Identification of the optic nerve head with genetic algorithms

Enrique J. Carmona[1*], Mariano Rincón[1], Julián García-Feijoó[2], José M. Martínez-de-la-Casa[2]

[1]*Departamento de Inteligencia Artificial, Escuela Técnica Superior de Ingeniería Informática, Universidad Nacional de Educación a Distancia. C/ Juan del Rosal 16, 28040. Madrid, Spain.*
{ecarmona, mrincon}@dia.uned.es

[2]*Departamento de Glaucoma. Servicio de Oftalmología. Hospital Clínico San Carlos. Instituto de Investigaciones Ramón Castroviejo. Universidad Complutense. Madrid, Spain.*
(jgarciafeijoo@hotmail.com, martinezcasa@ya.com)

*Corresponding author: ecarmona@dia.uned.es, Tfn: (34) 91 398 73 01, Fax : (34) 91 398 88 95.

Summary

Objective: This work proposes creating an automatic system to locate and segment the optic nerve head (ONH) in eye fundus photographic images using genetic algorithms.

Methods and material: Domain knowledge is used to create a set of heuristics that guide the various steps involved in the process. Initially, using an eye fundus colour image as input, a set of hypothesis points was obtained that exhibited geometric properties and intensity levels similar to the ONH contour pixels. Next, a genetic algorithm was used to find an ellipse containing the maximum number of hypothesis points in an offset of its perimeter, considering some constraints. The ellipse thus obtained is the approximation to the ONH. The segmentation method is tested in a sample of 110 eye fundus images, belonging to 55 patients with glaucoma (23.1%) and eye hypertension (76.9%) and random selected from an eye fundus image base belonging to the Ophthalmology Service at Miguel Servet Hospital, Saragossa (Spain).

Results and conclusions: The results obtained are competitive with those in the literature. The method's generalization capability is reinforced when it is applied to a different image base from the one used in our study and a discrepancy curve is obtained very similar to the one obtained in our image base. In addition, the robustness of the method proposed can be seen in the high percentage of images obtained with a discrepancy $\delta<5$ (96% and 99% in our and a different image base, respectively). The results also confirm the hypothesis that the ONH contour can be properly approached with a non-deformable ellipse. Another important aspect of the method is that it directly provides the parameters characterising the shape of the papilla: lengths of its major and minor axes, its centre of location and its orientation with regard to the horizontal position.

Keywords: Optic nerve head segmentation; Genetic algorithm; Constraint handling; Ellipse fitting; Glaucoma.

1. Introduction

The optic nerve head (ONH), also known as the papilla, is the area where all the retina nerve fibres converge to form the start of the optic nerve that leaves the back of the eyeball. In the centre of the papilla there is an area without any nerve fibres called the excavation, and around this area, another area slightly more raised than the rest of the retina, called the neuroretinal ring, whose external perimeter delimits the papillary contour. In a two-dimensional image, the ONH exhibits a lighter colour than the retina, is oval in shape and fragmented into different subregions by the blood vessels that emerge from the excavation to the retina. Moreover, inside the excavation there is an even brighter area called pallor.

Analysis of the ONH is one of the most important tests to diagnose a number of eye diseases. Thus, for example, the excavation/papilla vertical relation and the area of the neuroretinal ring are important parameters to consider in the diagnosis of glaucomatous papillae [1]. Furthermore, the location of the ONH is also interesting in retina studies

[2-5]. However, in the latter instance, the requirements for locating the papillary contour are less demanding (an approximate location is sufficient) because the final objective is to filter the papillary region to limit the analysis of the image to just the retina region. Several techniques are used for studying the state of the optic nerve and the nerve fibre layer [6]: photographic techniques (analogical and digital), laser polarimetry (GDx system), confocal laser scanning system (HRT system), optical coherence tomography (OCT) and scanning laser tomography (SLT). With any of these last four techniques, the expert extracts a series of details that enable him to mentally reconstruct the topography of the eye fundus and accordingly determine the papillary and excavation contour.

Current support systems for the diagnosis of eye diseases based on eye fundus information imply manual or semi-automatic tracing of the papillary contour and other ONH structures. Thus, from this initial information, the system determines a set of measures that are used together with other findings to do the diagnosis. Nevertheless, in the field of research, systems are beginning to appear which try to minimise human intervention in the process for segmenting and identifying different structures of the ONH. In this respect, detection of the papillary contour has aroused special interest. Specifically, the first semi-automatic strategies were based on geometric properties of the image pixels and their intensity level values. In [7], the expert marks four points with a high probability of belonging to the papillary contour and the algorithm traces an ellipse containing these four points as an initial approach and, then, the contour is refined from the grey levels close to this ellipse. Later, [8] presented a semi-automatic method based on the grey level gradient properties.

Active contour-based strategies have also been shown to be useful. A characteristic property of this type of technique is that it is highly dependent on a preliminary stage of contour initialisation, from which the final solution is refined. In other words, if the initialisation is bad, it is highly likely that the solution will not be achieved. In some instances, this initialisation stage has been done manually, [9-12], and in other instances automatically [4, 5, 12, 13]. We focus on the latter. Thus, in [5] an automatic system is proposed which uses the *template matching* technique to detect the centre of the papilla, and a circumference is defined around the centre which constitutes the first approach to the contour and its radius is estimated with an algorithm based on the least squares method. This circumference will be the starting point for the precise location of the papilla using the active contour technique. Conversely, in [13], segmentation was done using *fuzzy c-means clustering* to obtain a better initial approach to the ONH contour. The same as in [12], the *gradient vector flow field* was used to define the external energy of the active contour. In [4] template matching was also used to locate the centre of the ONH and, moreover, for the precise identification of the contour, controlled active contours were used with domain knowledge to progressively approach the solution.

This paper presents a method for locating and segmenting the ONH in eye fundus photographic images, without any intervention by the user. Thus the method identifies the ONH automatically and selects a point located approximately in its centre. For segmentation, a genetic algorithm is used that searches for the ellipse that best approaches the papillary contour. In the literature there is no record of the use of genetic algorithms to solve this kind of problems. However the results obtained here are shown to be highly competitive with the methods mentioned earlier. Basically, the idea is to

transform the solution space of the problem under analysis, formed by all the papillary contours possible, into a new representation space where the application of the genetic approach is feasible. In this transformation process, we will explain how to encode the possible solutions of the original space of the problem and how to evaluate each one. Besides, we will show how to inject the available domain knowledge into different stages of the method to obtain the solution.

Characteristic	No. images
Cataract (severe or moderate)	0
Light artifacts	3
Some of rim blurred or missing	5
Moderate peripapillary atrophy	16
Concentric peripapillary atrophy/artifacts	20
Strong pallor distractor	6

Table 1. Image base visual characteristics.

We used a random sample of 124 eye fundus images selected from an eye fundus image base belonging to the Ophthalmology Service at Miguel Servet Hospital, Saragossa (Spain). From this initial image base, we eliminated all those eyes that had some type of cataract (severe and moderate) and were left with a base of 110 images. The mean age of the patients was 53.0 years (s.d. 13.05), with 46.2% male and 53.8% female and all of them were Caucasian ethnicity. 23.1% patients had chronic simple glaucoma and 76.9% eye hypertension. The images were acquired with a colour analogical fundus camera, approximately centred on the ONH and they were stored in slide format. In order to have the images in digital format, they were digitised using a HP-PhotoSmart-S20 high-resolution scanner, RGB format, resolution 600x400 and 8 bits/pixel. Table 1 shows, for the 110 images selected, all those visual characteristics related to potential problems that may distort the detection process of the papillary contour.

The article is organised as follows: section 2 describes the different stages of the method proposed. Section 3 focuses on the genetic algorithm-based approach with which it possible to solve the problem under study. Section 4 evaluates the method results and compares them with other methods. Finally, Section 5 summarises the conclusions of this work.

2. Method Description

Currently, the easiest way of obtaining information on the state of the ONH without using very costly equipment (desirable property in screening campaigns) is the two-dimensional photo. The aim of this work is to locate and segment the ONH using photographs of eye fundus as input information. Accordingly, the method that we propose consists of three stages. In the first stage a pre-processing of the eye fundus image is done. Next, in the second stage, a set of points, called set of *hypothesis points* (HPs), is obtained. HPs exhibit intensity levels characteristic of the points forming the papillary contour. For this, the property that the papilla pixels have an intensity level higher than retina pixels (see Figure 1.a) is used. Therefore, if the intensity level of the pixels in a segment, whose extremes are located inside the papilla and retinal area, respectively (see Figure 1.a), is plotted as a function of the distance from each pixel to the segment extreme located in the papilla, the papillary contour pixels will always be situated in positions near to a change in intensity level with a strong negative slope (Figure 1.b). This process is repeated for other segments that are angularly equispaced and form a bundle. From now on, each of these segments will be called *radial segment* (RS) and the plot of intensity levels of the points belonging to an RS will be called

intensity level profile. Finally, the last stage consists of selecting the most suitable HPs from the initial set and joining them properly with an ellipse to form a solution papillary contour. Based on a set of constraints and initial hypotheses, we will use a genetic algorithm (GA) to search for the optimum ellipse. Figure 2 summarises in a block diagram all the stages involved in the whole process. The following subsections describe each of these stages in detail.

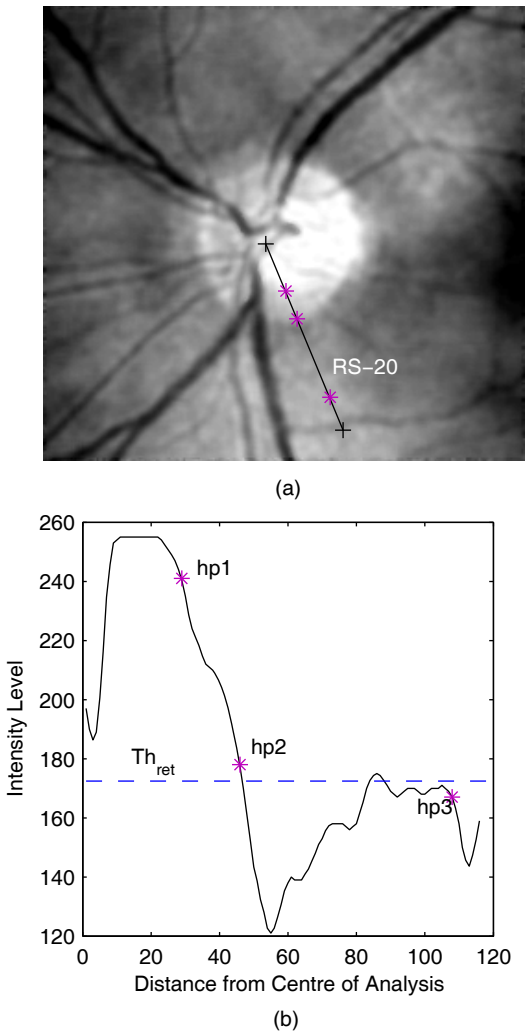


Figure 1. Obtaining hypothesis points (HPs) from the intensity level profile (ILP) of a radial segment (RS): a) example of RS-20 located in the normalised red channel of an eye fundus image; b) example of HPs located on strong negative slope sections of the ILP associated with RS-20.

2.1. Pre-processing

In order to reduce the computational cost associated with image processing, this stage begins by extracting a subwindow of the original image, of fixed size (261x261), where the papilla is approximately in the centre of the image. Bearing in mind that each photographic image is taken so that the ONH is near to the centre of the scene, this task is relatively simple: just thresholding in the three RGB channels is sufficient [14] to find the approximate position of the papilla and then centre the new analysis subwindow with respect to the centre of the raw segmented papilla. In order to smooth the image and eliminate spurious intensity values, a 5x5 Wiener filter was applied to the three RGB channels. Moreover, owing to the variability of intensity level range of the

different images, a RGB-normalisation was done to rescale images to the interval [0, 255] on each of the three channels associated with an image.

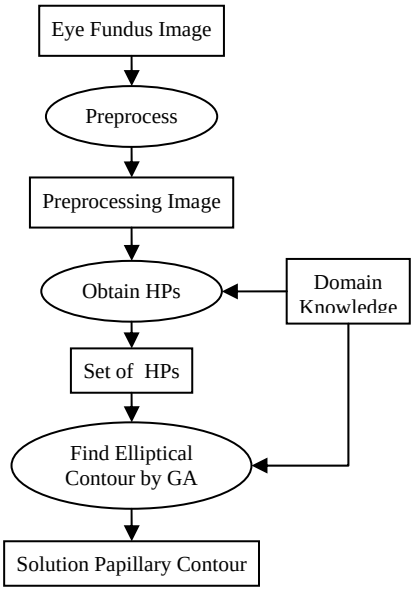
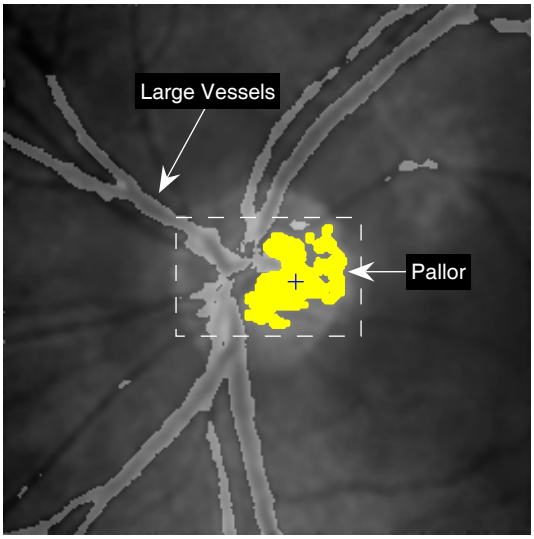


Figure 2. Stages involved in obtaining the papillary contour.

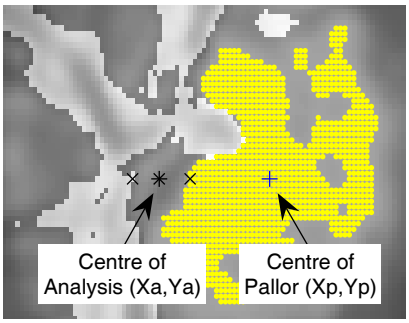
The point situated approximately in the centre of the papilla defining the centre of the bundle from which the RSs emerge is called the *centre of analysis* (CA). It is obtained automatically from the relative position of the pallor and the large blood vessels that leave the papilla at the top and bottom. Therefore, previously, we need to segment these two structures. Thus, the approximate segmentation of the pallor is achieved with thresholding (Th_p=253) of the normalised grey image, followed by selecting that region with the biggest area. Here the property of the pallor having high luminosity is used. Similarly, to locate the large vessels, a morphological bottom-hat filtering is performed, with a disc-shaped structuring element (radius=15), followed by thresholding, with a threshold equal to 15% of maximum value of normalised image intensity. Finally, to locate the coordinates of CA, (x_a, y_a), we used the following procedure: if (x_p, y_p) represents the coordinates of the geometric centre of the segmented pallor, then y_a=y_p; the coordinate x_a is calculated as the middle point of the shortest segment parallel to the x-axis at the height of y_p connecting the boundary of the segmented pallor with the boundary of the segmented large blood vessels (see Figure 3).

To simplify the processing of the two subsequent stages, all the right eye images are converted to the left eye using a horizontal reflection (it could also have been done the other way round). To determine the eye type, the relative position of the pallor with respect to the blood vessels is again used: in the left eye, the pallor is always to the right of the large vessels (see Figure 3.a).

In order to make the size of each RS independent of the image analysed, a fixed length, L_{RS}, is always used, which ensures that the opposite end to the CA is always situated in the retina. The value of L_{RS} is chosen depending on the resolution selected for the input images and is obtained from the statistical study of the size of a representative set of papillae.



(a)



(b)

Figure 3. Centre of analysis: (a) relative position between the pallor and the large blood vessels; (b) detail of the area of interest (see dotted rectangle in (a)).

Finally, Table 2 collects the value used for each of the parameters associated with the pre-processing stage. It is important to stress that these values (except for the normalisation range) are dependent on the resolution of the original image and they were heuristically selected for best performance. In most cases, the algorithm's overall performance is not very sensitive to changes in these parameters.

Pre-processing Parameters	**Value**
Subwindow Sise	261x261 pixels
Normalisation Range	[0, 255]
Wiener Filter	5x5 pixels
Pallor Threshold	253
Bottom-hat Radius	15 pixels
Bottom-hat Threshold	40
L_{RS} (Length of a RS)	110 pixels
N_{RS} (Number of RSs)	100

Table 2. Pre-processing stage parameter values.

2.2. Obtaining Hypothesis Points

In principle, the characterisation of a hypothesis point is based on the property mentioned at the beginning of this section: a point of the image which belongs to the papillary contour and presents a change in its intensity level with a strong negative slope. However, this property represents a necessary but not sufficient condition, since other causes may exist that give rise to the appearance of negative slopes in the intensity level profile. For example, the presence of blood vessels in the papilla, the imperfect uniformity of the colour of the retina and papilla or the existence of peripapillary atrophy areas. This is the reason why the set of points thus obtained are called hypothesis points, since this set may consist not only of solution points but also of other points which, generically, will be associated with noise. Next, a process for obtaining these points is described.

Let (p_1, ..., p_n) be the set of all the points belonging to an RS and ordered from less to greater according to their distance from the CA. An *interval of strong negative slope* (ISNS) is defined for an RS as a set of consecutive points, (p_r, ..., p_s), which satisfy the four conditions shown in Eq. (1), where *tan(x)* is the tangent operator on a curve at point *x*, α=-3 (71.6º) is a configurable threshold that marks the limit of the minimum tangent of the all points belonging to an ISNS, β=4 is a positive and configurable number that establishes the minimum number of points necessary to form an ISNS, and i=1,...,*m*, where *m* is the number of ISNSs in a RS. The idea is to obtain a representative point for each ISNS, called *strong negative slope hypothesis point* (SNS_HP). Specifically, to calculate the SNS_HP associated with the i-th ISNS in an RS, Eq. (2) is used, where *second_point_of* (*X*) is an operator that obtains the second point of the ordinate set *X*.

$$ISNS_{RS}^{i} = \left\{ \begin{array}{l} \left(p_r^i, ..., p_s^i\right) \subset \left(p_1, ..., p_n\right): \\ \left[\begin{array}{l} (s-r+1) \geq \beta \text{ and} \\ \mathbf{tan}(p_{r-1}^i) < \alpha, \text{if } p_r^i \neq \mathrm{p}_1, \text{and} \\ \mathbf{tan}(p_t^i) > \alpha, \forall t : r \leq t \leq s \text{ and} \\ \mathbf{tan}(p_{s+1}^i) < \alpha, \text{if } p_s^i \neq p_n \end{array} \right] \end{array} \right\} \quad (1)$$

$$SNS\_HP_{RS}^{i} = \left\{ p \in ISNS_{RS}^{i} : p = \text{second\_point\_of}\{ISNS_{RS}^{i}\} \right\} \quad (2)$$

Once the SNS_HPs associated with an RS have been obtained, a stage is applied to eliminate retina noise. Remember that the length of each RS was sufficiently large to reach the retina from the CA and, thus, ensure the inclusion of the points belonging to the papilla-retina boundary in the RS. However, imperfect uniformity of the intensity level of the retina can cause the existence of ISNSs on the RS section included in the retinal area and, consequently, of noisy SNS_HPs. In order to filter this noise, an intensity level threshold is defined, which is called the *retina threshold*, Th^{ret}, so that all the SNS_HPs obtained whose intensity level is below this value will be rejected (see hp_3 in Figure 1.b). This threshold is determined for each RS and is defined as the mean intensity level of its 50 pixels furthest from the centre of analysis, plus their standard deviation (Figure 1.b). The SNS_HPs not removed for the retina threshold are now called *filter hypothesis points* (F_HPs).

Indeed, the set of F_HPs are obtained for each colour channel using the same CA. It is important to analyse the three channels independently because the information obtained in each one is not always redundant and, therefore, using just one channel or the fusion of three in a grey level image could imply a loss of information. Since the number of F_HPs obtained, when handling the three channels, are many and very often close to one another, they can group together and form clusters along each RS. If each cluster is replaced by a new point, a large number of F_HPs will be removed, but the non-redundant information in each channel will be preserved. Thus, given the set of all the filter hypothesis points, (fhp_1, ..., fhp_k), obtained as a result of applying Eq. (2) and Th^{ret} in the same RS and in the three channels, and ordered from less to greater according to their distance from the centre of analysis, a *radial cluster*, $C_{RS,}$ is defined as the set of consecutive F_HPs, (fhp_r, ..., fhp_s), which satisfy the three conditions shown in Eq. (3), where $d(x,y)$ is the Euclidean distance operator between the points x and y, and λ=2 is a configurable threshold that establishes the maximum distance between two consecutive points belonging to the same cluster. Then, each cluster is transformed into a new point, called *cluster hypothesis point* (C_HP). Specifically, to calculate the C_HP associated with the j-th cluster in an RS, Eq. (4) is used, where the operator *mean* (X) represents the mean value of the points contained in the set X, and j=1,...,l, where l is the number of clusters in the RS.

$$C_{RS}^{j} = \left\{ \begin{bmatrix} \left(fhp_r^j, \ldots, fhp_s^j\right) \subset \left(fhp_1, \ldots, fhp_k\right): \\ d\left(fhp_{r-1}^j, fhp_r^j\right) > \lambda, \text{if } fhp_r^j \neq fhp_1, \text{and} \\ d\left(fhp_t^j, fhp_{t+1}^j\right) < \lambda, \forall t: r \leq t < s \text{ and} \\ d\left(fhp_s^j, fhp_{s+1}^j\right) > \lambda, \text{if } fhp_s^j \neq fhp_k \end{bmatrix} \right\} \tag{3}$$

$$C\_HP_{RS}^{j} = \left\{ p_j : p_j = \mathbf{mean}\left(C_{RS}^{j}\right) \right\} \tag{4}$$

The entire process followed here to obtain the C_HPs for an RS is repeated for the set of N_{RS}=100 radial segments in the bundle whose origin is always situated in the CA. As a summary, Figure 4 shows the pseudocode algorithm used to obtain the final hypothesis points, hereafter HPs. As an example, Figure 5 shows the HPs obtained from applying this algorithm to an eye fundus image. In relation to parameter values used in Eqs. (1)-(4), observe that the greater the values chosen for α (in absolute value) β and λ, the fewer the number of HPs associated with an RS will be, and vice versa. Therefore, we must find a compromise solution between not detecting the HP solution belonging to the real contour of the papilla in an RS (high values of α, β and λ) and obtaining too many HPs in this RS, including the solution point (very low values of α, β and λ). The values of α, β, λ reported throughout the paper were heuristically selected in this sense for best performance.

HP ={∅}
FOR each RS_j, j=1,..., N_{RS}
  FHP_j={∅}
  FOR each Ch^i, i∈{R,G,B}
    Obtains set $ISNS^i_j$, from (1)
    Obtains set $SNS\_HP^i_j = f(ISNS^i_j)$, from (2)
    Obtains set $F\_HP^i_j = f(SNS\_HP^i_j, Th_j^{ret})$
    $F\_HP_j = F\_HP_j \cup F\_HP^i_j$
  END for
  Obtains set $C_j = f(F\_HP_j)$, from (3)
  Obtains set $C\_HP_j = f(C_j)$, from (4)
  HP={HP, {C_HP_j}}
END for

Figure 4. Pseudocode algorithm to obtain the final set of hypothesis points, HPs, where RS_j is the j-th radial segment, Ch^i the i-th colour channel, $ISNS^i_j$ is the set of intervals of strong negative slope existing in the RS_j and associated with the Ch^i, $SNS\_HP^i_j$ is the set of strong negative slope points obtained from $ISNS^i_j$, $F\_HP^i_j$ is the set of filter hypothesis points obtained from $SNS\_HP^i_j$ and using the retina threshold, Th_j^{ret}, F_HP_j is the set of all the filter hypothesis points in RS_j from the three colour channels, C_j is the set of radial clusters existing in RS_j, and C_HP_j is the set of cluster hypothesis points associated with C_j. The final set of HPs stores the set of C_HPs associated with each RS separately, i.e., HP={{C_HP_1},..., {C_HP_{Nrs}}}.

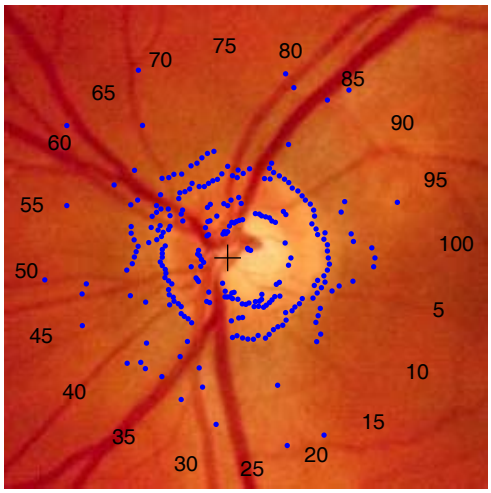


Figure 5. Eye fundus image final set of hypothesis points.

2.3 Obtaining the Papillary Contour

Since the human papillary contour always has a slightly oval aspect, this work proposes searching for the solution by approaching this contour with an ellipse, as done in other works, e.g. [4, 6]. For this, a genetic algorithm will be suitably instantiated (see details in Section 3), whose main objective will be to search for an ellipse containing the maximum number of HPs in its perimeter. From this instant, it is assumed that: 1) the geometric shape of the papilla can be approached by a non deformable ellipse, and 2) all those ellipses that are not a solution will contain few HPs, or what is the same, the more HPs that the ellipse contains, the more likely this ellipse will approach the real contour of the papilla. Later, by analysing the experiment results it will be possible to confirm the validity of these two hypotheses.

3. Genetic Algorithm

Genetic algorithms (GA) are one of the paradigms most frequently used in evolutionary computation and they owe their inspiration to the biological process of evolution: natural selection and survival of the fittest individuals [15]. Thus, given a problem of a specific domain, these algorithms code potential solutions using a structure of data like a chromosome (individual) where the genes are parameters of the problem proposed. The approach to the solution of each chromosome is calculated using a fitness function. Finally, by applying selection and variation operators (mutation and recombination), a population of chromosomes evolves to the optimum solution until a finalisation criterion is achieved. The modelling of a problem using GAs can be divided into four stages. The first consists of defining the most appropriate way of representing individuals or solutions. The second implies constructing the fitness function to assess the goodness of these individuals. The third consists of choosing or constructing the most appropriate selection and reproduction operators, in accordance with the coding of individuals and the nature of the problem to be solved. Finally, the fourth implies choosing the values of the parameters defining the intrinsic operating capacity of the GA (for example, maximum number of generations and size of population) or the operating capacity of the operators (for example, probability of mutation and probability of crossover).

Under the two hypothesis raised in the previous section, here we shall use the genetic paradigm to find that ellipse containing the maximum HPs in its contour. In order to manage the process of counting the number of HPs belonging to the perimeter of an ellipse, we work in fact with the HPs contained in an elliptic crown concentric to this ellipse and whose width is small enough to include only the HPs that actually belong to the papillary contour and large enough for the real papillary contour, which is not a perfect ellipse, to be included inside it. Nevertheless, the genetic search can produce local minima due to the presence of HPs associated with noise. The ellipse thus obtained could contain pallor points or another type of distractors that would prevent the contour proposed from representing the optimum solution. It is true that, thanks to the mutation operator and the population solution that a GA handles, it is possible to avoid local minima to a large extent. However to further minimise this possibility, and given that the CA is always chosen at a point very near to the pallor (see Figure 3), the HPs in each RS actually belonging to the pallor contour will always be nearer to the CA than those actually belonging to the papillary contour. Taking advantage of this property, a circumference is used, called *filter circumference*, with its centre in the CA, which aims to eliminate all those HPs that are inside the circumference and are associated with the pallor contour. There is a clear compromise when choosing the radius value of this circumference: it must be large enough to eliminate the large number of points belonging to the pallor and small enough not to filter papillary contour points. As will be seen in the next subsection, the optimum value of this parameter will be automatically obtained during the evolutionary process associated with executing the GA.

3.1 Coding of Individuals

From the description above, it can be established that the phenotypic space, solution space of the original problem, consists of the elliptic crown space defined from the infinite ellipses that can be traced in the image. However, as indicated at the beginning of this section, we are also interested in searching for the most appropriate elliptic

crown width and filter circumference radius value. Therefore, these two parameters will also form part of the problem search space. To code this type of solutions, the phenotypic space is transformed into a genotypic space consisting of real vectors of 7 variables [*Xe, Ye, φ, a, b, r, δ*]. With the first five components of this vector it is possible to code just one *hypothesis ellipse*. Thus, *(Xe,Ye)* represents the centre of this ellipse, *φ* the angle that its major axis forms with the x-axis, and *(a, b)* the magnitudes of its major and minor semi-axis, respectively. The sixth component, *r*, codes the radius of the filter circumference. It is not necessary to code the coordinates of the centre of this circumference because its centre always is CA. The last component, δ, represents the offset with which it is possible to define an elliptic crown concentric to the hypothesis ellipse and delimited by an external ellipse, (*Xe, Ye, φ, a+δ, b+δ*), and an internal ellipse, (*Xe, Ye, φ, a-δ, b-δ*).

3.2 Fitness Function

In order to assess the goodness of each individual, a fitness function is defined that assigns a numerical value to each of the solutions proposed by the GA in the genotypic space. Given that, for this, we are interested in counting the number of hypothesis points contained in each elliptic crown proposed, it is necessary to determine, as a preliminary stage, which points fall within this crown (see Figure 6). For example, let [Xe_i, Ye_i, φ_i, a_i, b_i, r_i, δ_i] be an individual of the population, then the following steps are applied:

1. Decoding of [Xe_i, Ye_i, φ_i, a_i, b_i, r_i, δ_i], (see Figure 6.a).

2. All the hypothesis points inside the filter circumference are eliminated (see Figure 6.b). Each remaining HP is called *circumference filter hypothesis point* (HP_{CF}).

3. The distance of each HP_{CF} to the hypothesis ellipse is calculated.

4. For each RS, a HP_{CF} is selected that is at a minimum distance from the hypothesis ellipse and we denote it by *minimum distance hypothesis point* (HP_{MD}), as shown in Figure 6.c.

5. Only the HP_{MD}s inside the elliptic crown are selected (see Figure 6.d). Each of these points is called *elliptic crown hypothesis point* (HP_{EC}).

6. Finally, the number of HP_{EC}s is counted.

To facilitate the optimum elliptic crown search, a new heuristic is introduced based on the idea of contour continuity. This heuristic establishes not only maximising the number of hypothesis points contained in the elliptic crown but also reducing the number of discontinuities. This will aid the convergence process, since, for example, given two crowns containing the same number of points, the one will survive that has the points distributed more uniformly along the contour. In order to achieve this aim, and bearing in mind that all the images were transformed to the left eye, the elliptic crown is divided into four regions labelled as the temporal region (the region closest to the temporal bone and defined in the range [-50º, 50º]), superior region, belonging to the range (50º, 130º), nasal region, the region closest to the nose and defined in the range [130º, 230º], and the inferior region, belonging to the range (230º, 310º). Therefore, in a first approach, the fitness function, F_1, is the one indicated in Eq. (5), where NP_X

denotes the number of points in region X of the elliptic crown and ω_i is the weighting factor of each of the addends.

$$F_1 = \omega_1 \cdot NP_T + \omega_2 \cdot NP_S + \omega_3 \cdot NP_N + \omega_4 \cdot NP_I \tag{5}$$

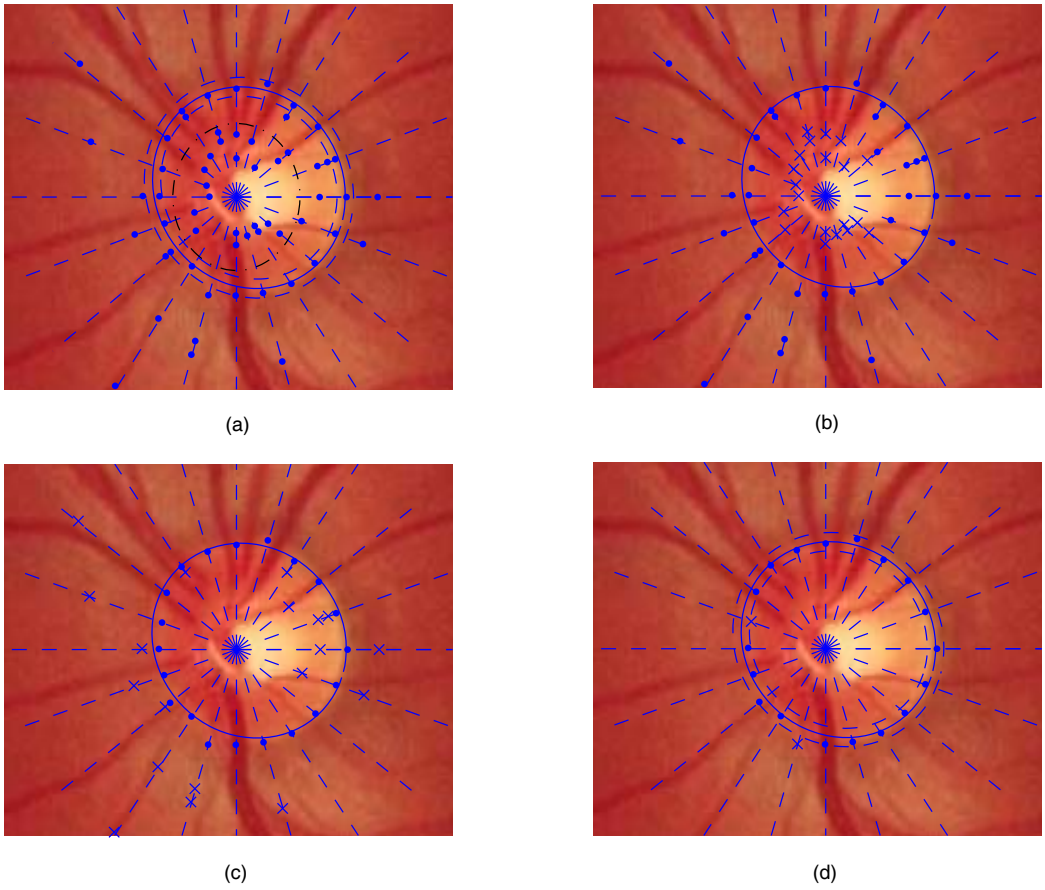


Figure 6. Stages in the process of counting hypothesis points in the elliptic crown: (a) HPs (•) and elements resulting from the decoding of an individual: ellipse (—), elliptic crown (- - -), filter circumference (— - —), (b) obtaining the HP_{CF}s (•) that are outside of the filter circumference (× HPs eliminated), (c) obtaining, in each RS, the HP_{MD}s (•) that are at a minimum distance from the genetic ellipse (× HPs eliminated) and (d) obtaining the HP_{EC}s (•) that are inside the elliptic crown (× HPs eliminated). In each figure only 20% of the RSs used are shown for clarity.

Another interesting phenomenon is related to the result of approaching the papillary contours traced by the experts with ellipses. Although the definition domain of the angle, φ, forming the major axis of an ellipse with the x-axis is $|\varphi| \in [0, 90]^o$, statistically it is shown that most papillary ellipses exhibit a major axis closer to the vertical than the horizontal axis. This makes it possible to define a new heuristic where the ellipse solution search is biased towards ellipses with $\varphi \rightarrow 90^o$. Moreover, we must remember that we are interested in maximising the filter circumference radius, r, and minimising the width of the elliptic crown, δ. Therefore, the idea is to include these variables in the fitness function in order to combine all these objectives. The new fitness function is the one indicated in Eq. (6).

$$F_2 = \omega_1 \cdot NP_T + \omega_2 \cdot NP_S + \omega_3 \cdot NP_N + \omega_4 \cdot NP_I + \omega_5 \cdot r + \omega_6 \cdot \delta + \omega_7 \cdot \varphi \quad (6)$$

Since some of the subobjectives considered in Eq. **(6**) need to be minimised and others maximised, it is necessary to redefine the fitness function to be able to consider the convergence in the same direction. Specifically, the final objective, without losing generality, was considered as a minimisation process. For this, all the subfunctions included in Eq. (6) and aimed towards a maximisation objective are transformed into new subfunctions where the new objective is now to minimise. Moreover, the range of variation of all of them is normalised in the interval [0,1].

Thus, the objective of maximising the number of points contained in each of the four regions defined earlier is transformed into a process of minimisation, as shown in Eq. (7), where $NP_X^{(min)}$ denotes the new subfunction to be minimised and NP_{Xmax} the maximum number of points that may be in region X.

$$NP_X^{(\mathbf{min})} = \frac{NP_{Xmax} - NP_X}{NP_{Xmax}} \quad (7)$$

Similarly, maximisation of the filter circumference radius is transformed too into a process of minimisation, as shown in Eq. (8), where $r^{(min)}$ is the new subfunction to be minimised and R_{max} and R_{min} are the maximum and minimum values that the filter circumference radius can have.

$$r^{(\mathbf{min})} = \frac{R_{\mathbf{max}} - r}{R_{\mathbf{max}} - R_{\mathbf{min}}} \quad (8)$$

The initial subobjective of minimising the width of the elliptic crown coincides with the minimisation strategy considered, so it will only be necessary to normalise, as indicated in Eq. (9), where $\delta^{(min)}$ is the new subfunction to be minimised and Δ_{max} and Δ_{min} are the maximum and minimum values that the width of the elliptic crown can have. Observe that as the value of δ is always defined with regard to the ellipse, the actual width of the elliptic crown is defined in the interval $[2\Delta_{min}, 2\Delta_{max}]$.

$$\delta^{(\mathbf{min})} = \frac{\delta - \Delta_{\mathbf{min}}}{\Delta_{\mathbf{max}} - \Delta_{\mathbf{min}}} \quad (9)$$

Finally, maximisation of the angle formed by the major axis of the ellipse proposed and the x-axis is now transformed into a process of minimisation, as shown in Eq. (10), where $\varphi^{(min)}$ is the new subfunction to be minimised, φ_{max} and φ_{min} are the maximum and minimum angles, in absolute value, of the major axis with the x-axis.

$$\varphi^{(\mathbf{min})} = \frac{\varphi_{\mathbf{max}} - |\varphi|}{\varphi_{\mathbf{max}} - \varphi_{\mathbf{min}}} \quad (10)$$

In short, the fitness function that the genetic algorithm must minimise is given by Eq. (11). The value of the weights, ω_i, is obtained experimentally by testing different configurations. Nevertheless, since the priority is to maximise the number of points for

each candidate ellipse crown, the magnitude of the weights associated with the $NP_i^{(\min)}$ will be weighted to a greater extent with regard to the rest of the weights, i.e. ω_1, ω_2, ω_3, $\omega_4 > \omega_5$, ω_6, ω_7. Moreover, from an anatomical point of view, a large number of the blood vessels, especially those with a large diameter emerge from the centre of the papilla and cross the papillary contour through the superior and inferior regions (see Figure 3.a). In these sections of the contour, there will be no transition between the papilla and retina and, consequently, there will be no associated hypothesis points. Consequently, the number of points found in the nasal and temporal region will be weighted more with regard to those found in the superior and inferior region. Furthermore, if it is admitted that quasi-symmetry exists for each coordinate axis, then $\omega_1=\omega_3 > \omega_2=\omega_4$. Next, the weights associated with the filter circumference radius and the elliptic crown offset will be weighted in importance and to an equal extent, i.e. $\omega_5=\omega_6 > \omega_7$. The smallest value is assigned to the weight ω_7.

$$F_3 = \omega_1 \cdot NP_T^{(\min)} + \omega_2 \cdot NP_S^{(\min)} + \omega_3 \cdot NP_N^{(\min)} + \ldots + \omega_4 \cdot NP_I^{(\min)} + \omega_5 \cdot r^{(\min)} + \omega_6 \cdot \delta^{(\min)} + \omega_7 \cdot \varphi^{(\min)} \quad (11)$$

Table 3 collects the value used for each of the parameters related to the fitness function. More importantly, the values of NP_{Xmax} depend, in addition to the angular range of each region $X \in \{T,N,S,I\}$, on the value of N_{RS}, namely, $\Sigma_x(NP_{Xmax})=N_{RS}$. The values of R_x and Δ_x depend on the resolution of the original image. The values of R_x, Δ_x and φ_x were obtained statistically from a representative sample of papillary contours when we approached the contour drawn by each expert with an ellipse. Finally, the values of ω_i, i=1,...,7, were obtained after trying several combinations and taking into account the constraints outlined in this subsection.

See eq.	Fitness Function Parameters	Values
(7)	$\{NP_{Tmax}, NP_{Nmax}, NP_{Smax}, NP_{Imax}\}$	{27, 27, 23, 23}
(8)	$\{R_{max}, R_{min}\}$	{40, 5} pixels
(9)	$\{\Delta_{max}, \Delta_{min}\}$	{4, 1} pixels
(10)	$\{\varphi_{max}, \varphi_{min}\}$	{90, 15} degrees
(11)	$\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7\}$	$\{2.5, 1.5, 2.5, 1.5, 0.8, 0.8, 0.4\}\cdot 10^{-1}$

Table 3. Value of each of the parameters associated with the fitness function.

3.3 Knowledge Injection and Constraint Handling

As well as the local knowledge injected in the different heuristics used in section 2 to obtain the set of HPs, modelling of the problem with the genetic approach is done using global knowledge. On the one hand, from the heuristics mentioned in the previous subsection to construct the fitness function and, on the other hand, from different heuristics used to initialise the population of individuals. The use of an ellipse as an approach to the solution, the choice of four regions in the papillary contour to account for discontinuities in the contour, the use of the filter circumference to eliminate false hypothesis points and the vertical position trend of the major semi-axis of the papillary ellipses belong to the first case. The choice of definition domains, limited above and below for each of the genes forming an individual, belong to the second case. In particular, the range of variation for each definition domain is based on statistical data obtained from the result of approaching the papillary contours traced by the two experts with an ellipse (see Table 4).

Variable	Statistics	Value
Major axis	(Min, Max)	(41.7, 60.3)
	Mean ± std	50.6 ± 4.1
Minor axis	(Min, Max)	(37.2, 54.9)
	Mean ± std	45.9 ± 3.9
|Angle| of major axis with X-axis	(Min, Max)	(2.9, 89.9)
	Mean ± std	63.3 ± 20.5
Ratio major-minor axis	(Min, Max)	(1.01, 1.30)
	Mean ± std	1.10 ± 0.056

Table 4. Statistics obtained when approaching each contour papillary drawn by the experts with an ellipse.

There is also a series of constraints based on the definition of the variables involved. Thus, for example, the solution ellipse always has to contain the CA, the minor semi-axis cannot be greater than the major semi-axis, the filter circumference radius cannot be greater then the minor semi-axis. Other constraints arise as a result of imposing the definition domains associated with each gene. Others appear as a consequence of observing certain structural parametric relations: for example, a threshold is imposed on the minor-major axis relation below which the associated ellipse would not be a solution. This fact is based on the low eccentricity associated with the solution ellipses (see Table 4).

Finally, constraint handling is done using a penalty function, *P(x)*, defined according to Eq. (12), where $\phi(\boldsymbol{x})$ is a boolean function, $R^n \rightarrow \{0,1\}$, which returns *true* if and only if the vector $\boldsymbol{x}$ satisfies all the *n* constraints and *k* is a constant much greater than zero, in our case $k=10^3$. The idea is to transform the problem of minimising the fitness function described in Eq. (11), subject to satisfying the set of constraints indicated above, into a new minimisation problem without constraints. For this, just incorporating the penalty function into the original fitness function and obtaining the new fitness function to be minimised, indicated in Eq. (13), is sufficient.

$$P(x) = \begin{cases} 0, & \text{if } \phi(x) = true \\ k >> 0, & \text{otherwise} \end{cases} \tag{12}$$

$$F(x) = F_3(x) + P(x) \tag{13}$$

3.4 Implementation of GA

The coding of the GA was done from the MATLAB toolbox (*Genetic Algorithm and Direct Search, v.1.0.1*). Specifically, the initialisation of the population was done randomly on the definition domain chosen for each gene, as shown in Table 5. The values in this table (except for the φ gene) are dependent on the resolution of the original image and all of them are obtained from the statistical data shown in Table 4. Each individual in the population was represented with a real vector of fixed dimension. This type of representation, which arose as a direct consequence of the problem formulation, has the advantage of facilitating the use of standard recombination and mutation operators. Thus, for the first case, the well-known one-point crossover operator applied probabilistically in accordance with its *crossover ratio* was used. For the second case, the non-uniform mutation operator with Gaussian distribution was used [16]. This latter operator was applied with a probability of 1 per gene. However, the magnitude of the mutation applied to each gene decreases in each new generation and is

proportional to the standard deviation of a zero-mean Gaussian distribution. This decrease is controlled using two parameters called *scale* and *shrink*. The first controls the standard deviation of the mutation in the first generation, and the second controls the ratio with which the mean mutation magnitude decreases. The standard deviation decreases linearly so that its final value is (1–*shrink*) times its initial value in the first generation. For example, if *shrink*=1, the mutation magnitude in the ultimate generation would be zero. A *generational model* was chosen as the population model. This choice directly determines the type of survival selection method, based on age. In principle, this selection method was combined with an *elitism scheme*, in an attempt to prevent losing the current fittest member in the population. However, as we will see in the next section, the experiment results showed that the population of elite reduced the performance. The parent selection is done by stochastic uniform sampling. Table 6 summarises this configuration.

Gene	Definition Domain
$[Xe_{min}, Xe_{max}]$	[80, 180]
$[Ye_{min}, Ye_{max}]$	[80, 180]
$[\lvert\varphi_{min}\rvert,\lvert\varphi_{max}\rvert]$	[15º, 90º]
$[a_{min}, a_{max}]$	[30, 70]
$[b_{min}, b_{max}]$	[30, 65]
$[r_{min}, r_{max}]$	[10, 40]
$[\delta_{min}, \delta_{max}]$	[1, 4]

Table 5. Definition domain of each of the genes in a chromosome.

Initialisation	*Random*
Representation	*Real vector (dim 1x7)*
Recombination	*One-point crossover*
Mutation	*Non-uniform with Gaussian distribution*
Parent selection	*Stochastic uniform*
Survival Selection	*Generational with elite*

Table 6. Summary of the GA configuration used in our approach.

4. Testing and Results

To measure the features of our algorithm we used 110 images obtained from the initial image base after eliminating all the cases of eyes with cataracts. In order to do the result of the evaluation quantitatively reproducible, we measured the mean discrepancy between the points of the contour obtained with the segmentation method and a gold standard (GS). For each image, the GS was defined from a contour that was the result of averaging two contours, each of them traced by an expert. Here we shall use the discrepancy definition, δ, done in [4], whose formula is determined by Eq. (14), where:

δ^j is the discrepancy measurement for the image *j*.

$i = 1..N$, where N is the number of angularly equispaced RSs.

m_i^j is the length of the radius defining the *i*-th point of the ellipse proposed for the image *j*.

μ_i^j and σ_i^j are the mean and typical deviation, respectively, of the lengths of the radii defining the *i*-th point of the contours traced by the experts and belonging to the image *j*.

$\varepsilon = 0.5$ is a small factor to prevent division by zero where the experts are in exact agreement.

$$\delta^j = \frac{\sum_{i=1}^{N} \left(\frac{|\, m_i^j - \mu_i^j \,|}{\sigma_i^j + \varepsilon} \right)}{N} \tag{14}$$

For visualisation purposes, the ogive of discrepancy is plotted, namely, the number of images with discrepancy less than δ, $|\{j : \delta^j < \delta\}|$, versus δ. Thus, the number (percentage) of images that fit any given level of accuracy can be read off the y-axis. Owing to the stochastic nature associated with the GA, each ogive of discrepancy represented in the different graphs depicted in the next subsection, corresponds in fact to the result of averaging five ogives of discrepancy, obtained as a result of executing the GA five times with the same parameter configuration.

4.1 Experiment Results

Different experiment results are shown in Figure 7 which aim to appropriately characterise the values associated with the intrinsic parameters of the GA proposed: population size (*P*), number of generations (*G*), crossover ratio (*C*), *shrink* value (*S*), associated to the mutation operator, and elite population size (*E*).

Figure 7.a shows the effect of the population size when the rest of the GA parameters remain constant: (*E*=2, *C*=0.5, *S*=0.5), and finalisation condition (*Cf*), obtained on reaching 200 generations, i.e., *Cf*=200G. It is observed that as *P* increases, the global performance of the discrepancy curve improves. Since a small population size may require more generations to produce a performance comparable to that of a larger population, an experiment was done where the product *P*x*G* was constant. The results are shown in Figure 7.b. It can be observed that, although the ogive of discrepancy for small populations improves when *G* increases (compare with Figure 7.a), there is a saturation value, *P*=400, above which significant improvement does not occur.

The variation in the size of the elite population, considering the rest of the fixed GA parameters (*P*=100, *C*=0.5, *S*=0.5), showed an interesting result (Figure 7.c): although no significant changes occur when the size of $E \in \{2, 4, 8\}$ varies, the results improve when no elite population is used in the evolutionary process (*E*=0). Observe that, in order to prevent these results from being biased because of a large population, a value of *P*=100 was used which, as in the previous experiment (Figure 7.b), exhibited a curve of intermediate performance. Similarly, in order to reduce the computational load associated with a fixed size of generations, the finalisation condition of the GA was made more flexible: reaching 200G maximum or 50G consecutively in which there is no improvement in fitness of the best individual, *Cf*=200*G*50.

Analysis of the effect of crossover and mutation operators was done in three experiments (see Figure 7.d-f). In each experiment, the *S* value remained fixed and only the *C* value varied. The rest of the parameters always remained constant (*P*=100, *E*=0, *Cf*=200*G*50). The experiments revealed the following: for a small crossover value (*C*=0.2) the results improve as the *S* value increases and the best result was obtained for *S*=0.8. Conversely, a high crossover value (*C*=0.8), irrespective of the *S* value, showed some similar discrepancy curves but had worse performance than that obtained for

C=0.2. Consequently, the optimum combination would be a high *S* value together with a low *C* value.

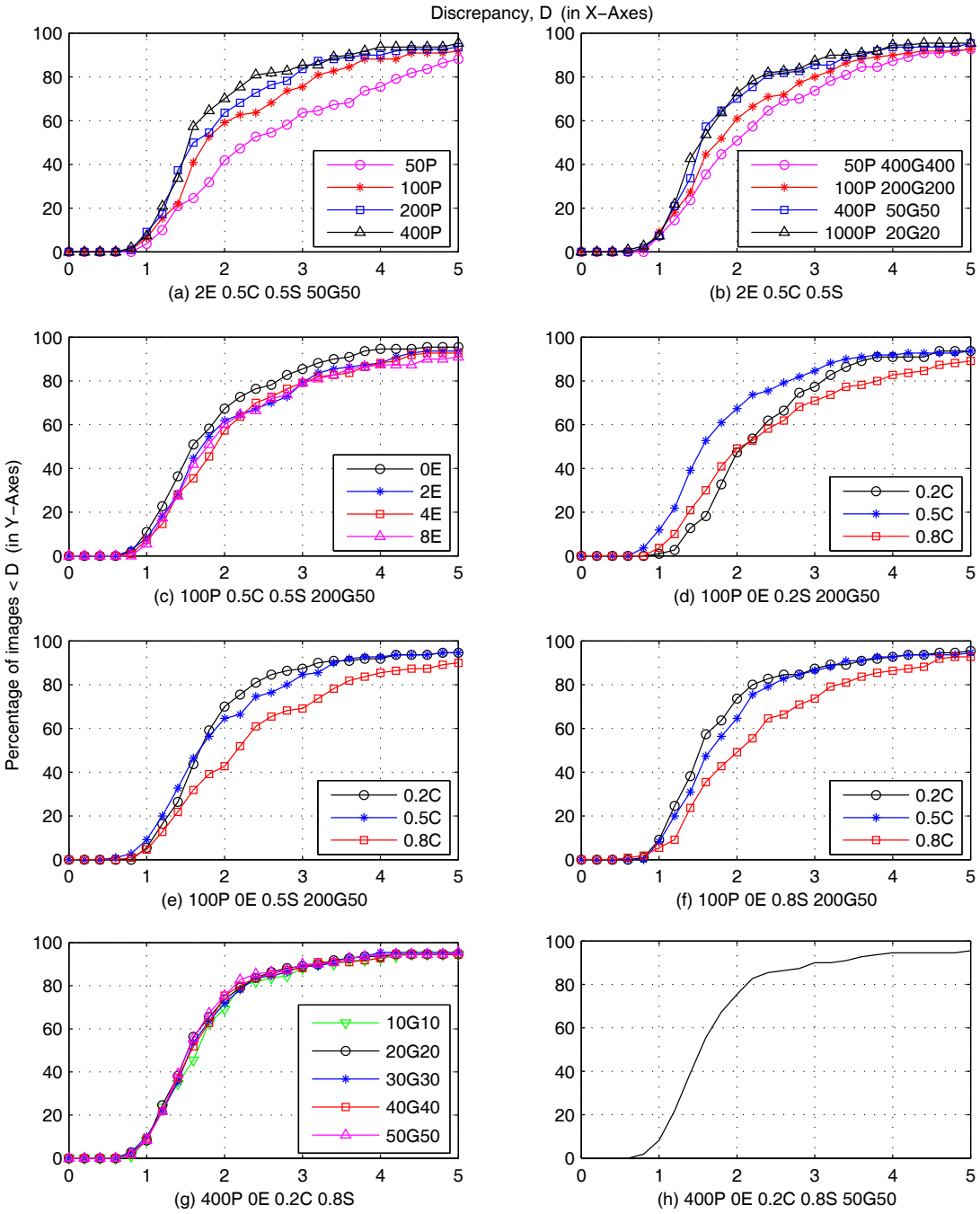


Figure 7. Accumulated mean discrepancy results for different GA configurations (population size=α_1 (α_1P), elite=α_2 (α_2E), crossover ratio=α_3 (α_3C), shrink=α_4 (α_4S), number of generations(G)=α_5 and stop condition if there is no improvement in α_6 consecutive generations (α_5Gα_6)): (a) for different population sizes, (b) for PxG=$2{\cdot}10^4$, (c) for different elite sizes, (d), (e) and (f) for different crossover values (C$\in${0.2, 0.5, 0.8}) with S=0.2, 0.5 and 0.8, respectively, (g) for different generation sizes in GA final configuration (400P, 0E,0.2C,0.8S) and (h) for GA final configuration (400P, 0E,0.2C,0.8S, 50G50).

4.2 Evaluation of Results

In the light of the results obtained in the experiments, the following final parameter configuration was chosen for the GA. A population size, *P*=400, because higher values

did not imply significant improvements in the accumulated discrepancy curve and, moreover, it would have meant increasing the computational cost unnecessarily. Owing to the good performance of the discrepancy curve obtained for *E*=0, there was no elite population in the evolutionary process. The crossover fraction and shrink values, *C*=0.2 and *S*=0.8, were chosen in accordance with the conclusion obtained in the preceding section. Note that a value of *C*=0.2 implies that, in each generation, 20% of the offspring is obtained with crossover, the other 80%, with mutation. However, owing to the type of mutation operator used, it should be pointed out that the magnitude of the mutation decreases with the number of generations. Specifically, a value of *S*=0.8 indicates that the standard deviation value of the Gaussian distribution decreases linearly until 0.2 times the initial value. Therefore, the high shrink value only implies a high mutation value in the initial generations, guaranteeing the extensive search in all the genotypic space during the initial period. In subsequent generations, the gradual decrease in the magnitude of the mutation will guarantee convergence and the fact that this magnitude never reaches zero in final generations will guarantee the diversity of the population. Moreover, the crossover value used is only apparently low, bear in mind that 20% of a population of 400 individuals means obtaining 80 offspring by crossover in each generation, i.e., improved population due to genetic mixing is also guaranteed.

To study the dependency of the computational cost regarding the number of generations, a final experiment was done where the final configuration chosen remained fixed (400P, 0E, 0.2C, 0.8S), and *G* was varied. The results (see Figure 7.g) show, except for G=10, an almost total overlapping of the curves for the four remaining G values. Obviously, if we are interested in minimising the computational cost, a value of G=20 would be the most appropriate without losing virtually any performance. Finally, Figure 7.h shows the discrepancy obtained for the final GA configuration chosen.

4.3 Comparison with other Algorithms

In the bibliography on automatic identification of ONH, the most usual way of evaluating results have been subjective, where the expert determines whether the contour obtained is valid for subsequent treatment. However, some quantitative evaluations have also been done. From all of these, the most objective is the direct comparison between the contour proposed and a gold standard, generated from averaging different contours traced by different experts. In this comparison the variability of the experts can be taken into account [4] or not [17]. This is an important issue because there is an inter-observer variability among different experts tracing the contour of a papilla and also intra-observer variability in each expert when he traces the contour of the same papilla at different instants of time [18, 19]. In another quantitative evaluation method [5], the comparison is made using characteristic parameters of the ellipse: vertical diameter, area, etc. Finally, other works have evaluated the contour proposed using indirect measurements, like [20], where the diagnosis of glaucoma obtained with the automatic contour is compared to that obtained with the contour traced by an expert.

In the bibliography consulted by the authors and discarding the evaluations done subjectively and not quantitatively, we believe that the Lowell *et al.* method [4] is the one that has obtained the best results to date. Therefore, we shall use this approach as a reference to compare our results. Thus, Figure 8.a shows the discrepancy curves obtained from applying the two methods to our image base and Figure 8.b shows the comparison of the two methods on the image base used in [4]. Given that the Lowell *et*

al. (LOW) method is highly dependent on the initial radius values, *r*, and aspect ratio, *a*, applied in its temporal lock phase, we use different values for these parameters, $r \in [40, 50]$ and $a \in [1.0, 1.1]$ when applying the LOW's method to our image base. The range of variation of these parameters was taken bearing in mind the papilla sizes of our image base (see Table 4). From the different discrepancy curves obtained for each value, in Figure 8.a we only show the best discrepancy curve obtained for r=47 and a=1.03. We believe that worse discrepancy results obtained with the LOW's method in our image base in comparison with his image base is because of the strong temporal lock phase dependency of his method with regard to the parameters *r* and *a*. Thus, the comments already made in this respect in [4] are confirmed. On the other hand, it can be observed that both discrepancy curves obtained with our method are very similar in the two image bases. This invariability provides evidence of our method's generalization capability, i.e., suggesting similar results for different image bases.

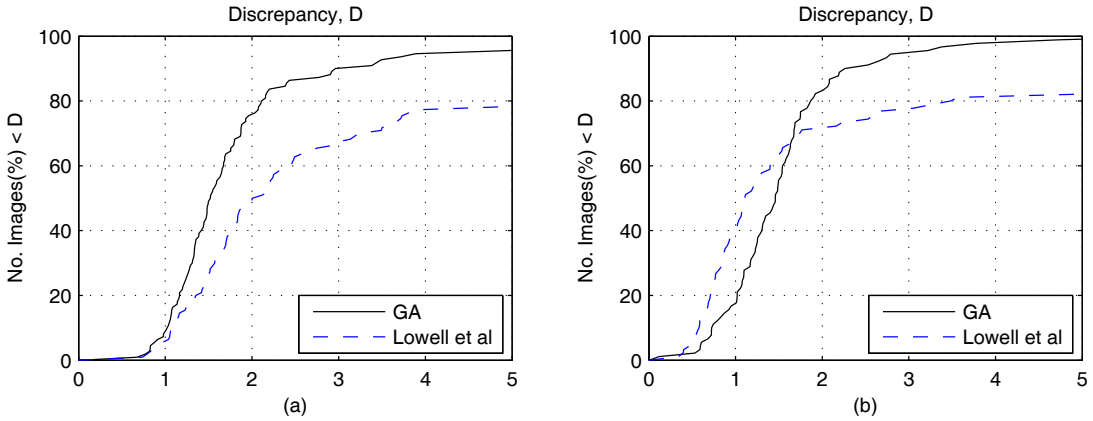


Figure 8. Accumulated discrepancy results for our GA versus the Lowell *et al.* method: (a) in our image base, (b) in Lowell's image base.

Direct comparison of the discrepancy curve obtained by both methods in the LOW's image base (see Figure 8.b) reveals two very different areas with opposite behaviours whose boundary is marked by the discrepancy value δ=1.6. Below this value, LOW's results are better than our proposal, and above this value, precisely the opposite occurs. We think that the explanation for the results obtained in the first discrepancy stretch is that while our method approaches the solution using a non deformable ellipse, the local deformation phase in LOW's method makes it possible to do a slight deformation of the ellipse obtained in the phase immediately before (global fit) and, thus, it better approaches the experts' real trace. However, the change in trend on the second stretch of the curves reveals the strong robustness of our method. In fact, while LOW's method only obtains 82% of images below a discrepancy δ=5, our method obtains 99%. This highlights that in LOW's method a bad approach done by the global fit cannot be offset by the local fit. Instead, from the two discrepancy curves obtained from our database (see Figure 8.a), better behaviour of our method is observed in all the discrepancy range. That might reveal that the generalization capability of the LOW's method is lower than our method: 96% of images with δ<5, compared with 79% in LOW's method. The robustness of our method is explained by the GA property of avoiding those solutions associated with local minima and because this property is reinforced by the injection domain knowledge in the constraint and fitness function definition which guides the GA search towards the real solution.

Finally, Figure 9 shows different examples of papillary contours obtained with our method. Although qualitatively Lowell's group defined four categories: *excellent*, *good*, *fair* and *poor*, containing images with discrepancy up to one, two, five, or more, respectively, it should be borne in mind that the discrepancy value is also a function of inter-observer variability, since it affects typical deviation in Eq. (4). In other words, the greater the experts' variability when tracing the contour of a papilla, the smoother the discrepancy error obtained. This fact is observed when the contours of the papillae in positions (3,1) and (1,2) in Figure 9 are compared. In fact, contour C_{31}, in spite of having a greater discrepancy than the contour C_{12} (δ_{31}=1.33, δ_{12}=0.65) subjectively presents a quality with regard to its gold standard, GS_{31}, comparable to the quality that C_{12} has with GS_{12}. The explanation lies in the fact that the mean standard deviation of C_{31} (σ_{31}=2.07) is greater than that of C_{12} (σ_{12}=0.92). Obviously, this type of circumstance only is significant in small discrepancies. Our experts estimated it in $\delta \in [0, 3]$. Otherwise it is impossible to obtain contours with high discrepancies that fit the GS well as a consequence of high inter-observer variability. The reason for this is that the experts' variability is always limited.

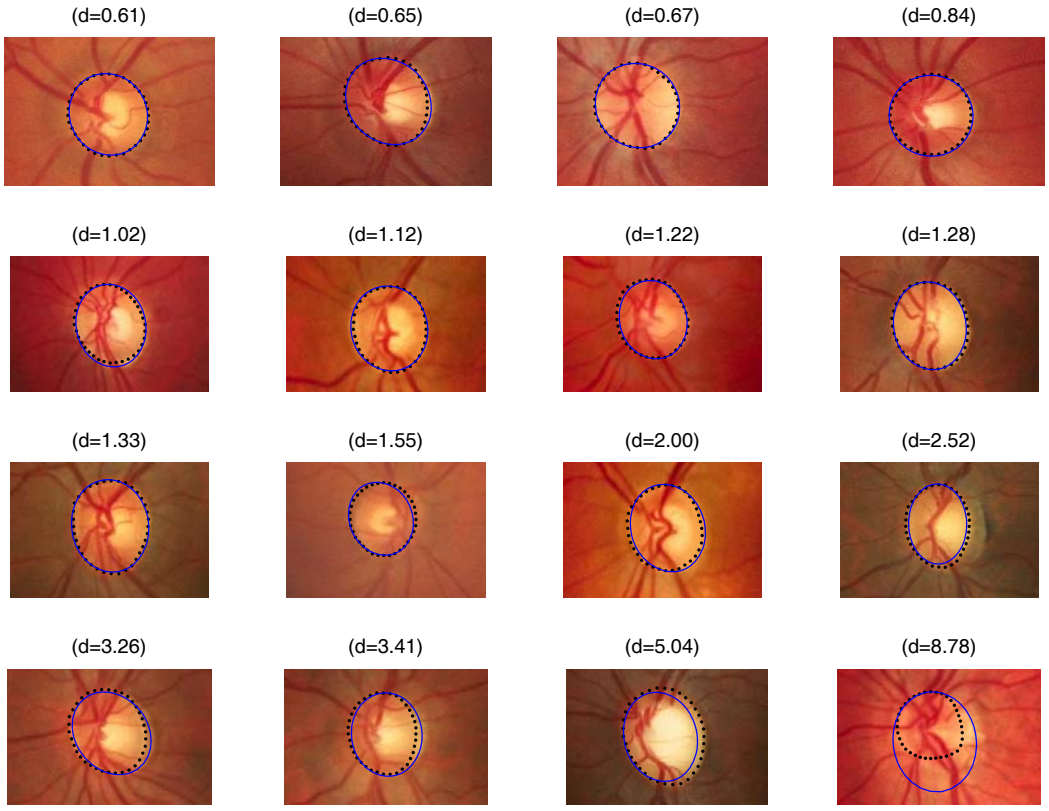


Figure 9. Several examples of segmentation obtained with our segmentation method: algorithm (solid line), gold standard (dotted line).

5. Conclusions

Although the enormous usefulness of active contours must be stressed for detecting contours in medical images and their frequent use in systems for locating the ONH, in this work they have not been used, precisely because they are highly dependent on the

initialisation stage. The great variability in the images used makes an approximate initialisation of the active contour to the papillary contour virtually impossible in all instances. Thus if we wish to construct a robust system, it is necessary to stress more the precise location stage of the contour and use domain knowledge to guide the process. In our case, local knowledge was initially used to obtain a first approach to the papillary contour using the set of hypothesis points: set of pixels with geometric characteristics and intensity levels similar to those of the papillary contour points. Then, global knowledge was injected into the GA. Thus, it was not only possible to construct a fitness function that appropriately guided the evolutionary process but also to minimise obtaining suboptimum solutions (local minima) associated with the presence of different types of distractors.

When our method is applied to a different image base from the one used in our study, similar discrepancy curve results are obtained, providing evidence of its generalization capability. In addition, the robustness of the method proposed can be seen in the high percentage of images obtained with a discrepancy $\delta<5$ (96% and 99% in our and the Lowell *et al*. image base, respectively). With this result it is possible to confirm the validity of the two initial hypotheses on which the construction of the fitness function is based: the geometric shape of the papilla can be approached with a non deformable ellipse and every ellipse that is not a solution contains few hypothesis points in an offset of its perimeter. Confirmation of the first hypothesis may be closely related to the fact that intra- and inter-observer variability does not explain the local fit, which follows the global fit, typical of active contour-based approaches.

Another important aspect of the method is that it directly provides the parameters characterising the shape of the papilla: lengths of its major and minor axes, its centre of location and its orientation with regard to the horizontal position. These parameters or new ones obtained as a result of combining them, together with other morphological magnitudes of the papilla, are frequently used by the expert in the diagnostic process of eye pathologies.

Number of Generations	10	20	30	40	50
Average time per image (s)	16.94	33.46	49.37	65.30	80.88

Table 7. Average time used per image to obtain a solution according to the number of generations (see Figure 7.g). The rest of the GA parameters remained fixed (400P, 0E, 0.2C, 0.8S).

One of the disadvantages of the method proposed is related to the computational cost of using GAs. Bearing in mind that all the automatic detection system of the papillary contour was programmed in MATLAB and this was executed in an Intel Pentium M processor, 1.80 GHz, the average time per image for obtaining the papillary contour, for the final configuration of GA parameters and, according to the number of generations, is as shown in Table 7. Nevertheless, it is necessary to highlight that the time obtained thus is additionally penalised due to the fact that MATLAB is an interpreted language. Both the use of the compilable language and the parallelisation of the GA are obvious solutions to reduce this cost. This will be the aim of future work.

Acknowledgments

We would like to thank the Ophthalmology Service at Miguel Servet Hospital (Saragossa, Spain) for providing us with the image base used in this study. We would also like to thank the Lowell *et al.* research team for allowing us access to their image base [4] and the MATLAB implementation of their ONH segmentation method [4]. Finally, the authors would also like to thank *Fundación Mutua Madrileña del Automóvil* for their support with this work.

References

[1] T. Damms and F. Dannheim, Sensitivity and specificity of optic disk parameters in chronic glaucoma, *Invest. Ophtalmol 34*, (1993), 2246-2250.

[2] G. Zahlmann, B. Kochner, I. Ugi, D. Schuhmann, B. Liesenfeld, A. Wegner, M. Obermaier, and M. Mertz, Hybrid fuzzy image processing for situation assessment, *IEEE Eng Med Biol Mag 19 (1)*, (2000), 76-83.

[3] A. Hoover and M. Goldbaum, Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels, *IEEE Transactions on Medical Imaging, 22(8)*, (2003), 951-958.

[4] J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, E. Fletcher and L. Kennedy, Optic nerve head segmentation, *IEEE Transactions on Medical Imaging 23 (2)*, (2004), 256-264.

[5] A. Osareh, M. Mirmehdi, B. Thomas and R. Markham, Colour Morphology and Snakes for Optic Disc Localisation, in: A Houston and R Zwiggelaar (eds.), 6th Medical Image Understanding and Analysis Conference, BMVA Press (2002), 21-24.

[6] M.B. Shields, *Color Atlas of Glaucoma*. Baltimore: Williams and Wilkins, 1998.

[7] J. Cox and I. Wood, Computer-assisted optic nerve head assessment, *Ophthal. Physiol. Opt. 11*, (1991), 27-35.

[8] T. Morris and I. Wood, The Automatic Extraction of the Optic Nerve Head, in: American Academy of Optometrists, Biennial European Meeting in Amsterdam, (1994), 11-22.

[9] T. Morris and C. Donnison, Identifying the neuroretinal rim boundary using dynamic contours, *Image and Vision Computing 17*, (1999), 169-174.

[10] T. Morris and Z. Newell, Segmentation of retinal images guided by the wavelet transform, in: B.G. Mertzios and P. Liatsis (eds.), Proceedings of the IWISP '96, Elsevier, Amsterdam (2002), 303-306.

[11] T. Morris and Z. Newell, Location of the optic nerve head boundary, in: Proceedings Medical image understanding and analysis (MIUA'98), Leeds, (1998).

[12] F. Mendels, C. Heneghan and J.P. Thiran, Identification of the optic disk boundary in retinal images using active contours, in: Proceedings of the Irish Machine Vision and Image Processing Conference (IMVIP'99), Dublin, Ireland, (1999), 103-115.

[13] A. Osareh, M. Mirmehdi, B. Thomas and R. Markham, Classification and Localisation of Diabetic-Related Eye Disease, in: A. Heyden, G. Sparr, M. Nielsen and P. Johansen (eds.), 7th European Conference on Computer Vision, LNCS 2353, Springer-Verlag (2002), 502-516.

[14] M. Bachiller, M. Rincón, J. Mira and J. García-Feijoó, An Automatic System for the Location of the Optic Nerve Head from 2D Images, in: J. Mira and A. Prieto, (eds.), Proceedings of the IWAN'01, LNCS, vol. 2085, Springer-Verlag, London (2001), 319-327.

[15] J.H. Holland, *Adaption in Natural and Artificial Systems*, (MIT Press, Cambridge, 1992), (1st edition: The University of Michigan Press, Ann Arbor, 1975).

[16] A.E. Eiben and J.E. Smith, *Introduction to Evolutionary Computing* (Springer-Verlag, Berlin, 2003).

[17] H. Li and O. Chutatape, Boundary detection of optic disk by a modified ASM method, *Pattern Recognition 36 (9)*, (2003), 2093-2104.

[18] R. Varma, W.C. Steinmann and I.U. Scott, Expert agreement in evaluating the optic disc for glaucoma, *Ophthalmology 99 (2)*, (1992), 215-221.

[19] J.M. Tielsch, J. Katz, H.A. Quigley, N.R. Miller and A. Sommer, Intraobserver and interobserver agreement in measurement of optic disc characteristics, *Ophthalmology 95(3)*, (1988), 350-356.

[20] R. Chrastek, M. Wolf, K. Donath, H. Niemann, D. Paulus, T. Hothorn, B. Lausen, R. Lammer, C.Y. Mardin and G. Michelson, Automated segmentation of the optic nerve head for diagnosis of glaucoma, *Medical Image Analysis 9 (4)*, (2005), 297-314.