www.icgst.com

# Survey of Retinal Image Segmentation and Registration

Mai S. Mabrouk[1], Nahed H. Solouma[2] and Yasser M. Kadah[3]
[1]*Biomedical Engineering, Misr University*
[2]*Laser in Engineering, NILES, Cairo University*
[3]*Biomedical Engineering, Cairo University*

## Abstract
Diagnosis and treatment of several disorders affecting the retina and the choroid behind it require capturing a sequence of fundus images using the fundus camera. These images are to be processed for better diagnosis and planning of treatment. Retinal image segmentation is greatly required to extract certain features that may help in diagnosis and treatment. Also registration of retinal images is very useful in extracting the motion parameters that help in composing a complete map for the retina as well as in retinal tracking. This paper introduces a survey for the segmentation and registration techniques that were reported as being well for retinal images.
***Keywords*** *Retinal images, Registration, Segmentation, Motion Parameter Estimation, Real Time Tracking.*

## 1. Introduction
Retinal image processing is greatly required in diagnosing and treatment of many diseases affecting the retina and the choroid behind it [1], [2]. Diabetic retinopathy is one of the complications of diabetes mellitus affecting the retina and the choroid. In this condition, a network of small blood vessels, called choroidal neovascularization (CNV), arises in the choroid and taking a portion of the blood supplying the retina. As the amount of blood supplying the retina is decreased, the sight may be degraded and in the severe cases, blindness may occur. The physicians try to treat this dangerous disorder by applying optical energy to photocoagulate the neovascularization. Argon laser is used in photocoagulation purposes to cauterize the small vessels which increases the amount of blood supplying the retina and thus maintaining the sight. This treatment modality is achieved in many sessions. The physician asks the patient to fixate his/her eye to be able to direct the laser beam to the affected area. The current success rate of this procedure is below 50% for eradication of CNV following one treatment session with a recurrence and/or persistence rate of about 50%. The latter condition requires repeating the treatment. Each treatment repetition in turn has a 50% failure rate. Moreover, several studies indicate that incomplete treatment was associated with poorer prognosis than no treatment [3], [4]. Consequently, the need to develop an automated laser system to treat the whole retina in one session has become a necessity. This system is intended to scan the retina and track it applying the laser energy to whole area except the sensitive objects that may be damaged by the laser energy. The system is assumed to do this by capturing the retinal images using a fundus camera. These images are to be accurately segmented to extract the sensitive objects in the retina such as the blood vessel tree, the optic disk, the macula and the region between the optic disk and the macula. The positions of laser shots are to be distributed in the rest of the retina. Also a robust registration technique is to be applied to detect the motion parameters of the retina to update the positions of laser shots accordingly. [5]
Moreover, the fundus camera can only provide an image for a portion the retina but not the whole retina. The physician sometimes needs to have a complete image for the retina to be able to have a reliable diagnosis and hence plan for good treatment. This problem may be overcome by some image processing algorithms to build a complete map for the retina. [6].
In this paper, we collect most of the image segmentation and registration techniques that were reported to be the best for retinal images. These algorithms were applied to our images to confirm its validity and accuracy. The paper is divided into four main sections: section I is an introduction, in section II, we describe the materials used in this work and give an explanation for the methods of retinal image segmentation and registration. Section III is a conclusion and Section IV is an acknowledgement. Finally a list of references is provided in section V.

## 2. Materials and Methods

The retinal images used in this work were acquired using a TOPCON TRC-501A fundus camera. All images collected for this work were acquired with the camera adjusted at 50˚ mode (this corresponds to the field of the eye seen by the camera) and with the red-free illumination mode. A Sony Charge Coupled Device (CCD) video camera was attached to the eyepiece of the fundus camera to make the images collected using the fundus camera available in standard video format. The video output from the CCD camera was interfaced to a Micron PC (2000 MHz processor, 512 Mbytes of RAM) through a video digitizer card. The retinal images captured by this setup may be one of the three possible types. The first type is gray level images in the normal case, (i.e., no injected dye). In this type of images, the blood vessels appear as dark objects on a bright background. The second type is gray level images with brighter blood vessels relative to a darker background. This type of images is obtained when the patient is injected with the Indo-Cyanine Green dye. The third type is the colored images that appear with the true colors of the retina. The images were captured to the memory of the computer system by means of a software package called Image Net. This package can capture either individual frames or real time video sequences. The size of the frame is 640x480 pixels for grey level images and 640x480x3 for colored images. The images were saved on a hard disk for further processing.

## 2. a Retinal Image Segmentation

In the retinal images acquired using the setup described above, the problem of inaccurate segmentation comes mainly as a result of non-uniform illumination of the background. Moreover, the variable distance of different retinal areas from the camera causes further degradation because of the expected loss of focus in some areas. These effects cause gray level variations within the same image and gray level differences between different images. This results in significant complications in detecting the blood vessel tree accurately [17]. The retinal images have many objects to be extracted and fully recognized because of their importance in diagnosing and treatment of retinal disorders. These objects are: the blood vessel tree, the optic disc, the macula, the region between the macula and the optic disc and the exudates if present. The following section describes the recommended methods for segmenting these objects and gives the results of applying these methods on our images.

### Segmenting the Blood Vessel Tree

The blood vessel tree appears as dark structure in brighter background in the normal image (i.e., the images with no injected dye). If the patient is injected with an Indo-Cyanine Green (ICG), the blood vessel tree appears as bright structure in a darker background.

Several Studies have been conducted in the area of blood vessel extraction from retinal images as well as; from other medical images as extracting the coronary artery in the cardiac images. Generally speaking, these studies can be classified into two main categories:

1. Detection of Blood Vessel Boundaries and
2. Extraction of the Core Area of the Blood Vessel Tree by Tracing Vessel Centers.

### Detection of Blood Vessel Boundaries by the Difference Operators

In digital form, derivatives may be implemented in several ways based on obtaining the differences in intensities of a pixel and its surrounding pixels. There are many masks that can achieve this task and they are called the *difference operators*. These masks include the "**Roberts** masks", the "**Perwitt** masks" and the "**Sobel** masks".

However, the **Sobel** operators have the advantage of providing both a differencing and a smoothing effect. Because derivatives enhance noise, the smoothing effect is a particularly attractive feature of the Sobel operators [9].

The **Canny** edge detector is a promising method in detecting the boundaries of the blood vessels. Canny edge detector first smoothes the image by a Gaussian filter to eliminate noise. It then finds the image gradient to highlight regions with high spatial derivatives. An edge point is defined to be a point whose strength is locally maximum in the direction of the gradient. The edge points determined give rise to ridges in the gradient magnitude image. The algorithm then tracks along the top of these ridges and set to zero all pixels that are not actually on the ridge top so as to give a thin line in the output, a process known as *nonmaximal suppression*. The ridge pixels are then thresholded using two thresholds $T_1$ and $T_2$, with $T_1<T_2$. Ridge pixels with values greater than $T_2$ are said to be strong edge pixels. Ridge pixels with values between $T_1$ and $T_2$ are said to be week edge pixels [10]. Finally the algorithm performs edge linking by incorporating the week pixels that are 8-connected to the strong pixels. Results of applying the these operators are shown in figure.1

### Boundary Detection Using Image Statistics

Image statistics such as the mean and the standard deviation can be used in boundary detection. The algorithm explained below is called DBDED, which stands for decision-based directional edge detector [11], used for automatic target recognition (ATR) based on the image statistics. In this work it is used to detect the boundaries of the blood vessel tree in the retinal images. Each point that passes a local threshold is investigated as an edge candidate. Then, the point $(x,y)$ is to be a one-dimensional edge

candidate from the east direction if it satisfies the following equation:

$$I(x,y) \geq Av[2I(x+1,y), I(x+2,y), I(x+3,y)] + Sd[2I(x+1,y), I(x+2,y), I(x+3,y)] + \eta$$

Where $I(x,y)$ is the intensity of an image at coordinates $x$ and $y$, $\eta$ is a constant threshold and $Av(.)$ and $Sd(.)$ denote average and standard deviation operations respectively. After following this procedure in all eight directions, the decision algorithm is applied and points are considered as 2-D edges if they satisfy the following conditions:

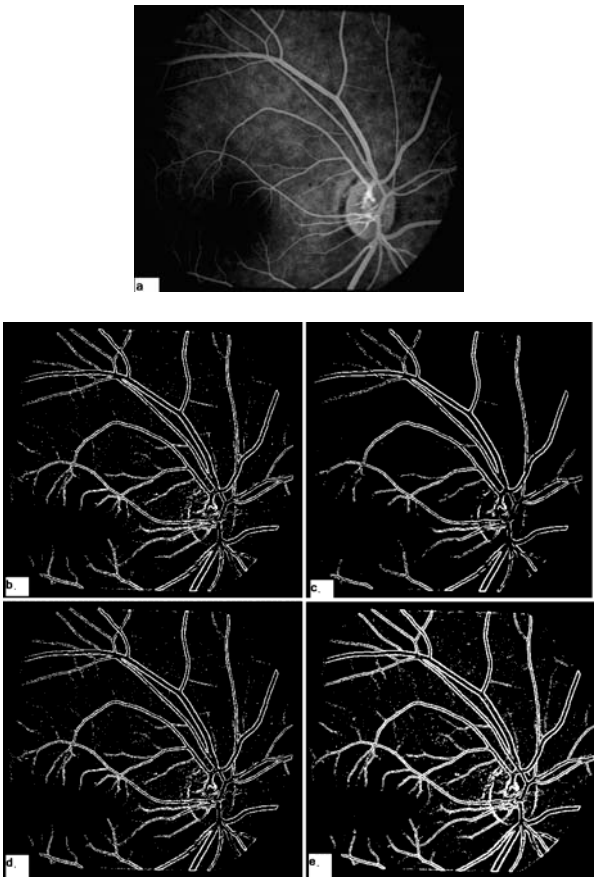


Figure 1: a. Original Image, b. Result of applying sobel operator, c. Result of applying prewitt operator, d. Result of applying Robert operator and e. Result of applying canny operator.

- They are 1-D edge candidates in at least two and at most seven directions.
- If $(x,y)$ is an edge candidate then, at least one of the immediate 8- neighboring Points:

$\{(x+\beta, y+\gamma) \mid \beta \in (0,1,-1), \gamma \in (0,1,-1), \beta+\beta\gamma+\gamma \neq 0\}$

is also an edge candidate. The result of applying this algorithm is shown in Figure.2.

Figure 2: Result of applying DEBDED on the image shown in figure 1.

**Extraction of Blood Vessel Boundaries Using Deformable Models**

The recent one of the methods of contour detection is deformable models or *snake.* [12-16]. A snake is an active contour model that is manually initiated near to the contour of interest. This contour model deforms according to some criteria and image features to finally stay to the actual contour(s) in the image. An energy function is formulated to obtain an estimate of the quality of the mode in terms of its internal shape, and external forces e.g. underlying image forces and user-constraint forces. The energy function integrates a weighted-linear combination of the internal and external forces of the contour:

$$1 \leq i \leq L-N+1, \qquad 1 \leq j \leq M-N+1$$

The internal energy of the contour with respect to elastic deformations and the bending of the snake:

$$\varepsilon_{internal}(v(s)) = \alpha_{elasticity}(s)\varepsilon_{elasticity}(v(s)) + \alpha_{bending}(s)\varepsilon_{bending}(v(s))$$
$$= \alpha_{elasticity}(s)\|v_s(s)\|^2 + \alpha_{bending}(s)\|v_{ss}(s)\|^2$$

The first order derivative term $v_s(s)$ make the snake behave like a membrane and represent the elastic energy of the contour. The second order derivative term $v_{ss}(s)$ makes the snake act like a thin plate and represents the contour bending energy. Decreasing $\alpha_{elasticity}$ allows the contour to develop gaps, while increasing $\alpha_{elasticity}$ increases the tension of the model by reducing its length. Decreasing $\alpha_{bending}$ allows the active contour model to develop corners, and increasing $\alpha_{bending}$ increases the bending rigidity, making the contour smoother and less flixible. Setting either of the weighting coefficients to zero permits first and second order discontinuities respectively [12].

The external energy term $\varepsilon_{image}$ represents the energy due to image forces like lines, edges and terminations of line segments and corners. In [13], the following image energy terms are suggested, consisting of a weighted-sum of the terms.

$$\varepsilon^*_{image}(v(s)) = \alpha_{line}(s)\varepsilon_{line}(v(s)) + \alpha_{edge}(s)\varepsilon_{edge}(v(s)) + \alpha_{term}(s)\varepsilon_{term}(v(s))$$

We studied the effect of the internal energies and compared it with the effect of the external energy (edge energy in this case) to make sure that the effect of the external energy will be dominant. The image was convolved by a Sobel operator in both the horizontal and the vertical directions. The magnitude

of the result, which represents the edge value, was taken to represent the external energy of the contour. Instead of selecting a single point having the minimum energy (i.e., maximum edge value) from the neighborhood of each snaxel, we search the neighboring points of each snaxel and extract all the points having a value above certain threshold as new points on the contour. This not only deforms the contour but also makes it grow. The obtained contour can be thinned to one pixel thickness, but in our work we need to obtain safety margins around the blood vessels [14, 15].

Approximately the same result was obtained when we let the constants $\alpha_{elasticity}$ and $\alpha_{bending}$ equal to zero. Also we neglect the term of user constraints because it does not have a reasonable effect on the images collected for this work. Hence, the computation requirements are reduced greatly and this accelerates the process and saves time. The snake now deforms according to the image energy (i.e., the edge energy). Neglecting the internal energy terms may lead to discontinuities in the contour, but we can overcome this problem by initiating multiple contours [16].

Instead of manual initialization of contours, the whole image was covered automatically with an initial set of small contours called *seed* contours. During the contour iteration convergence and growth, the seed contours that lie within the areas of poor edges shrink until vanishing [15]. The others will merge and/or split (the merging of the contours get ride of the discontinuities that may result) until recovering a continuous description of all edges that pass certain threshold. This process is repeated until no changes occur during a given iteration. This process leads to the complete detection of all small vessels in addition to the boundaries of the larger ones. Result of applying the deformable model is shown in figure.3.

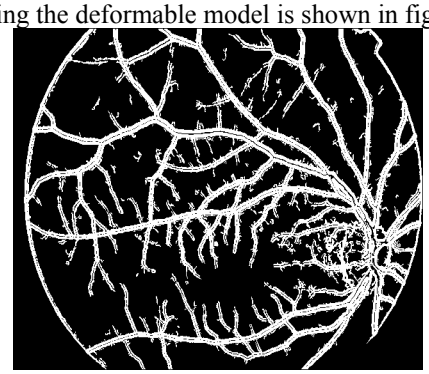


**Fig (3):** Result of applying deformable model on the image shown in fig.1

**Boundary Detection Using Morphological Gradient**

As mentioned above, dilation and erosion are the two basic methods of mathematical morphology. The morphological gradient of an image can be computed using the dilation and erosion as follows:

$$G = (f \oplus b) - (f \Theta b)$$

Where, $f \oplus b$ denotes the dilation of the image $f$ by the structuring element $b$ and $f \Theta b$ denotes the erosion of the image $f$ by the structuring element $b$. After subtracting the eroded image from the dilated image, a suitable threshold is applied to extract the edges [9]. The result of computing the morphological gradient of retinal images and taking the points that pass certain threshold is shown in Figure.4. The same problem of obtaining false edges in the background also exists.

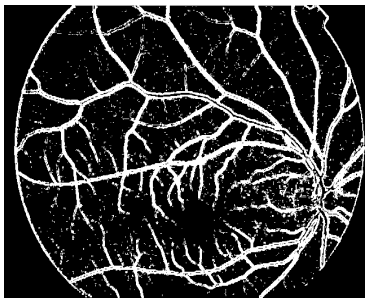


Figure 4: Result of applying morphological gradient on image shown in fig.1.

**Extraction of Blood Vessel Tree Using the Morphological Reconstruction**

Morphological reconstruction is to reconstruct an object in an image, called the marker image, containing at least one point belonging to that object from an image, called the mask image, containing that object and other objects and noise. Let $J$ and $I$ be two binary images defined on the same discrete domain D, and such that $J$ is a subset of $I$. In terms of mapping this means that for each pixel $p \in D$, $J(p) = 1 \rightarrow I(p) = 1$. $J$ is called the *marker* image and $I$ is the *mask*. Let $I_1, I_2 \ldots I_n$ be the connected components of $I$. Then the reconstruction of mask $I$ from marker $J$ denoted by $\rho_I(J)$ is the union of the connected components of $I$ ,which contain at least one pixel of $J$., it is as follows:

$$\rho_I(J) = \bigcup_{J \cap I_K \neq 0} I_K$$

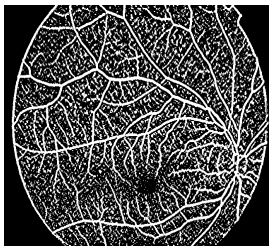
The following explanation concerned with both the binary and gray scale images, but the emphasis is put on binary reconstruction because it is the case used in our work. An efficient implementation of morphological reconstruction can be described as follows:

1. Label the connected components of the mask image, (i.e.) each of these components is assigned a unique number.
2. Determine the labels of the connected components, which contain at least a pixel of the marker image.
3. Remove all the connected components that are not of the previous ones [17]. Figure.5 shows the results of applying the morphological reconstruction.

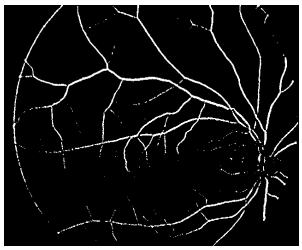
**Retinal Image Segmentation Using Watersheds**

Here, we applied the watershed algorithm which is known to be best applied on images with bulb-like structures but it was observed that it may give reasonable results with retinal images. In grey scale mathematical morphology the watershed transform is the method of choice for image segmentation. The watershed transform can be classified as a region-based segmentation approach. The intuitive idea underlying this method comes from geography: it is that of a landscape or topographic relief which is flooded by water, watersheds being the divide lines of the domains of attraction of rain falling over the region [9]. An alternative approach is to imagine the landscape being immersed in a lake, with holes pierced in local minima. Basins (also called `catchments basins') will fill up with water starting at these local minima, and, at points where water coming from different basins would meet, dams are built. When the water level has reached the highest peak in the landscape, the process is stopped. As a result, the landscape is partitioned into regions or basins separated by dams, called watershed lines or simply watersheds.

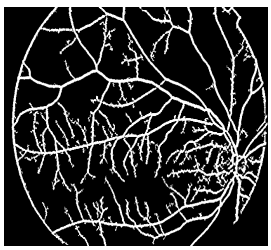
a. The original Image



b: The marker image



C: The mask image



d: The final result

Figure 5: Segmentation using morphological reconstruction

An algorithmic definition of the watershed transform by simulated immersion was given by Vincent and Soille [9]. Let f: D → N be a digital grey value image, with $h_{min}$ and $h_{max}$ the minimum and maximum value of f. Define a recursion with the grey level h increasing from $h_{min}$ to $h_{max}$, in which the basins associated with the minima of f are successively expanded. Let $X_h$ denote the union of the set of basins computed at level h. A connected component of the threshold set $T_h$+1 at level h + 1 can be either a new minimum, or an extension of a basin in $X_h$: in the latter case one computes the geodesic influence zone of $X_h$ within $T_{h+1}$, resulting in an update $X_{h+1}$. Let $min_h$ denote the union of all regional minima at altitude h. Define the following recursion:

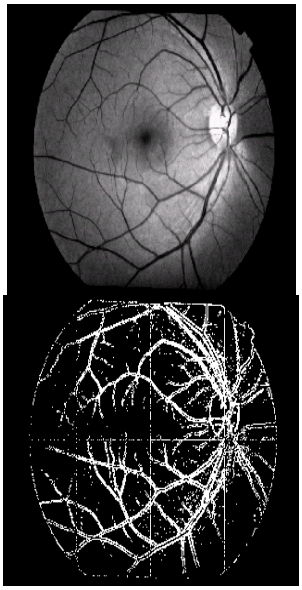
$$(X_{hmin} = \{p \in D \mid f(p) = h_{min}\} = T_{hmin}$$

$$X_{h+1} = min_{h+1} \cup IZ_{Th+1}(X_h), \quad h \in 2\ [h_{min}, h_{max})$$

The watershed Wshed(f) of f is the complement of $X_{hmax}$ in D:

$$Wshed(f) = D \setminus X_{hmax}$$

According to the recursion, it is the case that at level h + 1 all non-basin pixels (i.e. all pixels in $T_{h+1}$ except those in $X_h$) are potential candidates to get assigned to a catchment basin in step h + 1. Therefore, the definition allows that pixels with grey value $h_0 \leq h$ which are not yet part of a basin after processing level h, are merged with some basin at the higher level h + 1. Pixels which in a given iteration are equidistant to at least two nearest basins may be provisionally labeled as `watershed pixels' by assigning them the label W (we will refer to such pixels as W-pixels). However, in the next iteration this label may change again. A definitive labeling as watershed pixel can only happen after all levels have been processed. Figure.6 shows the results of applying watersheds [18].



a. Input image b. final result

Fig (6): Results of watershed algorithm

**Extraction of the Core Area of the Blood Vessel Tree by Matching the Image with Gaussian Filter:** Blood vessels have a special property that the two edges of a vessel always run parallel to each other. Such objects may be represented by piecewise linear directed segments of finite width. S. Chaudhuri et al

in [19] observed two interesting properties of the blood vessels in retinal images. First, since the blood vessels usually have small curvatures, the anti parallel pairs may be approximated by piecewise linear segments. Second, although the intensity profile varies by a small amount from vessel to vessel, it may be approximated by a Gaussian Curve:

$$f = A\{1 - k\exp(-d^2 / 2\sigma^2)\}$$

Where $d$ is the perpendicular distance between the point $(x, y)$ and the straight line passing through the center of the blood vessel in a direction along its length, σ defines the spread of the intensity profile, $A$ is the gray level intensity of the local background, and $k$ is a measure of reflectance of the blood vessel relative to its neighborhood.

Although the width of a vessel decreases as it travels radially outward from the optic disc, such a change in vessel caliber is a gradual one. The widths of the vessels are found to lie within a range of 2-10 pixels. For these calculations, however, the authors assumed that all the blood vessels in the image are of equal width 2σ.

So, they designed a two-dimensional Gaussian filter to match the blood vessel segments. Since the intensity profile can be assumed to be symmetrical about the straight line passing through the center of the vessel, the optimal filter must have the same shape as the intensity profile itself. In other words, the optimal filter is given by:

$$h_{opt}(d) = -\exp(-d2 / 2\sigma 2)$$

It may be appreciated that the vessel may be oriented at any angle $\theta$ ($0 \le \theta \le \pi$). The matched filter will have its peak response only when it is aligned at an angle $\theta \pm \pi/2$. Thus, to match the vessels in all possible directions, the filter needs to be rotated for all possible angles, the corresponding responses are to be compared and for each pixel only the maximum response is to be retained.

It has already been mentioned that the vessels may be considered as piecewise linear segments. Instead of matching a single intensity profile of the cross section of a vessel, a significant improvement can be achieved by matching a number of cross sections (of identical profile) along its length simultaneously. Such a kernel may be mathematically expressed as:

$$K(x,y) = -exp(-x^2/2\sigma^2) \qquad \text{For } | y | \le L/2$$

Where: $L$ is the length of the segment for which the vessel is assumed to have a fixed orientation. Here the direction of the vessel is assumed to be aligned along the $y$-axis. To match the vessels at different orientations, the kernel has to be rotated accordingly.

Taking (σ = 2) and (L = 9), the two-dimensional matched Gaussian filter in the discrete case is represented by a kernel of size 15*16 after subtracting the mean value of the filter to minimize false detection. Figure.7 shows the results of applying the Gaussian matched filter algorithm.

### Segmenting the Optic Disc

The optic disc is a well-delineated object with properly defined size and edges. [20] The nerve fibers and the blood vessels enter the retina through the optic disc. In the gray level retinal images, it always appears as a nearly circular structure that is brighter than the background of the retina. The authors in [21] used the Hough transform, which is well known as a robust method to extract circular objects, to detect the center and the radius of optic disc.

## Segmenting the Macula

The macula is the area of acute vision within the retina. It appears as the most homogeneous area near the optic disc. It can be segmented either by region growing algorithm or by morphological reconstruction. To segment the macula, we manually selected a point in the fovea (the center of the macula), then by a region-growing algorithm, the macula was extracted as the connected region around the fovea having intensities near the intensity of the central point. In reference [22], the radius of the macula was taken to be double the size of the radius of the optic disc in the $y$-direction and three times the same radius in the $x$-direction [1]. Figure.8, show the result of demarcating the optic disc and the macula in a normal image (no injected dye). The macula also can be segmented by the morphological reconstruction as in figure.8.

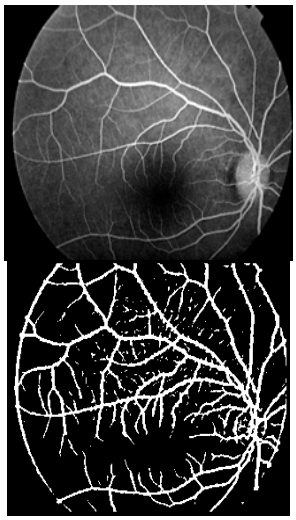


Fig (7): Core detection by Gaussian matched filter: a. Original image, b. final result.

### 2. b. Retinal Image Registration:

Image registration is to detect the similarity between different image frame by obtaining the motion parameters either to align different frames in one or more image or to restore the images to a reference

image. The motion parameters are generally represented by the vector $[T_x\ T_y\ \theta]^T$, where: $T_x\ T_y$ are the translation in the $x$ and $y$ directions and $\theta$ is the angle of rotation. The similar images are generally misregistered by rotation, scaling, and translation. Rotation is the angular twisting of one image with respect to the other, scaling is the magnification of one image with respect to the other, and translation is the displacement of the two images in the horizontal and/or the vertical directions. Other problems can complicate the registration, such as the nonlinear distortion in one image and not in the other, or when one object is viewed from two directions. In case of retinal images, the rotational movement is known to be very small (approaching and never exceeding 5°), and the scaling of one image with respect to the other can be controlled by asking the patient to fixate his eyes, and the head is hold to the fundus camera. The translation occurs in the vertical and in the horizontal directions and as there is no retinal detachment, the retina is known to move rigidly. The registration system can be characterized by its computation time and its probability of error [23].

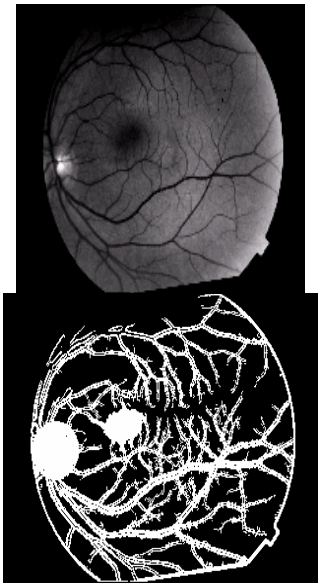


Fig (8): **S**egmenting the optic disc and the macula: a. original image, b. segmented image.

**Similarity Detection by Correlation:**

The method most widely used for the automatic determination of translation is correlation [24]. The elements of the non normalized cross-correlation surface $R(i, j)$ are defined to be:

$$R(i^*, j^*) = \sum_{l=1}^{N}\sum_{m=1}^{N} W^2(l,m), \quad \begin{matrix} 1 \le i \le L-N+1 \\ 1 \le j \le M-N+1 \end{matrix}$$

In the correlation scheme, a correlation output surface such as $R(i,j)$ is searched for a maximum at $(\hat{\imath}, \hat{\jmath})$. The procedure is successful if $(\hat{\imath}, \hat{\jmath})$ and $(i^*, j^*)$ are equivalent. As a counterexample, however, consider the non normalized cross correlation of (2) even in the ideal case where $W$ exactly matches some subimage; i.e. $W = S_N^{i^*j^*}$. Then:

$$R(i^*, j^*) = \sum_{l=1}^{N}\sum_{m=1}^{N} W^2(l,m)$$

Also for this ideal case, consider the non matching point $(\hat{\imath}, \hat{\jmath})$ where:

$$S_N^{\hat{i},\hat{j}}(l,m) = \max_{l,m} W(l,m) = W_N \qquad for \quad all \quad l,m$$

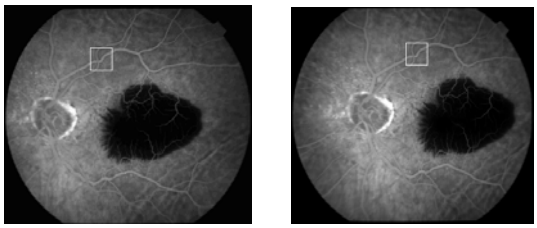
Clearly

$$R(\hat{i}, \hat{j}) = W_N \sum_{l=1}^{N}\sum_{m=1}^{N}, \qquad W(l,m) \ge R(i^*, j^*)$$

Therefore, even in the ideal case, a search for a maximum over $R(i,j)$ does not necessarily yield the registration point. Normalization is therefore necessary in even the simplest of cases [33]. The normalized cross correlation is given by:

$$R^2{}_N(i,j) = \frac{\left(\sum_{l=1}^{N}\sum_{m=1}^{N} W(l,m) S_N^{i,j}(l,m)\right)^2}{\left[\sum_{l=1}^{N}\sum_{m=1}^{N} W^2(l,m)\right]\left[\sum_{l=1}^{N}\sum_{m=1}^{N} S^2{}_N^{i,j}(l,m)\right]},$$
$$1 \le i \le L-N+1, 1 \le j \le M-N+1$$

Figure 9 shows the position of the window W selected from the reference image and the position of maximum response in the search image. Knowing the centers of the two windows, we can get the expected $T_x\ T_y$. The normalized correlation can also be used to obtain $\theta$, if we rotate the search image by a step angle and every time we apply the above equation and get the point of maximum response. Finally we have to extract the angle of maximum response from those maxima and the corresponding angle of rotation is estimated as $\theta$ [25].



a. The window chosen in the the reference image b. The position of window in the current image

Figure 9: Matching by correlation

## Registration Using Gabor filter

There are many features to consider, which can represent nature of image itself, such as lines, edges, colors and etc. One of the most interesting features is the orientation of the image. Since every image has its own orientation (for instance

an image with a lot of building has much more horizontal frequency than the vertical ones, or vice versa), main energy of an image is likely to spread to a particular direction. Gabor filter has been widely used in object recognition applications because filtering images with Gabor functions helps estimating directional blob or direction edge components [26]. The authors in [26], evaluates the directional features of a host image with 2-d Gabor kernel, which have their own energy spread into specific directions.

A 2-D Gabor function is the product of an elliptical Gaussian function with a complex exponential representing harmonic modulation. A general form of the 2-d Gabor function is given by:

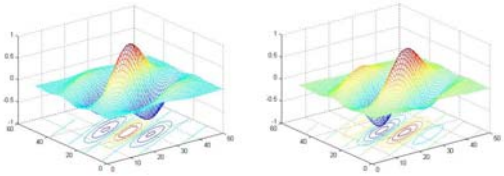
$$g(x,y) = e^{-[(x-x_o)^2/2\sigma_x^2+(y-y_o)^2/2\sigma_y^2]} * e^{-j(2\pi/\lambda)[(x-x_o)\cos(\theta)+(y-y_o)\sin(\theta)}$$

Where $(x_o,y_o)$ is the center of the elliptical Gaussian function, $\sigma_x^2, \sigma_y^2$ are the variances of the Gaussian function in the horizontal and vertical directions, and θ and λ are the orientation and wavelength of the harmonic modulation function, respectively. In addition assuming the unity aspect ratio $(\sigma_x/\sigma_y = 1)$, the function can be represented as a real Gabor function (RGF) and an imaginary Gabor function (IGF) as follows:

$$RGF(x,y) = \cos\left[\frac{2\pi}{\lambda}\{x\cos(\theta)+y\sin(\theta)\right]e^{\frac{-(x^2+y^2)}{2\sigma^2}}$$

$$IGF(x,y) = \sin\left[\frac{2\pi}{\lambda}\{x\cos(\theta)+y\sin(\theta)\right]e^{\frac{-(x^2+y^2)}{2\sigma^2}}$$

The shapes of these two functions are shown in figure 10. From the figure, the general usage of RGF is appropriate to detect object height and width, while IGF is to detect object edges [26].



a. 3-D plot of RGF b. 3-D plot of IGF

c. 2-D plot of RGF d. 2-D plot of IGF

Figure 10: Real and imaginary parts of -45˚ Gabor filter.

Since we are interested in Gabor kernels in that they can indicate some specific directions of images, we utilize them only as a directional estimator. Therefore it is expected that correlating with any image of the same direction will result in high correlation value [26]. The image shown in figure 1.a is rotated by 45˚ to investigate the applicability of Gabor filters in estimating the angle of rotation between two images. Correlating the two images with Gabor filters with different orientation from 0˚ to 180˚ in coarse steps of 10˚ resulted in highest correlation values for the first image in the range from 14˚ to 33˚ and for the second image in the range from 47˚ to 73˚. So we repeated the correlation with fine steps of 2˚in that range. The orientation of the first image was shown to be at 19˚ and that of the second was at 65˚ which means that the angle of rotation between the two images is 46˚. The estimated angle of rotation is very near to the actual rotation.

**Registration by Hit or Miss Transformation**

Hit or Miss Transformation is a morphological operation used to estimate the motion parameters between two misregistered image frames. This transformation tries to find a local pattern of pixels. The size of the pattern is determined by the size of the structuring element. The Hit or Miss transformation is given by :

$$A \otimes B = (A \Theta X) \cap [A^c \Theta (W - X)]$$

Where *A* is the image, *X* is the pattern to be searched for, *W* is a local background window taken around *X* and *B* is the structuring element which is given by: *B= (B1, B2)* where $B_1$ is the set formed by from elements of *B* associated with the pattern and $B_2$ is the set of elements of *B* associated with the corresponding background, i.e. $B_1 = X$ and $B_2 = (W-X)$. Hit or miss transformation is a template matching that finds a collection of pixels with the shape properties described by two structuring elements ($B_1$ the hit structuring element and $B_2$ the miss structuring element). Thus the set $A \otimes B$ will contain all the points at which, simultaneously, $B_1$ will find a match "hit" in *A* and $B_2$ will find a match in $A^c$ [9]. But the draw back of this method is that the shape of the structuring element must be found as it is in the search image to find a "hit", also it is applied on binary images only and hence the images are to be previously binerized . Result of finding the center of the structuring element in a two similar but shifted images is shown in figure 11.

**Registration by Geometrical (Rubber Sheet) Transformation**

If there is an image *f* with pixel coordinates *(x, y)* undergoes geometric distortion to produce an image g with coordinates *(x', y')*. This transformation may be expressed as:

$$x' = r(x,y)$$

And

$$y' = s(x, y)$$

Where r(x,y) and s(x,y) are the spatial transformations that produced the geometrically distorted image g(x',y'). The method used most frequently to overcome this difficulty is to formulate the spatial relocation of pixels by the use of tiepoints, which are a subset of pixels whose location in the input (distorted) and output (corrected) images is known precisely [9].

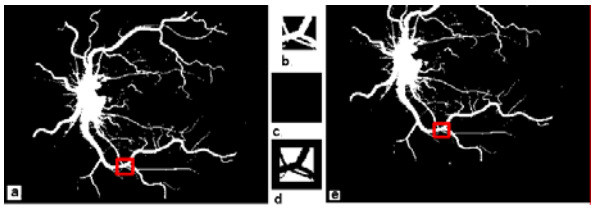



Figure 11: a. The original image, b. the selected pattern X, c. the background window W, d (W-X) and e. the position of the pattern in the shifted image.

Suppose that the geometric distortion process within the quadrilateral regions is modeled by a pair of bilinear equations so that:

$$r(x, y) = c_1 x + c_2 y + c_3 xy + c_4$$

And

$$s(x, y) = c_5 x + c_6 y + c_7 xy + c_8$$

Then, from the previous equations:

$$x' = c_1 x + c_2 y + c_3 xy + c_4$$

And

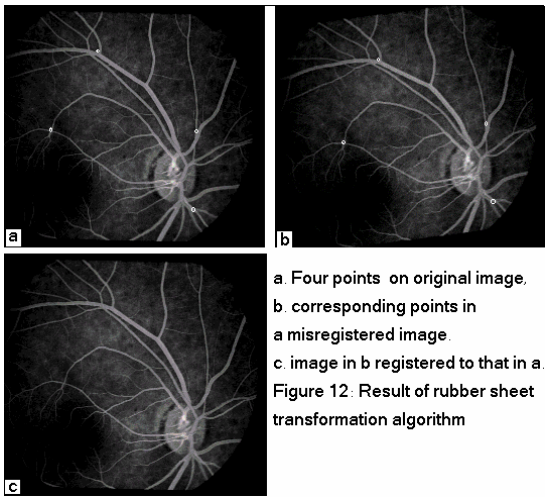
$$y' = c_5 x + c_6 y + c_7 xy + c_8$$

Since there are a total of eight known tiepoints, these equations can be solved for the eight coefficients $c_i$, $i = 1, 2 \ldots 8$. The coefficients constitute the geometric distortion model used to transform all pixels within the quadrilateral region defined by the tiepoints used to obtain the coefficients. In general, enough tiepoints are needed to generate a set of quadrilaterals that cover the entire image, with each quadrilateral having its own set of coefficients [27].

Once we have the coefficients, the procedure used to generate the corrected image is not difficult. The new positions of each pixel in the image can be estimated using *r(x,y)* and *s(x,y)*. Figure 10 shows results of rubber sheet transformation.

## 3. Discussion and Conclusion

In this work, a survey of retinal image segmentation and registration methods were applied on retinal images to be an assisting tool in the diagnosis and treatment of retinal disorders. At first, we applied different segmentation algorithms on the reference images to extract the sensitive objects including blood vessel tree, optic disc and macula and the region between them. Then, a binary image is composed which contains the sensitive objects as white objects in a dark background. Secondly, we applied many registration algorithms for the purpose of detection of motion parameters in retinal images considering the less time and accurate results**.** In this section we present a brief comparison among the proposed segmentation and registration algorithms in retinal images as shown in table.1 and table.2 respectively.




a. Four points on original image,
b. corresponding points in
a misregistered image.
c. image in b registered to that in a.
Figure 12: Result of rubber sheet transformation algorithm

The algorithm recommended to segment the contour of the blood vessel tree is a *deformable models* or snakes with a grid of seed contours and excluding the internal energy terms. The benefit of this algorithm is the higher accuracy with no user interaction. The proposed segmentation technique performs fast and accurate segmentation on retinal images . Its results are better than the other algorithms in that, it gives clear background and good appearance of small blood vessels the non user interaction reduces the execution time. But, it also has a disadvantage as it extracts the large blood vessels as splitted vessels; this problem is fixed by image dilation that will remove any splitting. It also extracts the small blood vessels and only the contour of wide one, a core extracting algorithm is used for this task.

The optic disc and the macula are segmented using the previously mentioned methods. The Hough transform, which is well known as a robust technique to segment circular objects is used to segment the optic disc. The macula is the most homogeneous area near the optic disc. The region growing algorithm is the most suitable methods to extract such a region. The dimensions to which this region will be extracted are optional to the physician because it depends upon the degree of the disease. According to table.2, the algorithm many of the algorithms can be reliable to be used in retinal image registration since we can assume that the translation is the only existing motion parameter. Rotation is minimal if present and can be compensated for. The scaling between images can be controlled. The hit or miss transformation is not recommended in any of the image registration purposes since it requires a very high similarity between the images regarding the shape and gray level relationships.

## 4. Acknowledgements

The authors wish to thank the team of National Institute of Laser Science for providing us by the retinal image. They also thank an anonymous reviewer for highlighting several works of relevance to this study.

| Algorithm | Accuracy | Edge continuity | Noise in background | Execution time in seconds |
|---|---|---|---|---|
| **Sobel Operator** | high | good | accepted | 1.582000 |
| **Robert Operator** | low | bad | accepted | 1.0509999 |
| **Prewitt Operator** | low | bad | accepted | 1.1120000 |
| **Canny Operator** | Very high | Very good | Fairly accepted | 7.099999 |
| **Decision Based Directional Edge Detector** | high | Very good | Not accepted | 1.3920000 |
| **Deformable models** | Very high | Very good | Highly accepted | 1.8599000 |
| **Morphological Gradient** | high | good | accepted | 0.8741000 |
| **Morphological Reconstruction** | high | Very good | Highly accepted | 2.1280000 |
| **Watersheds** | high | Fairly good | accepted | 15.9466000 |

**(Table.1),** a brief comparison among the proposed algorithms for retinal image segmentation.

| Algorithm | Accuracy | Atomicity | Execution time in seconds |
|---|---|---|---|
| **Similarity detection** | high | low | 0.035201000 |
| **Gabor Filter** | high | Fully automated | 0.90010001 |
| **Hit or Miss Transformation** | Very low | low | 0.01353400000 |
| **Rubber Sheet Transformation** | high | low | 0.024671000 |

**(Table. 2),** a brief comparison among the proposed algorithms for the purpose of detection of motion parameters on retinal images.

## 5. References

[1] B. JL. “Photo physical Processes in Recent Medical Laser Developments”. Lasers Med Sci., vol. 1, pp.47-66. 1986.

[2] B. GM. “Lasers in Medicine and Surgery”, JAMA, vol. 256, pp. 900-907, 1986.

[3] N. M. Bressler, S.B. Bressler, and E.S. Gragoudas, “Clincal characteristics of choroidal neovascular membranes,” *Arch. Ophthalmol*., vol. 105, pp. 209-213, 1987.

[4] P. N. Monahan, K. A. Gitter, and G. Cohen, “Evaluation of Persistence of Subretinal Neovascular Membranes Using Digitized Angiographic Analysis,” *Retina-J. Retinal, Vitreous Diseases*, vol. 13, pp. 196-201, 1993.

[5] S. Fine, “Observations Following Laser Treatment for Choroidal Neovascularization,” *Arch. Ophthalm.,* vol. 106, pp. 1524-1525, 1988.

[6] T. M. Clark, W. R. Freeman, and M. H. Goldbaum, "Digital Overlay of Flourescein Angiograms and Fundus Images for Treatment of Subretinal Neovascularization," *J. Retinal, Vitreous Diseases*, vol. 2, pp. 118-126, 1992.

[7] B. Kochner, D. Schuhmann, M. Michaelis, G. Mann, K.-H. Englmeier, "Course Tracking and Contour Extraction of Retinal Vessels from Color Fundus Photographs: Most Efficient Use of Steerable Filters for Model Based Image Analysis," *SPIE Conf. Image Processing*, no. 3338, pp. 755-761, 1998.

[8] W. T. Freeman, E.H. Adelson, "The Design and Use of Steerable Filters," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 13, no. 9, pp.891-906, 1991.

[9] Rafael C. Gonzalez, Richard E. Woods, "Digital image processing", second edition,Prentice Hall, pp 567-585, 2002.

[10] Bill Green, "Canny Edge Detection Tutorial", http://www.pages.drexel.edu/~weg22/can_tut.html, 2002.

[11] E. Aniram, H. Aydinoglu, and I. C. Goknar, "Decision Based Directional Edge Detector," *Signal Proc.*, vol. 35, pp. 149-156, 1993.

[12] T. McInerney and D. Terzopoulos, "Deformable Models in Medical Image Analysis: a survey," *Med Image Analysis*, vo. 1, no. 2, pp. 91-108, 1996.

[13] T. McInerney and D. Terzopoulos, "Topologically Adaptable Snakes," *Computer Vision (ICCV),* pp. 840-840, June 1995.

[14] J. Schnabel, "Active Contours, Snakes, or Deformable Curves," *Ph.D. Dissertation*, Universite Paris—Sud, 1997.

[15] J. Williams, M. Shah, "A Fast Algorithm for Active Contours and Curvature Estimation," CVGIP: IMAGE UNDERSTANDING, vol. 55, no. 1, pp. 14-26, 1992.

[16] M. Kass, D Terzopoulos, "Snakes: Active Contour Model," in Proceeding of First International Conference on Computer Vision, London, pp. 259-269, 1987.

[17] F. Meyer and S. Beucher, "Morphological Segmentation," *J. Visual Comm. Imag. Representation*, vol. 1, no. 1, pp. 21-46, 1990.

[18] J. B.T.M. Roerdink and Arnold Meijster, Arnold Meijster, " The Watershed Transform: Definitions, Algorithms and Parallelization Strategies", Institute for Mathematics and Computing Science University of Groningen. P.O Box 800, 9700 AV Groningen, Then Netherlands (187-228), 2001.

[19] S. Chaudhuri, C. Chatterjee, Norman Katz, Mark Nelson, And Michael Goldbaum, "Detection of Blood Vessels in Retinal Images Using Two-Dimensional Matched Filters," *IEEE Trans. Med. Imaging* , vol. 8, no. 3, pp.263-269, 1989.

[20] N. H. Solouma, A. B. Youssef, Y. A. Badr, Y. M. Kadah, "Robust Computer-Assisted Laser Treatment Using Real-Time Retinal Tracking," Proc. IEEE Engineering in Medicine and Biology Conference, October, 2001.

[21] A. Pinz, S. Bernogger, P. Datlinger, and A. Kruger, " Mapping the Human Retina," *IEEE Trans. Med. Imaging*, vol. 17, no. 4, pp. 606-619, 1998.

[22] N. H. Solouma, A. B. Youssef, Y. A. Badr, Y. M. Kadah, "Real-Time Retinal Tracking for Laser Treatment Planning and Administration," Proc. SPIE Med. Imag., San Diego, Feb., 2001.

[23] N. H. Solouma, Development of an Automated Laser System for Retinal Tracking", PhD thesis, Systems and Biomedical Department, Cairo University, 2001.

[24] B. Peli, R.A. Augliere and G.T. Timberlake, "Feature-Based Registration for Retinal Images," *IEEE Trans. Med. Imaging*, vol. MI-6, no. 3, pp.272-278, 1987.

[25] D. I. Barnea and H.F. Silverman, "A Class of Algorithms for Fast Digital Image Registration," *IEEE Trans. Computers*, vol. C-21, no.2, pp.179-186, 1972.

[26] Hiuk Jae Shin and Byeungwoo Jeon, "Rotation, scaling and Translation Robust Image watermarking Using Gabor Kernels", Proceeding of SPIE Vol. 4675, 2002.

[27] Q. Zheng, and R. Chellappa, "A Computational Vision Approach to Image Registration," *IEEE Trans. Image Processing*, vol. 2, no. 3, pp. 311-326, 1993.