Robust Image Registration using Adaptive Expectation Maximisation based PCA

Parminder Singh Reel†, Laurence S. Dooley†, K. C. P. Wong†, Anko Börner‡
†Department of Computing and Communications, The Open University, Milton Keynes, United Kingdom
‡Optical Sensor Systems, German Aerospace Center (DLR), Berlin, Germany

Email: †{parminder.reel, laurence.dooley, k.c.p.wong}@open.ac.uk, ‡anko.boerner@dlr.de

Abstract—Images having either the same or different modalities can be aligned using the systematic process of image registration. Inherent image characteristics including intensity non-uniformities in magnetic resonance images and large homogeneous non-vascular regions in retinal and other generic image types however, pose a significant challenge to their registration. This paper presents an adaptive expectation maximisation for principal component analysis with mutual information (aEMPCA-MI) similarity measure for image registration. It introduces a novel iterative process to adaptively select the most significant principal components using Kaiser rule and applies 4-pixel connectivity for feature extraction together with Wichard’s bin size selection in calculating the MI. Both quantitative and qualitative results on a diverse range of image datasets, conclusively demonstrate the superior image registration performance of aEMPCA-MI compared with existing MI-based similarity measures.

Index Terms—Principal component analysis.

I. INTRODUCTION

Registration of mono and multimodal images is a fundamental processing task for many computer vision and imaging based applications [1], such as in registering human brain and retinal images to assist in disease diagnosis and treatment planning. Brain and retinal image registration (IR) is particularly demanding due to the presence of inherent artefacts such as intensity non-uniformities (INU) in magnetic resonance images (MRI) [2] or the latent large homogeneous non-vascular regions [3] in retinal images. These collectively make the process of physically aligning a reference and sensed (source) image to maximise some predefined similarity measure (SM) [1] very challenging.

Various SM for IR have been proposed [1] which can be broadly categorized according to whether they are based on cross correlation, Fourier techniques and mutual information (MI) [4]. MI exploits the statistical relationship between the sensed and reference images and has been proven to be computationally efficient and well-established in the medical imaging domain [4]. It is however, very sensitive to interpolation artefacts in the presence of INU and can be seriously compromised when overlapping areas are small. In contrast, regional MI (RMI) [5] uses neighbourhood region features for MI by segmenting an image into several regions for feature extraction, so reducing the influence of INU upon IR quality. Its drawback however, is that it employs a covariance matrix for entropy approximation and as region expands, so commensurately does the computational complexity [5].

Recently, the expectation maximisation for principal component analysis with MI (EMPCA-MI) algorithm [6] has been shown to considerably lower the computational cost without loss of IR performance for dissimilar imaging modalities of the human anatomy [7]. It achieves dimensionality reduction by iteratively computing the first principal component instead of solving the complete covariance matrix as in conventional principal component analysis (PCA) [8]. It has subsequently been refined as modified EMPCA-MI (mEMPCA-MI) [9] which uses 8-pixel and 4-pixel connectivity for rearranging the neighbourhood data values to preserve both the spatial and intensity information, thus leading to superior IR performance. This gave the impetus to formalise the EMPCA-MI SM such that, it is able to adaptively determine the most appropriate number of principal components to represent prominent features in a particular dataset.

Various approaches to selecting the best subset of significant principal components have been proposed including Scree Graph, Broken-stick and their variants [8], which choose the subset by computing the cumulative variance of all components. In contrast, the Kaiser Rule [8] retains only those principal components whose eigenvalues are greater than one. It considers the variables in the image data to be independent and so their principal components are the same as variables, having unity variance in the case of the correlation matrix. This important attribute of the Kaiser Rule is image dependent [8], so it can automatically determine the number of significant principal components required to represent image features for a prescribed accuracy. Furthermore, it is well suited to EMPCA-MI since it iteratively determines the principal components in descending order of eigenvalues. These two distinct characteristics provided the motivation to incorporate the Kaiser Rule into the EMPCA-MI framework.

This paper presents an adaptive EMPCA-MI (aEMPCA-MI) SM which employs 4-pixel connectivity neighbourhood region for pre-processing in combination with a Kaiser Rule based, EMPCA computation of significant principal com-
components for both the reference and sensed images. These respective components are then used to calculate the MI. aEMPCA-MI enables better representation of image features compared to the conventional EMPCA-MI [6] which only used the first dominant principal component. Numerical and qualitative results confirm aEMPCA-MI provides lower IR errors in comparison with both EMPCA-MI and other SM for a variety of test image datasets.

The remainder of the paper is as follows: Section II describes the principles of IR and introduces the new aEMPCA-MI algorithm. Section III describes the experimental setup and test image datasets used together with a critical results analysis. Finally, Section IV provides some concluding comments.

II. ADAPTIVE EMPCA-MI BASED IMAGE REGISTRATION

A. IR Principles

IR is a multistage process which seeks to align a sensed image \( I_S \) with a reference image \( I_R \) [1] and generally involves the following key steps: i) transforming the coordinates of \( I_S \) with known transformation parameters in a given reference space; ii) generating a new interpolated source image \( I'_S \) in the reference space; iii) comparing \( I'_S \) with \( I_R \) using a SM; and iv) optimising the transformation parameters sequentially, to attain the best possible alignment, where the SM value is a maximum between the two images. The IR process for the proposed aEMPCA-MI is visualised in Fig.1(a) and is detailed in the following section.

B. aEMPCA-MI Similarity Measure

The new SM comprises three constituent steps (see Fig.1(a)). It first pre-processes the neighbourhood information for both \( I_S \) and \( I_R \), before adaptively computing the principal components, and finally calculating the MI.

1) Step I — Pre-processing: This derives a second-order representation of the image grayscale values by using 4-pixel connectivity neighbourhood as in [9]. This process helps to retain the local spatial neighbourhood relationship rather than pixel values as in [6], which leads to distinctive repetitive patterns. To illustrate this, consider the example of \( I_R \) in Fig.1(b), for a sliding window of radius \( r \) with \( c \)-pixel connectivity in neighbourhood \( B \) (with \( r=1 \) and \( c=4 \)). The resulting ith element, where \( 1 \leq i \leq c+1 \) of single dimension vector \( B^* \) of length \( c+1 \), can be expressed as:

\[
B^*_i = \begin{cases} 
B_i - B_{i+1} & \text{if } i \leq \frac{c}{2} + 1 \text{ and } i \geq \frac{c}{2} + 1 \\
B_i & \text{else}
\end{cases}
\]  

(1)

Each column vector \( B^* \) now represents the differential value of \( c \) connected pixels with respect to the centre pixel \( B_{i+1} \). These vectors are concatenated as \( Q_R \) with dimensions \( d \times N \) where \( d \) represents the dimensional space i.e., \( d=c+1 \) in the example in Fig.1(b) and \( N \) is the number of pixels scanned (excluding the boundary pixels). Similar preprocessing step are performed on \( I'_S \) to generate \( Q_S \).

2) Step II — Compute aEMPCA: This dimensionality reduction step for \( Q_R \) and \( Q_S \) uses Kaiser Rule [8] to adaptively find the significant principal components using EMPCA [6]. The flowchart in Fig.1(a) shows how the aEMPCA algorithm adaptively computes and retains the first \( p \) principal components from \( Q_R \) having eigenvalue \( \lambda_p \geq 1 \) (similarly compute \( p' \) components for \( Q_S \)). These components are then concatenated as \( X_R \) and \( X_S \) respectively. The order of computational complexity for aEMPCA is \( O(Np^2) \) which given that \( N \gg p, r \) is the same as that incurred by EMPCA i.e. \( O(N) \), though due to the iterative nature of the algorithm its overall time overhead will tend to be slightly higher. Note that when \( p=1 \), EMPCA [6] becomes a special case of aEMPCA.

3) Step III — Compute MI: This final step determines the MI [4] between the highest number of common principal components in both \( X_R \) and \( X_S \) of the reference and sensed

---

(a) Registration Framework for the aEMPCA-MI algorithm. (b) Illustration of the pre-processing Step I with 4-pixel region connectivity for an image pair size of \( 10 \times 10 \) pixels.
images. The Wichard approach [10] has been used to select the best bin size in computing the individual and joint probabilities, since it exploits the kurtosis measure of the data distribution rather than applying the traditional fixed bin-size approach for accurate MI computation [11].

### III. EXPERIMENTAL EVALUATION

#### A. Experimental Setup and Datasets

A series of IR experiments were undertaken to analyse the comparative performance of the aEMPICA-MI on a range of diverse image datasets. Multimodal MRI T1 and T2 datasets from the BrainWeb database [12] were chosen due to their challenging INU and noise artefacts with the corresponding parameter details being defined in Table I. To reflect the broad range of IR applications and evaluate the robustness of aEMPICA-MI, the well-known Lena (L) and Baboon (Bb) images were used in combination with an INU function [6], and additive Gaussian noise. In addition, a clinical monomodal retinal colour Fundus (F) IR was performed followed by multimodal F and Scanning Laser Ophthalmoscope (SLO) IR. These retinal images are inherently characterised by non-uniform illumination and large homogeneous non-vascular regions which render them especially challenging for IR, though neither INU nor noise was added to the retinal datasets because they are not prominent in the acquisition process [3]. Various IR experiments using the aforementioned datasets were conducted and classified into six Scenarios as shown in Table II. The transformation model consisted of \((X, Y, \theta, S)\), where \(X\) and \(Y\) are the respective translations along the \(x\) and \(y\) axis, \(\theta\) is the rotational motion and \(S\) is a scaling factor.

In Scenarios 1 — 4, \(S=1\) so \(\Delta S=0\), while in Scenarios 5 and 6, since these involve retinal images, the magnification changes can result from using different equipment or the motion in the direction of the optical axis which can vary the value of \(S\). To establish the requisite ground truth, all images were mis-registered by a known transformation, with the original images then being considered as \(I_S\). The \((X, Y, \theta, S)\) settings where the aEMPICA-MI is a maximum are then considered the best image alignment.

To critically analyse the IR performance, the percentage registration error (RE) [1] between the initial and final value for each parameter was determined. In order to equitably reflect the computational time incurred for each SM, the corresponding average run time (ART) was calculated for every IR iteration. The RE performance of aEMPICA-MI was evaluated against four existing MI-based SM for six Scenarios, with the results displayed in Table II, where all parameter settings are kept constant for each SM, except the adaptive number of principal components in aEMPICA-MI. The IR process used partial volume interpolation along with Powell’s optimization method [1]. All experiments were performed on Ubuntu 10.04 with 2.93 GHz Intel Core and 3GB RAM, and the assorted algorithms were all implemented in MATLAB.

#### B. Result Discussion

To clarify the nomenclature adopted in Table II; \(T1 + \alpha_{20}\) for example, represents a MRI T1 image slice with 20% INU, while \(Bb + Z + \beta\) is the Baboon image with INU and Gaussian noise artefacts. For aEMPICA-MI, the corresponding \(p\) value is displayed alongside the IR errors and signifies the number of adaptively determined principal components. Fig.2(a) displays the ART for three Scenario groups and while the ART is resource dependent, it is still an insightful time cost metric. The RE results confirm that the new aEMPICA-MI algorithm provides superior IR in all six Scenarios, including the challenging Scenario 6. For example, a significantly lower RE is observed for aEMPICA-MI in Scenario 2 for \(T1 + \alpha_{40} + \beta\) with T2 with a percentage RE of \((8.72, 2.95, 0.51, 0.00)\) contrasting with both EMPICA-MI [6] \((9.70, 4.30, 0.62, 0.00)\)

### Table I

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Resolution (pixels)</th>
<th>INU</th>
<th>Noise(β)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI T1 (T1)</td>
<td>[181 × 217 × 181]</td>
<td>α20=20% INU</td>
<td>Gaussian</td>
</tr>
<tr>
<td>MRI T2 (T2)</td>
<td>[181 × 217 × 181]</td>
<td>α40=40% INU</td>
<td>(μ=0.01,</td>
</tr>
<tr>
<td>Lena (L)</td>
<td>[256 × 256]</td>
<td>(Z_{(x,y)}=\frac{1}{3.2} (x+y))</td>
<td>(σ^2=0.01)</td>
</tr>
<tr>
<td>Baboon (Bb)</td>
<td>[256 × 256]</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Fundus (F)</td>
<td>[3888 × 2592]</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SLO (SLO)</td>
<td>[768 × 768]</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### Table II

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(T1 + \alpha_{20})</td>
<td>T1</td>
<td>9, 6.0, 0.5, 0.00</td>
<td>5.20, 4.00, 0.44, 0.00</td>
<td>2.00, 1.30, 0.36, 0.00</td>
<td>1.12, 0.93, 0.21, 0.00</td>
<td>1.02, 0.90, 0.18, 0.00</td>
</tr>
<tr>
<td>2</td>
<td>(T1 + \alpha_{40})</td>
<td>T1</td>
<td>10, 8.3, 0.7, 0.00</td>
<td>6.20, 5.30, 0.50, 0.00</td>
<td>4.50, 4.00, 0.42, 0.00</td>
<td>2.96, 3.04, 0.32, 0.00</td>
<td>2.25, 2.89, 0.27, 0.00</td>
</tr>
<tr>
<td>3</td>
<td>(T1 + \alpha_{40} + \beta)</td>
<td>T2</td>
<td>13, 16, 0.9, 0.00</td>
<td>10.5, 13.3, 0.60, 0.00</td>
<td>8.00, 10.0, 0.58, 0.00</td>
<td>7.45, 9.28, 0.46, 0.00</td>
<td>7.19, 9.01, 0.42, 0.00</td>
</tr>
<tr>
<td>4</td>
<td>(T1 + \alpha_{40} + \beta)</td>
<td>(L + Z)</td>
<td>12, 9.6, 0.7, 0.00</td>
<td>7.00, 7.00, 0.48, 0.00</td>
<td>2.60, 2.60, 0.42, 0.00</td>
<td>1.93, 1.71, 0.32, 0.00</td>
<td>1.80, 1.50, 0.29, 0.00</td>
</tr>
<tr>
<td>5</td>
<td>(Bb + Z + \beta)</td>
<td>(L + Z)</td>
<td>0.4, 0.5, 0.2, 0.00</td>
<td>0.20, 0.38, 0.21, 0.00</td>
<td>0.20, 0.32, 0.21, 0.00</td>
<td>0.16, 0.24, 0.17, 0.00</td>
<td>0.15, 0.23, 0.15, 0.00</td>
</tr>
<tr>
<td>6</td>
<td>(F)</td>
<td>(F)</td>
<td>12, 3.5, 0.2, 0.00</td>
<td>0.60, 1.06, 0.21, 0.00</td>
<td>0.45, 0.70, 0.21, 0.00</td>
<td>0.30, 0.60, 0.16, 0.00</td>
<td>0.28, 0.59, 0.14, 0.00</td>
</tr>
<tr>
<td>7</td>
<td>(F)</td>
<td>(SLO)</td>
<td>2.1, 1.9, 0.8, 0.3</td>
<td>0.32, 0.41, 0.49, 0.42</td>
<td>0.29, 0.37, 0.41, 0.39</td>
<td>0.25, 0.34, 0.38, 0.37</td>
<td>0.22, 0.31, 0.36, 0.34</td>
</tr>
</tbody>
</table>
and $m$EMPCA-MI [9] (8.99, 3.11, 0.56, 0.00). This is due to $a$EMPCA-MI using 4-pixel neighbourhood and $p=4$.

Conversely, both EMPCA-MI and $m$EMPCA-MI use just a single principal component, which can in certain circumstances limit the IR accuracy. This performance improvement for $a$EMPCA-MI however, is counterbalanced by a slower ART of 174ms due to its iterative nature as compared with 152ms and 95ms for EMPCA-MI and $m$EMPCA-MI respectively. Similar RE trends are evident in the Lena, Baboon and retinal datasets with $a$EMPCA-MI again consistently outperforming the comparators, due to more principal components being used in the MI computation, which ultimately leads to enhanced IR.

Particularly noteworthy is the IR performance of $a$EMPCA-MI in the challenging Scenarios 5 and 6 where $p=2$ is adaptively chosen due to the large homogeneous non-vascular regions in the Fundus and SLO images, providing the lowest RE in both Scenarios. Despite reducing the data dimensionality by using 4-pixel connectivity the performance offset for this improved IR is the higher ART of 320ms because of the high resolution of the Fundus images. A similar observation on the computational cost is apparent in other Scenarios, as an inevitable outcome of the adaptive process inherent in the new $a$EMPCA-MI.

Fig. 2(b) displays qualititative IR results using the checkerboard overlaying [1] technique for Fundus and Lena images. The perceptual IR results for the $F$ with SLO IR in Scenario 6 can be evaluated by checking the continuity of vessel structures (colour enhanced for higher visibility) and similarly with the positioning of the eyes for IR between $L+Z+\beta$ with $L$ in Scenario 3 (with the ground truth in red). These results palpably validate that the new $a$EMPCA-MI consistently provides superior IR and lower RE, while recognizing the computational cost trade-off required to achieve the improved IR robustness to INU, noise and large homogeneous features.

**IV. CONCLUSION**

This paper has presented an adaptive dimension to the existing Expectation Maximisation for Principal Component Analysis with MI (EMPCA-MI) similarity measure. Enhanced IR performance for clinical and generic images exhibiting a range of challenging features including INU, Gaussian noise and large homogeneous regions confirms the efficacy of the $a$EMPCA-MI algorithm. This is achieved by adaptively incorporating more significant principal components with a corresponding trade-off in terms of computational overheads.

**REFERENCES**