Multimodal retinal image registration using a fast principal component analysis hybrid-based similarity measure

Conference or Workshop Item

How to cite:

For guidance on citations see FAQs.

© 2013 IEEE

Version: Accepted Manuscript
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
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 $I'_S$ in the reference space; iii) comparing $I'_S$ with the reference image $I_R$ using a predefined SM; and iv) optimizing transformation $\mu$ to achieve the best alignment at $\mu_{reg}$. This multi-stage process can be generalised as a maximisation problem:

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

where $\mu_{reg}$ gives the best registration parameter settings.

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 multimodal 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.

2.2. 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
Table 1. Registration errors expressed as $\Delta x_1$, $\Delta t_y$, $\Delta \theta$, $\Delta S$ of the similarity transformation for multimodal RI pairs, with $\mu_{grd}$ 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.

<table>
<thead>
<tr>
<th>Image Pair No.</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu_{grd}=(0, 0, 60^\circ, 1)$</td>
<td>$\mu_{grd}=(10, 10, 30^\circ, 2)$</td>
<td>$\mu_{grd}=(8.7, 45^\circ, 0.8)$</td>
</tr>
<tr>
<td>RIR-BS [12]</td>
<td>EMPCA-MI</td>
<td>RIR-BS</td>
<td>EMPCA-MI</td>
</tr>
<tr>
<td>$P=1$, $r=1$</td>
<td>$1$, $0$, $0$, $0.048$, $0.005$</td>
<td>$1$, $0$, $0$, $0.035$, $0.002$</td>
<td>$1$, $0$, $0$, $0.058$, $0.006$</td>
</tr>
<tr>
<td></td>
<td>$2$, $0$, $0$, $0.24$, $0.013$</td>
<td>$1$, $0$, $0$, $0.075$, $0.009$</td>
<td>$1$, $0$, $0$, $0.195$, $0.018$</td>
</tr>
<tr>
<td></td>
<td>$3$, $0$, $0.62$, $0.025$, $0.057$, $0.043$</td>
<td>$1$, $0$, $0$, $0.022$, $0.032$</td>
<td>$1$, $0$, $0$, $1.049$, $0.085$</td>
</tr>
<tr>
<td></td>
<td>$4a$, $0$, $0.120$, $0.070$, $0.1260$, $0.071$</td>
<td>$1$, $0$, $0.192$, $0.060$</td>
<td>$1$, $0$, $0.340$, $0.022$</td>
</tr>
<tr>
<td></td>
<td>$4b$, $2$, $0$, $0.1219$, $0.553$</td>
<td>$1$, $0$, $0.1260$, $0.009$</td>
<td>$1$, $0$, $0.590$, $0.004$</td>
</tr>
<tr>
<td></td>
<td>$5$, $0$, $0.11$, $0.008$, $0.760$, $0.009$</td>
<td>$0$, $0$, $0.088$, $0.021$</td>
<td>$1$, $0$, $0.670$, $0.014$</td>
</tr>
</tbody>
</table>

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 $\mu_{grd}$, 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 $\mu_{grd}$ were considered. Scenario 1 is a straightforward registration case, with only the rotation ($\theta=60^\circ$) 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
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 ($\Delta S$) 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


