# MULTIMODAL RETINAL IMAGING: NEW STRATEGIES FOR THE DETECTION OF GLAUCOMA

*Paul L. Rosin and David Marshall*

Department of Computer Science
Cardiff University
UK
{Paul.Rosin,dave}@cs.cf.ac.uk

*James E. Morgan*

Department of Optometry and Vision Sciences
Cardiff University
UK
morganje3@cardiff.ac.uk

## ABSTRACT

Glaucoma is a serious worldwide disease whose treatment can be improved by early detection. As part of a new clinical approach this paper introduces some preliminary studies in the computerised detection of the disease. In particular we consider the problems of registering 3D laser data with a digital image. We introduce a new method based on windowed mutual information and show that it performs better than the standard mutual information technique.

## 1. INTRODUCTION

Glaucoma is the second commonest cause of blindness in the West and the commonest cause of blindness world wide [1]. Early detection of glaucoma forms the basis of effective management. If the disease is detected in its early stages, damage can be minimised and the long term prognosis for the patient improved.

Historically, clinical evaluation of patients suspected of having glaucoma has relied on the detection of visual field damage by automated perimetry and the detection of optic nerve damage by subjective retinal examination. Visual field analysis can detect visual field loss with high sensitivity and specificity in patients with moderate to advanced optic nerve damage, but it is less useful for the detection of early damage. Subjective examination of the optic nerve and retinal nerve fibre layer can detect this early damage but is limited because of the need for expert evaluation of the optic nerve appearance and the low level of agreement between observers for subjective evaluation of the optic disc.

These limitations have driven the development of digital imaging technologies such as digital fundus photography and scanning laser ophthalmoscopy (SLO) [2, 3] that can provide objective evidence of early optic nerve damage. Both these techniques provide high quality images of the optic nerve and, in the case of scanning laser ophthalmoscopy, detailed information on the surface topography of the retina. Statistical analysis of these images has been shown to detect early glaucomatous damage with high sensitivity and specificity – in many cases before the development of significant visual field damage [2].

In spite of these successes, up to 10% continue to be misclassified, even with optimised statistical algorithms. Since glaucoma affects up to 2% of the population (and up to 4% of those over 75 year of age) these errors can generate large numbers of patients with incorrect or late diagnoses. It is important that these techniques are developed to achieve further improvements in diagnostic precision. Preliminary work in our group has shown that when the two diagnostic techniques are used in combination, the sensitivity and specificity for the detection of early damage is further improved to achieve useful clinical reductions in the number of inaccurate diagnoses [4].

Since optic disc photographs and scanning laser tomographic analysis (see figure 1) provide complimentary data on the appearance of the optic nerve head, we have developed techniques for the combination of these images that should further improve the diagnostic precision of optic disc analysis. The use of these modalities is new. Current techniques for retinal image registration use more similar modalities, e.g. angiograms, that involve 2D image data, or registering the same modality over time. The principle of this method is that the images are combined to allow better demarcation of the boundary of the optic disc for topographic analysis of the retinal surface contour and to enable better correlation of topographic and visible optic nerve head damage. A two stage process is involved.

1. The alignment and registration of optic disc photographic and scanning laser tomographic images.

SLO and photographic images provide different views of the retina and optic nerve. Images obtained with the scanning laser ophthalmoscope indicate the degree of reflected light from the region of interest [5]. The extent to which light is reflected is determined by the surface topography of the structure being imaged. Areas of high slope change such as the edge of the optic nerve show up as areas of low

reflectivity. As such the images differ from those obtained photographically which are not as dependent on surface topography [6]. This paper focuses only on this stage.

2. Analysis of merged photographic and topographic images.

In the interpretation of the scanning laser ophthalmoscopic images this achieved by using the photographic information to demarcate the border of the optic disc. Many of the clinically useful parameters used to diagnose optic nerve head damage are determined on the basis of the location of the neuroretinal rim which is often difficult to define on the basis of tomographic images but can easily be discerned from optic disc photographs.

## 2. REGISTRATION

A tremendous amount of work on image registration has been carried out, and comprehensive surveys are given by [7, 8]. More specifically, many papers have been published on the registration of retinal images [9, 10, 11]. The two main approaches to registration can be categorised as either feature based or area based. The former first detects features such as edges, lines, junctions, etc., and these are then matched between images. Naturally this approach is dependent on good feature extraction, which is often difficult and unreliable, especially in the biological domain. To circumvent such problems the latter approach operates more directly on the image without extracting features. The two most common examples are probably correlation and Fourier-based methods. For the registration of multi-modal data sets mutual information (MI) based approaches have become very popular [12] since they can cope with substantial variations in appearance across modalities. Rooted in information theory, MI quantifies the similarity between two data sets based on the degree that one set predicts the other. It is calculated as

$$\sum_{i=1}^{M}\sum_{j=1}^{N} P(a_i,b_j)\log_2\left[\frac{P(a_i,b_j)}{P(a_i)P(b_j)}\right]$$

where there are $\{a_i; i = 1 \ldots N\}$ and $\{b_j; j = 1 \ldots M\}$ distinct pixel values in images $A$ and $B$ respectively; $P(a_i)$ and $P(b_j)$ are the proportions of pixel value $a_i$ in image $A$, and $b_j$ in image $B$, and $P(a_i, b_j)$ is the joint probability of their occurrence in the two images.

However, a limitation of the technique is that it is based on nominal values. Since the pixel values are continuous (although quantised) MI ignores much of the available information. For instance, even if there are many mappings between the images of a pair of values such as $5 \rightarrow 138$, if a similar mapping like $5 \rightarrow 137$ was infrequent it would register as a substantial change.

A second problem that we have found is that images containing large, homogeneous, noisy backgrounds can result in false matches caused by the background match response swamping the sparser features of interest.

In the following sections we describe some alternative algorithms to the basic MI approach that we have developed for our registration problem.

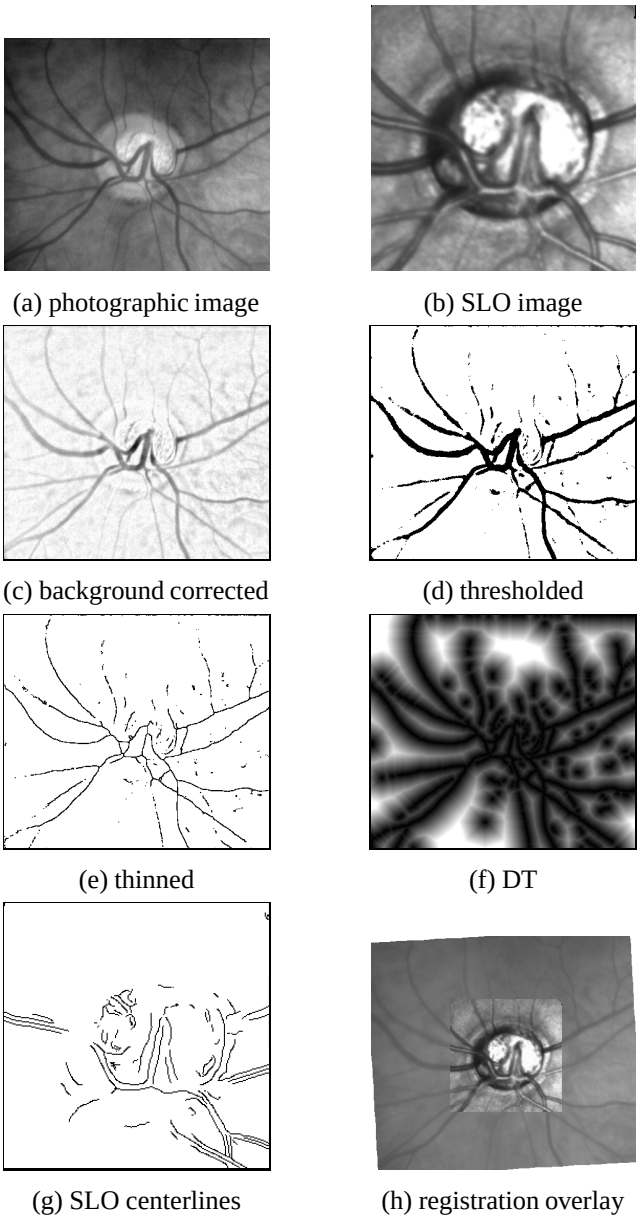
### 2.1. Feature Based Registration

The first approach is feature based. However, to limit the dependency on unreliable features such as junctions, simple thresholding is used as a basis for matching. First, the photographic image contains vignetting, and so to compensate for the decreasing intensity at the image borders a normalisation process is applied. A morphological closing operation is applied to eliminate the blood vessels, and the resultant is subtracted from the original image, as shown in figure 1c. To extract the blood vessels this is thresholded [13], and median filtered to remove some noise effects. After thinning the distance transform (using the chamfer algorithm) is run. Next, from the SLO image the blood vessel centerlines are extracted using a ridge detector [14]. Since vessels could appear either relatively dark or light against their background both ridges and troughs were extracted and combined (by a logical AND operation). Again, thresholding was applied, and this generates a binary feature map. The two images (figure 1f&g) are overlaid, and the DT values lying under the feature map are summed to provide a measure of registration error. Matching is performed by transforming the SLO image so as to minimise the error.

A variation on this scheme is to weight the DT error scores by the width of the blood vessel so that wide (and consequently important) vessels have more influence on the registration error. The width is computed by finding the distance transform of the inverted thresholded blood vessels and combining this with the thinned vessels to produce a gray-scale thinned image where values correspond to vessel widths (figure 2a). From these, the salience distance transform [15] generates the weighted DT (figure 2b).

### 2.2. Combining Features with MI

In an attempt to cope with the background swamping effect we have modified the basic MI to concentrate on salient areas in the image. Blood vessels are extracted by the procedure described above, expanded by applying a dilation operation, and are used as a mask to restrict computation of MI values. Precise segmentation is not important, and so errors in vessel extraction are expected and allowed. The aim is just to roughly identify the relevant features, match with them, and ignore most of the background.



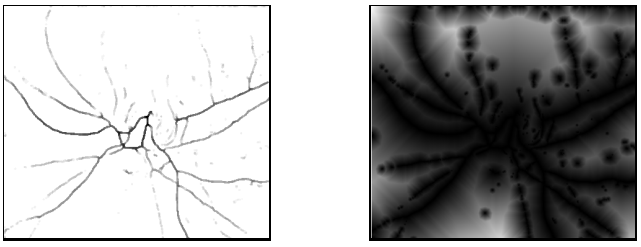
(a) photographic image (b) SLO image

(c) background corrected (d) thresholded

(e) thinned (f) DT

(g) SLO centerlines (h) registration overlay

**Fig. 1**. Example of feature based registration.



(a) width weighted skeleton (b) SDT

**Fig. 2**. Width weighted registration.

### 2.3. Variance Weighted MI

Although MI allows arbitrary mappings between images it seems reasonable to expect each intensity in one image to map onto a small range of intensities in the second image. This constraint is incorporated by weighting each term in the summation by the variances in the mappings

$$\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{P(a_i,b_j)}{\sigma_{a_i}+\sigma_{b_j}+1}\log_2\left[\frac{P(a_i,b_j)}{P(a_i)P(b_j)}\right]$$

where $\sigma_{a_i}$ is the variance of all the mapped values of $a_i$ in $B$, and $\sigma_{b_j}$ is the variance of all the mapped values of $b_j$ in $A$. Thus, the spread in values taken into account by the penalty term $\sigma_{a_i}+\sigma_{b_j}+1$.

### 2.4. Windowed MI

It can be seen that some areas of the image are relatively uninformative compared to others. For instance, some contain many blood vessels while others are restricted to background variations. Areas with distinct and discernible structures are expected to be more discriminating for matching.

Rather than attempt to perform segmentation (with the usual difficulties due to noise, poor contrast, etc.) the image was simply divided into a checkerboard of regular windows. It could then be expected that some of the windows would lie on background areas, others on areas mainly occupied by blood vessels, while the remainder would contain a small amount of image structure. In this way some element of spatial information can be introduced.

MI registration was applied to each window independently,[1] and the combined MI scores were maximised over the allowable transformation. The following combination schemes were tested: mean window score, median window score, and mean window score weighted by the window's intensity variance. The objective was to enable windows containing insufficient structure to allow for proper registration to be ignored, or at least downweighted.

## 3. RESULTS

Currently the images are registered using a similarity transformation based on the above methods. We use an exhaustive multi-scale search since we seek to determine the best image similarity measure. This avoids any bias in any heuristics present in more advanced search methods. Once this has been found future work will use faster search methods such as hill climbing or simulated annealing.

Twenty seven images were graded for alignment using a 4 point scale by a single experienced clinician studying

[1]Using windows as opposed to the complete image required the intensities to be quantised more coarsely to avoid the histograms becoming too sparse.

monochrome images: 0 none/poor, 1 moderate, 2 good, 3 excellent. The mean scores for the following algorithms are shown in table 1: blood vessel matching, width weighted vessel matching, MI, MI restricted to the blood vessel mask, variance weighted MI, MI calculated in 16 windows and combined using the mean, median, and variance weighted mean methods. It can be seen that the basic MI method performs very poorly. We found this surprising given its widespread use in the general literature. The explanation seems to be the substantial differences between the modalities, such that profiles across blood vessels are very different across the modalities, and some structures are not uniform within one or other modality. The modifications to the MI using masks or variance weighting did not significantly improve the results. Applying MI within windows produced much better results (at the good/excellent level), and was found to be the top ranking method. The feature based approach gives significantly worse means scores. Moreover, it is more prone to fail badly, and its standard deviation in score is 1.3 against the window MI whose standard deviation is 0.8.

## 4. CONCLUSIONS

We have considered the registration of two modalities which, from a clinical point of view, have great potential in diagnosing glaucoma. The first requirement in computer analysis is the registration of the two modalities. Due to their nature this is a difficult task. Conventional practice for the registration of retinal images does not work in this instance. We have developed alternative techniques which perform better. Although there is room for improvement (which is the subject of ongoing work) feedback from clinicians is very encouraging. Even though the research has focussed on retinal images we believe that then new windowed based MI approach is well suited to other difficult registration tasks.

| DT | SDT | MI | MI mask | MI var. | MI win. mean | MI win. med. | MI win. var. |
|---|---|---|---|---|---|---|---|
| 1.4 | 1.1 | 0.4 | 0.4 | 0.5 | 2.4 | 1.7 | 1.3 |

**Table 1**. Mean registration scores provided by clinician.

## 5. REFERENCES


[1] A. Sommer, "Doyne lecture. Glaucoma: facts and fancies," *Eye*, vol. 10, pp. 295–301, 1996.

[2] C.Y. Mardin, F.K. Horn, J.B. Jonas, and W.M. Budde, "Preperimetric glaucoma diagnosis by confocal scanning laser tomography of the optic disc," *Br. J. Ophthalmology*, vol. 83, pp. 299–304, 1999.

[3] A. Azuara-Blanco, G.L. Spaeth, J. Lanzl I.M. Nicholl, and Augsburger J.J., "Comparison between laser scanning tomography and computerised image analysis of the optic disc," *Br. J. Ophthalmology*, vol. 83, pp. 295–298, 1999.

[4] J.E. Morgan, N.J.L. Sheen, R. Goyal, and R.V. North, "Discrimination of glaucomatous optic neuropathy by segmental neuroretinal rim analysis," *in preparation*.

[5] R. Webb, G. Hughes, and F. Delori, "Confocal scanning laser ophthalmoscope," *Appl. Optics*, vol. 26, pp. 1492–1499, 1987.

[6] R. Webb and F. Delori, "How we see the retina," in *Laser Technology in Ophthalmology*, J. Marshall, Ed. Kugler and Ghedini, 1988.

[7] L.G. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325–376, 1992.

[8] J.B.A. Maintz and M.A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[9] G.K. Matsolpoulos, N.A. Mouravliansky, K.K. Delibasis, and K.S. Nikita, "Automatic retinal image registration scheme using global optimization techniques," *IEEE Trans. on Information Technology in Biomedicine*, vol. 3, no. 1, pp. 47–60, 1999.

[10] N. Ritter, R. Owens, J. Cooper, R.H. Eikelboom, and P.P. Van Saarloos, "Registration of stereo and temporal images of the retina," *MedImg*, vol. 18, no. 5, pp. 404–418, 1999.

[11] F. Zana and J.C. Klein, "A multimodal registration algorithm of eye fundus images using vessels detection and hough transform," *IEEE Trans. on Medical Imaging*, vol. 18, no. 5, pp. 419–428, 1999.

[12] W. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," 1996.

[13] W.H. Tsai, "Moment-preserving thresholding," *CVGIP*, vol. 29, pp. 377–393, 1985.

[14] C. Steger, "An unbiased detector of curvilinear structures," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 2, pp. 113–125, 1998.

[15] P.L. Rosin and G.A.W. West, "Salience distance transforms," *CVGIP: Graphical Models and Image Processing*, vol. 57, pp. 483–521, 1995.