Permissions & Reprints

Glaucoma risk index: Automated glaucoma detection from color fundus images

Rüdiger Bock[a, b,] ✉, Jörg Meier[a,] ✉, László G. Nyúl[c,] ✉, Joachim Hornegger[a, b,] ✉ and Georg Michelson[d, b, e,] ✉

a Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander-University Erlangen-Nuremberg, Germany

b Erlangen Graduate School in Advanced Optical Technologies (SAOT), Friedrich-Alexander-University Erlangen-Nuremberg, Germany

c Department of Image Processing and Computer Graphics, University of Szeged, Hungary

d Department of Ophthalmology, Friedrich-Alexander-University Erlangen-Nuremberg, Germany

e Interdisciplinary Center of Ophthalmic Preventive Medicine and Imaging, Friedrich-Alexander-University Erlangen-Nuremberg, Germany

Abstract

Glaucoma as a neurodegeneration of the optic nerve is one of the most common causes of blindness. Because revitalization of the degenerated nerve fibers of the optic nerve is impossible early detection of the disease is essential. This can be supported by a robust and automated mass-screening. We propose a novel automated glaucoma detection system that operates on inexpensive to acquire and widely used digital color fundus images. After a glaucoma specific preprocessing, different generic feature types are compressed by an appearance-based dimension reduction technique. Subsequently, a probabilistic two-stage classification scheme combines these features types to extract the novel Glaucoma Risk Index (GRI) that shows a reasonable glaucoma detection performance. On a sample set of 575 fundus images a classification accuracy of 80% has been achieved in a 5-fold cross-validation setup. The GRI gains a competitive area under ROC (AUC) of 88% compared to the established topography-based glaucoma probability score of scanning laser tomography with AUC of 87%. The proposed color fundus image-based GRI achieves a competitive and reliable detection performance on a low-priced modality by the statistical analysis of entire images of the optic nerve head.

Glaucoma Risk Index:
Automated glaucoma detection from color fundus images

Rüdiger Bock [a,b,*] Jörg Meier [a] László G. Nyúl [c] Joachim Hornegger [a,b] Georg Michelson [d,b,e]

[a] *Pattern Recognition Lab, Department of Computer Science,*
Friedrich-Alexander-University Erlangen-Nuremberg, Martensstr. 3, 91058 Erlangen, Germany
[b] *Erlangen Graduate School in Advanced Optical Technologies (SAOT), Friedrich-Alexander-University Erlangen-Nuremberg*
[c] *Department of Image Processing and Computer Graphics, University of Szeged, Árpád tér 2, 6720 Szeged, Hungary*
[d] *Department of Ophthalmology, Friedrich-Alexander-University Erlangen-Nuremberg,*
Schwabachanlage 6, 91054 Erlangen, Germany
[e] *Interdisciplinary Center of Ophthalmic Preventive Medicine and Imaging, Friedrich-Alexander-University*
Erlangen-Nuremberg, Schwabachanlage 6, 91054 Erlangen, Germany

Abstract

Glaucoma as a neurodegeneration of the optic nerve is one of the most common causes of blindness. Because revitalization of the degenerated nerve fibers of the optic nerve is impossible early detection of the disease is essential. This can be supported by a robust and automated mass-screening. We propose a novel automated glaucoma detection system that operates on inexpensive to acquire and widely used digital color fundus images. After a glaucoma specific preprocessing, different generic feature types are compressed by an appearance-based dimension reduction technique. Subsequently, a probabilistic two-stage classification scheme combines these features types to extract the novel Glaucoma Risk Index (GRI) that shows a reasonable glaucoma detection performance. On a sample set of 575 fundus images a classification accuracy of 80 % has been achieved in a 5-fold cross validation setup. The GRI gains a competitive area under ROC (AUC) of 88 % compared to the established topography-based Glaucoma Probability Score of scanning laser tomography with AUC of 87 %. The proposed color fundus image-based GRI achieves a competitive and reliable detection performance on a low-priced modality by the statistical analysis of entire images of the optic nerve head.

Key words: computer aided diagnosis, glaucoma, optic disk, appearance-based image analysis, linear principal component analysis

1. Introduction

* Corresponding author. Telephone: +49-9131-85-27775
Fax: +49-9131-303811
Email addresses:
ruediger.bock@informatik.uni-erlangen.de (Rüdiger Bock), joerg.meier@informatik.uni-erlangen.de (Jörg Meier), nyul@inf.u-szeged.hu (László G. Nyúl), joachim.hornegger@informatik.uni-erlangen.de (Joachim Hornegger), georg.michelson@uk-erlangen.de (Georg Michelson).

Glaucoma is one of the most common causes of blindness with a mean prevalence of 2.4 % for all ages and of 4.7 % for ages above 75 years (Klein et al. 1992). The disease is characterized by the progressive degeneration of optic nerve fibers and astrocytes showing a distinct pathogenetic image of the optic nerve head.

Glaucoma leads to (i) structural changes of the optic nerve head (ONH) and the nerve fiber layer

and (ii) a simultaneous functional failure of the visual field. The structural changes are manifested by a slowly diminishing neuroretinal rim indicating a degeneration of axons and astrocytes of the optic nerve (Fig. 1).

As lost capabilities of the optic nerve can not be recovered, early detection and subsequent treatment is essential for affected patients to preserve their vision (Michelson et al. 2008). Commonly, glaucoma diagnosis is based on the patient's medical history, intraocular pressure, visual field loss tests and the manual assessment of the ONH via ophthalmoscopy or stereo fundus imaging (Lin et al. 2007). To additionally objectify the glaucoma stage and its progression geometric parameters of the ONH are documented. These geometric parameters measure ONH structures that are changing in case of glaucoma disease: optic disk diameter, optic disk area, cup diameter, rim area, mean cup depth etc.

This contribution provides a data driven framework extracting a novel glaucoma parameter from fundus images. Contrary to the established detection techniques, it does not require accurate measurements of geometric ONH structures as it performs a statistical data mining technique on the image patterns themselves. The proposed methodology can be transferred to other domains and might be able to extract further parameters providing new insights to other ophthalmic questions.

2. Background

The glaucoma disease is characterized by the degeneration of optic nerve fibers and astrocytes that is often accompanied by an increased intraocular pressure. Due to the loss of nerve fibers the retinal nerve fiber layer (RNFL) thickness is descreasing. In the course of disease, the interconnection between the photoreceptors and the visual cortex is reduced. In the worst case, the visual information of the photoreceptors can no longer be transmitted to the brain and visual field loss up to blindness is threatening. The disappearance of axons and astrocytes affects the structural appearance of the ONH and causes a reduction of the functional capabilities of the retina. The ONH can be examined by ophthalmoscopy or by stereo fundus photography: in the course of the disease the neuroretinal rim gets thinner while the cup is expanding due to the loss of nerve fibers and astrocytes (Fig. 1).

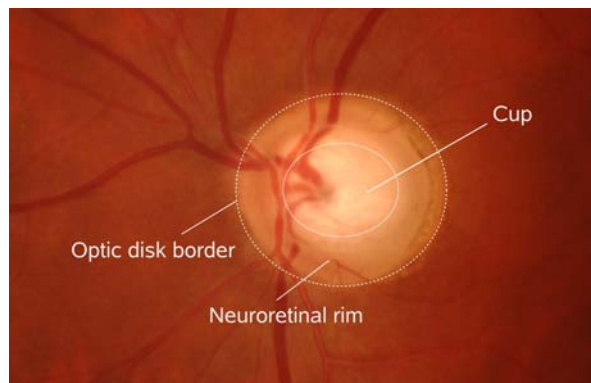The qualitative assessment of the ONH structure

Fig. 1. Major structures of the optic nerve head that are visible in color fundus photographs: The optic disk is margined by the optic disk border and can be divided into two major zones: (i) the neuroretinal rim is composed of astrocytes and nerve fibers while (ii) the brighter cup or excavation exclusively consists of supporting tissue.

and the functional abilities in addition with the patient's medical history and intraocular pressure are the common base for a reliable glaucoma diagnosis by ophthalmologists. This inherent subjectivity of the gained conclusion leads to a considerable inter- and intra-observer variability in differentiating between normal and glaucomatous ONHs (Varma et al. 1992).

However, quantitative parameters can help to make the qualitative assessment more objective, reproducible and lead to a reduction of the observer variability or to track glaucoma progression in patient follow-up. Such parameters can be gained manually or even by computer based technologies from several imaging modalities.

Stereoscopic images of the ONH are commonly used for documenting its cup shaped structure. Important ONH characteristics such as disk area, disk diameter, rim area, cup area or cup diameter can be extracted from the stereo image by planimetry (Betz et al. 1982) to also gain the well established cup-to-disk ratio. For the glaucomatous disease the cup-to-disk ratio measures the decrease of rim area while the disk area remains constant.

Although, this ratio is highly influenced by the disk size, it gives a general estimation whether the ONH shape is within its normal limits or it has to be considered conspicuous.

There exist several imaging modalities which provide quantitative parameters of the ONH in glaucoma: (i) Confocal Scanning Laser Ophthalmoscopy (CSLO), (ii) Scanning Laser Polarimetry (SLP) or (iii) Optical Coherence Tomography (OCT). CSLO,

commercially available as Heidelberg Retina Tomograph (HRT, Heidelberg Engineering, Heidelberg, Germany), provides a 2.5-dimensional topographic image of the ONH through the undilated pupil (Fig. 2(b)). After a manual outlining of the optic disk border, the device is able to generate geometric parameters such as the cup volume, cup depth, cup shape measure or even retinal height variations along the rim contour. Discriminant analysis (Moorfields Regression Analysis (MRA)) properly combines these geometric parameters. It shows a good classification performance that was validated by Miglior et al. (2003) and Wollstein et al. (1998). Due to the manual outlining of the optic disk border, the gained quantitative parameters are not fully objective. However, the Glaucoma Probability Score (GPS) of the latest version of HRT utilizes the parameters of a non-linear shape model of the topographic ONH shape for glaucoma classification. As this cup model is automatically fitted to original shape, this method overcomes the subjectivity of contour based methods while it shows a comparable overall diagnostic accuracy to the MRA (Burgansky-Eliash et al. 2007). In a series of consecutive acquisitions over years the progression of glaucomatous degeneration could be quantified. The temporal glaucomatous structural ONH changes are automatically located and quantified by the HRT Topographic Change Analysis (TCA) from consecutive acquisitions. Thus, this technique is promising to quantify a further progression of glaucoma (Chauhan et al. 2000).

Alongside the structural changing of the ONH, the degeneration of the nerve fibers is depicted by a thinning of the RNFL in the course of glaucoma disease, too. The thickness of the RNFL can be measured by SLP or OCT (Fig. 2(c)). In SLP the retina is illuminated by polarized light and RNFL thickness can be directly determined from the polarization change of the reflected light (Sehi et al. 2007). OCT provides complete depth profiles of the retina utilizing low-coherence interferometry of near infrared light. The desired structural RNFL can be extracted from these depth profiles utilizing manual or automated segmentation techniques (Fernández et al. 2005). From these thickness maps several global and sectoral geometric parameters such as average thickness, minimum thickness etc. can be extracted that differentiate between glaucomatous and control cases (Medeiros et al. 2004b).

Beside structural characteristics of the retina, the loss of functional capacities of the optic nerve is
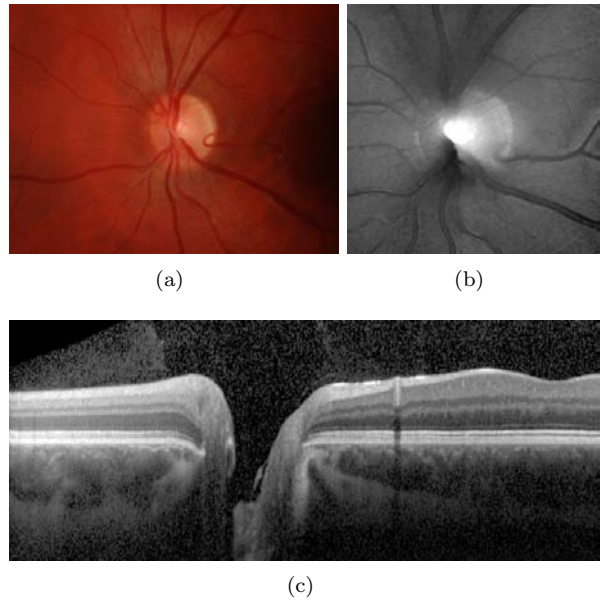

(a)　　　　　　　　　(b)

(c)

Fig. 2. Example images of the central retina: Optic nerve head (ONH) centered fundus photograph (a) is used for automated glaucoma detection by the proposed glaucoma risk index while glaucoma probability score utilizes HRT 2.5-dimensional topography images (b). OCT line scan (c) traversing the ONH illustrates different layers of the retina such as nerve fiber layer as the top one.

one major criteria for reliably diagnosing glaucoma. Modalities such as Standard Automated Perimetry (SAP), Short-wavelength Automated Perimetry (SWAP) or Frequency Doubling Technology (FDT) stimulate retinal regions to identify visual field defects. Quantitative medical parameters are derived from the number and the location of missing stimulated spots. Although this technology shows a high specificity for detecting glaucoma, it can only detect already occurred functional damage (Medeiros et al. 2004a).

Overall, these structural parameters of the optic nerve head are well established in the medical community and verified in several studies (Greaney et al. 2002; Sharma et al. 2008). As these parameters are derived from structural measurements to characterize structural glaucomatous changes they are very meaningful and intuitive to the physician.

We propose a data-driven framework that is able to extract a novel glaucoma parameter from ONH depicted on color fundus images. The proposed data mining technique is based on the idea of "Eigenimages" (Turk and Pentland 1991) that statistically analyzes the pixel input data to capture character-
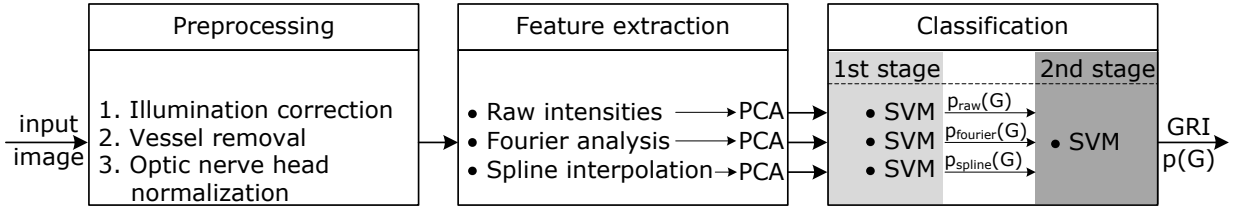
3

Fig. 3. Processing pipeline in detail: Glaucoma risk calculation consists of three steps: (i) Preprocessing eliminates the disease independent variations from the input image, (ii) Feature extraction transforms the preprocessed input data to characteristic and compact representation, and (iii) Classification generates the Glaucoma Risk Index (GRI) by a two-stage probabilistic SVM classifier.

istic variations that look promising for differentiating between glaucomatous optic nerve heads and controls. Due to the analysis of pure image data, the achieved parameters are influenced by the appearance of the whole ONH and not by structural measurements as it is the case for some established quantitative glaucoma parameters.

3. The concept of Eigenimages for glaucoma detection

Due to the high variability of the ONH appearance, the established determination of geometric ONH parameters utilized for glaucoma detection is difficult to automate.

We consider the described situation for automated glaucoma detection similar to that one stated by Turk and Pentland (1991) for early face detection methods. Before the 1990's face detection systems characterized faces by a set of geometric parameters such as normalized distances or ratios between characteristic facial landmarks (Bledsoe 1966). Based on this previous work Turk and Pentland (1991) concluded that the established methods "on automated face recognition [have] [. . .] ignored the issue of just what aspects of the face stimulus are important for identification". This time, they proposed an information theory approach that captures facial variations from a set of training images to gain a compact, but meaningful collection of parameters that are usable for classification purposes. As an information theory technique they proposed the Principal Component Analysis (PCA). This method models a linear transformation that projects the image space to a low dimensional feature space while a maximum of data variation is preserved. The captured variation between the images can then be represented by a set of Eigenimages. This so called appearance-based method is still consid-

ered as the baseline face recognition system.

In this contribution, the idea of appearance-based recognition is transferred to the domain of glaucoma recognition in order to get novel differentiating non-geometric parameters and to possibly gain new insights to the glaucoma disease. The major procedure illustrated in Fig. 3 consists of three steps:

(i) *Preprocessing*: The appearance-based techniques preserve the data variation in the low dimensional representation independent from its origin although it might not be related to the classification task. Variations such as illumination inhomogeneities are not linked to the glaucoma disease and have to be excluded from the image data beforehand.

(ii) *Feature extraction*: Beside the common Eigenimage approach on raw pixel intensities we propose further types of data representation in order to capture additional image information. These feature types are then compressed separately by PCA to gain a low dimensional image representation for classification.

(iii) *Classification*: In the last processing step, a probabilistic two-stage classifier scheme combines the different types of features to gain one single glaucoma prediction.

4. Image preprocessing

The proposed appearance-based approach analyzes the entire input image data to capture the glaucoma characteristics. To emphasize these desired characteristics in the input data, the variations not related to the glaucoma disease are excluded from the images in a preprocessing step. This includes variations due to image acquisition, such as inhomogeneous illumination or different optic nerve head

locations, but also retinal structures not directly related to glaucoma, e.g. the vessel tree.

Due to the reflection properties of the eye ground, the red channel of the fundus photos is often oversaturated (especially in the central, optic nerve head region), while the blue channel can be undersaturated and noisy. Therefore, we used the green channel for a proper image processing as only this channel shows a reliable saturation.

4.1. *Illumination correction*

The acquired images can be disturbed by bright speckles or inhomogeneous background e.g. due to different visual angle of the patients during image acquisition. Although, these interferences are not originated by glaucoma they affect the illumination of the ONH and would have an influence to the subsequent statistical analysis.

To avoid this behavior, we desire (i) a homogeneous lightning of the ONH and (ii) a similar illumination level among all images of the sample set. This can be achieved by global correction techniques (Youssif et al. 2006) applying a background correction. These methods subtract the estimated retinal background from the original image to gain a homogeneously illuminated fundus image. The estimation of a background can be done by average intensity filtering within a large neighborhood (Chrástek et al. 2005; Hoover et al. 2000) or polynomial surface fitting (Narasimha-Iyer et al. 2006).

We implemented a correction method similar to the one proposed by Narasimha-Iyer et al. (2006) as it does not require a time consuming image low pass filtering as in case of average intensity filtering. According to their lighting model, the observed intensity in each color channel of an input image $\mathbf{I} \in \mathbb{R}^{n \times m}$ is a pixel-wise product of a luminosity component $\mathbf{L}$ and a reflectance component $\mathbf{R}$. The first component is due to the source illumination, while the second one is related to the structures in the retina and their properties. Taking the logarithm, the pixel at (x, y) becomes an additive term:

$$
\begin{aligned}
\log(\mathbf{I}(x, y)) &= \log(\mathbf{L}(x, y) \cdot \mathbf{R}(x, y)) \\
&= \log(\mathbf{L}(x, y)) + \log(\mathbf{R}(x, y)).
\end{aligned} \quad (1)
$$

Low frequency changes of image intensities are considered as illumination inhomogeneities and can be well described by a 4th-order polynomial surface. This two dimensional polynomial is defined by the coefficients $\mathbf{c} \in \mathbb{R}^{15}$ which are determined by a least-square estimate from the following (weighted) linear equation system (Narasimha-Iyer et al. 2006):

$$
\log(\mathbf{i}) = (\mathbf{WS})\mathbf{c}, \quad (2)
$$

where $\mathbf{i}$ is the vector representation of $\mathbf{I}$. The matrix $\mathbf{S} \in \mathbb{R}^{N \times 15}$ with $N = n \cdot m$ stores the polynomial coefficients of the pixel locations. Retinal structures such as optic nerve head or vessel tree set apart from background and have to be excluded from the least square polynomial fitting. The diagonal matrix $\mathbf{W} \in \mathbb{R}^{N \times N}$ is used to mask out these regions from the computation. The diagonal elements are 1 where a pixel is considered as background and 0 for foreground structures.

We consider pixels having intensities between the 15th and 85th percentile of the image histogram as background while the remaining pixels mainly belong to retinal structures such as the bright ONH structure or the dark vessel structures. It is only of importance that the set of background pixels does not contain any structure pixels as only these pixels are utilizable to estimate the background. Therefore, the masking of the structures does not need to be very accurate.

The vector of the logarithmic reflectance component $\mathbf{r}$ is then recovered as

$$
\log(\mathbf{r}) = \log(\mathbf{i}) - \mathbf{Sc}. \quad (3)
$$

The reflectance component $\mathbf{R}$ is obtained by reshaping $\mathbf{r}_{\log}$ to matrix notation and transforming it back from logarithmic space. It shows a considerably reduced amount of illumination artifacts and intensity inhomogeneities (Fig. 4 second row).

4.2. *Vessel removal*

The glaucoma disease is mainly related to the optic nerve fibers and astrocytes. The vessel tree strongly varies among different patients, but the vessel location and the vessel diameters are minimally affected by glaucoma itself. The proposed appearance-based approach captures these variations of the used training sample set. In case of the utilization of images without removed vessel that would lead to the emphasis of the vessels and not of glaucoma (Meier et al. 2007). To avoid this behavior, the vessel structures in the eye ground are removed by (i) segmentation and subsequent (ii) inpainting of the detected vessel tree.

First, we perform a rough segmentation of the retinal vascular structure. In the literature, several automated vessel detection methods were proposed

over the last years (Niemeijer et al. 2004). Most of these techniques exploit the local appearance of the vessel including (i) edge-based (Al-Diri et al. 2009; Martinez-Perez et al. 2007; Sofka and Stewart 2006), template-matching (Ricci and Perfetti 2007; Sofka and Stewart 2006; Wang et al. 2007) or supervised approaches (Niemeijer et al. 2007; Ricci and Perfetti 2007; Soares et al. 2006; Staal et al. 2004) as well as a combination of those (Ricci and Perfetti 2007; Sofka and Stewart 2006). Based on the initial pixel-wise segmentation continuative methods ensure the global connectivity and topography of the vessel tree (Grisan et al. 2004; Niemeijer et al. 2009b). We combined edge-based and template-matching techniques to extract the vessel tree from the fundus images. Initially, we use an adaptive thresholding technique, wherein for each pixel, the median of its 15×15 neighborhood is taken as a threshold to separate foreground from background. The size of the neighborhood was determined to approximately match the size of the structures, i.e. the vessels. Then, the information of this mask and a Canny edge map (Canny 1986) is combined. This generated mask is filtered such that small objects are removed and only structures that are bounded by parallel running pairs of edges are kept. These potential vessel parts are validated by gridding and a matched filter technique (Can et al. 1999) where edge templates are applied to the grid points at different orientations and distances. A morphological closing of the valid regions yields the final vessel mask (Fig. 4 third row).

Second, an iterative inpainting technique, used in photo restoration and video processing (Bertalmio et al. 2000; Shen and Chan 2002), replaces the invalid pixels of the vessel mask by those of the neighborhood in a visually pleasing way. In our implementation, the vessel regions are iteratively filled layer by layer from outside inwards. The missing pixels become a distance weighted average of the already valid neighboring values. Finally, we gain a vessel-free photography of the optic nerve head (Fig. 4 fourth row).

4.3. *Optic nerve head normalization*

For glaucoma detection the ONH is one of the most important structures for observing glaucomatous characteristics. As known from face detection (Turk and Pentland 1991), the proposed appearance-based method requires at least rough
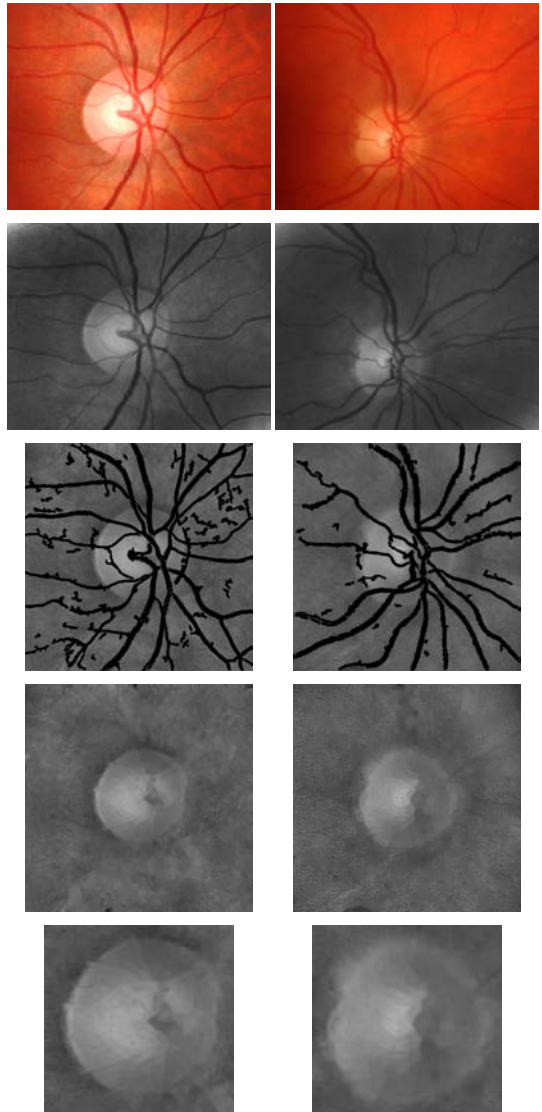

Fig. 4. Image preprocessing eliminates glaucoma disease independent variations and allows an appearance-based post-processing. Row-to-row: (i) Original fundus photos. (ii) Illumination corrected images with (iii) overlaid vessel mask which is then (iv) inpainted to hide the vessel tree. (v) Final normalized optic nerve head image used for feature extraction.

point correspondences to be able to gain a reasonable performance. Consequently, we normalize the rim area according to the optic nerve head border within distinct ranges of optic nerve head sizes.

First, the ONH rim has to be determined. In the literature, some methods restrict the segmented ONH rim to be circular or elliptical using e.g. Hough transform (Blanco et al. 2006; Zhu et al. 2009) while other techniques allow higher variabil-

6

ity of the ONH shape using parametric or free-form deformable models (Chrástek et al. 2005; Li and Chutatape 2003; Xu et al. 2007). The performance of the proposed segmentation techniques strongly rely on proper initialization with the location of the ONH: Proposed algorithms utilize ONH template-matching, intensity assumptions or the convergence of the vessel tree (Chrástek et al. 2005; Hoover and Goldbaum 2003; Lowell et al. 2004; Youssif et al. 2008). Besides, supervised techniques were also successfully applied for both purposes (Abràmoff et al. 2007; Merickel et al. 2007; Niemeijer et al. 2009a). As we are interested in a circular mapping of the ONH rims, we utilize the segmentation technique of Chrástek et al. (2005) that determines the circular rim contour of the ONH border as the first processing step: Assuming the ONH as the brightest spot in the fundus image, a center estimate is achieved by a strong intensity smoothing and further threshold probing. This estimation then restricts the subsequent circular Hough transform that is performed on the edge map to find the optic disk border.

Second, a square box of size three times the calculated ONH radius, centered at the ONH center is cropped and then scaled to the preprocessed image $\mathbf{P}$ of fixed reference size of 128×128 pixels (Fig. 4 last row).

This proposed procedure ensures images of the same dimension in addition with circularly mapped ONH rims that is required for a reasonable appearance-based feature computation.

5. Feature extraction

The performed image preprocessing emphasizes glaucomatous variations among the images and allows a generic and appearance-based feature extraction. The high dimensional preprocessed images $\mathbf{P}$ are statistically compressed by PCA to gain compact, and meaningful features $\mathbf{f}$. To capture complementary image information we propose three different generic image representations with different spatial and frequency resolution for feature extraction.

5.1. *Pixel intensity values*

Like the standard appearance-based approach (Turk and Pentland 1991), we serialized the preprocessed 2-dimensional images $\mathbf{P} \in \mathbb{R}^{128 \times 128}$ to an image vector $\mathbf{p} \in \mathbb{R}^{128 \cdot 128}$. This is then decomposed by a precalculated decomposition matrix to

low dimensional feature vector $\mathbf{f}_{\mathrm{raw}}$. According to Turk and Pentland (1991) this decomposition matrix was determined by an implicit computation of Eigen vectors from images of a training set. We have shown (Meier et al. 2007) that thirty principal components capture already at least $95\,\%$ of data variation. Therefore, the feature vector from raw pixel intensities was restricted to dimensionality of $\mathbf{f}_{\mathrm{raw}} \in \mathbb{R}^{30}$.

5.2. *FFT coefficients*

In contrast to the spatial pixel intensities, Fourier coefficients capture the image's global frequency information and are computable from an image by Fourier transform (FT). We apply a discrete version of FT, the Fast Fourier Transform (FFT), on the preprocessed image $\mathbf{P}$ and compress the real response of its Fourier coefficients by PCA to feature $\mathbf{f}_{\mathrm{fft}} \in \mathbb{R}^{30}$.

5.3. *B-spline coefficients*

In addition to the described feature types, B-spline coefficients decode spatial frequency information as they are defined by piecewise polynomials over a pixel neighborhood (Ibanez et al. 2005; Unser 1999; Unser et al. 1993a, b).

The discrete input image $\mathbf{P}$ of size $n \times m$ with $x = \{1, \ldots, n\}$ and $y = \{1, \ldots, m\}$ is transformed to B-spline coefficients $c(k, l)$ with $k = \{1, \ldots, n\}$ and $l = \{1, \ldots, m\}$.

$$p(x,y) = \sum_{k=0,l=0}^{n,m} c(k,l)\beta^d(x-k, y-l) \quad (4)$$

with d denoting the degree of the central B-spline $\beta^d(x, y)$. The $(d+1)$-fold convolution of the rectangular B-spline β^0 generates splines of higher degree which are symmetrical bell shaped (Unser 1999):

$$\beta^0(x,y) = \begin{cases} 1, & |(x,y)^T|_2 < \dfrac{1}{2} \\ \dfrac{1}{2}, & |(x,y)^T|_2 = \dfrac{1}{2} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

$$\beta^d(x,y) = \underbrace{\beta^0 \star \beta^0 \star \ldots \star \beta^0(x,y)}_{(d+1)\ \text{times}} \quad (6)$$

Because the number of input pixels $p(x, y)$ is equal to the used number of spline coefficients, equation

(4) defines a projection and no image information will be lost.

Our processing calculates coefficients $c(k,l)$ of B-splines of degree $n = 4$ from preprocessed images $\mathbf{P}$. The coefficients $\mathbf{C}$ are subsequently compressed by PCA and denoted as features $\mathbf{f}_{\text{spline}} \in \mathbb{R}^{30}$.

6. Classification

In the final classification step (Fig. 3), a glaucoma probability and the associated class label such as "glaucoma" or "not glaucoma" is computed from the three different feature types $\mathbf{f}$ that will be denoted as the *Glaucoma Risk Index (GRI)*.

6.1. *Classifier*

In general, classifiers achieve good results if their underlying separation model fits well to the distribution of the sample data.

In our previous investigations (Bock et al. 2007), we evaluated the glaucoma detection performance for three different kinds of classifiers in an appearance-based pipeline configuration. The Support Vector Machine classifier (SVM) shows similar results compared to other classifiers such as the naive Bayes classifier or k-nearest neighbor classifier. In this work, we decided to use the SVM as it is known that the SVM is less prone to a sparsely sampled feature space compared to the other classifiers as it is the case here. This SVM classifier determines a maximum-margin and soft hyperplane that best separates the considered classes in a kernel transformed feature space (Chen et al. 2005; Schölkopf et al. 2000).

A further improvement of classification performance could not be verified in our elaborate evaluations (Bock et al. 2007) that additionally utilized different classifier enhancement methods like AdaBoost or attribute selection in combination with SVM classifier. Therefore, a stand-alone SVM classification scheme is performed for the proposed classification purpose.

6.2. *Two-stage classification combines different feature types*

In order to be able to benefit from the complementary information captured by the three proposed feature types, they have to be combined. We propose a two-stage classification scheme that synthesizes one final result from them.

In the first stage, each feature type ($\mathbf{f}_{\text{raw}}$, $\mathbf{f}_{\text{fft}}$, $\mathbf{f}_{\text{spline}}$) obtained from feature extraction is classified separately. A probabilistic SVM classifier determines a probability $p(G)$ for the normalized PCA compressed feature (e.g. $p_{\text{raw}}(G)$ from raw pixel intensities $\mathbf{f}_{\text{raw}}$).

In the second stage, these probabilities are concatenated to one low-dimensional, common feature space by composing a feature vector:

$$\mathbf{f}_{\text{p}} = \big(p_{\text{raw}}(G), p_{\text{fft}}(G), p_{\text{spline}}(G)\big)^{\text{T}}. \qquad (7)$$

Probabilistic SVM now processes this generated vector of probabilities as feature $\mathbf{f}_{\text{p}}$ and outputs one common glaucoma probability $p(G)$.

7. Evaluation

Based on the presented fully automated processing procedure illustrated in Fig. 3, we achieved a novel probabilistic index that we refer to as *Glaucoma Risk Index* (GRI). In order to quantify its ability in detecting glaucoma from color fundus images the performance of GRI is first characterized in more detail by some key figures and a reliability analysis. Furthermore, its performance is compared to (i) that of glaucoma experts and (ii) to medically relevant and well established glaucoma parameters such as Glaucoma Probability Score (GPS) of the HRT III device (Swindale et al. 2000) and the cup-to-disk ratio (Betz et al. 1982).

Our evaluations showed that the proposed two-stage classification scheme outperforms configurations using one common single feature space and that it is competitive to common glaucoma parameters.

7.1. *Image data set and setup*

The used data set of fundus images was randomly selected from the Erlangen Glaucoma Registry (EGR) which contains more than 2,000 records of multi-modal fundus images of a long-term screening study. The gold standard diagnosis was given by an experienced ophthalmologist based on a complete ophthalmological examination with anamnesis, ophthalmoscopy, visual field test, intraocular pressure (IOP), and scanning laser tomography (Heidelberg Retina Tomograph, HRT II). The color fundus photos were acquired by a Kowa NonMyd alpha digital

fundus camera with a optic nerve head centered 22°
field of view and an image size of 1600 × 1216 pixels
(Fig. 9).

For the performance evaluation of the GRI we
used a data set of 575 ONH centered color fundus im-
ages from 358 persons with normal sized ONHs (av-
erage vertical ONH diameter: 1.8 ± 0.22 mm). The
mean age was 56.1 ± 11.4 years, 52.2 % female. The
samples had an unambiguous gold standard diagno-
sis with 239 glaucomatous images and 336 normals.
For comparison, the corresponding topography im-
ages were applied with HRT III device to calculate
the GPS. The linear cup-to-disk ratio was also ex-
tracted from HRT topography images based on the
manually outlined optic disk border. A subset of 240
images (160 normals, 80 glaucomatous) from the
data set was additionally evaluated by two ophthal-
mologists experienced in diagnosing glaucoma. The
findings were done on the fundus images exclusively
without using any anamnesis or further image data
to fairly rank the ability of fundus image based GRI.

We performed a 5-fold cross validation setup to
gain a robust statistical evaluation. Each fold con-
tained nearly the same number of subjects and a
similar ratio of glaucomatous versus controls. As
SVM classifier the libSVM implementation (EL-
Manzalawy and Honavar 2005) was utilized with
non-linear radial basis kernel and probabilistic out-
put. For the first stage classification, the PCA com-
pressed features $\mathbf{f}$ were normalized beforehand by
min-max normalization. The parameters of SVM
were manually optimized based on the classification
accuracy of the first fold.

For reliability analysis, we successively captured
images of 17 selected subjects. The consecutive ac-
quisitions were done at the same day, but with suffi-
cient relaxation time in-between to avoid a decline of
image quality due to miosis. For each eye three im-
ages were acquired and processed individually which
are referred to as one *image series*. Because of bad
image quality, seven of the 102 images had to be
excluded from the analysis as they showed obvious
processing failure such as missed vessel segmenta-
tion or wrong optic nerve head localization. Finally,
the analysis was done on 95 images linked to 32 eyes.

7.2. *Properties of the proposed Glaucoma Risk Index*

For the novel GRI, we proposed a two-stage clas-
sification scheme to combine the different feature
types ($\mathbf{f}_{raw}$, $\mathbf{f}_{fft}$, $\mathbf{f}_{spline}$). In principle, two alternative

Table 1
Performance of proposed classifiers indicated by classifica-
tion accuracy, area under Receiver Operating Characteris-
tic curve (AUC), sensitivity and specificity (%) for detect-
ing glaucoma: The performance of the proposed two-stage
Glaucoma Risk Index (GRI) is compared to (i) single-feature
classification ($\mathbf{f}_{raw}$, $\mathbf{f}_{fft}$, $\mathbf{f}_{spline}$), (ii) single-stage classifica-
tion setup and (iii) human experts and established quanti-
tative glaucoma parameters i.e. Glaucoma Probability Score
(GPS) and cup-to-disk ratio (CDR).
The p-values denote the statistical significance for a different
ROC compared to GRI: * $p \leq 0.05$.

Setup	Accuracy	AUC	Sensitivity	Specificity	p-value
GRI	**80**	**88**	**73**	**85**	–
$\mathbf{f}_{raw}$	80	87	69	88	0.07
$\mathbf{f}_{fft}$	79	86	71	85	0.01*
$\mathbf{f}_{spline}$	81	88	69	90	0.46
single-stage	79	86	70	85	0.08
expert 1	83	–	54	97	–
expert 2	83	–	51	98	–
GPS	78	87	88	72	0.96
CDR	68	88	93	50	0.51

configurations are also possible: (i) *Single-feature
classification*: Instead of combining the three differ-
ent feature types, only one single feature type is uti-
lized that provides the best detection performance.
(ii) *Single-stage classification*: To combine different
features, these feature are concatenated to one sin-
gle high dimensional feature space. This is then re-
duced by attribute selection to gain an expedient
feature dimensionality that can directly be used for
classification.

In this section, the properties of the GRI are char-
acterized in more detail by key numbers and a reli-
ability analysis in order to show the advantages of
the proposed two-stage scheme compared to alter-
native classification setups. For calculation of accu-
racy, sensitivity and specificity the decision thresh-
old was fixed at a 0.5 level.

7.2.1. *Performance figures*

The Receiver Operating Characteristic (ROC)
curve shows classifier performance for different
decision thresholds. Therefore, it provides infor-
mation on how to tune the decision threshold in
order to achieve the best tradeoff between sensi-
tivity and specificity for the desired application. A
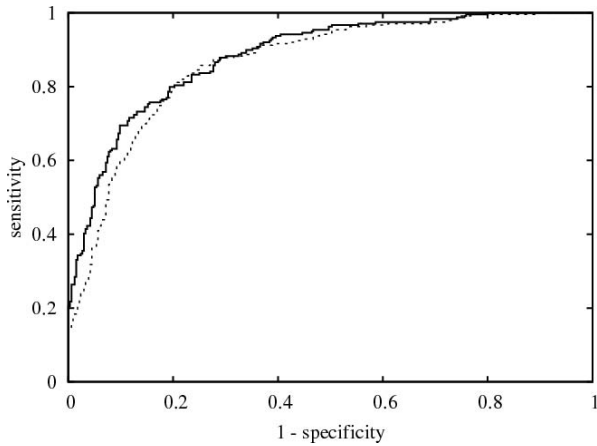non-parametric test based on Mann-Whitney-U-

Fig. 5. Receiver operating characteristic (ROC) curves for detecting glaucoma. In comparison to the ROC curve of the single-stage classification (- -), the ROC curve of GRI (—) shows higher sensitivities for specificities lower than 0.8. This improvement might be due to the repeated usage of class information that allows the determination of boundaries better separating the two classes.

statistics (DeLong et al. 1988; Vergara et al. 2008) is used to show the significance of the ROC curve differences (statistical significance level: $p \leq 0.05$).

The proposed two-stage GRI gains an area under ROC curve (AUC) of 88 % (Table 1) and a sensitivity of 73 % at a specificity of 85 % (Fig. 5).

Comparing the single-feature classifications among themselves, the classifications on $\mathbf{f}_{fft}$ or $\mathbf{f}_{raw}$ with an AUC of 86 % or 87 % respectively nearly reach the performance on $\mathbf{f}_{spline}$ with an AUC of 88 %. While the ROC curve from $\mathbf{f}_{spline}$ ($p = 0.46$) is comparable to the one of GRI, the performance of the remaining single-feature classifications show at least a trend to be significant inferior ($p \leq 0.07$).

The single-stage classification gains an AUC of 86 % and a sensitivity of 70 %. This setup is not competitive to GRI as the ROC curves of both setups are nearly significant different ($p \leq 0.08$). Especially, for high specificities, the sensitivity is decreasing as it is illustrated by the ROC curves shown in Fig. 5.

7.2.2. *Distribution of glaucoma probabilities*

The distribution of gained glaucoma probabilities $P(G)$ (specified by the histogram of $p(G)$ from the sample set) expresses the separability of the two different classes and the classifier's certainty (Fig. 6). As the data set (Section 7.1) consists of cases with a definitive diagnosis and thus without any suspicious cases, we expect an undoubtful classification with an high confidence level of the classification.

With a high confidence level (i.e., $p(G) \leq 0.1$
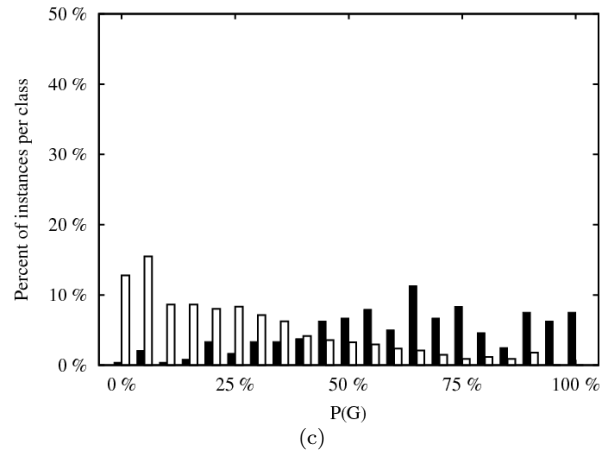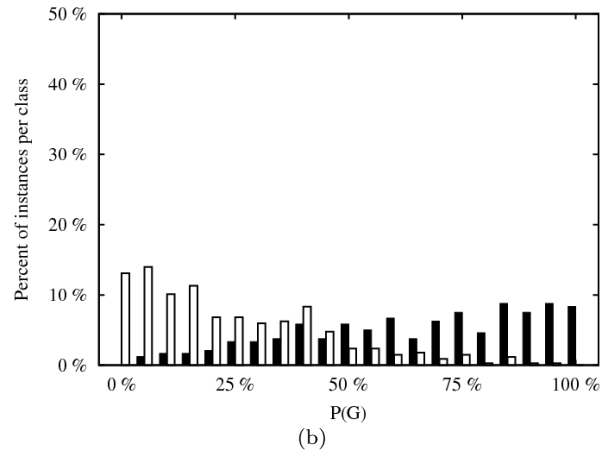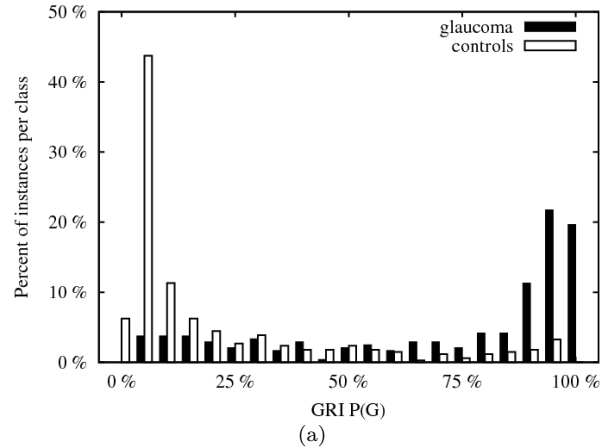

(a)

(b)

(c)

Fig. 6. The probability distributions $P(G)$ of control and glaucomatous cases. The two-stage GRI setup (a) shows two compact distributions with distinct peaks at the borders. This reflects the definitive disease stages of our samples. The two other configurations, namely single-feature classification represented by $\mathbf{f}_{spline}$ (b) and single-stage (c) are characterized by undesired widely scattered distributions.

10

or $p(G) \geq 0.9$) 58.6 % of the controls and 48.3 % of the glaucomatous cases are correctly classified with GRI. The histogram of the calculated glaucoma probabilities $p(G)$ (Fig. 6(a)) illustrates the desired distinct peaks at the borders. As representation for single-feature classification setup, we show the histogram for feature $\mathbf{f}_{\text{spline}}$ that provides a classification performance comparable to GRI. However, a correct classification at a high confidence level is achieved in only 31.5% of the controls and 23.3 % of the glaucomatous cases (Fig. 6(b)). Similar results are gained in case of a common single-stage classification where a high confidence level is only reached for 33.6 % of the controls and 17.8 % of the glaucomatous (Fig. 6(c)). These lower confidence levels are also reflected by their histograms of the calculated glaucoma probabilities $p(G)$ which do not show the desired instance accumulation around $p(G) = 0$ for the control class or $p(G) = 1$ for glaucomatous cases respectively. Although, the majority of clearly diagnosed subjects were correctly classified the distribution is widely scattered.

In conclusion, the proposed two-stage classification assembly of the GRI shows a superior performance compared to the alternative configurations. Only for $\mathbf{f}_{\text{splines}}$ the classification performance was comparable. However, only the GRI exclusively provides a high classification confidence level for the majority of the instances that well reflects the present definitive stages of disease of our data set. For this reason, we consider the two-stage classification setup of GRI as the standard procedure that is further validated in a reliability analysis and in comparison to other quantitative glaucoma parameters.

7.2.3. *Reliability*

Beside the system performance, the reliability of the proposed algorithm is also a relevant issue. Only if the system gains similar results for the same patient but from different acquisitions is it usable for reliable screening applications.

As quantitative reliability measure, we report the maximal absolute deviation from the median of the GRIs calculated from the three images (i.e. image series) acquired from the same subject and eye.

Figure 7 shows the histogram of the deviations with respect to the glaucoma probability which indicates that for the majority of the image series the intra-subject deviation is low. In 19 of 32 series (59 %) the maximal deviation was less than 2 % and the largest reported deviation was 11 % for one
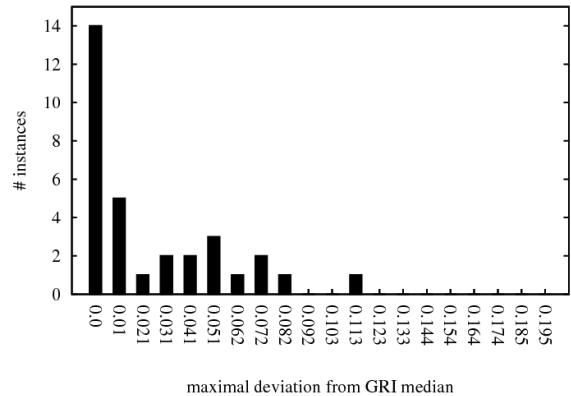

Fig. 7. Reliability measure as maximal absolute deviation from the median of the GRIs calculated from fundus images acquired in series. The occurrence of deviations illustrated by the histogram indicates that for the majority of the samples the intra-subject deviation is less than 2 %.

single series. Due to these reported deviations, we consider the processing procedure and classification scheme reliable.

7.3. *Comparing GRI to experts and glaucoma parameters*

Based on the preceding evaluations the GRI gains a reliable glaucoma detection performance with an AUC of 88 %. A comparison of GRI with (i) the rating of human experts on sole fundus images and (ii) established quantitative glaucoma parameters in the following section shows a competitive detection performance.

7.3.1. *Expert performance*

In contrast to the usual clinical glaucoma diagnosis utilizing medical history, different imaging modalities and preceding examinations, the glaucoma experts investigated the color fundus images on its own. This setup ensures a fair performance comparison as it provides the experts the same information that is also utilized by GRI.

Both experts classified 83 % of the instances correctly that is slightly superior compared to GRI (Table 1). Their performance is highly specific while still reaching a moderate sensitivity over 50 %. By putting theses values in relation to the ROC curve (Fig. 8), it is obvious that the human experts outperform the GRI and also GPS, however on a moderate level of sensitivity.

Based on the evaluation of digital color fundus images which is not conform to the daily clinical work-
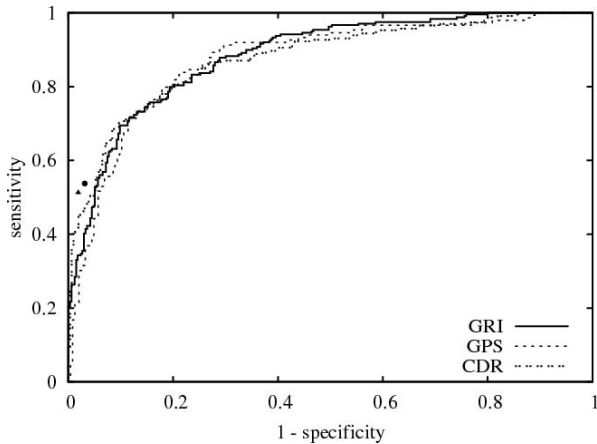
Fig. 8. The Receiver operating characteristic (ROC) curve of GRI (—) in comparison: (i) human experts (●, ▲) on single fundus images without medical history are very specific while the sensitivity reaches a moderate level. (ii) The GRI shows a competitive performance compared to Glaucoma Probability Score (GPS) and linear cup-to-disk ratio (CDR) in particular for a specificity around 0.8.

flow, our proposed GRI almost reaches the performance of the human experts.

7.3.2. *Glaucoma Probability Score (GPS)*

The HRT III software offers a module to determine the GPS from 2.5-dimensional topographic images of the ONH gained by scanning laser tomography technique. The score is established in the clinical workflow and was validated by several trials (Alencar et al. 2008; Burgansky-Eliash et al. 2007). The GPS is an objective measurement as it does not depend on manual outlining of the optic disk border and can be considered as a de facto standard in automated glaucoma detection from HRT topography images.

On our data set, the GPS achieves a classification accuracy of 78 % and an AUC of 87 % (Table 1). For glaucoma detection we get a sensitivity of 88 % at a specificity of 72 %.

The AUC for the GPS is similar to GRI (p = 0.96). This is also reflected by the ROC curves (Fig. 8) which behave quite similar especially for a specificity around 0.8.

Although, GRI and GPS were calculated from different imaging modalities, they show comparable detection performance.

7.3.3. *Cup-to-disk ratio*

The cup-to-disk ratio measures the ratio of the diameter of the cup to the diameter of the entire disk.

Commonly, this ratio can be determined by planimetry (Betz et al. 1982) from color fundus images by outlining the optic disk border and the cup. The HRT provides an equivalent ratio, the linear cup-to-disk ratio, from topography images that can be determined from manually outlined optic disk rim. In order to avoid the manual outlining of the cup, we compare the GRI to the linear cup-to-disk ratio of HRT.

On our data set, the cup-to-disk ratio, achieves a similar AUC of 88 % compared to GRI, however the accuracy drops down to 68 % assuming a decision threshold of 0.5 (Table 1). The ROC curves of cup-to-disk ratio and GRI (Fig. 8) are comparable with p = 0.51.

In summary, the GRI provides a detection performance on color fundus photographs that is competitive to human experts on single fundus images, the topography based GPS and cup-to-disk ratio. In particular for specificities around 80 % the markers show a competitive sensitivity. Thus, the novel GRI provides a new tool to automatically detect glaucoma from the low priced and widespread fundus camera.

8. Discussion

A reliable and competitive scheme for an automatic glaucoma detection was presented. The majority of the subjects were correctly classified as already shown. To give a better impression on the classification outcome, Fig. 9 shows examples of correctly classified and misclassified fundus images.

Figure 9 (a)-(c) illustrates correctly classified controls characterized by a typical small cup. In contrast, an expanded cup denotes the correctly classified images of glaucomatous eyes (Fig. 9 (d)-(f)) which corresponds to the common glaucoma disease pattern.

Some images of control and glaucomatous ONH were misclassified. In case of misclassified controls, shown in Fig. 9 (g)-(i), we suppose that the GRI calculation procedure might be misled by the pale neuroretinal rim or an increased disk size with increased cup area. The misclassified glaucomatous cases (Fig. 9 (j)-(l)) are marked by a low contrast between rim and cup and a greater optic disk size.

As described in Section 4.3, we perform a downsampling of larger ONH during normalization step to gain well mapped tissue structures that are required for appearance-based recognition. We are
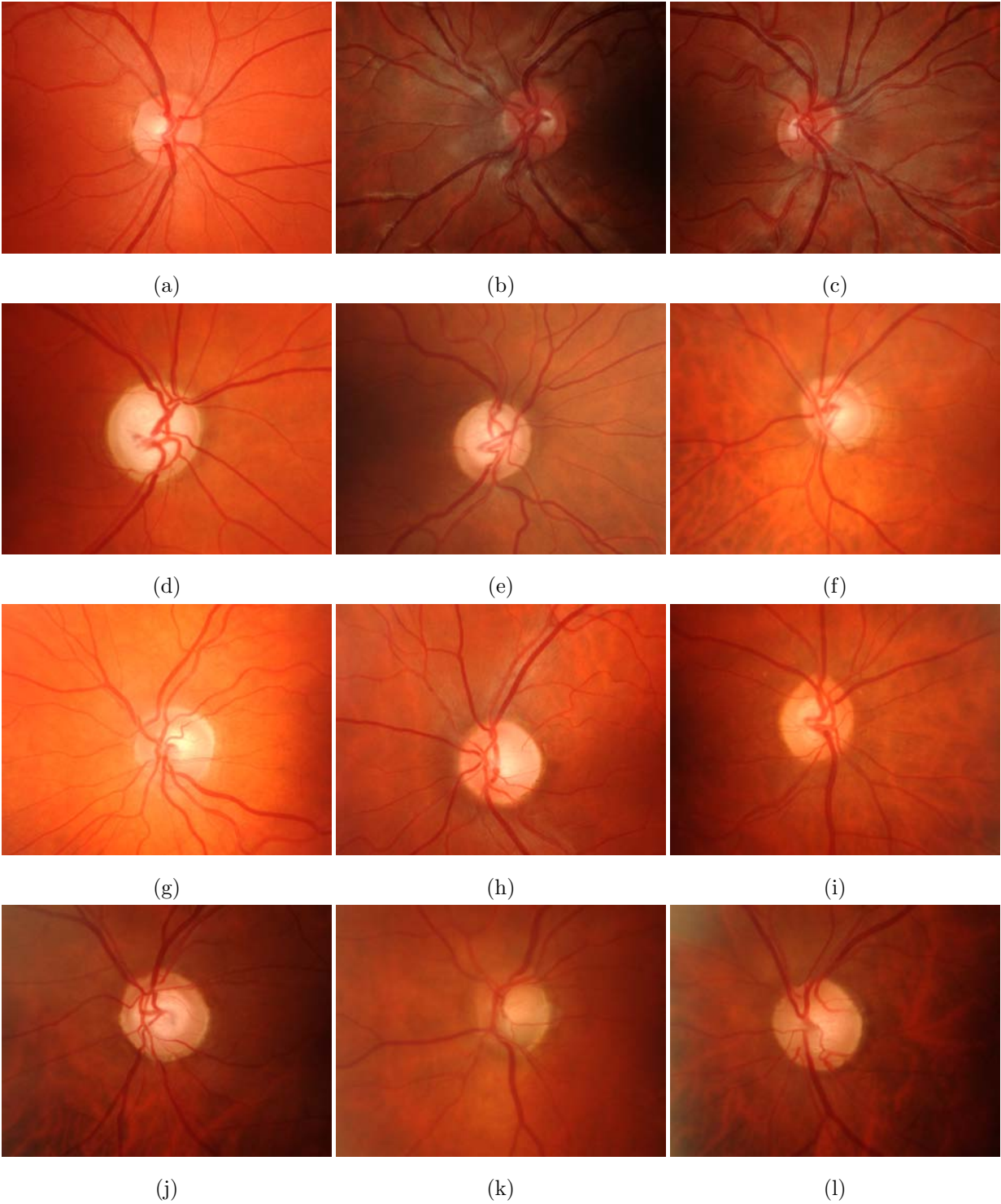
Fig. 9. Example images automatically assessed by GRI: (a)-(c) correctly classified controls, (d)-(f) correctly classified glaucomatous cases, (g)-(i) misclassified controls, (j)-(l) misclassified glaucomatous cases.

aware that this normalization can bias the pure glaucomatous variations and may decrease the classification performance. To account for this effect, we utilized a data set of limited ONH diameter variations (average vertical ONH diameter: 1.8 ± 0.22 mm) and of one race. Considering the wrong classified images (Fig. 9 (g)-(l)) it seems as if the GRI is sensitive to larger optic disk diameters. However, the GRI classification is able to properly handle different optic disk sizes as the correctly classified images (Fig. 9 (a)-(f)) show highly varying optic disk diameters. In order to apply the GRI in a widespread study, we propose to generate a separate model for each cluster of different ONH diameters and races.

As the GRI procedure is applied on color fundus images, the index can only capture glaucomatous signs that are visible on these images. Thus, it can only provide first indicators for the multilayered and progressive glaucoma disease. However, the digital color fundus camera is much cheaper than other ophthalmic imaging devices such as HRT or OCT. In the future, the GRI can provide a first low-priced glaucoma indication in order to possibly reduce the amount of false positives misrouted to the cost-intensive elaborate clinical investigations.

The proposed appearance-based technique known from face recognition statistically analyzes the entire image pixel pattern to emerge the competitive GRI and does not rely on accurately determined geometric parameters. Due to this approach, the GRI only provides a general statement, but does not indicate conspicuous retinal regions that might hinder its acceptance in clinical environment.

The required preprocessing of the fundus images performs a rough determination of the retinal vessel tree and the ONH rim for normalization. An accurate segmentation of the cup rim as for automated determination of cup-to-disk ratio can be omitted. Based on the shown high reliability, we conclude that the appearing unavoidable variations originated by segmentation only lead to negligible variations of the GRI.

9. Conclusion

Based on color fundus photos of the eye, the presented procedure for robust and reliable appearance-based feature extraction allows the automated quantification of the probability to suffer from glaucoma disease.

Due to glaucoma specific preprocessing and the appropriate combination of generic features, we are able to successfully apply the generic data-driven approach for this medical classification task. The proposed two-stage classification scheme helps to combine classifiers of different image inputs. The evaluation showed that this assembly improves the certainty of the classification and it makes the final decision more robust. Established methods early reduce the amount of feature dimensionality by using parametric models or structural measurements. In contrast, we compress the whole image data into discriminating features.

The obtained Glaucoma Risk Index (GRI) reached a classification accuracy of 80% in a two class problem (control vs. glaucomatous eyes) taking a gold standard diagnosis by ophthalmologists as a basis. The AUC was 88 % with a sensitivity of 73 % at a specificity of 85 % in detecting glaucoma. The results in our evaluation with 575 images were comparable to the commercial and established Glaucoma Probability Score (GPS) of the HRT III and cup-to-disk ratio.

Overall, this contribution provides a competitive, reliable and probabilistic glaucoma risk index from images of the low-cost digital color fundus camera as its performance is comparable to medical relevant glaucoma parameters. This proves, data-driven GRI is able to extract relevant glaucoma features. In the future, it might give a first, low-cost glaucoma indication to route the patients to more elaborate clinical trials only if necessary.

References

1 Abràmoff, M. D., Alward, W. L. M., Greenlee, E. C., Shuba, L., Kim, C. Y., Fingert, J. H., Kwon, Y. H., 2007. Automated segmentation of the optic disc from stereo color photographs using physiologically plausible features. Invest Ophthalmol Vis Sci 48 (4), 1665–1673.

2 Al-Diri, B., Hunter, A., Steel, D., 2009. An active contour model for segmenting and measuring retinal vessels. IEEE Trans Med Imaging 28 (9), 1488–1497.

3 Alencar, L. M., Bowd, C., Weinreb, R. N., Zangwill, L. M., Sample, P. A., Medeiros, F. A., 2008. Comparison of HRT-3 glaucoma probability score and subjective stereophotograph assessment for prediction of progression in glaucoma. Invest Ophthalmol Vis Sci 49 (5), 1898–1906.

4 Bertalmio, M., Sapiro, G., Caselles, V., Ballester, C., 2000. Image inpainting. In: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 2000, New Orleans, USA. pp. 417–424.

5 Betz, P., Camps, F., Collignon-Brach, J., Lavergne, G., Weekers, R., 1982. Biometric study of the disc cup in open-angle glaucoma. Graefes Arch Clin Exp Ophthalmol 218 (2), 70–74.

6 Blanco, M., Penedo, M. G., Barreira, N., Penas, M., Carreira, M. J., 2006. Localization and extraction of the optic disc using the fuzzy circular Hough transform. Lecture Notes in Computer Science 4029, 712–721.

7 Bledsoe, W. W., 1966. The model method in facial recognition. Tech. rep., Panoramic Research Inc., Palo Alto, CA, Rep. PRI:15.

8 Bock, R., Meier, J., Michelson, G., Nyúl, L. G., Hornegger, J., 2007. Classifying glaucoma with image-based features from fundus photographs. In: 9th Annual Symposium of the German Association for Pattern Recognition, DAGM. Lecture Notes in Computer Science (LNCS). Vol. 4713/2007. Berlin, pp. 355–365.

9 Burgansky-Eliash, Z., Wollstein, G., Bilonick, R. A., Ishikawa, H., Kagemann, L., Schuman, J. S., 2007. Glaucoma detection with the Heidelberg Retina Tomograph 3. Ophthalmology 114 (3), 466–471.

10 Can, A., Shen, H., Turner, J. N., Tanenbaum, H. L., Roysam, B., 1999. Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms. IEEE Trans Inf Technol Biomed 3 (2), 125–138.

11 Canny, J. F., 1986. A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 8 (6), 679–698.

12 Chauhan, B. C., Blanchard, J. W., Hamilton, D. C., LeBlanc, R. P., Mar 2000. Technique for detecting serial topographic changes in the optic disc and peripapillary retina using scanning laser tomography. Invest Ophthalmol Vis Sci 41 (3), 775–782.

13 Chen, P. H., Lin, C. J., Schölkopf, B., 2005. A tutorial on ν-support vector machines. Applied Stochastic Models in Business and Industry 21 (2), 111–136.

14 Chrástek, R., Wolf, M., Donath, K., Niemann, H., Paulus, D., Hothorn, T., Lausen, B., Lämmer, R., Mardin, C., Michelson, G., 2005. Automated segmentation of the optic nerve head for diagnosis of glaucoma. Med Image Anal 9 (4), 297–314.

15 DeLong, E., DeLong, D., Clarke-Pearson, D., 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. Biometrics 44, 837–845.

16 EL-Manzalawy, Y., Honavar, V., 2005. WLSVM: Integrating LibSVM into Weka Environment. Software available at http://www.cs.iastate.edu/~yasser/wlsvm.

17 Fernández, D. C., Salinas, H. M., Puliafito, C. A., 2005. Automated detection of retinal layer structures on optical coherence tomography images. Opt Express 13 (25), 10200–10216.

18 Greaney, M. J., Hoffman, D. C., Garway-Heath, D. F., Nakla, M., Coleman, A. L., Caprioli, J., 2002. Comparison of optic nerve imaging methods to distinguish normal eyes from those with glaucoma. Invest Ophthalmol Vis Sci 43 (1), 140–145.

19 Grisan, E., Pesce, A., Giani, A., Foracchia, M., Ruggeri, A., 2004. A new tracking system for the robust extraction of retinal vessel structure. In: Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE. Vol. 1. pp. 1620–1623.

20 Hoover, A., Goldbaum, M., 2003. Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. IEEE Trans Med Imaging 22 (8), 951–958.

21 Hoover, A., Kouznetsova, V., Goldbaum, M., 2000. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. IEEE Trans Med Imaging 19 (3), 203–210.

22 Ibanez, L., Schroeder, W., Ng, L., Cates, J., 2005. The ITK Software Guide. Kitware, Inc. ISBN 1-930934-15-7, http://www.itk.org/ItkSoftwareGuide.pdf, 2nd Edition.

15

23 Klein, B. E., Klein, R., Sponsel, W. E., Franke, T., Cantor, L. B., Martone, J., Menage, M. J., 1992. Prevalence of glaucoma. The Beaver Dam Eye Study. Ophthalmology 99 (10), 1499–1504.

24 Li, H., Chutatape, O., 2003. Boundary detection of optic disk by a modified ASM method. Pattern Recognition 36 (9), 2093–2104.

25 Lin, S. C., Singh, K., Jampel, H. D., Hodapp, E. A., Smith, S. D., Francis, B. A., Dueker, D. K., Fechtner, R. D., Samples, J. S., Schuman, J. S., Minckler, D. S., 2007. Optic nerve head and retinal nerve fiber layer analysis: a report by the American Academy of Ophthalmology. Ophthalmology 114 (10), 1937–1949.

26 Lowell, J., Hunter, A., Steel, D., Basu, A., Ryder, R., Fletcher, E., Kennedy, L., 2004. Optic nerve head segmentation. IEEE Trans Med Imaging 23 (2), 256–264.

27 Martinez-Perez, M. E., Hughes, A. D., Thom, S. A., Bharath, A. A., Parker, K. H., 2007. Segmentation of blood vessels from red-free and fluorescein retinal images. Med Image Anal 11 (1), 47–61.

28 Medeiros, F. A., Sample, P. A., Weinreb, R. N., 2004a. Frequency doubling technology perimetry abnormalities as predictors of glaucomatous visual field loss. Am J Ophthalmol 137 (5), 863–871.

29 Medeiros, F. A., Zangwill, L. M., Bowd, C., Weinreb., R. N., 2004b. Comparison of the GDx VCC scanning laser polarimeter, HRT II confocal scanning laser ophthalmoscope, and stratus OCT optical coherence tomograph for the detection of glaucoma. Arch Ophthalmol 122 (6), 827–837.

30 Meier, J., Bock, R., Michelson, G., Nyúl, L. G., Hornegger, J., 2007. Effects of preprocessing eye fundus images on appearance based glaucoma classification. In: 12th International Conference on Computer Analysis of Images and Patterns, CAIP. Lecture Notes in Computer Science (LNCS). Vol. 4673/2007. Berlin, pp. 165–173.

31 Merickel, M. B. J., Abràmoff, M. D., Sonka, M., Wu, X., 2007. Segmentation of the optic nerve head combining pixel classification and graph search. Proceedings of SPIE 6512 (1), 651215.

32 Michelson, G., Wärntges, S., Hornegger, J., Lausen, B., 2008. The papilla as screening parameter for early diagnosis of glaucoma. Dtsch Arztebl Int 105 (34-35), 583–589.

33 Miglior, S., Guareschi, M., Albe', E., Gomarasca, S., Vavassori, M., Orzalesi, N., 2003. Detection of glaucomatous visual field changes using the Moorfields regression analysis of the Heidelberg retina tomograph. Am J Ophthalmol 136 (1), 26–33.

34 Narasimha-Iyer, H., Can, A., Roysam, B., Stewart, C. V., Tanenbaum, H. L., Majorovics, A., Singh, H., 2006. Robust detection and classification of longitudinal changes in color retinal fundus images for monitoring diabetic retinopathy. IEEE Trans Biomed Eng 53 (6), 1084–1098.

35 Niemeijer, M., Abràmoff, M., van Ginneken, B., 2007. Segmentation of the optic disc, macula and vascular arch in fundus photographs. IEEE Trans Med Imaging 26 (1), 116–127.

36 Niemeijer, M., Abràmoff, M. D., van Ginneken, B., 2009a. Fast detection of the optic disc and fovea in color fundus photographs. Med Image Anal 13 (6), 859–870.

37 Niemeijer, M., Staal, J., van Ginneken, B., Loog, M., Abràmoff, M. D., 2004. Comparative study of retinal vessel segmentation methods on a new publicly available database. In: Proceedings of SPIE. Vol. 5370. p. 648.

38 Niemeijer, M., van Ginneken, B., Abràmoff, M. D., 2009b. A linking framework for pixel classification based retinal vessel segmentation. In: Proceedings of SPIE. Vol. 7262. p. 726216.

39 Ricci, E., Perfetti, R., 2007. Retinal blood vessel segmentation using line operators and support vector classification. IEEE Trans Med Imaging 26 (10), 1357–1365.

40 Schölkopf, B., Smola, A. J., Williamson, R. C., Bartlett, P. L., 2000. New support vector algorithms. Neural Comput 12 (5), 1207–1245.

41 Sehi, M., Guaqueta, D. C., Feuer, W. J., Greenfield, D. S., 2007. Scanning laser polarimetry with variable and enhanced corneal compensation in normal and glaucomatous eyes. Am J Ophthalmol 143 (2), 272–279.

42 Sharma, P., Sample, P. A., Zangwill, L. M., Schuman, J. S., Nov 2008. Diagnostic tools for glaucoma detection and management. Surv Ophthalmol 53 Suppl1, S17–S32.

43 Shen, J., Chan, T. F., 2002. Mathematical models for local nontexture inpaintings. SIAM J Appl Math 62 (3), 1019–1043.

44 Soares, J. V., Leandro, J. J., Jr, R. M. C., Jelinek, H. F., Cree, M. J., 2006. Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification. IEEE Trans Med Imaging 25 (9), 1214–1222.

45 Sofka, M., Stewart, C. V., 2006. Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures. IEEE Trans Med Imaging 25 (12), 1531–1546.

46 Staal, J., Abràmoff, M., Niemeijer, M., Viergever,

M., van Ginneken, B., 2004. Ridge-based vessel segmentation in color images of the retina. IEEE Trans Med Imaging 23 (4), 501–509.

47 Swindale, N. V., Stjepanovic, G., Chin, A., Mikelberg, F. S., 2000. Automated analysis of normal and glaucomatous optic nerve head topography images. Invest Ophthalmol Vis Sci 41 (7), 1730–1742.

48 Turk, M., Pentland, A., 1991. Eigenfaces for recognition. J Cogn Neurosci 3 (1), 71–86.

49 Unser, M., 1999. Splines: A perfect fit for signal and image processing. IEEE Signal Processing Magazine 16 (6), 22–38.

50 Unser, M., Aldroubi, A., Eden, M., 1993a. B-spline signal processing. I. Theory. IEEE Trans Signal Process 41 (2), 821–833.

51 Unser, M., Aldroubi, A., Eden, M., 1993b. B-spline signal processing. II. Efficiency design and applications. IEEE Trans Signal Process 41 (2), 834–848.

52 Varma, R., Steinmann, W. C., Scott, I. U., 1992. Expert agreement in evaluating the optic disc for glaucoma. Ophthalmology 99 (2), 215–221.

53 Vergara, I., Norambuena, T., Ferrada, E., Slater, A., Melo, F., 2008. StAR: a simple tool for the statistical comparison of ROC curves. BMC Bioinformatics 9, 265–269.

54 Wang, L., Bhalerao, A., Wilson, R., 2007. Analysis of retinal vasculature using a multiresolution hermite model. IEEE Transactions on Medical Imaging 26 (2), 137–152.

55 Wollstein, G., Garway-Heath, D. F., Hitchings, R. A., 1998. Identification of early glaucoma cases with the scanning laser ophthalmoscope. Ophthalmology 105 (8), 1557–1563.

56 Xu, J., Chutatape, O., Sung, E., Zheng, C., Kuan, P. C. T., 2007. Optic disk feature extraction via modified deformable model technique for glaucoma analysis. Pattern recognition 40 (7), 2063–2076.

57 Youssif, A. A., Ghalwash, A. Z., Ghoneim, A., 2008. Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter. IEEE Transactions on Medical Imaging 27 (1), 11–18.

58 Youssif, A. A. A., Ghalwash, A. Z., Ghoneim, A. S., 2006. Comparative study of contrast enhancement and illumination equalization methods for retinal vasculature segmentation. In: Proceedings of the Third Cairo International Biomedical Engineering Conference (CIBEC'06). pp. 1–5.

59 Zhu, X., Rangayyan, R., Ells, A., 2009. Detection of the optic nerve head in fundus images of the retina using the hough transform for circles. Journal of Digital Imaging published online: February, 24th 2009, `http://dx.doi.org/10.1007/s10278-009-9189-5`.