

University of Dundee

Accurate and reliable segmentation of the optic disc in digital fundus images

Giachetti, Andrea; Ballerini, Lucia; Trucco, Emanuele

Published in:
Journal of Medical Imaging

DOI:
[10.1117/1.JMI.1.2.024001](https://doi.org/10.1117/1.JMI.1.2.024001)

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):
Giachetti, A., Ballerini, L., & Trucco, E. (2014). Accurate and reliable segmentation of the optic disc in digital fundus images. *Journal of Medical Imaging*, 1(2), Article 024001. <https://doi.org/10.1117/1.JMI.1.2.024001>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

Accurate and reliable segmentation of the optic disc in digital fundus images

Andrea Giachetti
Lucia Ballerini
Emanuele Trucco

SPIE.

Accurate and reliable segmentation of the optic disc in digital fundus images

Andrea Giachetti,^{a,*} Lucia Ballerini,^b and Emanuele Trucco^b

^aUniversità di Verona, Dipartimento di Informatica, Strada Le Grazie 15 Verona 37134, Italy

^bUniversity of Dundee, VAMPIRE, School of Computing, School of Computing, Queen Mother Building, Balfour Street, Dundee DD1 4HN, United Kingdom

Abstract. We describe a complete pipeline for the detection and accurate automatic segmentation of the optic disc in digital fundus images. This procedure provides separation of vascular information and accurate inpainting of vessel-removed images, symmetry-based optic disc localization, and fitting of incrementally complex contour models at increasing resolutions using information related to inpainted images and vessel masks. Validation experiments, performed on a large dataset of images of healthy and pathological eyes, annotated by experts and partially graded with a quality label, demonstrate the good performances of the proposed approach. The method is able to detect the optic disc and trace its contours better than the other systems presented in the literature and tested on the same data. The average error in the obtained contour masks is reasonably close to the interoperator errors and suitable for practical applications. The optic disc segmentation pipeline is currently integrated in a complete software suite for the semiautomatic quantification of retinal vessel properties from fundus camera images (VAMPIRE). © 2014 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.1.2.024001]

Keywords: optic disc; inpainting; multiresolution; radial symmetry; active contours; validation.

Paper 13008RR received Oct. 15, 2013; revised manuscript received May 27, 2014; accepted for publication Jun. 16, 2014; published online Jul. 14, 2014.

1 Introduction

Locating and segmenting the optic disc (OD) is important in retinal image analysis. OD is a fundamental landmark to establish retinal coordinates, and its dimensions and relative position with respect to other landmarks are sometimes used to calibrate measurements. OD analysis is used to determine the severity of some diseases, most importantly, glaucoma. The disc can also be used as a starting point for vessel tracking methods.

OD localization and segmentation in digital fundus images may seem an easy task, due to the fact that the OD appears in most of the images as the brightest spot, approximately circular. This means that common segmentation methods, such as thresholding and pixel classification model fitting, should, in principle, provide sufficiently good results. However, even the most recent algorithms proposed in the literature, if tested on large datasets of retinal images, are not always able to trace boundaries or to locate the OD. Failures may be caused by the fact that images are sometimes heavily inhomogeneous, the OD is covered by vessels irregular and variable in shape, and lesions may create false targets or change the expected OD features, especially near its borders. The relevant variability of anomalous cases makes it difficult to find methods working well in general.

The recent survey by Winder et al.¹ cites 38 papers on localization of OD and identification of its boundary. Localization and segmentation are usually two separate tasks in the literature. Several authors define the location of OD as its center, specifying either its estimated coordinates or a likelihood mask located on the center estimate. Segmentation of OD usually refers to the subsequent task of determining the contour.¹ OD is usually

the brightest component of the fundus; therefore, early methods based on the identification of clusters of bright pixels proved simple and effective with images of healthy retinas; for instance, Sinthanayothin et al.² employed simply the intensity variance as a localizing feature. However, algorithms that rely solely on identifying the brightest region often fail in images containing white lesions or other confounding anomalies. Lalonde et al.³ used a pyramidal decomposition and Hausdorff-based template matching. Yu et al.⁴ used three simple detectors to roughly localize the OD center, showing that a simple circular binary mask performs as well as more complex ones for OD detection with template matching.

Many robust methods for OD localization are based on context, e.g., the presence of vascular structures and their orientation. For instance, Hoover and Goldbaum⁵ proposed the computation of a map describing a “fuzzy convergence of blood vessels” to generate candidate positions for the optic nerve center. Foracchia et al.⁶ exploited the convergence of the retinal vasculature through a parametric model estimation. A similar method, using two-dimensional Gaussian matched filters, was presented by Abdel-Razik et al.⁷ Rangayyan et al.⁸ detected the blood vessels using Gabor filters and the phase portrait analysis to detect points of convergence of the vessels. A different, effective approach using multiple cues in a K-NN regression model has been proposed by Niemeijer et al.⁹ In the work of Perez-Rovira and Trucco,¹⁰ several weak hypotheses for the location of arcades, fovea, and OD are computed and subsequently combined using anatomical constraints to obtain a robust OD location.

The main issue with contextual methods using the vascular structure is that they often require an accurate vessel segmentation, which is, in general, a difficult task. Furthermore, these

*Address all correspondence to: Andrea Giachetti, E-mail: andrea.giachetti@univr.it

methods include a set of control parameters that must be optimized using training images. Consequently, performance may depend on the similarity of the test images to those in the training set. Predicting this similarity, and ultimately performance given training and testing sets, is very difficult and remains unapproached so far. A recent paper by Duangate et al.¹¹ presented a parameter-free method to overcome this issue, which appears to be fast and robust, especially on blurred and noisy images or images having different characteristics.

Another popular method, investigated by many authors, is the Hough transform (HT), providing a parametric contour based on precomputed edges. Fleming et al.¹² deployed a generalized HT to detect the circular shape of OD. In Ref. 13, the OD is localized using the circular HT and the parabolic HT. HT is highly tolerant of gaps in feature contour descriptions and relatively unaffected by image noise; however, it tends to be sensitive to the image resolution. Aquino et al.,¹⁴ however, obtained very good results for the OD contour segmentation using circular HT after OD center rough estimation and morphological postprocessing of the image in the region of interest. This work is particularly interesting in that it shows that with a careful preprocessing, a simple circular model driven by edges actually provides lower OD mask overlapping error than complex deformable shape/appearance models.

Active contours have been largely applied for OD segmentation,¹⁵⁻¹⁷ even if the results are often not very good due to noise and anomalies, and algorithms require a preliminary manual or automatic identification of a region of interest containing the OD. Lowell et al.¹⁸ performed first OD location with a template matching approach and then segmentation with a constrained deformable model. Validation was done on a set of 100 images, with results graded by an ophthalmologist as excellent-fair in 83% of cases. Recent applications of active contours have been focused on region-based approaches. Joshi et al.¹⁹ proposed a segmentation method that integrates the local image information around each point of interest in a multidimensional feature space. Yu et al.⁴ presented a very fast and robust OD boundary segmentation technique. It employs a hybrid level set model, which combines region and local gradient information with simple automatic initialization. The average OD area intersection error obtained is, however, surprisingly higher than that obtained with Hough circles on the same images.¹⁴

The analysis of the literature reveals that both the problems of locating and segmenting the optic nerve head are more complex than they may appear. This is mainly due to the variable appearance of the OD in different subjects, making quite difficult to obtain robust automatic location and segmentation methods that can be applied in the clinical practice. Intensity-based segmentation approaches do not easily handle the problem of varying image color, unpredictable intersections of vessels with the bright region, and effects of different pathologies. Model-based approaches, like active shape models, face the impossibility of creating a reliable point-to-point correspondence in training contours, and also parametric models for vascularization and relative positions of features cannot predict new anomalous cases well. Training-based methods suffer from the impossibility of representing the variable aspect of anomalous spots due to pathologies and risk of creating overfitting related to anomalous cases in the training set.

For all these reasons, we tried to find a simpler approach to use contextual information, considering only qualitative characteristics that are consistently found in all the images of the various datasets we had the opportunity to work with.

The result is a method based on two simple assumptions about the OD.

- Its shape is approximately elliptic. It is not always the brightest part of the retina, but, even in many anomalous cases, it is the bright part with the highest radial (circular) symmetry.
- There is a high vessel density inside its contour. The structure of the vasculature may not be easy to model, but vessels can always be seen near/inside the OD and a rough segmentation of them can be used to estimate a local density.

These characteristics appear rather stable in the different datasets analyzed, more than other features commonly used for the task (highest brightness, color features matching a set of training examples, shape of vascular tree near OD).

Our OD location method is, therefore, based on a simple combination of a radial symmetry detector and a vessel density, and the accurate segmentation is based on an iteratively refined model based on contour search constrained by vessel density. Another simple observation that we used in our approach is that, as suggested in Ref. 20, instead of making a deformable model or an OD pattern sufficiently complex to handle relevant discontinuities or modeling the vascular paths, it is better to perform OD search and contour location on inpainted images obtained by removing the rough vessel mask and propagating neighboring information in the masked region.

In our opinion, the advantages of this approach are twofold. First, experimental results show that the method is able to accurately segment the OD in more images with respect to other state-of-the-art methods. Second, the simple assumptions made could, in principle, be checked as a first step in the processing pipeline, and alternative methods specifically designed for specific pathological cases could be applied accordingly.

2 Materials and Methods

We implemented our complete OD location and segmentation pipeline as a MATLAB® module to be included in the VAMPIRE software suite.²¹ To validate the segmentation results, we employed a large public dataset, MESSIDOR (Ref. 22), including images with pathologies (exudates, microaneurysms).

It consists of 1200 images acquired in three different centers using a color video 3CCD camera on a Topcon TRC NW6 non-mydiatic retinograph with a 45-deg field of view and varying image resolutions (2304 × 1536, 2240 × 1488, and 1440 × 960) and with a depth of 8 bits per color plane. Images are graded according to the risk of macular edema and retinopathy.

This dataset has been used in recent works^{4,14,23} and it is very useful if we consider that most results reported in the previous literature were tested on very small or nonpublic datasets (and cannot be, therefore, compared directly or significantly). A public archive with OD contours traced by an expert for these images is available at the University of Huelva.²⁴ We used these annotations for comparison with other methods (Sec. 3).

To compare automatic method performances with human ones (results reported in Sec. 3.2), we also collected different annotations on a subset of 300 images of the dataset. On these images, three experienced ophthalmologists traced the OD contour using the annotation tool included in the VAMPIRE software suite (Fig. 1).

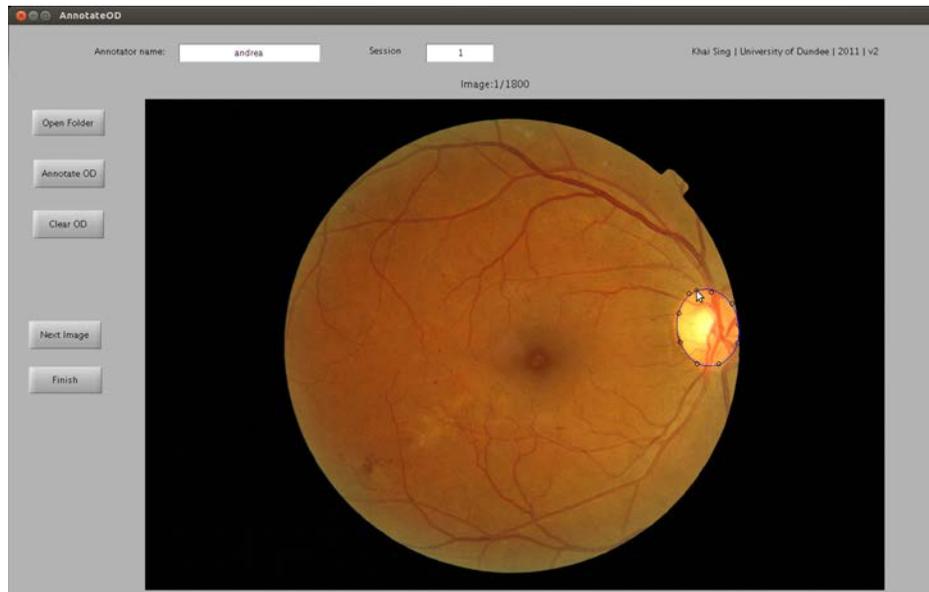


Fig. 1 The simple VAMPIRE user-friendly interface for elliptic contour annotation.

The tool draws parametric ellipses given 5 to 10 clicked points and allows a simple interactive modification of the contour. This choice was motivated by the feedback of experts who considered the annotation with freeform contours too slow and difficult to obtain.

These images have also been annotated by the experts with an image quality grade.

- Easy: disc contour clearly visible with high contrast; interrupted only by vessel widths; any peripapillary atrophy (PPA) or scleral rim clearly distinguishable from disc margin; no other obscuring pathology;
- Intermediate: in between other two categories; margin interrupted by vessels running obliquely or bifurcating;
- Hard: disc contour too hazy/blurred/indistinct to trace; subtle PPA or scleral rim; interrupted by other pathology; myelinated nerve fibers; disc margin appears to be composed of two or more discontinuous ellipses.

In the following, we describe our segmentation pipeline, while validation results are presented in Sec. 3.

2.1 Processing Pipeline

The procedure basically finds optimal OD contours in a coarse-to-fine refinement scheme. Important aspects of our approach are as follows:

- We use a radial symmetry prior derived by the work of Loy and Zelinsky²⁵ to locate the OD and initialize the segmentation.
- To drive the contour optimization, we use two separate maps: the grayscale converted and vessel-inpainted image and the density of the segmented vessels.
- In the coarse-to-fine scheme, we not only increase image resolution but also parametric model complexity starting with an initial circle and then using elliptic shapes. A free-form contour (snake) is used as a final step only for relatively small corrections using locally adaptive forces.

The scheme of the procedure is represented in Fig. 2. We first preprocess the image in order to obtain a decoupled vessel and OD information at different scales; then we localize and

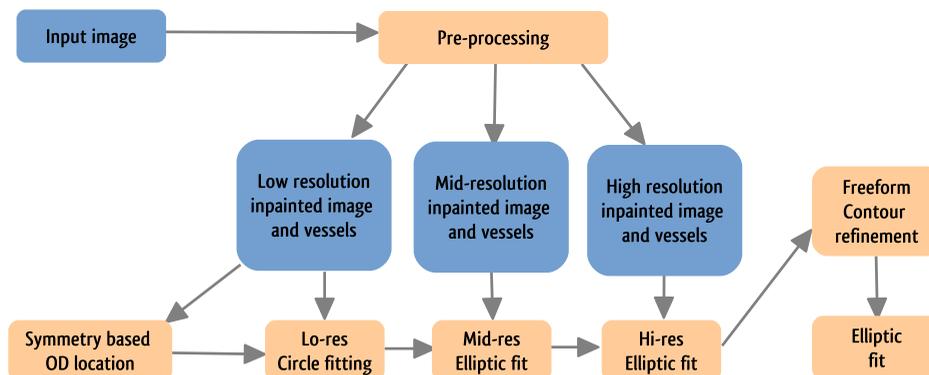


Fig. 2 Flow chart showing the steps of the multiscale optic disc localization and segmentation.

iteratively refine the OD contour with successive steps. Let us describe the entire procedure in more detail.

2.2 Preprocessing

First, we resize the image (using bicubic interpolation) so that the field of view (45 deg in MESSIDOR) spans a constant number of pixels (480). This avoids the necessity of adapting the parameters of the subsequent algorithms. The scale factor used can be obtained from the typology and field of view of the original image, and according to our tests performed on different datasets, it does not need to be very accurate to ensure a stable output.

On this resized image, we perform a rough vessel segmentation on the green channel by subtracting the top-hat filtered version from the original component image and thresholding the result using the Otsu algorithm.²⁶

The vessel mask is then processed with morphological operators (dilation and removal of small areas), and the vessel-removed grayscale image is finally obtained by taking the original grayscale-converted image and inpainting the mask pixels. Inpainting algorithms usually try to fill missing parts in an image by propagating external information so that structure continuity is preserved. Several approaches have been proposed for this task; popular ones use gradient information²⁷ or patch similarity.²⁸ For our application, we obtained good results with the following iterative procedure:

- Remove vessel pixel from the image. Select the border set composed of the empty pixels close to valued ones.
- For each border pixel, compute the median of the valued pixel in a 5×5 neighborhood and fill the pixel with this value.
- Recompute the border set and iterate until it is empty.

In our tests, the resulting images show a continuous OD border (Fig. 3), and experimental tests (see Sec. 3) show that curve-fitting algorithms perform better on inpainted images rather than on original or top-hat filtered images. This procedure is similar to that used in fundus images by Bock et al.,²⁰ but differs in the use of the median instead of the average of the valued pixels.

Grayscale-inpainted images and vessel masks are then decomposed into a Gaussian pyramid with three levels of detail, and the OD detection and segmentation is then performed starting from the coarsest scale.

The reduced image resolutions used for our analysis have been selected with preliminary experiments, showing that their choice does not affect the accuracy of results. Our findings are similar, in this sense, to those of other authors; for example, the finest resolution used by our segmentation method is slightly higher than that used by Aquino et al.¹⁴ on the same images, and also, the resolution used in the initial localization is close to that used in the same paper (slightly lower).

2.3 Robust OD Location

To obtain an approximate location of the OD center, we use both vascular and brightness related priors. But, instead of creating complex models to link OD center position to the feature data, we use a simple probabilistic approach based on combining the output of two simple detectors, one related to radial symmetry and the other related to vessel density. This approach is rather robust and can also be applied with few modifications for fovea detection.²⁹

The vessel density, avoiding false detections of bright spots in avascular regions, is obtained from the output of a vessel enhancement filter. This density, computed at the finest scale, is thresholded in order to capture only major vessels, with an adaptive threshold computed automatically in order to take a fixed percentage of vessel pixels. The result is then downsized at the coarsest resolution. The resulting map is then convolved with a disc-shaped kernel with the expected approximate size of the OD at that scale (10 pixels) and normalized dividing the result by its maximum. We assume that this resulting vessel kernel density estimation $OD_v(\vec{p})$ encodes a prior probability of OD location; due to the fact that the OD is located near the convergence of major vessels, a reasonable assumption is that the OD center is located where this vessel density is high.

The circular symmetry cue is obtained using the fast radial symmetry transform by Loy and Zelinsky,²⁵ an efficient and effective method that is widely used for circle detection. It is based on edge projection at selected distances R_i along directions perpendicular to the edge itself and the subsequent increment of a map in the neighborhood of the projected point, which depends on edge strength and orientation. The map is here tuned to detect only bright symmetrical regions with a radius in a range of distances R_i that correspond to the integer approximations of possible values of the OD radius at the input image resolution and is taken as the OD center symmetry-base likelihood $OD_s(\vec{p})$.

For each pixel location, we also store, in a separate map, the value of R_i giving the largest contribution to OD_s , which allows

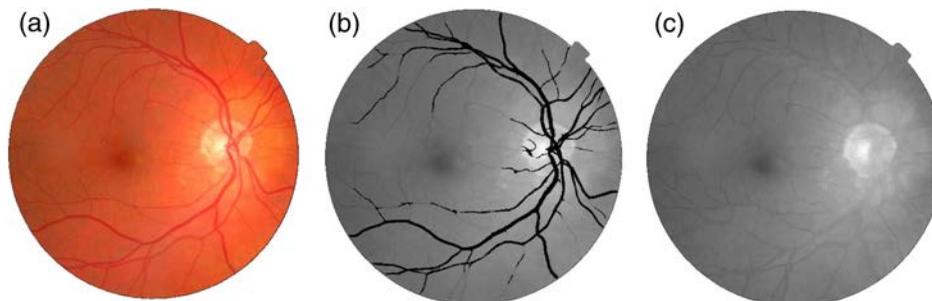


Fig. 3 Image preprocessing. (a) The original RGB image. (b) Estimated vessel mask. (c) Inpainted gray-scale image, in which the optic disc (OD) is clearly visible and not occluded or cluttered by vessels.

us to obtain a rough estimate of the OD radius. This value is used to initialize the subsequent circle fitting. The final approximate estimate of the OD position is obtained by finding (with subpixel accuracy) the maximum of the combined likelihood

$$p_{\text{OD}}(\vec{p}) = \text{OD}_s(\vec{p}) \cdot \max[0.1, \text{OD}_v(\vec{p})], \quad (1)$$

where a small minimum value (0.1) for the vessel probability has been added to give a preference to the symmetry prior when no regions with relevant values of both OD_v and OD_s are found. Figure 4 shows the full rough OD localization procedure. The vessel density map [Fig. 4(d)] here acts similar to the “fuzzy vessel convergence” proposed by Hoover and Goldbaum,⁵ being maximal near the vessel convergence, and can roughly identify the OD region. The combination of this cue with the radial symmetry cue results in a localization algorithm that relevantly outperforms the combination of vessel density and brightness maps or template matching.²⁹

2.4 Multiresolution Ellipse Optimization

At each level, we refine the image, the contour model, and the objective (cost) function used by the optimization procedure to find the contour. At the coarsest level, the parametric contour is a circle described by its center and radius. The optimal circle is found by initializing circles in the previously detected OD position, running the deterministic Nelder-Mead optimizer; the overall final result is the solution with minimum value of the cost function. To build the cost function, we sample the contour at N discrete positions, with N depending on the image resolution (30 at the coarsest scale). For each point $\vec{p}(i) = (x_i, y_i)$, we consider the unit vector $\vec{n}(i)$ perpendicular to the contour and we sample S equally spaced internal points $C_{\text{in}}(i, k) = \vec{p}(i) - k\alpha\vec{n}(i)$ and S equally spaced external points $C_{\text{out}}(i, k) = \vec{p}(i) + k\beta\vec{n}(i)$ with $k = 1 \dots S$, $\alpha = R/S$, $\beta = 1.5R/S$. R here is the currently estimated radius and S was taken equal to 8 at the coarsest scale.

Sampled values are used to build the function

$$F(C) = \sum_{i=1}^N \sum_{k=1}^S w(i) \min\{I[C_{\text{in}}(i, k)] - I[C_{\text{out}}(i, k), D]\}, \quad (2)$$

where $w(i)$ are weights that are maximal for small i to enhance the effect of edges near the contour border (Fig. 5), and D is a constant introduced in order to remove the effect of outliers and was set equal to 25, representing a reasonable difference between inner intensity and outer intensity in normal cases. The idea behind these choices is to make the contour attracted by discontinuities near the border and by global differences between the internal and the external values without requiring a specific model for the internal intensity distribution (e.g., constant as done in Chan-Vese region-based active contours³⁰).

We add to the objective functions an additional term, adding a large fixed penalty when the percentage of vessels in the central part of the disc ($r < 3R/4$) is $< 10\%$ of the total OD pixels. This choice introduces in the method the high vessel density constraint without introducing a bias on contour position. The threshold has been set by trial and error working on different images. A similar choice has been introduced to penalize very small OD radii (less than half the approximately expected value).

After the first circular fitting, the final value of the objective function is checked: if the result is too large, we assume that the OD initial location failed and restart the procedure from multiple tentative locations corresponding to most frequent positions in fundus images (as done in Ref. 31), keeping at the end the result with the lowest value of the function.

At each new (finer) resolution level, we initialize the contour using the result obtained at the previous level and start a new optimization procedure. Increasing the resolution, we also change the contour model, adding a vertical elongation term. At the highest resolution, we also add a rotation term, so that

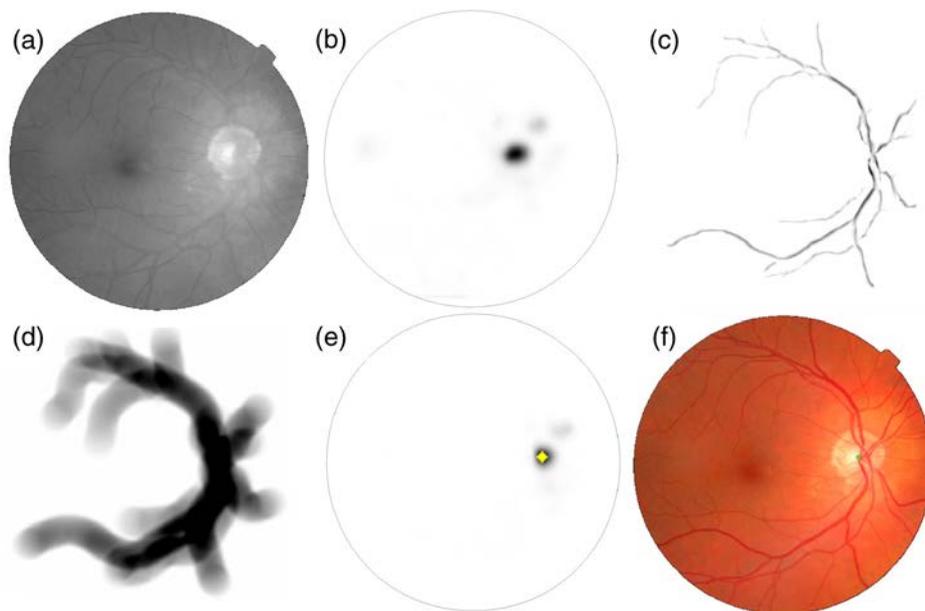


Fig. 4 Approximate OD location. (a) Grayscale converted image. (b) Radial symmetry map tuned for OD-like structures detection. (c) Major vessels extracted. (d) Kernel smoothed vascular density. (e) Final OD center location probability estimate. (f) Final OD center estimate superimposed on the original image.

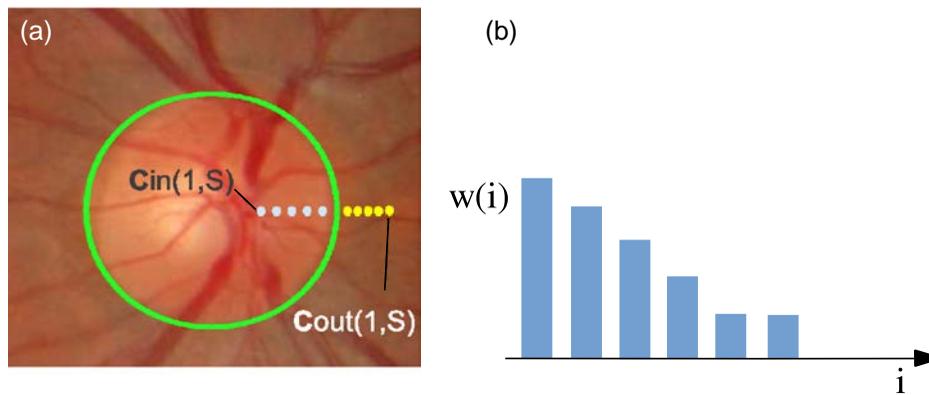


Fig. 5 The objective functions sum, on selected points of the circular/elliptic contour, values sampled on the inpainted image (a) along the normal direction, multiplied by weights enhancing discontinuities near the border (b).

the finally extracted OD is a generic ellipse described by five parameters (Fig. 6). Apart from the change in the contour model, the objective function used in the following steps is the same, with the addition of a term penalizing extremal OD eccentricity measures.

2.5 Snake-Based Contour Refinement and Ellipse Refitting

The goal of the procedure is to estimate a freeform contour that can be used to analyze the optic nerve and to search for related biomarkers. As shown in Refs. 4 and 14, the active contour approach alone, even if region-based and constrained, is relevantly affected by noise and does not provide very good results.

At the finest scale, we then refine the segmented contour with a snake.³² We have chosen a simple explicit model because our interest is only to refine locally the previously computed contour points according to the local image structure. We sample the elliptic contour with a fixed number of equally spaced points and let them move iteratively driven by the classical elastic and bending forces and two image forces, a classical gradient-based attraction and an additional force adapted to the local image structure.

This force is defined as a repulsive force with direction perpendicular to the contour and intensity estimated as follows: we sample points along each line perpendicular to the initial contour in a distance range of half the expected OD radius

and we check the intensity profile. If the local variability of the profile is smaller than a threshold (set to 8 in our experiments), we consider that there is no visible step edge and set a force value that exactly compensates for the shrinking effect due to the local curvature. Otherwise a two-means clustering on the sampled values is performed to classify OD and background. In this case, an inflating term is added if points outside the contour are classified as OD.

The snake evolution is stopped with the following criteria: when a point that changes inward/outward direction is detected and labeled as “oscillating” and when at least three of the five points in its neighbor are oscillating, the point is stopped. When $>80\%$ of the points are labeled as oscillating, the evolution is stopped. The thresholds used have been set by trial and error on images of a different dataset, so that the results could be slightly improved with a specific optimization.

It is worth noting that if we restore the elliptical shape of the contour by using a standard ellipse fitting method³³ on the extracted contour points, as we will see in the Experimental Results, the accuracy of the resulting elliptic segmentation is (slightly) better than that obtained with the multiscale optimization. This is probably due to the limits of the deterministic optimizer that may end in local minima.

The extraction of the best possible elliptic output for the segmentation is useful for practical applications, because, for example, vascular biomarkers are often computed in regions defined by OD-based coordinates (e.g., the arterial to venular ratio³⁴).

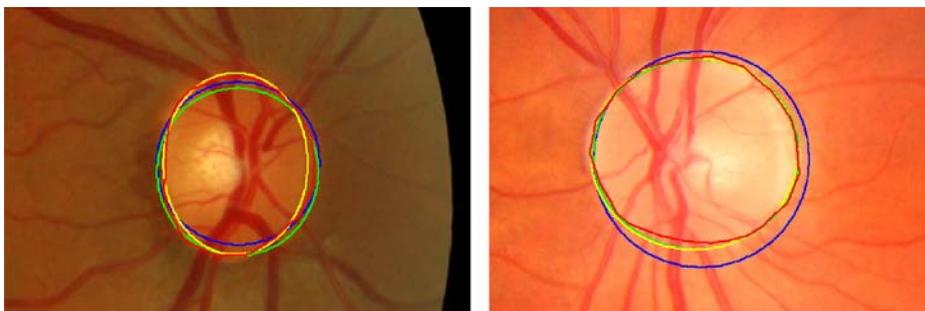


Fig. 6 Example of contours detected at multiple scales. Black line (blue in the electronic version): initial circular shape. Light gray line (green in the electronic version): intermediate elliptic contour. White line (yellow in the electronic version): final ellipse fitting result. Dark gray line (red in the electronic version): freeform contour computed with the snake-based refinement algorithm.

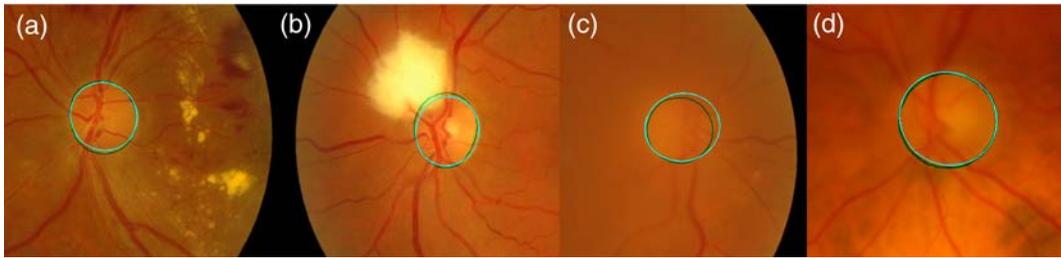


Fig. 7 Examples of automatically computed contours (bright lines, cyan in the electronic version) versus manual annotations (dark lines, green in the electronic version). Results are here rather good even if the image quality is poor. Overlap scores are $S = 93.7$, $S = 90.1$, $S = 83.8$, and $S = 94.2$ for images (a), (b), (c), and (d), respectively.

3 Experimental Results

Most of the OD contours extracted on most images appear close to the manually annotated ones even if the images are visually quite different. Figure 7 shows examples of detected contours that are very close to the manually annotated ones even if the image quality is poor.

To compare the accuracy of the OD segmentation with other methods on the MESSIDOR data, we measured the superimposition of the extracted regions with the ground truth masks using the Jaccard index.

$$S = \frac{OD_{\text{segm}} \cap OD_{\text{ref}}}{OD_{\text{segm}} \cup OD_{\text{ref}}}$$

In the following we will call “overlap error” with respect to the reference annotation the value $1 - S$.

Table 1 shows the average overlap scores obtained with our method on the MESSIDOR dataset with the Huelva annotations, compared with the results obtained with the other methods tested in the literature on these data.

It is possible to see that the multiscale fitting procedure outperforms other techniques and the contour adjustment steps slightly improve the accuracy.

Figure 8 shows graphically the overlap error reduction in the different intermediate steps of our procedure, e.g., simple low-resolution radial symmetry estimation (with associated radius), circle fitting, ellipse fitting, snake-based correction, and ellipse

Table 1 Our techniques provide a higher average overlap score than other methods on the complete MESSIDOR dataset annotated at the University of Huelva. They also provide relevantly higher percentages of contours highly overlapped with the manual annotation.

	Ell. fit	Snake	Refitted ell.	Ref. 14	Ref. 4
S	0.87	0.88	0.88	0.86	0.84
$S \geq 0.70$	92%	94%	94%	93%	—
$S \geq 0.75$	89%	92%	92%	90%	—
$S \geq 0.80$	85%	89%	88%	84%	—
$S \geq 0.85$	78%	81%	82%	73%	—
$S \geq 0.90$	62%	56%	59%	46%	—
$S \geq 0.95$	12%	7%	13%	7%	—

refitting: it is possible to see that even the simple symmetry-based circle detection provides a reasonable overlap and that the optimization procedure provides a good result, even before the snake-based adjustment.

Improvements provided by the snake-based refinement and problems of the different steps are also shown in Fig. 9. The top row shows a correctly placed elliptic contour in cyan, visually close to the manually traced contour shown in green [Fig. 9(a)] and the improved result obtained with the refinement [Fig. 9(b)]. In some cases, however, the accuracy of the contour is not similarly good and a local refinement is not always able to fix problems or improve the result. The middle row shows a poor elliptic fit [Fig. 9(c), $S \sim 0.6$], where the local correction can improve the result, but only in the lower part of the disc [Fig. 9(d)]. The bottom row shows a poor elliptic fit, where the local correction actually decreases the overlap score [Figs. 9(e) and 9(f)].

Table 2 shows the overlap scores obtained for the images labeled with all the possible combinations of retinopathy and macular edema grades given by experts. It appears that the overlap is decreasing only very slowly with retinopathy grade, while the risk of macular edema is probably not correlated with segmentation quality. This seems reasonable, as neither condition affects the OD severely.

Apart from the improved overlap score, our method provides further comparative improvements. The method of Aquino et al.,¹⁴ in fact, is based on edge fitting and cannot provide

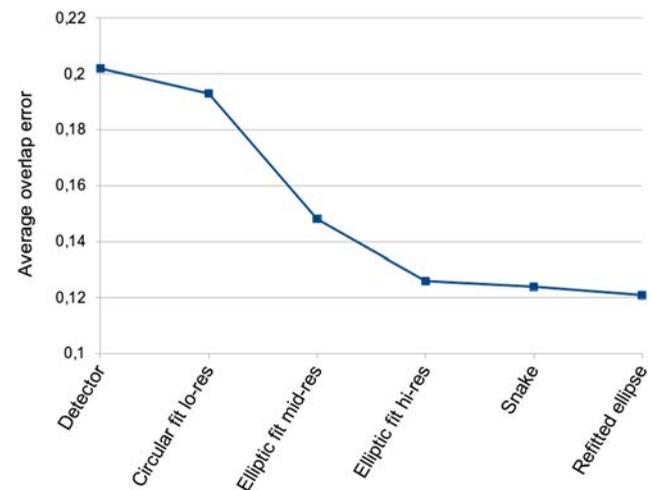


Fig. 8 The average overlap error is decreased during the different optimization steps. The result of the ellipse refitting is slightly better than the snake curves due to the elliptic ground truth constraint.

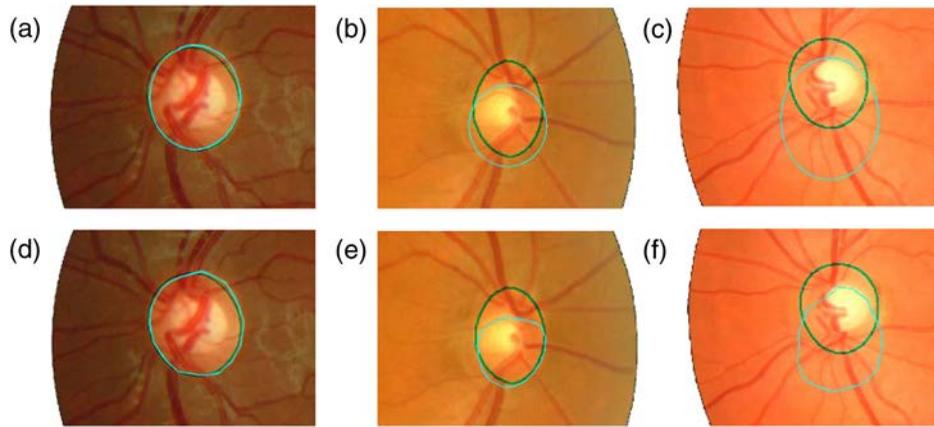


Fig. 9 When the elliptic OD fit is sufficiently good, (a) the snake-based refinement increases the accuracy (d). Estimated contours are the bright ones (cyan in the electronic version), while manual ones are darker (green). In some cases, the refinement can fix at least partially large differences between corresponding manually (b) and automatically traced (e) contours. When the information is poor, the correction can actually decrease the overlap score [(c) and (f)].

Table 2 Overlap scores appear approximately independent on the diagnostic annotations provided by the MESSIDOR project. Only a slight decrease of the score with increasing retinopathy grade is observed.

		Retinopathy grade				
		0	1	2	3	Any
Risk of macular edema	0	0.884	0.884	0.876	0.873	0.881
	1	—	0.864	0.842	0.869	0.858
	2	—	0.878	0.887	0.868	0.873
	Any	0.884	0.883	0.873	0.870	0.879

OD shape feature (they show that classical ellipse-fitting methods applied on their edge extraction provide poor results). The method of Yu et al.⁴ can provide freeform or elliptic contours, but has a much lower overlap score.

The method can also be, therefore, used to estimate shape parameters with good accuracy. Figure 10 shows the Bland-Altman plot representing differences versus sums of manual and automatic estimations of the OD area (normalized by the

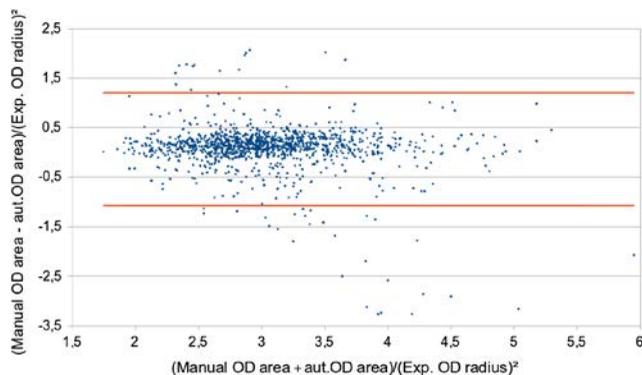


Fig. 10 Bland-Altman plot representing the difference versus average OD area estimations obtained with manual and automatic procedure.

expected OD radius) for the 1200 MESSIDOR images. The plot demonstrates that, apart from a few outliers, the measurement procedures are in accordance. A Welch's *t* test on the two sets shows that the measurements are the same within the 1% confidence level.

3.1 Evaluation of Algorithmic Choices

3.1.1 OD detection

An extremely effective component of our pipeline is the symmetry-based automatic OD localization technique. Using the same rule used in compared papers, where the OD localization is considered successful if the distance between the estimated OD center and the ground truth OD center is less than the expected OD diameter, our localization method fails only in two images of the set of 1200. On the same dataset, the OD localization technique proposed by Yu et al.,⁴ based on a simple template matching, fails on 11 images, while the method proposed by Aquino et al., based on the combination of simple detectors (low-pass filter, maximum difference, maximum variance) fails in 14 cases.¹⁴ It is also worth noting that in this last paper, in the case of localization failure, the results of OD segmentation are obtained with manual initialization. Our segmentation results reported in Table 1 are obtained with a completely automatic procedure, so that for the two images where the initialization failed, the overlap score is zero. Figure 11 shows these two images where our detection failed. It is easy to understand why our method fails: in these cases, the OD is darker than a larger surrounding area, and the hypotheses used for detection and segmentation are false.

3.1.2 Inpainting and vessel mask

To evaluate the effect of the inpainting procedure applied, we compared the accuracy of the segmentation that we obtained replacing our method with a simple morphological filter or not using it at all. Table 3 shows that the inpainting method proposed is effective in reducing both the missed detections and the overlap error, while a morphological filter with a kernel sufficiently large to remove vessels creates corrupted images and results in decreased detection rates and overlap scores. This



Fig. 11 The only two images of the 1200 where our algorithm fails: the OD region is darker than the surroundings.

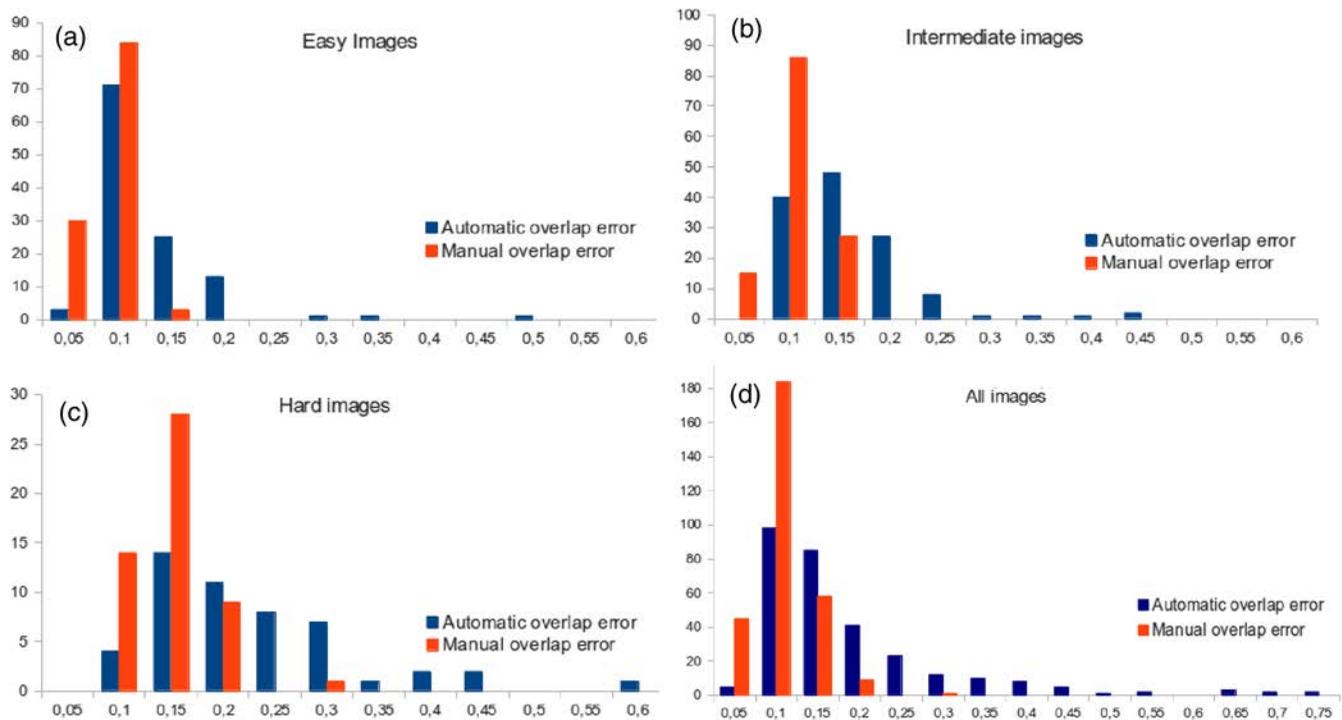


Fig. 12 Histograms showing the number of automatic contours (dark gray, blue in the electronic version) and manual ones (light gray, red in the electronic version) within increasing area overlap error ranges, for different image quality labels. (a) Easy images, (b) intermediate images, (c) hard images, (d) all images. It is possible to see that the behavior is similar (error increasing with the expected difficulty) with a slightly lower error for the manual procedure and a few outliers for the automatic detector making the average errors in this case larger.

Table 3 The inpainting procedure reduces the number of failed detections and improves the segmentation accuracy. Also, the symmetry-based vessel extraction is effective in improving the detector performances.

	Vessel removal			Vessel estimation	
	None	Morph.	Inpainting	Morph.	Symmetry
Failed det.	8	5	2	6	2
<i>S</i>	0.865	0.807	0.879	0.875	0.879

suggests that a good vessel removal does improve the segmentation procedure and that the method proposed is sufficiently reliable.

We similarly tested the method used to extract the vessel mask: in Table 3, it is also possible to see the effect of different choices on our OD detection and segmentation method. The use of the symmetry-based vessel mask does not improve the segmentation phase, but removes four detection failures, due to the higher specificity of the method with respect to simple morphological procedure (difference between image and top-hat filtered image and thresholding).

3.2 Comparison with Inter- and Intraobserver Errors

Unlike the MESSIDOR diagnostic annotation, the quality grade given by the experts is strongly correlated with both the quality of the automatic segmentation and interobserver errors, as shown in Fig. 12. This may indicate that the visual cues used by human observers are similar to those applied in the automatic algorithm.

Interoperator errors obtained by our experts are lower than those reported in Ref. 4 on a smaller (100 images) subset of the complete dataset. Interoperator variability between our experts are similar to interoperator variability between our experts and the expert from Huelva.

Table 4 shows overlap errors and Hausdorff distances (e.g., the maximal value of the distance between a point of one contour and its closest point in the other one) comparing interoperator contour tracing, intraoperator contour tracing, and automatically versus manually traced contours. It is possible to see that even if our system outperforms the other methods tested in the literature, differences between annotators are still lower, and this is particularly evident in easy images. This fact suggests that it is still possible to improve the performance of the automatic system. Results presented so far in the literature show, however, that it is really difficult to create more effective appearance model of the contour, for example, using learning-based approaches, due to the huge variability of image cues and contextual information. A possible alternative way to improve the contour segmentation on which we plan to concentrate our efforts consists of learning from the analysis of human heuristics rather than simply on traced contours. To reach this goal, we plan to develop *ad hoc* protocols and software tools to record annotation procedures and analyze how the expert ophthalmologists trace the OD boundaries. The understanding of what is actually annotated by the experts as OD contour and

Table 4 Average area overlap errors and Hausdorff distances for the automatic procedure compared with average intraoperator and interoperator errors.

	Avg auto versus human	Avg interop.	Avg intraop.
All images			
Overlap err. (std. dev.)	13.9 (11.0)	8.0 (3.3)	6.8 (2.8)
Hauss. dist. (std. dev.)	18.8 (18.7)	9.3 (3.7)	8.7 (3.9)
Easy			
Overlap err. (std. dev.)	13.1 (11.4)	6.2 (1.9)	5.5 (1.7)
Hauss. dist. (std. dev.)	17.1 (18.5)	7.3 (2.3)	7.2 (2.9)
Intermediate			
Overlap err. (std. dev.)	13.8 (10.1)	7.9 (2.6)	6.7 (2.4)
Hauss. dist. (std. dev.)	18.0 (14.7)	9.1 (3.0)	8.5 (3.4)
Hard			
Overlap err. (std. dev.)	15.7 (12.7)	12.3 (9.9)	9.9 (2.9)
Hauss. dist. (std. dev.)	20.7 (20.7)	14.3 (3.6)	12.3 (4.6)

how the same result could be obtained automatically is particularly interesting, also considering that recent research works based on the use of three-dimensional optical coherence tomography (OCT) revealed that the optic disc margins seen in fundus images are actually composed of different anatomical structures.³⁵

3.3 Limits

The algorithm proposed provides good results on the vast majority of the MESSIDOR set, including the case of lesions due to the pathologies considered in MESSIDOR; however, performance decreases when the assumptions are violated. The failure of our hypotheses is sufficiently easy to detect, enabling the application of disease-specific detectors.

The algorithm, currently, is designed mainly for population studies, e.g., retinal biomarkers for a variety of eye and systemic conditions (a field attracting a high volume of research). In these cases, and in cases where pathologies affect part of the eye other than the OD, our tool seems appropriate and effective. Obviously, it would need adapting in case of OD-specific alterations. Notice that Table 2 reports results with imaging graded at various levels of diabetic retinopathy and risk of macular edema, although neither condition affects the OD contour severely.

Another limit is that the accuracy of the contour detection, compared with annotations, is still inferior to the interannotator consistency. We notice that this is a common situation in several retinal image analysis algorithms, e.g., Ref. 36. Multimodal analysis of the OD borders and manual annotation procedures will be investigated in the future to improve the quality of the segmentation procedure.

3.4 Computational Complexity

The complete segmentation procedure has been implemented in MATLAB® and included in the VAMPIRE software suite. A single complete segmentation takes ~8 s for a 2304 × 1536 image on a DELL XPS 17 laptop (CPU Intel Core i7-740QM), including ~2 s for preprocessing and inpainting and ~2 s for the vessel segmentation. This time is already suitable for interactive use on the VAMPIRE software tool,³⁷ even if it could be easily reduced by optimizing and compiling the code.

4 Conclusion

In this paper, we presented a complete pipeline for the localization and accurate segmentation of the OD in digital fundus images and the results of its validation performed on a large set of images, including examples of different pathologies. The results obtained show that the proposed method provides a robust localization, and a segmentation of the contour close to the manual accuracy in most cases, allowing a reasonable estimate of shape parameters. On the largest annotated dataset used in recent works, our method outperforms the ones demonstrated in the literature for both detection rate and segmentation accuracy. Even if our tests performed with three expert annotators on a large subset show that the interoperator contour overlap error is still lower than the automatic one, the use of the automatic algorithms, now fully included in the interactive VAMPIRE software suite,³⁷ is extremely useful even if we need a very high precision. Users, in fact, can perform automatic segmentation and manually correct the few bad results obtained, saving a huge amount of time.

Acknowledgments

We gratefully thank Dave Knight, James Welch, and Peter Wilson for providing the manual annotations of optic disc contours. This project is partially supported by Leverhulme Trust Grant RP-419.

References

1. R. Winder et al., "Algorithms for digital image processing in diabetic retinopathy," *Comput. Med. Imaging Graph.* **33**(8), 608–622 (2009).
2. C. Sinthanayothin et al., "Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images," *Br. J. Ophthalmol.* **83**(8), 902–910 (1999).
3. M. Lalonde, M. Beaulieu, and L. Gagnon, "Fast and robust optic disc detection using pyramid decomposition and Hausdorff-based template matching," *IEEE Trans. Med. Imaging* **20**(11), 1193–1200 (2001).
4. H. Yu et al., "Fast localization and segmentation of optic disk in retinal images using directional matched filtering and level sets," *IEEE Trans. Inf. Technol. Biomed.* **16**(4), 644–657 (2012).
5. A. Hoover and M. Goldbaum, "Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels," *IEEE Trans. Med. Imaging* **22**(8), 951–958 (2003).
6. M. Foracchia, E. Grisan, and A. Ruggeri, "Detection of optic disc in retinal images by means of a geometrical model of vessel structure," *IEEE Trans. Med. Imaging* **23**(10), 1189–1195 (2004).
7. A.-H. A.-R. Youssif, A. Ghalwash, and A. A.-R. Ghoneim, "Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter," *IEEE Trans. Med. Imaging* **27**(1), 11–18 (2008).
8. R. Rangayyan et al., "Detection of the optic nerve head in fundus images of the retina with Gabor filters and phase portrait analysis," *J. Digit. Imaging* **23**(4), 438–453 (2010).
9. M. Niemeijer, M. Abrámoff, and B. van Ginneken, "Fast detection of the optic disc and fovea in color fundus photographs," *Med. Image Anal.* **13**(6), 859–870 (2009).
10. A. Perez-Rovira and E. Trucco, "Robust optic disc location via combination of weak detectors," in *Annual Int. Conf. of the IEEE Engineering in Medicine and Biological Society*, pp. 3542–3545, IEEE (2008).
11. C. Duangate et al., "Parameter-free optic disc detection," *Comput. Med. Imaging Graph.* **35**(1), 51–63 (2011).
12. A. D. Fleming et al., "Automatic detection of retinal anatomy to assist diabetic retinopathy screening," *Phys. Med. Biol.* **52**(2), 331 (2007).
13. S. Sekhar, W. Al-Nuaimy, and A. K. Nandi, "Automated localisation of retinal optic disk using Hough transform," in *IEEE Int. Symp. on Biomedical Imaging*, pp. 1577–1580, IEEE (2008).
14. A. Aquino, M. Gegundez-Arias, and D. Marin, "Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques," *IEEE Trans. Med. Imaging* **29**(11), 1860–1869 (2010).
15. F. Mendels, C. Heneghan, and J. Thiran, "Identification of the optic disc boundary in retinal images using active contours," in *Proc. of the Irish Machine Vision and Image Processing Conf.*, pp. 103–115 (1999).
16. H. L. H. Li and O. Chutatape, "Automated feature extraction in color retinal images by a model based approach," *IEEE Trans. Biomed. Eng.* **51**(2), 246–254 (2004).
17. J. Xu, O. Chutatape, and P. Chew, "Automated optic disk boundary detection by modified active contour model," *IEEE Trans. Biomed. Eng.* **54**(3), 473–482 (2007).
18. J. Lowell et al., "Optic nerve head segmentation," *IEEE Trans. Med. Imaging* **23**(2), 256–264 (2004).
19. G. Joshi, J. Sivaswamy, and S. Krishnadas, "Optic disk and cup segmentation from monocular color retinal images for glaucoma assessment," *IEEE Trans. Med. Imaging* **30**(6), 1192–1205 (2011).
20. R. Bock et al., "Glaucoma risk index: automated glaucoma detection from color fundus images," *Med. Image Anal.* **14**(3), 471–481 (2010).
21. A. Perez-Rovira et al., "Vampire: vessel assessment and measurement platform for images of the retina," in *Proc. of the IEEE Engineering in Medicine and Biology Society*, pp. 3391–3394, IEEE (2011).
22. "Méthodes d'Evaluation de Systèmes de Segmentation et d'Indexation Dédiées à l'Ophthalmologie Rétinienne," 13 March 2014, <http://messidor.crihan.fr> (2 July 2014).
23. H. Yu et al., "Fast localization of optic disc and fovea in retinal images for eye disease screening," *Proc. SPIE* **7963**, 796317 (2011).
24. "Expert system for early automatic detection of diabetic retinopathy by analysis of digital retinal images," 2012, <http://www.uhu.es/retinopathy> (31 October 2012).
25. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Sys., Man., Cyber.* **9**(1), 62–66 (1979).
26. N. Otsu, "A threshold selection method from gray-level histograms," *Automatica* **11**(285–296), 23–27 (1975).
27. M. Bertalmío et al., "Image inpainting," in *SIGGRAPH*, pp. 417–424, ACM Press/Addison-Wesley Publishing Co. (2000).
28. A. Criminisi, P. Prez, and K. Toyama, "Object removal by exemplar-based inpainting," in *Conf. on Computer Vision and Pattern Recognition*, Vol. 2, pp. 721–728, IEEE (2003).
29. A. Giachetti et al., "The use of radial symmetry to localize retinal landmarks," *Comput. Med. Imaging Graph.* **37**(5), 369–376 (2013).
30. T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.* **10**(2), 266–277 (2001).
31. A. Giachetti et al., "Multiresolution localization and segmentation of the optical disc in fundus images using inpainted background and vessel information," in *IEEE Int. Conf. on Image Process.*, pp. 2145–2148, IEEE (2011).
32. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *Int. J. Comput. Vis.* **1**(4), 321–331 (1988).
33. M. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least-squares fitting of ellipses," *IEEE Trans. Pattern Anal. Mach. Intell.* **21**(5), 476–480 (1999).
34. T. Y. Wong et al., "Retinal microvascular abnormalities and incident stroke: the atherosclerosis risk in communities study," *Lancet* **358** (9288), 1134–1140 (2001).
35. A. S. Reis et al., "Optic disc margin anatomy in patients with glaucoma and normal controls with spectral domain optical coherence tomography," *Ophthalmology* **119**(4), 738–747 (2012).
36. C. A. Lupaşcu, D. Tegolo, and E. Trucco, "Accurate estimation of retinal vessel width using bagged decision trees and an extended multiresolution hermite model," *Med. Image Anal.* **17**(8), 1164–1180 (2013).
37. E. Trucco et al., "Novel VAMPIRE algorithms for quantitative analysis of the retinal vasculature," in *Proc. of the 4th IEEE Biosignals and Biorobotics Conf.*, pp. 1–4, IEEE (2013).

Andrea Giachetti is an associate professor at the Department of Computer Science of the University of Verona. He received his PhD in physics from the University of Genova in 1997. He also worked with CRS4 (Center for Advanced Studies, Research and Development in Sardinia), where he was head of the Medical Image Processing area, and key staff of EU funded projects. His main research interests are in the field of medical image processing, computer vision, and computational geometry.

Lucia Ballerini received her PhD degree in bioengineering from the University of Florence. She is currently a member of the VAMPIRE team at the University of Dundee, United Kingdom. She is the co-author of more than 70 peer-reviewed scientific publications. Her main research interests are in the field of image processing and analysis. She is working both with theoretical methods and application-oriented ones.

Emanuele Trucco is the NRP professor of computational vision in the School of Computing, University of Dundee. He is a principal or co-investigator in medical image processing projects on retinal image analysis, colonoscopy, and whole-body MR angiography. He coordinates the VAMPIRE project with Dr. Tom MacGillivray (University of Edinburgh), with partners in the United States, Europe, Asia, and Australia. VAMPIRE aims to develop usable and accurate software to assist clinical- and biomarker-oriented research (vampire.computing.dundee.ac.uk).