Comparison of Image Registration Methods for Composing Spectral Retinal Images

Lauri Laaksonen[1], Ela Claridge[2], Pauli Fält[3], Markku Hauta-Kasari[3], Hannu Uusitalo[4], and Lasse Lensu[1]

[1]Machine Vision and Pattern Recognition Laboratory
Department of Mathematics and Physics
Lappeenranta University of Technology
PO Box 20, FI-53851 Lappeenranta, Finland
lauri.laaksonen@lut.fi
[2]School of Computer Science, University of Birmingham, United Kingdom
[3]School of Computing, University of Eastern Finland
[4]Department of Ophthalmology, SILK, University of Tampere, Finland

Abstract. Spectral retinal images have significant potential for improving the early detection and visualization of subtle changes due to eye diseases and many systemic diseases. High resolution in both the spatial and the spectral domain can be achieved by capturing a set of narrow-band channel images from which the spectral images are composed. With imaging techniques where the eye movement between the acquisition of the images is unavoidable, image registration is required. In this paper, the applicability of the state-of-the-art image registration methods for the composition of spectral retinal images is studied. The registration methods are quantitatively compared using synthetic channel image data of an eye phantom and semisynthetic set of retinal channel images subjected to known transformations. The experiments show that Generalized dual-bootstrap iterative closest point method outperforms the other evaluated methods in registration accuracy and the number of successful registrations.

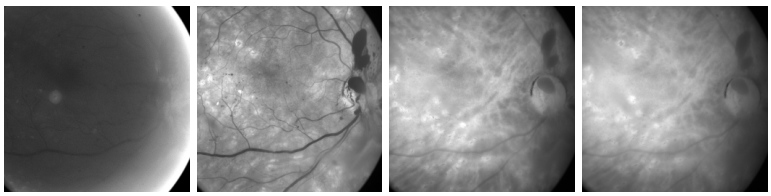
Keywords: Image registration, spectral imaging, retinal imaging, fundus imaging, quantitative evaluation

1 Introduction

Eye diseases such as Diabetic retinopathy (DR), glaucoma and Age-related macular degeneration (AMD), or complications of many systemic diseases like diabetes and Systemic hypertension (SH), cause structural changes in the eye fundus. Early detection of the retinal changes, their progression and risk factor analysis allow better and more cost-effective treatment as most diseases can be successfully treated if diagnosed early and monitored regularly.

Retinal imaging provides a non-invasive view into the eye and its vascular bed, and is the standard practice to screen, document, diagnose and monitor eye diseases. Greyscale or RGB images with high spatial resolution are commonly

used in the diagnosis, complemented with more advanced eye imaging methods when necessary. To support further development of the diagnostic tools, methods for spectral reflectance measurements, especially spectral imaging of the retina, have been developed [3,4,5,10]. The aim has been to significantly improve the possibility to detect and visualize different parts of the retina and lesions related to the eye diseases. In [3], an imaging system for measuring spectral images of the retina using a set of 30 bandpass interference filters is presented. Two example images acquired with the setup are shown in Fig. 1.



(a) $\lambda = 442$ nm (b) $\lambda = 589$ nm (c) $\lambda = 650$ nm (d) $\lambda = 690$ nm

Fig. 1: Example channel images captured with the system in [3]. The images are enhanced for visualization.

Depending on the imaging technology, composing spectral images based on a set of channel images may require image registration at some stage. The purpose of the registration is to find the geometric transformation needed to align the floating image (the image to be registered) with the base image (the registration target). Manually selecting corresponding points in image pairs becomes infeasible when the number of individual channels increases or when salient features become very different between images. To solve the problem of image alignment, there exists a significant body of work in the field of image registration.

This paper presents a comparison of image registration approaches for the composition of spectral retinal images. The registration methods are quantitatively evaluated on a set of channel images of an eye phantom and a retinal image set, with each image deformed by a known transformation.

2 Methods

In local similarity based registration, the floating image is deformed in such manner that a set of local similarities are maximized. In the framework used in this work, the nodes of an $n \times n$ grid are iteratively moved based on local similarity of the base and floating image. The final transformation is obtained by b-spline interpolation using the grid nodes as control points. Several similarity measures were quantitatively evaluated in this framework: Correlation coefficient (CC), Similarity measure by Myronenko et al. (CD2) [8], Similarity measure by Cohen and Dinstein (MS) [2], Mutual information (MI) [15], Minimization of

residual complexity (RC) [7], Sum of absolute differences (SAD) and Sum of squared differences (SSD).

Thirion [11] has presented a registration approach (Demons) where the floating image is considered as a deformable grid model diffusing through semi-permeable membranes defined by, e.g., the gradients of the base image. The floating image is transformed by a grid that is deformed by internal forces (relations between grid points) and demons that locally push a diffusing model (grid node) based on the underlying base image. Vercauteren et al. [13,14] have extended the demons framework Log-demons to the space of diffeomorphic transformations by Lie groups, and further extend the diffeomorphic demons framework into log-domain to guarantee the existence of and access to the inverse transformation.

Generalized dual-bootstrap iterative closest point (GDB-ICP) [16] finds a transformation aligning two images by starting from a small area of overlap (bootstrap region) between the images and a locally stable similarity transformation. An initial transformation derived from a Scale-invariant feature transform (SIFT) descriptor match is refined and validated by feeding edge and corner points inside a growing bootstrap region to a robust Iterative closest point (ICP) algorithm. Edge-driven dual-bootstrap iterative closest point (ED-DB-ICP) [12] is a modification of GDB-ICP designed for the registration of multimodal Fluorescein angiogram (FA) sequences. The main differences of ED-DB-ICP with respect to GDB-ICP are the use of gradient magnitude images instead of intensity images and extending the SIFT-descriptors with a shape context descriptor presented in detail in [6]. Publicly available implementations were used for all the methods.

3 Experiments and results

3.1 Datasets and performance evaluation

To quantitatively evaluate the performance of each method with the data acquired with the system described in [3], five channel image sets were used to directly estimate the image registration error. A fully aligned spectral image consisting of 30 channels with spatial resolution of 1024x1024 of a phantom (artificial eye) was used as the basis, with a known deformation of increasing degree applied to each set. The produced sets are referred to as the synthetic test set.

To evaluate the registration performance on real medical data, a semisynthetic test set was generated by deforming spectral retinal images taken with the system in [1]. The system captures a set of six channel images in approximately 0.5 s to avoid inter-channel eye movement [1]. The retina was illuminated at six selected wavelengths {507, 525, 552, 585, 596, 611} nm [10] using a halogen white-light source filtered through a liquid crystal tunable filter. During the experiments, the maximum inter-channel displacement was found to be 2.3 pixels (referred to as the system error).

For both the synthetic and semisynthetic test sets, five image sets were generated from the original images by transforming each channel by a known transfor-

mation with transformation parameters sampled from a parameter distribution. Based on the experiments in [9], a quadratic transformation was deemed appropriate to represent the deformations due to eye movement during retinal image acquisition. The parameter distribution for the transformations was determined from a previously registered true retinal channel image set (with 1 442 images and successful registration confirmed by an expert) by using kernel density estimation. The the quadratic term parameters were multiplied by an increasing weighting constant k for each image set to simulate increasingly distorted images. For each set, a channel image near the middle of the wavelength range ($\lambda = 540$ nm for the synthetic and $\lambda = 552$ nm for the semisynthetic, respectively) was chosen as the base image for each registration. Examples of the test set images and the corresponding transformations are shown in Fig. 2.

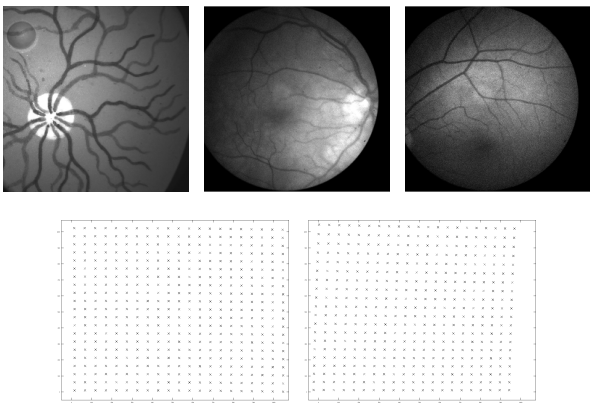


Fig. 2: Examples of synthetic and semisynthetic test sets (enhanced for visualization). The images from left to right and top to bottom are as follows: an original image of the synthetic set ($\lambda = 600$ nm); examples of Set 1 and Set 5 of the semisynthetic test set ($\lambda = 585$ nm); transformation grid examples for synthetic Set 1 and Set 5.

The parameters for each registration method were systematically selected by registering a subset of the test images whilst varying the parameter values. The parameter combination that produced the smallest error on the images was chosen. All registration errors were measured as the Euclidean distance between the grid points of the registered image and the base image. In the semisynthetic test set, it is possible that the registration method, in addition to estimating the synthetic transform, corrects some of the system error. This would manifest itself as increased error despite the more accurate registration. Therefore, error values below the system error of 2.3 pixels are considered as zero for the semisynthetic test set.

3.2 Results

The results of the image registration performance comparison are summarized in Tables 1 and 2. While the most relevant results are discussed in this section, visualizations of imagewise errors for most of the methods are omitted for brevity.

Table 1: Median (Med) and standard deviation (Std) of registration error for the synthetic set. Base stands for the initial error before registration. Corr is the correlation with baseline error.

		Base	CC	CD2	MI	MS	RC	SAD	SSD	Demons	Log-demons	GDB-ICP	ED-DB-ICP
Med.	Set 1	33.7	2.9	64.4	2.3	154.1	2.0	128.8	93.9	110.4	33.5	1.0	0.7
	Set 2	47.3	2.6	101.3	7.2	128.6	2.0	144.3	81.0	323.8	48.4	1.1	0.7
	Set 3	28.2	3.0	105.8	3.6	233.8	2.2	177.4	114.4	143.4	28.9	1.1	0.7
	Set 4	28.9	4.2	29.4	11.9	257.2	4.3	115.3	55.8	18.7	32.8	1.1	0.7
	Set 5	43.0	6.1	105.0	11.8	200.1	5.4	147.4	126.6	139.9	43.3	1.0	0.7
Std.	Set 1	6.6	3.0	46.4	1.9	74.9	4.9	67.8	51.2	50.8	10.3	0.4	0.4
	Set 2	18.6	6.4	40.4	9.3	52.4	4.6	47.4	40.0	210.9	18.7	0.4	0.3
	Set 3	9.8	3.5	22.3	5.4	53.8	6.1	47.1	42.4	63.8	15.2	0.3	0.2
	Set 4	14.6	8.6	28.2	19.4	69.3	18.9	50.7	30.6	18.9	17.4	0.4	0.4
	Set 5	14.5	8.3	20.7	9.0	41.5	8.9	35.2	44.0	75.4	18.4	0.4	0.3
Corr.		1.00	0.16	0.53	0.25	-0.74	0.09	0.01	0.20	0.78	0.98	-0.33	0.05

CC performed relatively well for both the synthetic and semisynthetic sets, achieving a median registration error below five pixels for most images. The variation of error for some images was high, especially for synthetic sets 4 and 5, however. Excluding some individual images, RC produced low median errors for all sets for both synthetic and semisynthetic images. However, the error variance was high, especially for the fourth synthetic set. MI produced relatively good results for the first synthetic set, but showed sensitivity to the increasing level of deformation with the successive sets. For the semisynthetic set MI performed comparably to CC. CD2, MS, SAD and SSD managed to, at least in part, successfully register the images at wavelengths near the base image, but produced high errors elsewhere.

Demons showed very high standard deviation in error and produced globally unrealistic transformations. The method, however, registered local image areas with high accuracy, resulting in error minimum close to zero for most images and low median errors for the wavelengths close to the base image. Log-Demons performed similarly to Demons, but produced less extreme errors and smaller standard deviations for most images. Furthermore, log-demons showed high correlation between the registration and baseline error, suggesting that the method is sensitive to the initial configuration.

ED-DB-ICP was hindered because of using the gradient information, especially in the semisynthetic test set. While the successful registrations were highly

Table 2: Median (Med) and standard deviation (Std) of registration error for the semisynthetic set. Base stands for the initial error before registration. Corr is the correlation with baseline error.

		Base	CC	CD2	MI	MS	RC	SAD	SSD	Demons	Log-demons	GDB-ICP	ED-DB-ICP
Med	Set 1	13.8	4.5	36.8	5.8	50.8	4.0	56.5	33.0	39.2	16.1	0.0	3.5
	Set 2	51.4	0.0	18.4	0.0	21.3	9.9	44.7	15.1	15.2	39.0	0.0	0.0
	Set 3	25.7	0.0	32.6	0.0	10.9	0.0	61.4	3.7	15.2	23.3	0.0	-
	Set 4	58.7	4.6	167.7	4.5	6.4	5.1	17.1	4.8	8.7	34.2	0.0	34.7
	Set 5	33.2	3.4	41.3	3.5	11.5	4.2	29.4	10.2	12.4	31.8	0.0	332.2
Std	Set 1	5.1	3.8	14.6	21.6	21.6	11.3	21.4	13.1	14.5	6.0	0.2	2.3
	Set 2	1.3	0.3	7.9	10.9	10.9	14.1	21.5	6.3	6.7	4.4	0.0	0.0
	Set 3	3.5	1.1	2.4	5.6	5.6	5.0	27.6	2.2	5.3	3.5	0.0	-
	Set 4	9.2	2.0	6.7	4.8	4.8	2.4	11.3	2.3	4.8	9.6	0.0	13.0
	Set 5	13.6	3.2	19.9	8.4	8.4	6.4	16.9	8.0	9.6	11.1	0.0	117.6
Corr.		1.00	-0.07	0.59	-0.23	-0.65	0.59	-0.74	-0.57	-0.76	0.91	0.00	-0.15

accurate, the method failed (i.e., the method found no stable transformation) with several images in all sets. GDB-ICP showed very good performance for both the synthetic and semisynthetic test sets. With the exception of failed registrations for the two shortest wavelength images of the synthetic set, and two partially successful registration in the semisynthetic set, the method achieved a close to sub-pixel accuracy with minimal standard deviation. The channel-wise performances for Set 1 of the test sets are shown in Fig. 3.

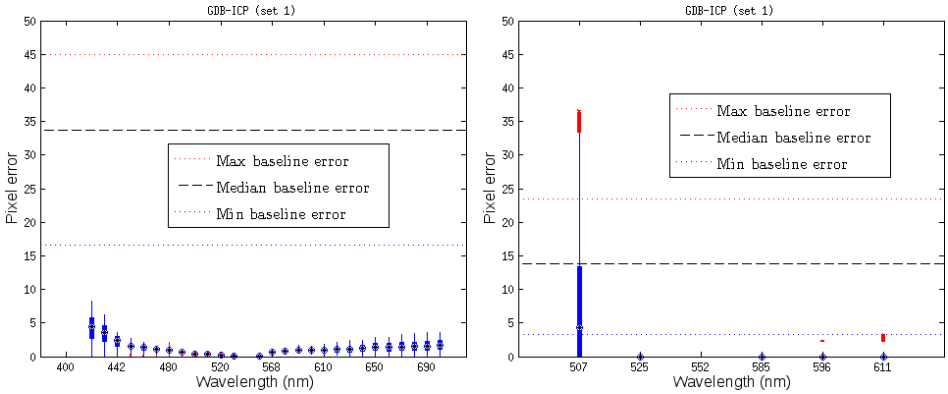
As expected, the errors tend to increase with the wavelength difference of the registered images as the prominent image features change (see Fig. 1). The channel-wise errors of the best-performing methods are visualized in Fig. 4.

4 Conclusion

This paper focused on the comparison of image registration methods for composing spectral retinal images. The experiments on synthetic and semi-synthetic data showed that the registration error increases with increasing wavelength difference between the floating and base image. GDB-ICP outperformed the other methods in both the number of successful registrations and registration accuracy. The other well-performing methods were CC, MI and RC, but they could not match the accuracy and success rate of GDB-ICP.

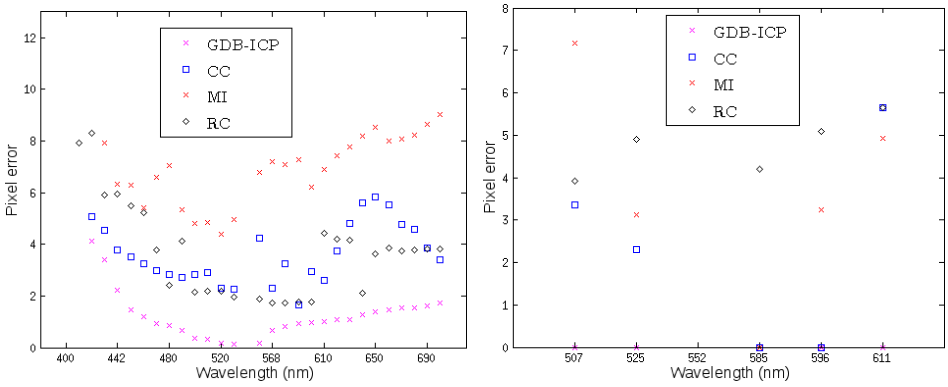
Acknowledgements

The authors would like to thank the Academy of Finland for the financial support of the ReVision project (No. 259560).



(a) Synthetic (set 1) (b) Semisynthetic (set 1)

Fig. 3: Generalized dual-bootstrap iterative closest point (GDB-ICP) errors; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually.



(a) Synthetic images (b) Semisynthetic images

Fig. 4: Wavelength-wise registration error; Median errors over the sets for each wavelength.

References

1. Calcagni, A., Gibson, J.M., Styles, I.B., Claridge, E., Orihuela-Espina, F.: Multispectral retinal image analysis: a novel non-invasive tool for retinal imaging. Eye 25(12) (2011)
2. Cohen, B., Dinstein, I.: New maximum likelihood motion estimation schemes for noisy ultrasound images. Pattern Recognition 35(2), 455–463 (2002)
3. Fält, P., Hiltunen, J., Hauta-Kasari, M., Sorri, I., Kalesnykiene, V., Pietilä, J., Uusitalo, H.: Spectral imaging of the human retina and computationally determined optimal illuminants for diabetic retinopathy lesion detection. Journal of Imaging Science and Technology 55(3), 030509–1–030509–10 (2011)
4. Gao, L., Smith, R.T., Tkaczyk, T.S.: Snapshot hyperspectral retinal camera with the image mapping spectrometer (ims). Biomed. Opt. Express 3(1), 48–54 (Jan 2012), http://www.opticsinfobase.org/boe/abstract.cfm?URI=boe-3-1-48
5. Johnson, W.R., Wilson, D.W., Fink, W., Humayun, M., Bearman, G.: Snapshot hyperspectral imaging in ophthalmology. J. Biomed. Opt. 12(1), 014036 (2007)
6. Mortensen, E., Deng, H., Shapiro, L.: A sift descriptor with global context. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. pp. 184 – 190. San Diego, USA (Jun 2005)
7. Myronenko, A., Song, X.: Intensity-based image registration by minimizing residual complexity. IEEE Transactions on Medical Imaging 29(11), 1882–1891 (2010)
8. Myronenko, A., Song, X., Sahn, D.: Maximum likelihood motion estimation in 3d echocardiography through non-rigid registration in spherical coordinates. Functional Imaging and Modeling of the Heart 5528, 427–436 (2009)
9. Stewart, C., Tsai, C.L., Roysam, B.: The dual-bootstrap iterative closest point algorithm with application to retinal image registration. IEEE Transactions on Medical Imaging 22(11), 1379 – 1394 (2003)
10. Styles, I., Calcagni, A., Claridge, E., Orihuela-Espina, F., Gibson, J.: Quantitative analysis of multi-spectral fundus images. Medical Image Analysis 10(4), 578–597 (August 2006)
11. Thirion, J.P.: Image matching as a diffusion process: an analogy with maxwell's demons. Medical Image Analysis 2 3, 243 – 260 (1998)
12. Tsai, C.L., Li, C.Y., Yang, G., Lin, K.S.: The edge-driven dual-bootstrap iterative closest point algorithm for registration of multimodal fluorescein angiogram sequence. IEEE Transactions on Medical Imaging 29(3), 636 – 649 (2010)
13. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Non-parametric diffeomorphic image registration with the demons algorithm. In: Proceedings of Medical Image Computing and Computer-assisted Intervention – MICCAI 2007. pp. 319–326. Brisbane, Australia (Oct 2007)
14. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Symmetric log-domain diffeomorphic registration: a demons-based approach. In: Proceedings of Medical Image Computing and Computer-assisted Intervention – MICCAI 2008. pp. 754–761. New York, USA (Sep 2008)
15. Viola, P., Wells III, W.: Alignment by maximization of mutual information. International Journal of Computer Vision 24(2), 137–154 (1997)
16. Yang, G., Stewart, C., Sofka, M., Tsai, C.L.: Registration of challenging image pairs: Initialization, estimation and decision. IEEE Transactions on Pattern Analysis and Machine Intelligence 29(11), 1973 – 1989 (2007)