Robust retinal image registration using expectation maximisation with mutual information

Conference Item

How to cite:

For guidance on citations see FAQs.

© 2013 IEEE
Version: Accepted Manuscript

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.

oro.open.ac.uk
ABSTRACT

Retinal images (RI) are widely used to diagnose a variety of eye conditions and diseases such as myopia and diabetic retinopathy. They are inherently characterised by having non-uniform illumination and low-contrast homogeneous regions which represent a unique set of challenges for retinal image registration (RIR). This paper investigates using the expectation maximization for principal component analysis based mutual information (EMPCA-MI) algorithm in RIR. It combines spatial features with mutual information to efficiently achieve improved registration performance. Experimental results for mono-modal RI datasets verify that EMPCA-MI together with Powell-Brent optimization affords superior robustness in comparison with existing RIR methods, including the geometrical features method.

Index Terms—Image registration, principal component analysis, mutual information, expectation-maximization algorithms, retinopathy.

1. INTRODUCTION

Image registration is an integral procedure in many computer vision and image processing applications [1], [2]. For example, registering medical images in conjunction with image fusion facilitates disease diagnosis and treatment planning [3] in many areas of human anatomy. Retinal images (RI) of a patient need to be registered in order to diagnose eye conditions and diseases such as myopia, glaucoma and diabetic retinopathy [4]. Retinal image registration (RIR) aligns the vessel structures of the retina to assist in ophthalmology, particularly in the tracking and analysis of the advancement of these diseases.

RIR is challenging because the images exhibit spatially non-uniform contrasts and intensity distributions, as well as having large homogeneous non-vascular regions. These collective characteristics can be further compromised by degradations in various pathologies [4]. Existing intensity and feature-based RIR techniques are not always robust to effectively manage RI and have been shown to be dependent on the quality of features extracted, while hybrid approaches which combine intensity and feature techniques are not computationally efficient. This paper investigates the application of a new hybrid-based similarity measure [5], which incorporates RI spatial information with mutual information (MI). The corresponding RIR performance is compared with an alternative current feature-based method for a series of mono-modal RI test scenarios and found to be more robust.

2. PREVIOUS WORK

RIR can be broadly classified into feature, intensity and hybrid based techniques [6]. Feature-based approaches primarily use extracted vessel structure and landmark points from the RI, while intensity techniques focus solely on pixel intensity information. They can also be combined effectively to form hybrid techniques [1]. Utilising features in RIR is analogous to manual registration, as key structures are extracted from the RI such as optical disk [7], fovea [8] and vascular structural features [9]. An objective function is constructed from the extracted features, which is subsequently optimised to secure the best alignment position of the RI. Dual-bootstrap iterative closest point [10] and its scale-invariant feature transform variants [11], use vascular features of small regions to grow a bootstrap region to accomplish registration. While these methods are computationally efficient, their registration performance is very dependent on the quality of segmentation and degree of overlapping of the extracted features.

The geometrical features method (GFM) [12] is a recent feature-based method which has been shown to provide better registration results compared with other existing RIR techniques. By applying affine moment invariant (AMI) descriptors, GFM performs closed-boundary region extraction.
The interpolation method used for transforming I registration errors [2].

In contrast, intensity-based techniques use a similarity measure (SM) such as cross correlation, least mean square error, phase correlation or MI [13], [14] to compare and match RI by applying an optimization strategy to maximise the SM within transformation search space. MI has been widely adopted in both mono and multi-modal medical image registration of various parts of the human body. Instead of having pre-processing steps such as segmentation and feature extraction, MI establishes a statistical relationship between the intensity values of the images, though because of the challenging characteristics of RI, MI does not perform very effectively [15].

Hybrid approaches combine various aspects of the aforementioned feature and intensity methods, such as for example, using extracted vascular structures [16] together with spatial information as in regional MI [17] and feature neighbourhood MI [18], [19]. While these techniques use a covariance matrix to reduce data complexity [17] instead of high-dimensional histograms, as the spatial information expands, it does lead to a commensurate increase in computational time. EMPCA-MI [5], [20] is a hybrid-based algorithm used for registering brain computed tomography and magnetic resonance images by combining spatial information with MI. It has been shown to exhibit superior robustness in its registration performance in the presence of non-uniform intensity and noise, and given the innately challenging characteristics of RI this provided the impetus to investigate applying EMPCA-MI [5] within a RIR context.

## 3. PROPOSED REGISTRATION FRAMEWORK

### 3.1. Principles of Retinal Image Registration (RIR)

RIR consists of aligning a reference image IR with a sensed image IS in a multi-step process [2]: i) transforming the coordinates of IS in a known reference space; ii) generating a new interpolated image IS in the reference space; iii) comparing IS with the reference image IR using a predefined SM; and iv) optimizing transformation µ to achieve the best alignment at µreg. This can be generalised as a maximisation problem:

\[
\mu_{reg} = \arg \max_{\mu} SM(I_R, \mu(I_S))
\]

where \(\mu_{reg}\) gives the best registration parameter settings. The interpolation method used for transforming IS plays a central role in RIR, since interpolation artefacts can considerably impact upon SM performance leading to increased registration errors [2].

### 3.2. EMPCA-MI as a Similarity Measure

EMPCA-MI [5], [20] 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(a), it comprises three steps involving pre-processing (image data rearrangement), EMPCA and MI calculation. Note the block colours in Fig. 1(a) symbolise pixels for pre-processing and not the actual pixel itself. IR and IS are both pre-processed (Step I) in vector form for a given neighbourhood radius \(r\), so 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 in Step II, instead of solving the covariance matrix. The final MI value is calculated between \(X_R\) and \(X_S\) in Step III, with a higher MI value meaning the two images are better aligned. In the EMPCA-MI algorithm, only the first principal component is considered, i.e., \(P=1\) since this is the direction of highest variance and represents the most dominant feature. EMPCA-MI can be formally expressed as:

\[
EMPCA-MI(I_R, I_S) = \sum_{X_R, X_S} p(X_R, X_S) \log \frac{p(X_R, X_S)}{p(X_R)p(X_S)}
\]

where \(p(X_R)\) and \(p(X_S)\) are the individual probabilities of \(X_R\) and \(X_S\) respectively, while \(p(X_R, X_S)\) is their joint probability.

### 3.3. Transformation and other registration settings

RI acquisition inherently leads to a distortion between IS and IR which can be modelled as a similarity transformation [12], [18], [19]. This special form of the global affine transform [21] represents RI distortion as either eye or camera (x-y translational and rotational \(\theta\)) motion, with magnification changes due to using different equipment or the motion in the direction of the optical axis being modelled as a uniform scaling S [22]. These four parameters can be formulated as:

\[
\begin{bmatrix}
    x' \\
    y'
\end{bmatrix} = \begin{bmatrix}
    t_x & S \cos \theta & -S \sin \theta \\
    t_y & S \sin \theta & S \cos \theta
\end{bmatrix} \begin{bmatrix}
    1 \\
    x \\
    y
\end{bmatrix}
\]

(3)

Where \((x, y)\) and \((x', y')\) are the original and transformed pixel positions respectively. The notation \(\mu (t_x, t_y, \theta, S)\) is adopted to represent the four key registration parameters: the x translation, the y translation, the rotation \(\theta\) and scaling factor \(S\) of transform \(\mu\). To automatically determine these parameters, Powell’s multidimensional direction set method is applied along with Brent optimization [23] for line minimization, because this provides a local search which is accurate, fast and particularly suited to RIR [2].

## 4. EXPERIMENTAL SET-UP AND RESULTS

To evaluate the RIR, the clinical dataset DRIVE [24], from a Dutch diabetic retinopathy screening program was used.
This comprises twenty colour images with spatial resolution 768 by 584 pixels and 45° circular fields of view. All RI contain non-uniform illumination and low contrast, which accentuates the registration challenge. Since reference images were not available for this database, they were simulated to establish the requisite ground truth, by mis-registering them by a known transformation $\mu_{grd}$, with the original images considered as the sensed images. RIR was undertaken only on the green channel, since this has the highest contrast compared with the red and blue channels, which can often be saturated and contain acquisition noise [12].

For optimization, the tolerance thresholds for the Powell and Brent criteria were set to $10^{-5}$ and $10^{-3}$ respectively [2], with the maximum number of iterations being 200. As delineated in Section 2, GFM [12] is one of the leading feature-based solutions, so this was chosen as the comparator in the following RIR experiments.

4.1. Results Discussion

As alluded in Section 3.1, interpolation plays key role in the registration process and can significantly perturb the performance of a SM. Before analysing the performance of EMPCA-MI for RIR, the influence of different interpolation techniques including nearest neighbour (NN), Bilinear and Bicubic on EMPCA-MI was examined. Various simulation runs for EMPCA-MI using different Image Pairs from the dataset [24] were performed using Bicubic, Bilinear and NN interpolation. It was observed that Bicubic interpolation consistently provided the lowest registration error as evidenced in Fig 1(b), which shows the respective EMPCA-MI traces for Image Pair 3, with $\mu_{grd}=(0, 10, 0, 1)$ for all three interpolators. As both the Bilinear and NN interpolators generated higher errors, the Bicubic interpolator was considered for the various RIR experiments. Fig. 1(c) shows the EMPCA-MI results for Image Pair 6, with $\mu_{grd}=(6, 50, 1.2)$ where all
**Table 1.** Registration errors for four scenarios of retinal image pairs. $\mu_{grd}$ is ground truth transform and $\Delta t_x$, $\Delta t_y$, $\Delta \theta$, $\Delta S$ are registration errors for the similarity transformation.

<table>
<thead>
<tr>
<th>Image Pair No.</th>
<th>Scenario 1 $\mu_{grd}=(0, 0, 5^\circ, 1)$</th>
<th>Scenario 2 $\mu_{grd}=(0, 0, 60^\circ, 1)$</th>
<th>Scenario 3 $\mu_{grd}=(5, 5, 20^\circ, 2.8)$</th>
<th>Scenario 4 $\mu_{grd}=(8, 9, 45^\circ, 0.8)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta t_x$, $\Delta t_y$, $\Delta \theta$, $\Delta S$</td>
<td>GFM [12]</td>
<td>EMPCA-MI $P=1$, $r=1$</td>
<td>GFM [12]</td>
<td>EMPCA-MI $P=1$, $r=1$</td>
</tr>
<tr>
<td>1</td>
<td>0, 0, 0.002, 0</td>
<td>0, 0, 0.001, 0</td>
<td>0, 0, 0.001, 0</td>
<td>0, 0, 0.001, 0</td>
</tr>
<tr>
<td>2</td>
<td>0, 0, 0.005, 0</td>
<td>0, 0, 0.003, 0</td>
<td>0, 0, 0.001, 0</td>
<td>0, 0, 0.001, 0</td>
</tr>
<tr>
<td>3</td>
<td>0, 0, 0.02, 0</td>
<td>0, 0, 0.06, 0</td>
<td>0, 0, 0.02, 0</td>
<td>0, 0, 0.16, 0</td>
</tr>
<tr>
<td>4</td>
<td>0, 0, 0.00, 0</td>
<td>0, 0, 0.00, 0</td>
<td>0, 0, 0.09, 0</td>
<td>0, 0, 0.08, 0</td>
</tr>
<tr>
<td>5</td>
<td>0, 0, 0.10, 0</td>
<td>0, 0, 0.08, 0</td>
<td>0, 0, 0.17, 0</td>
<td>0, 0, 0.08, 0</td>
</tr>
<tr>
<td>6</td>
<td>0, 0, 0.00, 0</td>
<td>0, 0, 0.00, 0</td>
<td>0, 0, 0.09, 0</td>
<td>0, 0, 0.08, 0</td>
</tr>
<tr>
<td>7</td>
<td>0, 0, 0.05, 0</td>
<td>0, 0, 0.04, 0</td>
<td>0, 0, 0.02, 0</td>
<td>0, 0, 0.01, 0</td>
</tr>
<tr>
<td>8</td>
<td>0, 0, 0.11, 0</td>
<td>0, 0, 0.08, 0</td>
<td>0, 0, 0.12, 0</td>
<td>0, 0, 0.11, 0</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>0, 0, 0.03, 0</td>
<td>-</td>
<td>0, 0, 0.15, 0</td>
</tr>
<tr>
<td>10</td>
<td>0, 0, 0.02, 0</td>
<td>0, 0, 0.01, 0</td>
<td>0, 0, 0.04, 0</td>
<td>0, 0, 0.03, 0</td>
</tr>
</tbody>
</table>

This paper has applied a new similarity measure using Expectation Maximization for Principal Component Analysis with Mutual Information (EMPCA-MI) to successfully achieve retinal image registration (RIR). Retinal images are especially challenging because of their inherent characteristics of low contrast, non-uniform illumination and large number of homogeneous regions. Quantitative results for monomodal clinical datasets confirm EMPCA-MI consistently out-performed the geometric features method in terms of RIR robustness, while concomitantly providing analogous registration accuracy.
6. REFERENCES


