MULTIMODAL RETINAL IMAGE REGISTRATION USING A FAST PRINCIPAL COMPONENT ANALYSIS HYBRID-BASED SIMILARITY MEASURE

Parminder Singh Reel[†] *Laurence S. Dooley*[†] *K.C.P Wong*[†] *Anko Börner*[⋆]

[†]Department of Communication and Systems, The Open University, Milton Keynes, United Kingdom
[⋆]Optical Sensor Systems, German Aerospace Center (DLR), Berlin, Germany
Email: [†]{*p.s.reel, l.s.dooley, k.c.p.wong*}*@open.ac.uk*, [⋆]*anko.boerner@dlr.de*

ABSTRACT

Multimodal *retinal images* (RI) are extensively used for analysing various eye diseases and conditions such as myopia and diabetic retinopathy. The incorporation of either two or more RI modalities provides complementary structure information in the presence of non-uniform illumination and low-contrast homogeneous regions. It also presents significant challenges for *retinal image registration* (RIR). This paper investigates how the *Expectation Maximization for Principal Component Analysis with Mutual Information* (EMPCA-MI) algorithm can effectively achieve multimodal RIR. This iterative hybrid-based similarity measure combines spatial features with mutual information to provide enhanced registration without recourse to either segmentation or feature extraction. Experimental results for clinical multimodal RI datasets comprising colour fundus and scanning laser ophthalmoscope images confirm EMPCA-MI is able to consistently afford superior numerical and qualitative registration performance compared with existing RIR techniques, such as the *bifurcation structures* method.

***Index Terms*—** Image registration, ophthalmological image processing, principal component analysis, mutual information, expectation-maximization algorithms.

1. INTRODUCTION

Image registration is a vital procedure in many computer vision and image processing applications [1], [2]. In medical imaging for example, images firstly need to be registered before image fusion can occur to facilitate disease diagnosis and treatment planning [3]. Multimodal *retinal images* (RI) like colour fundus and *scanning laser ophthalmoscope* (SLO) have specific clinical applications for identifying eye conditions and diseases like myopia, glaucoma and diabetic retinopathy [4], [5]. The modalities provide complementary information with the colour fundus images showing the boundary of the optic nerve head, while the near infrared SLO images reveal deep layer reflectivity and surface topography of the optic nerve head and retina [5].

Retinal image registration (RIR) seeks to align the vessel structures of the retina in these images, to assist in ophthalmology, especially in the tracking and analysis of the advancement of various eye diseases [6]. Multimodal RIR is especially challenging compared to single modality RIR because while providing complementary details on retina structure, the images can have different sizes and resolutions, exhibit different reflectivity levels in their non-uniform contrasts and have large homogeneous non-vascular regions. The quality of a particular image modality can be further compromised by degradations in various pathologies [4].

RIR can be categorised into *feature*, *intensity* and *hybrid* based techniques [6]. Feature-based approaches use extracted vessel structure and landmark points from the RI, while intensity techniques focus solely on pixel intensity information. Using features for RIR is analogous to manual registration, as key structures like optic disk [7] and vascular structural features [8], [9] are extracted from RI. Bifurcation point matching approaches such as dual-bootstrap iterative closest point [10] and its variant [11] accomplish registration by using vascular features to grow a bootstrap region. While these methods are computationally fast, their performance is very dependent on the segmentation quality and degree of overlap of the extracted features.

Vascular *bifurcation structures* (BS) is a recent method for multimodal RIR [12] which has been shown to perform satisfactorily for linear and affine transformation models. RIR-BS defines a master bifurcation point with up to three connected neighbours, with each structure vector having both a normalised branch angle and length, which estimates the transformation for image alignment.

In contrast, intensity-based techniques use a *similarity measure* (SM) such as cross correlation, phase correlation [13] or *mutual information* (MI) [14], [15] to align the RI by applying an optimization strategy to maximise the SM within the transformation search space. MI has been widely adopted in multimodal medical image registration of various parts of the human anatomy including the brain, lungs and eyes [1]. It establishes a statistical relationship between the intensity values of images, though the innately challenging characteristics of RI can lead to a rapid degradation in RIR performance when MI is applied directly [16].

Hybrid approaches [1] combine different aspects of the aforesaid feature and intensity methods, as for example, in using extracted vascular structures [17] together with spatial information in regional MI [18] and feature neighbourhood MI [19], [20]. These techniques use covariance matrices to reduce the data complexity [18] instead of high-dimensional histograms, though as the spatial information increases, so commensurately does the corresponding computational cost. EMPCA-MI [21], [22] is a recent hybrid-based algorithm which has been developed for efficiently registering brain computed

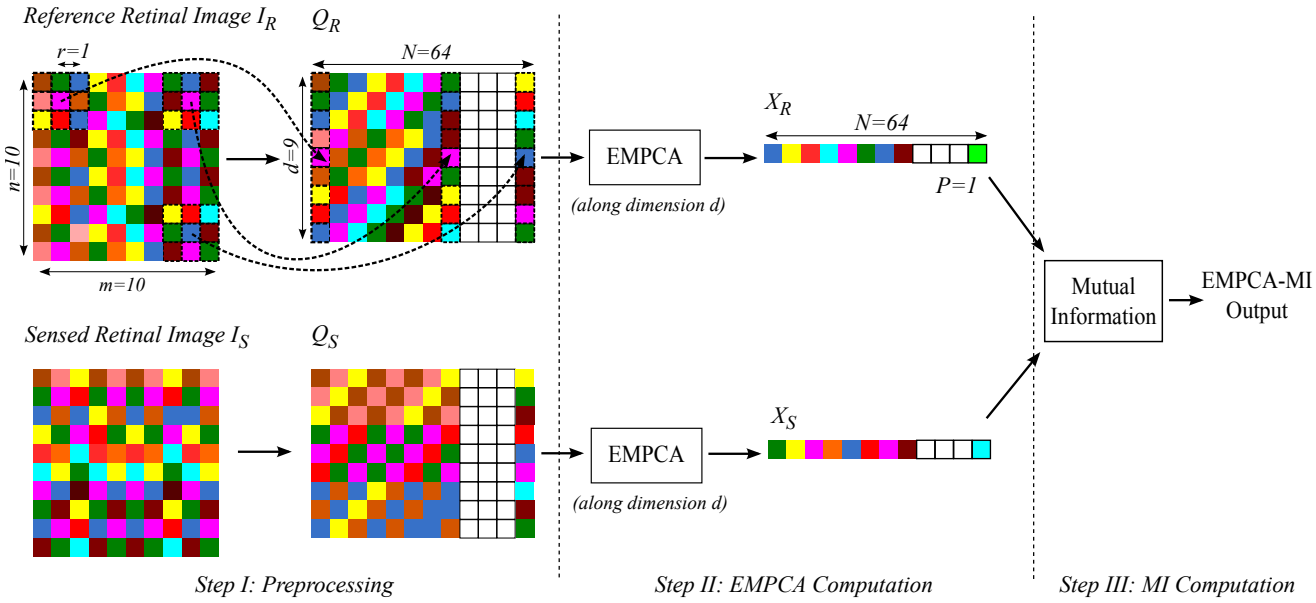

Fig. 1. Example of EMPCA-MI computation using a multimodal RI pair of size *10* × *10* pixels (with *P*=1 and *r*=1).

tomography and magnetic resonance images by combining spatial information with MI. It has subsequently been successfully deployed in RIR, though only for single modal images [23]. EMPCA-MI exhibits superior robustness in image registration performance in the presence of both non-uniform intensity and noise, and this provided the impetus to appraise its suitability to the very challenging multi-modal RIR domain.

This paper analyses the performance of the hybrid-based EMPCA-MI similarity measure [21] for multimodal RI datasets. The corresponding RIR performance has been compared with the recently proposed feature-based *bifurcation structures* method, with both quantitative and perceptual results consistently validating the superior registration performance of the EMPCA-MI algorithm.

The remainder of the paper is organised as follows: Section 2 presents the EMPCA-MI-based multimodal RIR framework, while Section 3 details the experimental set-up and analyses the registration performance. Finally, some concluding comments are provided in Section 4.

2. PROPOSED REGISTRATION FRAMEWORK

2.1. Principles of Multimodal RIR

Multimodal RIR consists of aligning a reference image I_R of one modality with a sensed image I_S of a different modality, in a four-stage process [2]: *i)* transforming the coordinates of I_S in a known reference space; *ii)* generating a new interpolated image $I_S^\star$ in the reference space; *iii)* comparing $I_S^\star$ with the reference image I_R using a predefined SM; and *iv)* optimizing transformation μ to achieve the best alignment at μ_{reg}. This multi-stage process can be generalised as a maximisation problem:

$$\mu_{reg} = \arg\max_{\mu} SM(I_R, \mu(I_S)) \quad (1)$$

where μ_{reg} gives the best registration parameter settings.

2.2. EMPCA-MI as a Similarity Measure

EMPCA-MI [21], [22] is a recently introduced SM for image registration that efficiently incorporates spatial information together with MI without incurring high computational overheads. As illustrated in the example in Fig. 1, it comprises three steps which involve pre-processing (image information rearrangement), followed by EMPCA [24] and MI [21] calculations. Note the block colours in Fig. 1 represent the pixels for pre-processing and not actual pixel values. The full EMPCA-MI model is detailed in **Algorithm 1**. I_R and I_S are the respective duo-modal RI which are pre-processed in *Step I* (Lines 1-5 in **Algorithm 1**) into vector forms Q_R and Q_S for a given neighbourhood radius *r*, so that the spatial and intensity information is preserved. The first *P* principal components X_R and X_S of the respective reference and sensed images are then iteratively computed by EMPCA along dimension *d* in *Step II* (Line 6), instead of solving the covariance matrix. The final MI value is calculated between X_R and X_S in *Step III* (Line 7), with a higher MI value reflecting that I_R and I_S are better aligned.

The EMPCA-MI framework in [21] considered only the first principal component, i.e., *P=1* as this represents the direction of the highest variance and the most dominant feature in a particular region. In this paper, the iterative nature of the EMPCA-MI model is exploited in two RIR case studies where additional principal components are mandated to improve the RIR quality. These scenarios represent a major challenge for multimodal RIR, because the images are characterised by large homogeneous regions. This means using *P=1* for RIR does not sufficiently represent the overall variance of the image information, with the corollary being higher registration errors. By including supplementary principal components i.e., *P=2*, EMPCA-MI addresses this limitation. For instance, the second principal component has the highest variance in the same region, subject to it being orthogonal to the first principal component, so the effect is to enhance the overall RIR process as more accurate image

Image Pair No.	*Scenario 1*		*Scenario 2*			*Scenario 3*		
	μ_{grd}=(0, 0, 60°, 1)		μ_{grd}=(10, 10, 30°, 2)			μ_{grd}=(8, 7, 45°, 0.8)		
	RIR-BS [12]	EMPCA-MI *P=1, r=1*	RIR-BS [12]	EMPCA-MI (*r=1*)		RIR-BS [12]	EMPCA-MI (*r=1*)	
				P			*P*	
1	0, 0, 0.02, 0	0, 0, 0.01, 0	0, 0, 0.048, 0.005	1	0, 0, 0.035, 0.002	0, 0, 0.058, 0.006	1	0, 0, 0.048, 0.008
2	0, 0, 0.24, 0	0, 0, 0.09, 0	0, 0, 0.163, 0.013	1	0, 0, 0.075, 0.009	0, 0, 0.195, 0.018	1	0, 0, 0.127, 0.016
3	0, 0, 0.62, 0	0, 0, 0.25, 0	0, 0, 0.057, 0.043	1	0, 0, 0.022, 0.032	0, 0, 1.049, 0.085	1	0, 0, 1.023, 0.064
4a	0, 0, 1.20, 0	0, 0, 0.70, 0	0, 0, 1.260, 0.071	1	0, 0, 1.920, 0.060	0, 0, 0.340, 0.022	1	0, 0, 0.380, 0.025
4b				2	0, 0, 1.219, 0.053		2	0, 0, 0.218, 0.016
5	0, 0, 0.11, 0	0, 0, 0.08, 0	0, 0, 0.760, 0.009	1	0, 0, 0.590, 0.004	0, 0, 0.880, 0.021	1	0, 0, 0.670, 0.014

Table 1. Registration errors expressed as Δt_x, Δt_y, $\Delta\theta$, ΔS of the similarity transformation for multimodal RI pairs, with μ_{grd} being the ground truth. *Image Pair 4a* and *4b* are the same RI *Image Pair*, but with EMPCA-MI using *P=1* and *P=2* respectively.

information is available in both the MI computation and optimization steps of **Algorithm 1** (Lines 6 and 7). This issue will be elaborated further in Section 3.1.

2.3. Transformation and other registration settings

Multimodal RI acquisition intrinsically introduces distortion, which is normally modelled by a *similarity transformation* [19], [20]. This special form of the global affine transform [25], which represents RI distortion as either *i)* changes in magnification to reflect that different equipment has been used, or *ii)* motion in the direction of the optical axis modelled as a uniform scaling *S* allied with (*x*-*y*) translational and rotational θ) components, to reflect eye and/or camera motion [26]. The notation adopted to represent the four key registration parameters is μ *(*t_x, t_y, θ, *S)* with the t_x translation, the t_y translation, θ rotation and scaling factor *S* of transform μ. To automatically determine these parameters, bicubic interpolation is used for the RI transformation and Powell's multidimensional direction set method along with Brent optimization [27], since it exhibits fast and accurate local search performance and is well-suited to RIR [2].

Algorithm 1: EMPCA-MI *for* RIR

Inputs: Multimodal RI I_R and I_S with spatial resolution of $m \times n$ *pixels*; $r-$ neighbourhood radius. $P-$ number of principal components.

Variables: $d-$ dimensional space; $N-$ total no of pixels; *i, j, k*$-$ indexes; Q_R, Q_S- rearranged matrices for I_R and I_S; X_R, X_S- principal components using EMPCA [24] for I_R and I_S.

Output: EMPCA-MI value

1: Initialise $d = (2r+1)^2$, $N = (m-2r)(n-2r)$ and $k = 1$;
2: While $k < N$ with $1 + r \leq i \leq m - r$, $1 + r \leq j \leq n - r$
3: Define $Q_{R([1\ldots d])(k)} = I_{R([(i-r)\ldots(i+r)])([(j-r)\ldots(j+r)])}$
4: $k = k + 1$
5: REPEAT Steps 2 to 4 for I_S to produce Q_S
6: Calculate X_R and X_S using Q_R and Q_R for given *P* [24]
7: Calculate MI(X_R, X_S) [21]
8: STOP

3. EXPERIMENTAL SET-UP AND RESULTS

To rigorously evaluate the performance of EMPCA-MI for RIR, a multimodal clinical dataset consisting of ten pairs of colour fundus and gray-scale SLO images was used [13], [28]. Each colour fundus image had a spatial resolution of *3888* $\times$ *2592* pixels and was acquired from a Canon CR-1, while the SLO images had a resolution of *768* $\times$ *768* pixels and were obtained from a SLO system (Spectralis, Heidelberg Engineering). Each RI modality contained non-uniform illumination, low contrast and differently sized homogeneous regions with varying reflectivity levels to accentuate the registration challenge. To establish the ground truth, since one was not available for this clinical dataset, selected images were misregistered by a known random transformation μ_{grd}, with the original images considered as the sensed images I_S. RIR was performed on only the green channel of the colour fundus images, since it has the highest contrast compared with the other channels, which are often saturated and contain acquisition noise [15], [17]. The tolerance thresholds for the Powell and Brent optimization were set to 10^{-5} and 10^{-3} respectively as in [2], [27] with the maximum permitted number of iterations being 200. RIR-BS [12] was chosen as a performance comparator in the experiments as it represents one of the most promising contemporary feature-based RIR techniques.

3.1. Results Discussion

To analyse the registration performance, three diverse scenarios with randomly generated μ_{grd} were considered. *Scenario 1* is a straightforward registration case, with only the rotation (θ=60°) parameter varied with the other three parameters being fixed, while *Scenarios 2* and *3* reflect more challenging cases with all four registration parameters being varied. In these two latter scenarios, the value of *S* reflects that the RI have been acquired with different equipment magnifications, while the t_x and t_y translations are pragmatically kept low as these motions are generally minimal during the image acquisition process [10]. While RIR using both EMPCA-MI and RIR-BS was undertaken for all ten *Image Pairs*, due to space limitations **Table 1** only displays the comparative registration errors for the first five pairs in *Scenarios 1, 2* and *3*.

The results conclusively reveal that for all *Image Pairs* EMPCA-MI provided consistently superior and more robust RIR performance compared with RIR-BS for each *Scenario*, with the notable exception of *Image Pair 4a* in *Scenarios 2* and *3*. *Image Pair 4* is the most challenging pair in the dataset as it has pathology alongside low contrast and very large homogeneous regions, which compromises

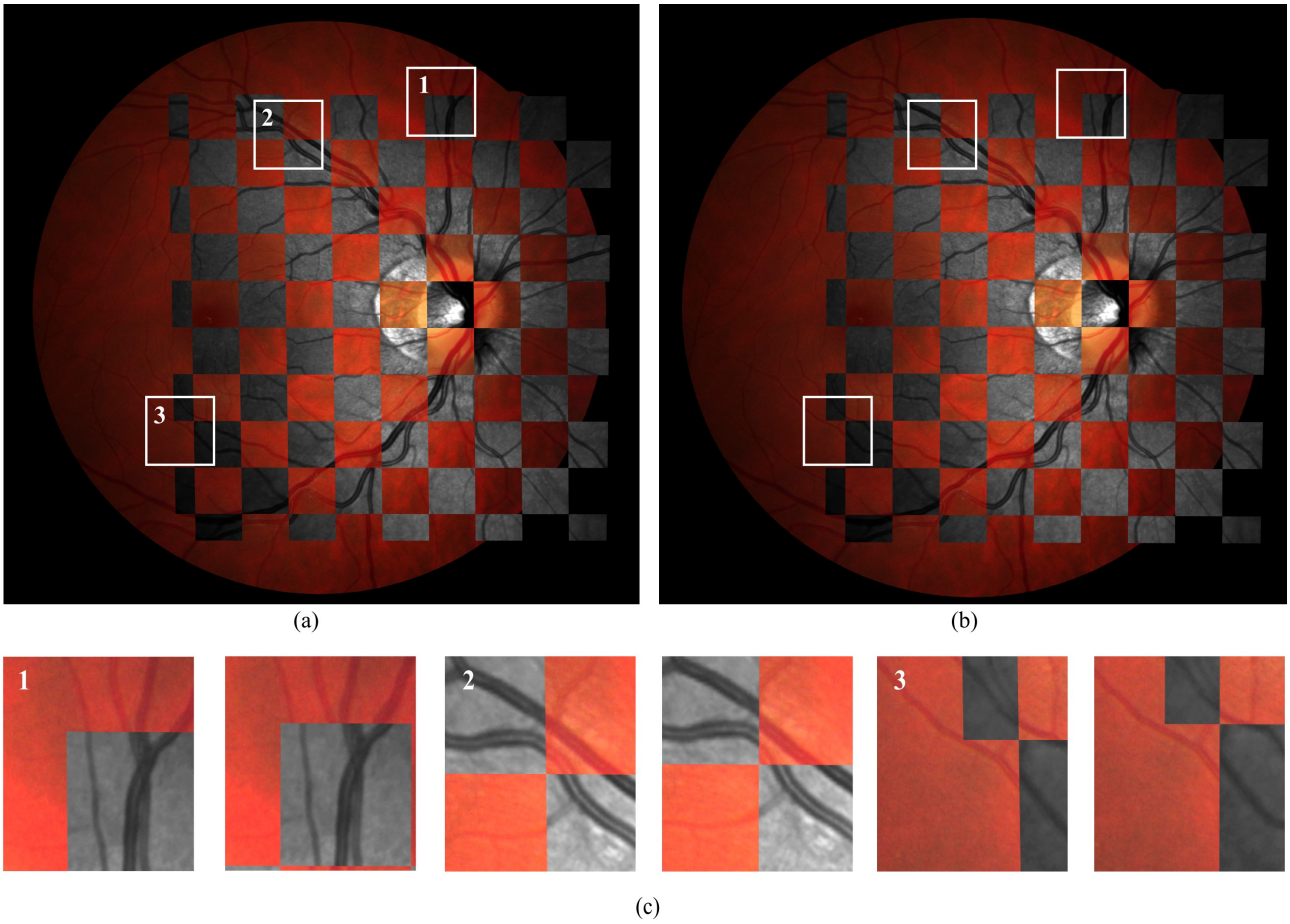

Fig. 2. Checkerboard overlay [2] illustration for *Image Pair 4b* in *Scenario 3* with three zoomed areas for (a) EMPCA-MI and (b) RIR-BS; (c) zoomed-in vascular structure details for three highlighted areas in both pairs.

EMPCA-MI performance when only one principal component is used. In contrast, since the BS method uses segmentation, RIR-BS quality is influenced much more by low contrast rather than by the presence of large homogeneous regions in the RI. The versatility of the EMPCA-MI model however, is illustrated when in order to secure improved RIR accuracy, the number of principal components is iteratively increased. This is evinced for *Image Pair 4b* in both *Scenarios 2* and *3*, where *P=2* is used so the overall variance of the RI is now represented by two orthogonal components. The corresponding results confirm this, with the EMPCA-MI rotational ($\Delta\theta$) and scaling (ΔS) registration errors for *Image Pair 4b* being reduced by more than *36%* and *11%* respectively in comparison with *Image Pair 4a* for *Scenario 2*, and by *43%* and *36%* for *Scenario 3*. Note the neighbourhood radius *r=1* has not been increased because due to the large homogeneities in this RI pair, using a larger *r* provides no benefit in terms of RIR performance.

To appreciate the corresponding perceptual RIR performance, Fig. 2 shows an example of the RIR results for *Image Pair 4b*, using the checkerboard overlaying method [2], with I_R and I_S being displayed in colour and gray-scale respectively. Some of the vascular structure areas in the resulting RIR for EMPCA-MI and RIR-BS are respectively highlighted in Fig. 2(a) and (b). Fig. 2(c) provides zoomed-in versions for three specific areas in this *Image Pair*, where the enhanced continuity in the vessel structures of the EMPCA-MI registration is clearly evident in boxes 1, 2 and 3 compared with the corresponding RIR-BS registered areas. These subjective results allied with the numerical results in **Table 1** corroborate both the superior performance and flexibility of the EMPCA-MI algorithm in registering multimodal RI.

4. CONCLUSION

This paper has analysed the performance of the hybrid-based *expectation maximization for principal component analysis with mutual information* (EMPCA-MI) algorithm in providing enhanced multimodal RIR, by combining spatial features with mutual information. Multimodal RIR is a very challenging problem due to the latent characteristics of low contrast, non-uniform illumination, large homogeneous regions and variable reflectivity levels in the different modalities. Both quantitative and subjective results for multimodal clinical datasets confirm the versatility of the EMPCA-MI model in being able to iteratively increase the number of principal components to furnish lower registration errors and superior RIR quality compared with the existing feature-based *bifurcation structures* method.

5. REFERENCES

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

[2] B. Zitová and J. Flusser, "Image registration methods: a survey," *Image Vision Comput.*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.

[3] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," *Phys. Med. Biol.*, vol. 46, no. 3, pp. R1–R45, Mar. 2001.

[4] C. Sanchez-Galeana, C. Bowd, E. Z. Blumenthal, P. A. Gokhale, L. M. Zangwill, and R. N. Weinreb, "Using optical imaging summary data to detect glaucoma," *Ophthalmology*, vol. 108, no. 10, pp. 1812–1818, Oct. 2001.

[5] P. J. Saine and M. E. Tyler, *Ophthalmic photography : retinal photography, angiography, and electronic imaging*, Butterworth-Heinemann, Boston, 2002.

[6] F. Laliberte, L. Gagnon, and Y. Sheng, "Registration and fusion of retinal images-an evaluation study," *IEEE Trans. Med. Imag.*, vol. 22, no. 5, pp. 661 –673, May 2003.

[7] J. Xu, O. Chutatape, E. Sung, C. Zheng, and P. Chew Tec Kuan, "Optic disk feature extraction via modified deformable model technique for glaucoma analysis," *Pattern Recogn.*, vol. 40, no. 7, pp. 2063–2076, July 2007.

[8] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[9] M. Sofka and C.V. Stewart, "Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures," *IEEE Trans. Med. Imag.*, vol. 25, no. 12, pp. 1531 –1546, Dec. 2006.

[10] C.V. Stewart, C.-L. Tsai, and B. Roysam, "The dual-bootstrap iterative closest point algorithm with application to retinal image registration," *IEEE Trans. Med. Imag.*, vol. 22, no. 11, pp. 1379 –1394, Nov. 2003.

[11] C. Pereira, N. Martins, L. Goncalves, and M. Ferreira, "Registration of retinal images by a MAS-ICP approach - a preliminary study," in *(ENBENG 2012)*, Feb. 2012, pp. 1 –4.

[12] L. Chen, Y. Xiang, Y. Chen, and X. Zhang, "Retinal image registration using bifurcation structures," in *2011 18th IEEE International Conference on Image Processing (ICIP)*, Sept. 2011, pp. 2169 –2172.

[13] R. Kolar, V. Harabis, and J. Odstrcilik, "Hybrid retinal image registration using phase correlation," *Imaging Sci. J.*, 2012.

[14] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187 –198, Apr. 1997.

[15] M. Skokan, A. Skoupy, and J. Jan, "Registration of multimodal images of retina," in *(EMBS/BMES 2002)*, 2002, vol. 2, pp. 1094 – 1096.

[16] L. Kubecka and J. Jan, "Registration of bimodal retinal images - improving modifications," in *26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2004. IEMBS '04*, Sept. 2004, vol. 1, pp. 1695 –1698.

[17] T. Chanwimaluang, G. Fan, and S.R. Fransen, "Hybrid retinal image registration," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 129 –142, Jan. 2006.

[18] D. B. Russakoff, C. Tomasi, T. Rohlfing, and C. R. Maurer, "Image similarity using mutual information of regions," in *(ECCV 2004)*, vol. 3023, pp. 596–607. 2004.

[19] P. A. Legg, P. L. Rosin, D. Marshall, and J. E. Morgan, "Incorporating neighbourhood feature derivatives with mutual information to improve accuracy of multi-modal image registration.," in *MIUA*, 2008, pp. 39–43.

[20] P. A. Legg, P. L. Rosin, D. Marshall, and J. E. Morgan, "A robust solution to multi-modal image registration by combining mutual information with multi-scale derivatives," in *(MICCAI 2009)*, number 5761, pp. 616–623. Jan. 2009.

[21] P. S. Reel, L. S. Dooley, and K. C. P. Wong, "A new mutual information based similarity measure for medical image registration," in *(IET IPR 2012)*, July 2012, pp. 1 –6.

[22] P. S. Reel, L. S. Dooley, and K. C. P. Wong, "Efficient image registration using fast principal component analysis," in *2012 19th IEEE International Conference on Image Processing (ICIP)*, Sept. 2012, pp. 1661–1664.

[23] P. S. Reel, L. S. Dooley, K. C. P. Wong, and Anko Börner, "Robust retinal image registration using expectation maximisation with mutual information," in *2013 38th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2013, pp. 1118–1122.

[24] S. Roweis, "EM algorithms for PCA and SPCA," in *1997 Conference on Advances in Neural Information Processing Systems (NIPS)*, 1998, pp. 626–632.

[25] L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surv.*, vol. 24, no. 4, pp. 325–376, Dec. 1992.

[26] N. Ryan, C. Heneghan, and P. de Chazal, "Registration of digital retinal images using landmark correspondence by expectation maximization," *Image Vision Comput.*, vol. 22, no. 11, pp. 883–898, Sept. 2004.

[27] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, Cambridge University Press, 3 edition, Sept. 2007.

[28] R. Kolar and P. Tasevsky, "Registration of 3D retinal optical coherence tomography data and 2D fundus images," in *Biomedical Image Registration*, B. Fischer, B. M. Dawant, and C. Lorenz, Eds., number 6204 in Lecture Notes in Computer Science, pp. 72–82. Springer Berlin Heidelberg, Jan. 2010.