathological retina images based on the divergence of vector fields, *IEEE Transaction on Medical Imaging* **27**(2): 237–246.

Leavers, V. F. (1992). *Shape Detection in Computer Vision Using the Hough Transform*, Springer-Verlag New York, Inc. Secaucus, NJ, USA.

Lee, S. & Wang, Y. (1999). Automatic retinal image quality assessment and enhancement., *Proceedings of SPIE Image Processing*, pp. 1581–1590.

Luzio, S., Hatcher, S., Zahlmann, G., Mazik, L., Morgan, M. & Liesenfeld, B. (2004). Feasibility of using the tosca telescreening procedures for diabetic retinopathy, *Diabetic Medicine* **21**: 1121.

Mann, G. (1997). Ophtel project, *Technical report*, European Union.

Niemeijer, M., Abramoff, M. D. & van Ginneken, B. (2006). Image structure clustering for image quality verification of color retina images in diabetic retinopathy screening., *Medical Image Analysis* **10**(6): 888–898.

Niemeijer, M., Abramoff, M. D. & van Ginneken, B. (2007). Segmentation of the optic disc, macula and vascular arch in fundus photographs, *IEEE Trans Med Imag* **26**(1): 116–127.

Ohlander, R., Price, K. & Reddy, D. R. (1978). Picture segmentation using a recursive region splitting methods, *Computer Graphic and Image Processing* **8**: 313–333.

Ojala, T. & Pietikainen, M. (1996). A comparative study of texture measures with classification based on feature distribution, *Pattern Recognition* **29**: 51–59.

Patton, N., Aslam, T. M., MacGillivray, T., Deary, I. J., Dhillon, B., Eikelboom, R. H., Yogesan, K. & Constable, I. J. (2006). Retinal image analysis: concepts, applications and potential, *Progress in retinal and eye research* **25**(1): 99–127.

Ricci, E. & Perfetti, R. (2007). Retinal blood vessel segmentation using line operators and support vector classification, *IEEE Transaction on Medical Imaging* **26**(10): 1357–1365.

Rosin, P. L. (1993). Ellipse fitting by accumulating five-point fits, *Pattern Recognition Letters*, Vol. 14, pp. 661–699.

Rudnisky, C. J., Tennant, M. T. S., Weis, E., Ting, A., Hinz, B. J. & Greve, M. D. J. (2007). Web-based grading of compressed stereoscopic digital photography versus stan-

dard slide film photography for the diagnosis of diabetic retinopathy, *Ophthalmology* **114**(9): 1748–1754.

Sheikh, H. R., Sabir, M. F., Bovik, A. C., Sheikh, H., Sabir, M. & Bovik, A. (2006). A statistical evaluation of recent full reference image quality assessment algorithms, *IEEE Transactions on image processing* **15**(11): 3440–3451.

Sivic, J., Russell, B., Efros, A., Zisserman, A. & Freeman, W. (2005). Discovering object categories in image collections, *Proceedings of International Conference Computer Vision, Beijing*.

Teng, T., Lefley, M. & Claremont, D. (2002). Progress towards automated diabetic ocular screening: a review of image analysis and intelligent systems for diabetic retinopathy, *Medical and Biological Engineering and Computing* **40**(1): 2–13.

ter Haar Romeny, B. M. (2003). *Front-End Vision and Multi-Scale Image Analysis*, 1st edn, Springer.

Tierney, L. M., McPhee, S. J. & Papadakis, M. A. (2002). *Current medical Diagnosis & Treatment. International edition.*, New York: Lange Medical Books/McGraw-Hill.

Tobin, K. W., Chaum, E., Govindasamy, V. P., Karnowski, T. P. & Sezer, O. (2006). Characterization of the optic disc in retinal imagery using a probabilistic approach, *Proceedings of SPIE*, Vol. 6144.

Usher, D., Himaga, M. & Dumskyj, M. (2003). Automated assessment of digital fundus image quality using detected vessel area, *Proceedings of Medical Image Understanding and Analysis*, British Machine Vision Association (BMVA), pp. 81–84.

Vincent, L. (1993). Morphological grayscale reconstruction in image analysis: applications and efficient algorithms, *IEEE Journal of Image Processing* **2**(2): 176–201.

Wang, Y., Tan, W. & S C Lee, S. (2001). Illumination normalization of retinal images using sampling and interpolation, *Proceedings of SPIE*, Vol. 4322.

Wyszecki, G. & Stiles, W. S. (2000). *Color science: Concepts and methods, quantitative data and formulae.*, 2nd edn, New York, NY: John Wiley & Sons.

Zana, F. & Klein, J. C. (2001). Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation, *IEEE Transaction on Image Processing* **10**(7): 1010–1019.

New Developments in Biomedical Engineering
Edited by Domenico Campolo

ISBN 978-953-7619-57-2
Hard cover, 714 pages
Publisher InTech
Published online 01, January, 2010
Published in print edition January, 2010

Biomedical Engineering is a highly interdisciplinary and well established discipline spanning across engineering, medicine and biology. A single definition of Biomedical Engineering is hardly unanimously accepted but it is often easier to identify what activities are included in it. This volume collects works on recent advances in Biomedical Engineering and provides a bird-view on a very broad field, ranging from purely theoretical frameworks to clinical applications and from diagnosis to treatment.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen