Chapter 3

Registration of retinographies

The method developed in the previous chapter made use of the creaseness of CT and MR images as anatomical landmarks for alignment. It seemed natural to investigate if the same algorithm would work also for images where creases appear in a similar fashion. One such type of images are the retinographies, where the vascular tree appears as a network of creases. In this chapter we demonstrate the validity of the algorithm for a large database of images, including multi-modality and temporal series, and for unrestricted conditions of noise and contrast. Most of the work, however, will focus on a particular type of retinography, namely, *SLO* video sequences.

3.1 Introduction

Non-invasive techniques make use of the reflectance properties of the pupil of the eye to examine the retina, which is the portion of the eye visible to an external observer. The light is sent through the lens in the iris to the fundus of the retina, where it is reflected back and composes the image (figure 3.1). The light suffers multiple transformations as it travels through the cornea, the lens and the layers of the fundus, and therefore suitable refined techniques had to be developed. The origins of the first instruments, called ophthalmoscopes, reside in Charles Babbage (1847) and Helmholtz (1851). More complex, electronic ophthalmoscopes, were first developed a century later by Harold Ridley. The last advances were achieved by R. Webb [111, 110] with the employment of laser devices to scan the retina (figure 3.3 shows two examinations).

Many different acquisition protocols exist, each highlighting a particular feature of the eye. For a number of diseases, e.g. diabetic retinopathy, it is convenient to track the evolution through a period of time, while others may benefit from fusing information from different modalities.

Today ophthalmoscope images have a wide range of applications in normal clinical practise. Standard uses include: retinal densitometry, macular pigment measurements and infrared or red imaging of subretinal structures [12]. Many specialised applica-

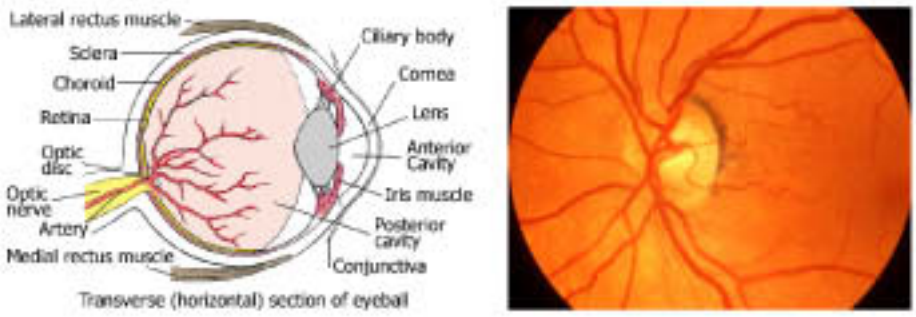

Figure 3.1: Left: schematic section of an eyeball. Right: retinography showing vascular tree.

tion are continuously being developed; for instance, the 3-D reconstruction from the sub-retinal tissues to the optical disc [67], automatic mapping of the retina [75], or tomographic reconstruction of the retina [105].

The following devices are used to acquire the images:

- **retinographies** are taken using the ophthalmoscope under natural light, and often they have been applied a green filter (**green images**) to discard the dominant red component. In a sense, the modalities in the following paragraphs are also retinographies as they depict the retina. However, from now on we will narrow its meaning to the later definition.

- **Scanning Laser Ophthalmoscope (SLO)**: a low power laser is directed through the retina , and the reflected light is descanned and detected by a photodiode. Since the laser scans the retina in a raster fashion, the resulting reflectance readings must be composed with the aid of a computer. Compared to the previous camera pictures, this device provides better resolution, controllable excitation intensity and variable depth of focus. *SLO* cameras are classified according to the laser emission: infrared, He-Ne, Argon-blue and many others.

- **Confocal Scanning Laser Ophthalmoscope(CSLO)** is a refinement on the *SLO* :employs an additional aperture in a plane conjugate to the sample by means of a small pinhole placed in front of the photodiode, thus improving the control of the origin of light. Since the depth of the sample plane can be modified at will, it is considered a tomographic modality [105].

These acquisition devices are able to extract different types of information:

- **anatomical information** regarding the position and pattern of the fundus structures.

- **Fluorescein angiographies** sense the fluorescence emitted from the vessels of the retina after the injection of a contrast dye. Resulting images are acquired by an normal ophthalmoscope or by *SLO* .

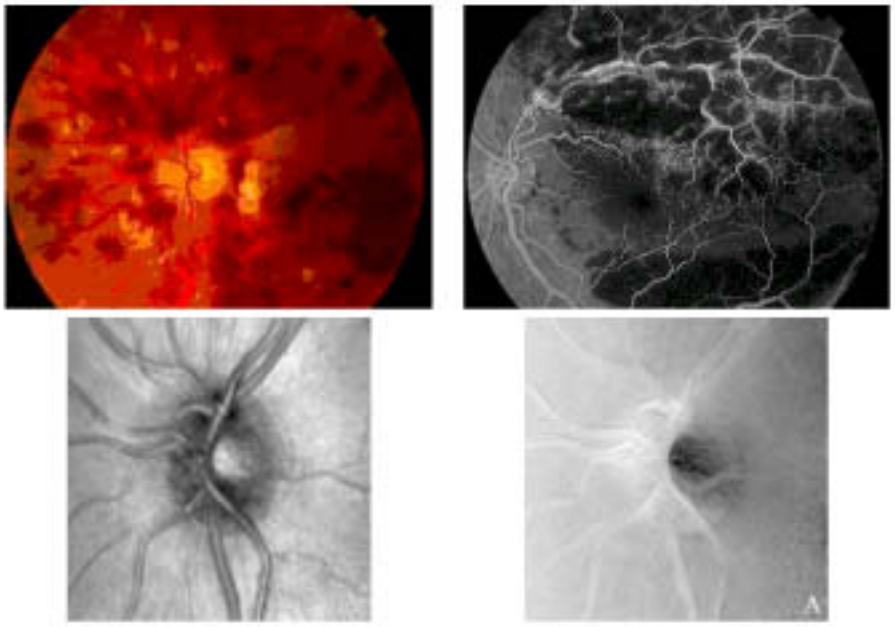

Figure 3.2: From left to right, top to bottom: a) retinography without the green filter, b) angiography of the same region, c) SLO, d) confocal SLO of the same region as c).

- **Static scotometry (SC)** takes two simultaneous pictures: an anatomical of the fundus and a functional measure of the response of the macula.

In addition, all these modalities can be combined to produce simultaneous **stereo pairs** or sequential **temporal series** of the same eye, in order to study the optic disk for signs of an illness progression. Martínez-Costa [65], for instance, seeks high increments in grey level at the foveal centre in a series of temporal images to detect macular leakage due to retinal vein obstructions.

When more than one image is used for diagnosis, it is necessary to align corresponding features in order to compare them. The movements of the eye are involuntary and are cause by the cardiac pulse and microsaccades [3] of the iris muscles. Very small movements have a high impact in the imaging process due to the small field of view of the camera. Therefore, the problem of registering the ophthalmoscopic images arises.

Many papers devoted to the ophthalmologic field are based on comparisons between two or more images, and therefore they must deal with the registration problem prior to any further step. However, in some of them the registration step is described poorly, sometimes with a short number of test images, while others seem to apply only for a particularly well suited modality. In section 3.2 we make a full report of the state of the art.

The vascular tree is a common landmark for many of the methods. Indeed, it

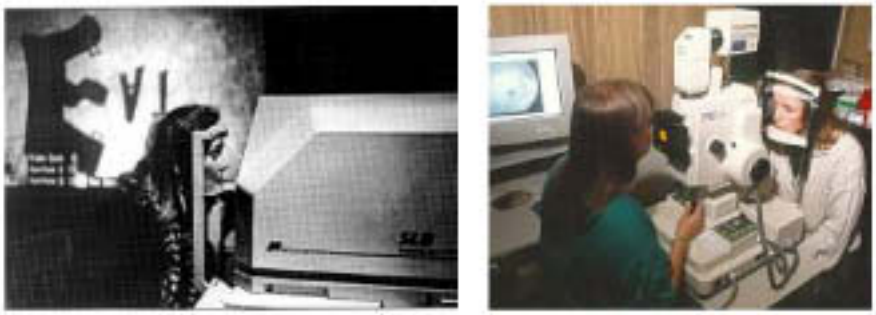

Figure 3.3: Ophthalmologic devices used to acquire the *SLO* images.

seems to be the only structure automatically detectable. These methods are then classified as landmark-based. Others work with the intensity values of the images, thus being classified as voxel-based methods. However, for these methods to be of any use, they must count for the high variability of the images, and the large variation of the translation parameters.

Similarly to what happened with the head images, our algorithms evolved as new images became available to us. We did not own any ophthalmologic equipment at our site, so we had to ask other research groups for them. This is the list, with acknowledgement, of the source of our images: F. Zana at the Centre de Morphologie Mathématique, Ecole des Mines de Paris, E. de Ves at Instituto de Robótica, Univ. de València and J. Pinós at Hospital General Universitario de Valencia. Also, C. Barry and N. Ritter at the McCusker Glaucoma Foundation let us use their large image database.

With the images from those groups, we presented first results in a paper at [48]. The algorithm worked basically for pairs of images, either mono or multimodality, some of them comprising one year between both observations. Shortly after the paper was finished, a group from Universidade da Coruña and Universidade da Santiago, in Galicia, headed by Manuel F. González Penedo and M.J. Carreira, offered us to collaborate with a clinical-related project they had commenced. The purpose of the project was the registration of large series of *SLO* images. Since that was a natural extension of the work already done, we agreed and a member of the group, Castor Pérez Mariño, came to our site. This happened to be a fruitful collaboration, and a good chance to test our basic assumptions for real medical requirements. We submitted the final results as a paper in [45] and [64].

3.2 Short review of registration methods for ophthalmologic images

Following we present a short review of some papers describing the algorithm to register images. It is interesting to see their evolution as computers gained CPU speed, and the requirements from the medical counterpart grew higher.

Year	Author	Ref	Modality	Number of images	Time	Computer
94	Noack	[70]	SLO	1 seq.	< 1 s	Sun Sparc 1
98	Costa	[65]	Angio.	100 pairs	2 min	HP-735
98	Pinz	[75]	SLO	seq. of 100	$5 - 6$ min †	SGI O2+KBVision
98	Wade	[106]	cSLO	seq. of 256	1 sec	P 133MHz + Matrox
99	Zana	[118]	Angio, green	29 pairs	5-7 min	P 150MHz
99	Ritter	[81]	Retino.	100 pairs	3-38 sec ‡	P 200Mhz

Table 3.1: Comparison of the methods found in literature for registration of fundus images. Time must be weighted according to three factors: the machine employed, the assumptions on the nature of the images and the accuracy required.
† full map
‡ depending on the accuracy demanded

Noack and Sutton [70] made first attempts to register large sequences of *SLO* images. They recognised the problem of the large grey level difference between frames a few seconds apart, and proposed to compensate it with the extraction of binary thresholded gradient images obtained using Prewitt kernels. A linear correlation measured the alignment between binary reference templates and the images. Their design emphasised real applicability with a fast low-level software design, but they did not provide much detail about the sensitivity of the algorithm to the election of the templates, or to varying acquisition conditions. No actual clinical results were described either.

Wade and Fitzke [106] worked in a similar fashion, but with the aid of a modern dedicated signal processor (Matrox board). A window of the reference image is matched against the other images taking the cross-correlation as the alignment measure. They make full use of the hardware capabilities. For instance, a Gaussian filter to eliminate noise, or the full refinements of the in-box algorithmic search. The most relevant success of their paper is to demonstrate the capabilities of a dedicated image processing hardware for the registration of ophthalmologic images, but they do not report in detail its success for sequences other than the original, neither their analysis of the results is statistically significant.

Pinz and coworkers present a complete method including a medical application assessment. In early papers [76, 77] they describe a generic method for image registration which uses the so-called tokens, which are a symbolic representation based on structures extracted from the image and do not require exact correspondence to work. The algorithm defines an exponentially decreasing distance function and a hierarchical structure is built with the dilate morphological operator, which proves to be robust and suitable for a broad range of images.

The last paper from the same group, [75], remarkably widens its scope to propose a method for a full map of the human retina. They register the images by means of the previous method, but carry out the automatic detection of the vessels by following the stages: a) extracting the edge elements along the boundary, b) grouping and search for cross sections and c) combinations of cross sections to tubes. The rest of the paper

is devoted to the detection of several anatomical and pathological features: optic disc, fovea and foveola, skotoma and subretinal leakage. The complex and heuristic detection algorithm seems successful at taking into account variations in the source images, and results are plausibly suitable for the required medical purposes.

Two other methods employ a similar approach consisting on robust matching of fork points detected at the vessel tree. Domingo, Martínez–Costa and coworkers [9, 65] apply Förstner's operator to detect them as the corners of the vessels modelled as lines. They adapt the operator to the various saturation levels of the image by means of a local histogram equalisation. In order to guide the corresponding search, each detected landmark point is associated to a value invariant to rotation and translation; they choose a measure related to a cross-correlation function of a local neighbourhood. An stochastic search associates pairs of points belonging to each image, and a further iterative optimisation refines the result.

A similar method is described by Zana and Klein in [118]. The general scheme is similar, with different algorithmic choices: first, vessels are detected after a long sequence of operations, including morphological and Laplacian transforms. Then, bifurcation points are extracted applying pattern matching to a set of possible configurations. For each point, as the previous method does, an invariant measure is computed, in this case related to the local geometry of the vessels. Finally, a Bayesian Hough transformation selects the best candidate for a further iterative transformation cycle, where bifurcation points are matched accurately until the desired convergence is achieved.

In our opinion, the methods from both [65, 118] seem to be highly dependent on the results of the extraction step. Whenever the number of detected points is too low, or the corresponding sets too separate, their algorithm is prone to fail.

For completeness we include some related papers which appeared in ophthalmologic journals. They make some elemental use of registration algorithms, but emphasise mostly its medical applications: in [67], Masters reconstructs the retina in 3–D by acquiring consecutive tomographic images and registering them. Elsner [12] makes a review of applications to reflectometry. Vieira, in [105], examines the algorithm to construct *cSLO* images, and demonstrates the advantage of use a double Gaussian to model the double reflection phenomenon of the two structures in the retina.

Tolias and Panas [99] concentrate on the proper detection of ocular fundus vessels. Their algorithm is relevant as it seems robust to noise, accurate, parameter-free and makes no explicit model for the morphology of the vessels, thus being suitable as the first detection process for further registration purposes. In short, they apply fuzzy clustering to grey level profiles to determine vessel and non-vessel regions, with heuristics to determine forks, crosses and terminations.

The paper by Deguchi et al [7] was very important to our research because it provided an optics-based explanation for some geometrical mismatches we observed in our own registered images. It models the image acquisition as formed by the fundus sphere, the eye lens and the camera coordinates (see section 3.5). Any translation of the camera relative to the fundus changes the point of view in a non-linear fashion, due to the particularities of the optics. At their paper, they aim at the reconstruction and 3D display of the fundus from a set of partial images, taken at several view angles. Although in our experiments the angle is meant to be kept stable, in reality small

variations exists, which count for the distortions observed.

Another well known approach in medical image registration are the mutual information methods. These methods, applied firstly on the registration of head images in 1995 [112, 54], do not extract corresponding features but make use of all the information available. They measure the statistical dependence or information redundancy between the image intensities of voxels corresponding at both images.

Ritter has report success in a method based on this approach in [81] for fluorescein angiographies and *SLO* pairs of images. In her paper, she provides with an excellent review of existing methodologies, and she stresses the importance of a good database of images to test the algorithms under realistic assumptions. This is a proper remark, for some papers take into account only translations, or accept as good a ten-pixels misregistration result. Most of the paper is dedicated on the problem of local maxima. They occur when the function to measure the alignment gets stuck into a value which is the highest for a local neighbourhood but not for the whole parameter space. Because the optimisation is a function of 3 or 5 parameters (two translation plus one rotation angle, plus sometimes two scaling factors) an exhaustive search is not feasible. Thus, she uses a scheme based on pyramidal sampling and the simulated annealing search algorithm. She reports success on her experiments on an about 100 images database.

Our major criticism to her work regards to the proposed search method. Although successful, the simulated annealing algorithm is highly dependent on a large number of parameters, 6 global plus 7 for each layer of the pyramidal search. These parameters are hidden deep into the code, making it impossible for a user to tune to any change in the image constraints. Also, our experience with her database compared to other *SLO* images suggests that hers is a small subset of the full range of expected images, for whom her method still has to be assessed.

In sum, literature presents a plethora of methods for ophthalmologic images registration. They can be broadly divided into three sets: those using landmarks, those using cross-correlation of the images normalised in some fashion, and those based on mutual information. Most of them are restricted to one modality, sometimes even to images acquired with a single device, which makes their results highly dependent on the particular setting. Others present a good methodology and prove their applicability for a wider scope.

When comparing visually different types of ophthalmologies, sequences of *SLO* images probably would stand for the most difficult choice. Because of their large volume, up to thousands of frames, they include images of very different quality, contrast, and noise, and therefore general methods working with standard modalities are, in our opinion, not likely to succeed.

We propose a method which deals favourably with the characteristics of the *SLO* sequences. This chapter is organised as follows: first, we describe the registration algorithm, divided in two parts: vessel extraction and alignment scheme. Next, we investigate the robustness of the algorithm with regard to several choices of parameters. Results motivate the choice of one particular combination as the best suited. Then, we evaluate the accuracy and, after explaining a small unobserved misalignment particular to this modality, we propose a refinement which works in a local fashion. Finally, we compare one of our results with that of manually obtained by an expert.