

Total correlation-based groupwise image registration for quantitative MRI

Jean-Marie Guyader¹, Wyke Huizinga¹, Valerio Fortunati¹, Dirk H. Poot^{1,2}, Matthijs van Kranenburg³, Jifke F. Veenland¹, Margarethus M. Paulides⁴, Wiro J. Niessen^{1,2}, Stefan Klein¹

¹Biomedical Imaging Group Rotterdam, Departments of Radiology and Medical Informatics, Erasmus MC, Rotterdam, the Netherlands

²Imaging Science and Technology, Faculty of Applied Sciences, TU Delft, the Netherlands

³Departments of Radiology and Cardiology, Erasmus MC, Rotterdam, the Netherlands

⁴Hyperthermia Unit, Department of Radiation Oncology, Erasmus MC Cancer Institute, Rotterdam, the Netherlands

Abstract

In quantitative magnetic resonance imaging (qMRI), quantitative tissue properties can be estimated by fitting a signal model to the voxel intensities of a series of images acquired with different settings. To obtain reliable quantitative measures, it is necessary that the qMRI images are spatially aligned so that a given voxel corresponds in all images to the same anatomical location. The objective of the present study is to describe and evaluate a novel automatic groupwise registration technique using a dissimilarity metric based on an approximated form of total correlation. The proposed registration method is applied to five qMRI datasets of various anatomical locations, and the obtained registration performances are compared to those of a conventional pairwise registration based on mutual information. The results show that groupwise total correlation yields better registration performances than pairwise mutual information. This study also establishes that the formulation of approximated total correlation is quite analogous to two other groupwise metrics based on principal component analysis (PCA). Registration performances of total correlation and these two PCA-based techniques are therefore compared. The results show that total correlation yields performances that are analogous to those of the PCA-based techniques. However, compared to these PCA-based metrics, total correlation has two main advantages. Firstly, it is directly derived from a multivariate form of mutual information, while the PCA-based metrics were obtained empirically. Secondly, total correlation has the advantage of requiring no user-defined parameter.

1. Introduction

Based on a series of images obtained with different acquisition settings, quantitative magnetic resonance imaging (qMRI) allows the computation of quantitative imaging fea-

tures that characterize tissue properties. One condition is essential to ensure that the computed features are reliable: the images from which they are derived should be spatially aligned. In practice, there are multiple possible causes of misalignment of the acquired qMRI images, such as the motion of the subject during the acquisition or geometric distortions caused by the acquisition. In these cases, automatic image registration can be used to compensate for misalignments within qMRI datasets.

Pairwise image registration based on mutual information is a technique that is commonly used for aligning images characterized by different contrasts [26, 17]. It consists of aligning a moving image to a fixed reference image. When more than two images have to be aligned, this pairwise paradigm has two main drawbacks. Firstly, the choice of fixed reference image may impact registration accuracy [8]. To overcome this first drawback, a solution based on multiple pairwise registrations has been proposed [21]. The technique consists of applying pairwise registration between all the possible pairs of images, and then of combining the obtained transformations to align all the images into a mean space. Using mutual information as dissimilarity metric, such a method was used in different studies [21, 9]. Secondly, pairwise registration does not allow to register all images within a single optimization procedure, even when applying the method presented in [21].

The aim of groupwise registration techniques is to account for all image information within a single optimization procedure, and thereby to simultaneously register multiple images. In this paper, we aim to derive such a groupwise registration technique based on the concept of mutual information. Though the formulation of mutual information for two images is unique, several multivariate versions have been proposed for its generalization to two or more images. In [16], a metric called interaction information was proposed that expresses the amount of information shared by all images. Total correlation, a metric expressing the amount of information shared between any subset of im-

ages among the images to register, was proposed in [29], and will be adopted in this work. Two main reasons guided our choice towards total correlation rather than interaction information. Firstly, total correlation is theoretically better adapted than interaction information for the registration of multiple images. This is explained by the fact that interaction information has the undesirable property that it equals zero as soon as one image does not share information with all the other images, whereas total correlation becomes trivial only when all images share no information [7, 24]. Secondly, total correlation can easily be approximated, for computational purposes. This is of particular interest because calculating total correlation requires computing the joint entropy of the images to register. However, increasing the number of images may complicate the computation of the joint entropy due to a phenomenon called curse of dimensionality [4]. To solve that issue, a simplified formula for joint entropy was proposed, based on the condition that the image intensities are jointly normally distributed [1]. In this paper, we propose to incorporate this approximation within the formula of total correlation proposed by [29]. This approximated form of total correlation was implemented within an existing parametric registration framework.

Registration performance was evaluated on five real qMRI datasets. Registration results obtained with the proposed approximated groupwise total correlation metric were compared with results of other metrics, including pairwise registration based on mutual information, and on groupwise registration using the PCA-based metrics of [10].

2. Method

2.1. Mutual information

Let us consider M_g , a series of G images that we want to register, with $g \in \{1 \dots G\}$. Each image M_g (N voxels per image) can be represented as one column of a $N \times G$ matrix \mathbf{M} containing the intensities of all images. A row of \mathbf{M} can be considered as a data point in a G -dimensional space. To quantify how well the G images are aligned, a dissimilarity metric has to be defined.

When $G = 2$, the negated mutual information (\mathcal{D}_{MI}) has been shown to be a robust dissimilarity metric for image registration [17]. For two images M_1 and M_2 , the negated mutual information can be written as:

$$\mathcal{D}_{\text{MI}}(M_1, M_2) = H(M_1, M_2) - H(M_1) - H(M_2) \quad (1)$$

with $H(M_1)$ the entropy [22] of image M_1 , $H(M_2)$ the entropy of image M_2 , and $H(M_1, M_2)$ the joint entropy of M_1 and M_2 .

2.2. Total correlation

For cases with $G \geq 2$ images, two main multivariate generalizations of mutual information have been proposed [16, 29]. The first is known as interaction information [16], denoted \mathcal{D}_{Inf} , and measures the amount of information shared by all the images. For G images $\{M_1, \dots, M_G\}$, the negated interaction information is written as follows:

$$\mathcal{D}_{\text{Inf}}(M_1, \dots, M_G) = - \sum_{T \subseteq \{M_1, \dots, M_G\}} (-1)^{G-|T|} H(T) \quad (2)$$

with T any subset of images among $\{M_1, \dots, M_G\}$, $|T|$ the number of images in the corresponding subset and $H(T)$ the joint entropy of subset T . Interaction information quantifies the amount of information that all images participate in. This means that if at least one of the images M_1, \dots, M_G shares no information with all other images, then the interaction information will be zero as shown by [7, 3].

The second form, called total correlation [29], measures the amount of information shared between any subset of the images $\{M_1, \dots, M_G\}$. The negated total correlation is written as follows:

$$\mathcal{D}_{\text{TC}}(M_1, \dots, M_G) = H(M_1, \dots, M_G) - \left[\sum_{g=1}^G H(M_g) \right] \quad (3)$$

with $H(M_1, \dots, M_G)$ the joint entropy of images M_1, \dots, M_G . Total correlation is able to quantify the amount of shared information between all possible combinations of the images, while interaction information only quantifies the amount of information shared by all the images [7, 24]. In the context of image registration, total correlation is therefore more flexible than interaction information. We therefore did not consider interaction information in our study.

As Equation (3) shows, computing total correlation implies evaluating the joint entropy $H(M_1, \dots, M_G)$. Doing so typically requires to build a sparsely filled G -dimensional joint histogram, which is computationally challenging because it is subject to the curse of dimensionality [4]. We therefore propose to approximate the expression of the joint entropy $H(M_1, \dots, M_G)$ by a simplified form. Under the assumption that the intensities of the images M_1, \dots, M_G are jointly normally distributed, Ali Ahmed et al. [1] showed that the expression of the joint entropy becomes:

$$H(M_1, \dots, M_G) = \frac{G}{2} + \frac{G}{2} \ln(2\pi) + \frac{1}{2} \ln(\det(\mathbf{C})) \quad (4)$$

with $\det(\cdot)$ the determinant operator, and \mathbf{C} the $G \times G$ matrix of covariances between the images M_g . To make the registration method robust to linear intensity scalings and

offsets, we choose to incorporate an intensity standardization (i.e. z-score) within the definition of the metric. This implies that the marginal entropies $H(M_g)$ are constant for all $g = 1..G$, and that the covariance matrix \mathbf{C} is equal to the correlation matrix \mathbf{K} :

$$\mathbf{K} = \frac{1}{N-1} \mathbf{\Sigma}^{-1} (\mathbf{M} - \overline{\mathbf{M}})^T (\mathbf{M} - \overline{\mathbf{M}}) \mathbf{\Sigma}^{-1} \quad (5)$$

where $\mathbf{\Sigma}$ is a diagonal matrix containing the standard deviations of the columns of \mathbf{M} as its diagonal elements, and $\overline{\mathbf{M}}$ is a matrix with in each column the column-wise average of \mathbf{M} . By combining Equations (3) and (5), the expression \mathcal{D}_{TC} of total correlation becomes:

$$\mathcal{D}_{\text{TC}}(M_1, \dots, M_G) = \frac{1}{2} \ln(\det(\mathbf{K})) = \frac{1}{2} \sum_{j=1}^G \ln \lambda_j \quad (6)$$

using $\det(\mathbf{K}) = \prod_{j=1}^G \lambda_j$, with λ_j the j^{th} eigenvalue of \mathbf{K} , and $\lambda_j > \lambda_{j+1}$.

2.3. Groupwise registration framework

In our groupwise registration framework, the images M_g are simultaneously brought to a mid-point space by means of a transformation $\mathbf{T}(\boldsymbol{\mu})$, where $\boldsymbol{\mu}$ is a vector containing the transformations $\mathbf{T}_g(\boldsymbol{\mu}_g)$ related to each image M_g , and $\boldsymbol{\mu}_g$ their corresponding parameters. In the groupwise scheme, the metric \mathcal{D} quantifies the dissimilarity between all transformed images $M_g(\mathbf{T}_g(\boldsymbol{\mu}_g))$. Groupwise registration can therefore be formulated as the constrained minimization of the dissimilarity metric \mathcal{D} with respect to $\boldsymbol{\mu}$:

$$\hat{\boldsymbol{\mu}} = \arg \min_{\boldsymbol{\mu}} \mathcal{D}(M_1(\mathbf{T}_1(\boldsymbol{\mu}_1)), \dots, M_G(\mathbf{T}_G(\boldsymbol{\mu}_G))) \quad (7)$$

subject to:

$$\sum_{g=1}^G \boldsymbol{\mu}_g = \mathbf{0} \quad (8)$$

where the constraint of Equation (8) serves to define a mid-point space [2].

2.4. Relation with existing work

Huizinga et al. [10] previously proposed groupwise metrics for the groupwise registration of qMRI datasets. They assumed that the images M_1, \dots, M_G of a qMRI dataset follow a low-dimensional model m_g , such that the value of the image M_g at position \mathbf{x} equals:

$$M_g(\mathbf{x}) = m_g(\boldsymbol{\theta}(\mathbf{x})) + \epsilon(\mathbf{x}) \quad (9)$$

with $\boldsymbol{\theta}$ a vector of dimension $\Gamma < G$ (hence the model is called low-dimensional), containing the parameters of the

low-dimensional model, and ϵ the noise at coordinate \mathbf{x} . An example of such low-dimensional models is the monoexponential model [9] used with diffusion-weighted MR images (DW-MRI):

$$m_g(\boldsymbol{\theta}) = S_0 \exp(-b_g \times \mathbf{u}_g^T \mathbf{D} \mathbf{u}_g) \quad (10)$$

with $\boldsymbol{\theta} = (S_0, D_{11}, D_{22}, D_{33})$, \mathbf{u}_g the direction vector of the diffusion gradient, \mathbf{D} a 3×3 symmetric diffusion tensor, and b the so-called b-value. The ADC is given by $\text{tr}(\mathbf{D})/3$.

Huizinga's metrics are based on principal component analysis (PCA) and rely on the idea that an aligned set of qMRI images can be described by a small number of eigenvalues, since the underlying model m_g is low-dimensional. A misaligned set of images would, on the contrary, be characterized by a flatter eigenvalue spectrum. The first metric proposed in [10], denoted \mathcal{D}_{PCA} , measures the difference between the sum of all eigenvalues and the sum of the first few eigenvalues:

$$\mathcal{D}_{\text{PCA}}(M_1, \dots, M_G) = \sum_{j=1}^G \lambda_j - \sum_{j=1}^L \lambda_j = \sum_{j=L+1}^G \lambda_j \quad (11)$$

with L a user-defined constant with $1 \leq L \leq G$, and $\sum_{j=1}^G \lambda_j = \text{tr}(\mathbf{K}) = G$. This means that \mathcal{D}_{PCA} is the sum of the last (i.e. with lower values) $G - L$ eigenvalues.

Contrary to \mathcal{D}_{PCA} , the second metric, denoted $\mathcal{D}_{\text{PCA2}}$, does not require the selection of an arbitrary cut-off L . It consists of weighting the last eigenvalues more than the first ones:

$$\mathcal{D}_{\text{PCA2}}(M_1, \dots, M_G) = \sum_{j=1}^G j \times \lambda_j \quad (12)$$

The expressions of \mathcal{D}_{TC} , \mathcal{D}_{PCA} and $\mathcal{D}_{\text{PCA2}}$, respectively given in Equations (6), (11) and (12), are quite similar: all of them consists of a sum of weighted eigenvalues. The main theoretical disadvantage of Huizinga's \mathcal{D}_{PCA} is that it requires to choose the cut-off L . In $\mathcal{D}_{\text{PCA2}}$, this user-defined constant is avoided, but the weights j in Equation (12) are actually still chosen arbitrarily. In principle, one could propose any metric of the following form:

$$\mathcal{D}_{\text{PCA3}}(M_1, \dots, M_G) = \sum_{j=1}^G f(j) \times \lambda_j \quad (13)$$

with f a monotonically increasing function. The main theoretical advantage of the total correlation metric \mathcal{D}_{TC} that we propose is that the contribution of each eigenvalue follows naturally from the derivation of mutual information: the influence of each eigenvalue is automatically calibrated.

3. Datasets and experiments

Huizinga et al. [10] evaluated their PCA-based group-wise registration techniques \mathcal{D}_{PCA} and \mathcal{D}_{PCA2} on five qMRI studies. The aim of the experiments is to use the proposed total correlation metric \mathcal{D}_{TC} to register these same datasets, and subsequently compare the obtained results with these of [10].

3.1. Description of the five datasets

The first qMRI study, denoted T1MOLLI-HEART [10] consists of nine T_1 -weighted image datasets of porcine hearts with transmural myocardial infection of the lateral wall. For each of the nine datasets, $G = 11$ two-dimensional images were acquired. For each registration case, a voxelwise curve fitting was applied to the registered images, producing quantitative T_1 maps.

The second study, denoted T1VFA-CAROTID [5], involves MR images of the carotid arteries. For each of the eight patients, $G = 5$ three-dimensional images were acquired. For each patient, the acquired images were registered and fitted to obtain quantitative T_1 maps.

The third study consists of diffusion-weighted MR images (DW-MRIs) of the abdominal region, and is denoted ADC-ABDOMEN [9]. Five datasets, each of them including $G = 19$ three-dimensional images, were registered and fitted to produce ADC maps.

The fourth qMRI study is denoted DTI-BRAIN [15, 6, 28, 25, 19] and consists, for each of the 5 considered datasets, of registering diffusion tensor images (DTI) of the brain. The number of images to register varied between $G = 33$ and $G = 70$ for each dataset (see [10]). In this fourth study, the fitted parameter is the mean diffusivity (MD).

The fifth qMRI study involves DCE images of the abdomen. Five DCE-ABDOMEN [11] datasets were acquired, each of them containing $G = 160$ three-dimensional images. The fitted parameter of interest considered in this study is K^{trans} .

The full descriptions of the fitting models are provided in [10].

3.2. Registration characteristics

The four dissimilarity metrics (\mathcal{D}_{MI} , \mathcal{D}_{PCA} , \mathcal{D}_{PCA2} and \mathcal{D}_{TC}) were implemented in the elastix toolbox [13]. The adaptive stochastic gradient descent (ASGD) proposed by [12] was used as optimisation method for image registration. For all registrations, we used two resolutions, 1000 iterations per resolution, and 2048 random coordinate samples per resolution. For comparison purposes, we performed all the registrations using the sets of parameters reported in [10]. In particular, when applying \mathcal{D}_{PCA} , the value of L was 3 for T1MOLLI-HEART, 1 for

T1VFA-CAROTID, 4 for ADC-ABDOMEN, 7 for DTI-BRAIN, and 4 for DCE-ABDOMEN. For the DTI-BRAIN dataset, we used an affine transformation model. Similar to Wachinger and Navab [27], we used an exponential mapping of the affine matrix for parametrization. For all the other datasets, non-rigid transformation models in which deformations are modelled by cubic B-splines were chosen [20]. For each dataset, [10] reported results for various B-spline grid spacings. In this study, we compare results only for the intermediate values of the spacings, i.e. 64 mm for T1MOLLI-HEART, 16 mm for T1VFA-CAROTID, 64 mm for ADC-ABDOMEN and 64 mm for DCE-ABDOMEN. An affine transformation model was used for the DTI-BRAIN dataset.

3.3. Evaluation measures

No ground truth alignment was available for any of the five datasets we considered. Nevertheless, registration performance was evaluated based on four different measures. These measures are described in [10], which is why they are only succinctly described in this section.

3.3.1 Landmark correspondence and overlap of volumes of interest

Landmarks were manually defined on images of the T1VFA-CAROTID and DCE-ABDOMEN datasets. The correspondence between the corresponding landmarks was evaluated by computing a mean target registration error (mTRE).

In the T1MOLLI-HEART case, segmentations of the myocardium were outlined on between 6 and 9 images per patient. In the ADC-ABDOMEN case, the spleen was manually delineated on 8 images. For these two cases, the overlap between the segmented structures was then evaluated using a Dice coefficient.

For the DTI-BRAIN study, neither landmarks nor structures could be reliably identified on the diffusion weighted images, which is why no overlap or point correspondence was calculated.

3.3.2 Smoothness of the transformation

Extreme and non-smooth deformations are unexpected for the experiments we conducted. The smoothness of the deformation field can therefore be used to identify such undesirable transformations. A quantification of smoothness can be obtained by computing the standard deviation of the determinant of $\partial \mathbf{T}_g / \partial \mathbf{x}$ over all \mathbf{x} for all images: $\text{STD}_{\det(\partial \mathbf{T}_g / \partial \mathbf{x})}$. This smoothness quantifier was computed for all datasets except DTI-BRAIN (an affine transformation was used in that last case, which is why smoothness was not computed).

	T1MOLLI-HEART	T1VFA-CAROTID	ADC-ABDOMEN	DTI-BRAIN	DCE-ABDOMEN
	<i>Dice</i> [%]	<i>mTRE</i> [mm]	<i>Dice</i> [%]	-	<i>mTRE</i> [mm]
Initial	48 ± 8	1.47 ± 0.54	70 ± 4	-	8.49 ± 4.54
\mathcal{D}_{MI}	37 ± 11	1.22 ± 0.43	64 ± 16	-	6.46 ± 2.32
\mathcal{D}_{PCA}	53 ± 7	1.11 ± 0.42	71 ± 5	-	6.11 ± 2.33
\mathcal{D}_{PCA2}	52 ± 11	1.08 ± 0.39	75 ± 5	-	5.99 ± 2.18
\mathcal{D}_{TC}	53 ± 11	1.09 ± 0.40	74 ± 5	-	6.18 ± 2.40

Table 1. Dice coefficients or mTRE values (mean value ± standard deviation)

	T1MOLLI-HEART	T1VFA-CAROTID	ADC-ABDOMEN	DTI-BRAIN	DCE-ABDOMEN
Initial	0 ± 0	0 ± 0	0 ± 0	-	0 ± 0
\mathcal{D}_{MI}	7 ± 2	2 ± 0	8 ± 3	-	4 ± 2
\mathcal{D}_{PCA}	2 ± 1	2 ± 1	3 ± 2	-	4 ± 2
\mathcal{D}_{PCA2}	1 ± 1	1 ± 0	3 ± 1	-	2 ± 1
\mathcal{D}_{TC}	2 ± 1	1 ± 0	5 ± 2	-	4 ± 2

Table 2. Transformation smoothness: $STD_{\det(\partial T_g / \partial \mathbf{x})}$ [%] (mean value ± standard deviation)

	T1MOLLI-HEART	T1VFA-CAROTID	ADC-ABDOMEN	DTI-BRAIN	DCE-ABDOMEN
	T_1 [ms]	T_1 [ms]	ADC [$\mu\text{m}^2/\text{ms}$]	MD [$\mu\text{m}^2/\text{ms}$]	K^{trans} [min^{-1}]
Initial	92 ± 19	> 1000	1.37 ± 0.83	0.096 ± 0.029	2.84 ± 2.30
\mathcal{D}_{MI}	97 ± 16	501 ± 83	0.25 ± 0.05	0.084 ± 0.028	3.64 ± 4.13
\mathcal{D}_{PCA}	87 ± 16	498 ± 93	0.23 ± 0.06	0.085 ± 0.029	1.52 ± 1.18
\mathcal{D}_{PCA2}	83 ± 12	510 ± 110	0.27 ± 0.05	0.084 ± 0.028	1.27 ± 0.92
\mathcal{D}_{TC}	77 ± 13	500 ± 96	0.32 ± 0.05	0.085 ± 0.029	1.87 ± 1.79

Table 3. Uncertainty estimation: $90^{\text{th}}\sqrt{\text{CRLB}}$ of the fitted parameters (mean value ± standard deviation)

3.3.3 Uncertainty estimation of the qMRI fit

For all datasets, curve fittings were performed to respectively generate T_1 , T_1 , ADC , MD and K^{trans} quantitative maps. The qMRI models were fitted using a maximum likelihood (ML) estimator that takes into account the Rician characteristic of the noise in MR data. We used the fitting same method as [10], based on the work of [18]. The uncertainty of these fitted qMRI model parameters can be quantified by the 90^{th} percentile of the square root of Cramér-Rao lower bound (CRLB), which provides a lower bound for the variance of the maximum likelihood parameters. This uncertainty estimate is denoted $90^{\text{th}}\sqrt{\text{CRLB}}$.

3.4. Joint normality experiment

As mentioned in the Method section, the presented total correlation metric assumes that the image intensities are jointly normally distributed. However, this condition is not necessarily fulfilled. The aim of this experiment is to check whether the condition of joint normality is, in practice, required for obtaining good alignment results.

The joint normality of two images can be easily assessed by computing and visualizing their joint histogram. In the present case, however, joint normality has to be studied on

more images (see section 3.1), which requires more sophisticated methods. A possible graphical approach to analyze multivariate joint normality is based on quantile-quantile (Q-Q) plots [23]. In such plots, the quantiles of the Mahalanobis distances to the mean obtained for each voxel location are plotted against the quantiles of a χ^2 distribution with degree of freedom G . If the data is jointly normally distributed, the plotted points should follow the line $y = x$. The implementation we used is based on the R package MVN [14].

4. Results

4.1. Registration performance

Registration performances in terms of landmark correspondence (mTRE) or overlap of volumes of interest (Dice coefficient) are given in Table 1. For all dataset, better alignments (i.e. lower mTRE) or overlaps (i.e. higher Dice coefficients) were obtained with groupwise total correlation \mathcal{D}_{TC} than with pairwise mutual information \mathcal{D}_{MI} . Table 2 provides values of the transformation smoothness $STD_{\det(\partial T_g / \partial \mathbf{x})}$. In all cases, \mathcal{D}_{TC} yields lower (i.e. better) values of $STD_{\det(\partial T_g / \partial \mathbf{x})}$ than \mathcal{D}_{MI} . Table 3 provides

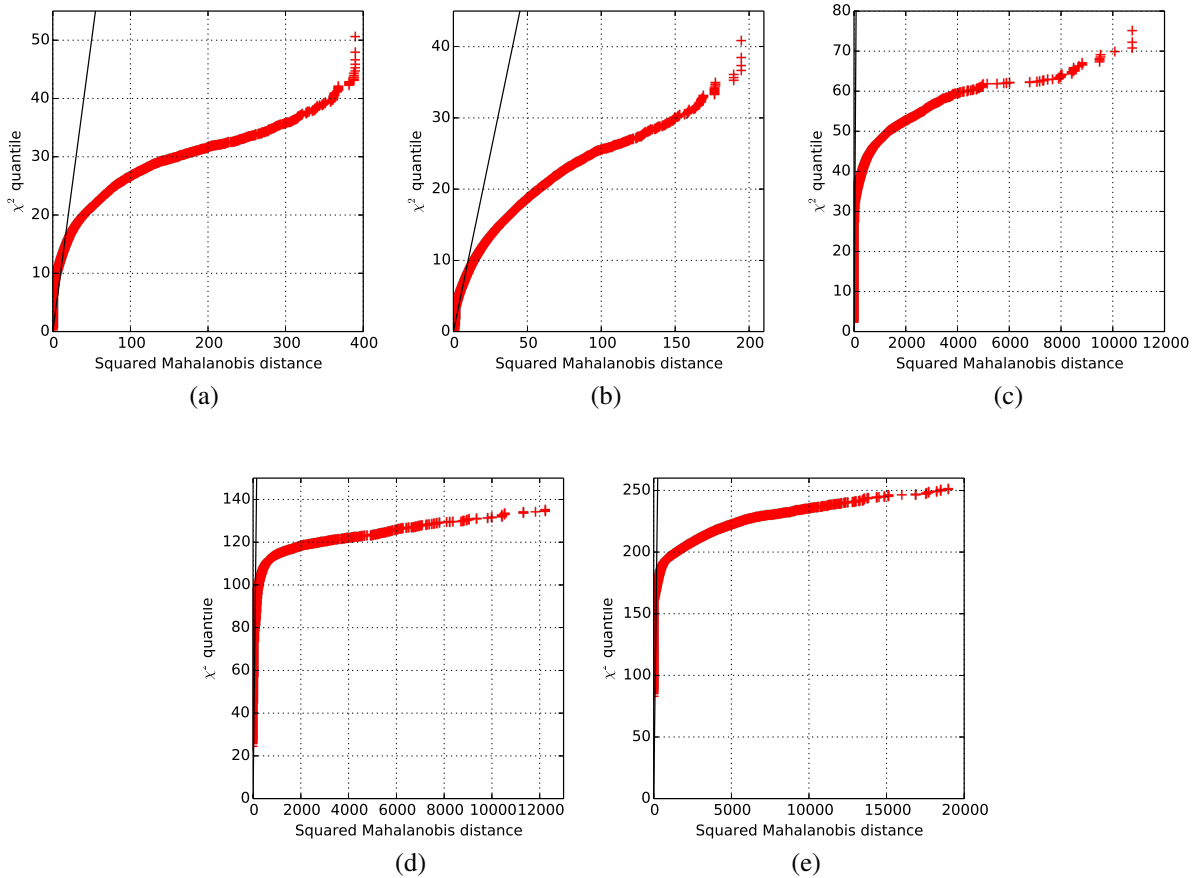


Figure 1. Q-Q plots: χ^2 distribution (degree of freedom = G) versus squared Mahalanobis distance. (a) TIMOLLI-HEART; (b) T1VFA-CAROTID; (c) ADC-ABDOMEN; (d) DTI-BRAIN; (e) DCE-ABDOMEN

estimations of the goodness of fit ($90^{\text{th}}\sqrt{\text{CRLB}}$) for each dataset. The results indicate that the values of $90^{\text{th}}\sqrt{\text{CRLB}}$ are lower (i.e. better) with \mathcal{D}_{TC} than with \mathcal{D}_{MI} for the TIMOLLI-HEART and DCE-ABDOMEN datasets, while they are quite similar for T1VFA-CAROTID and DTI-BRAIN, and higher (i.e. worse) for the ADC-ABDOMEN dataset.

For all experiments, the registration results obtained with our groupwise total correlation metric \mathcal{D}_{TC} are quite similar to the results obtained with the groupwise methods \mathcal{D}_{PCA} and $\mathcal{D}_{\text{PCA2}}$, previously proposed in [10].

4.2. Results of the joint normality experiment

For each of the five qMRI datasets, the quantiles of a χ^2 distribution of degree of freedom G are plotted against the quantiles of all Mahalanobis distances. The obtained Q-Q plots are shown in Figure 1. As mentioned in the Method section, if the points of the obtained Q-Q plot follow a joint distribution that is normal, then they should follow the line $y = x$. This is not the case for any of the five datasets that

are considered in this study. The joint intensity distributions in the considered datasets can therefore not be considered as multivariate normal distributions.

5. Discussion and conclusion

Results obtained on five quantitative MRI datasets show that the proposed method based on approximated total correlation yields better results than pairwise mutual information, and comparable results to two PCA-based methods of proposed by Huizinga et al. [10].

The total correlation method we describe in this study, \mathcal{D}_{TC} , is based on the assumption of joint normality of the image intensities. The results indicate that even though this condition of joint normality is not fulfilled in practice, approximated total correlation yields better registration results than the conventional pairwise mutual information method.

Furthermore, the results indicate that total correlation provides registration performances that are similar to the PCA-based metrics of Huizinga et al. However, compared to \mathcal{D}_{PCA} and $\mathcal{D}_{\text{PCA2}}$, \mathcal{D}_{TC} has two main advantages.

Firstly, it is directly derived from a multivariate form of mutual information, while the PCA-based metrics were obtained empirically. Secondly, total correlation has the advantage of requiring no user-defined parameter.

6. Future work

The total correlation method we described in this study is based on the concept of mutual information. Since mutual information is commonly used for the registration of pairs of multimodal images, a natural follow-up study will be to apply groupwise total correlation to multimodal data.

Acknowledgements

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking (<http://www.imi.europa.eu>) under grant agreement nr. 115151 (QuIC-ConCePT project), resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) (EU-FP7) and EFPIA companies in kind contribution. Funding was also provided by EU-FP7 under grant agreement no. 601055, VPH-DARE@IT. The authors would also like to thank:

- H. M. M. van Beusekom, M. van Kranenburg, R. J. M. van Geuns and A. Uitterdijk for providing the TIMOLLI-HEART data. The acquisition of the TIMOLLI-HEART data was financially supported by Agentschap NL (SENER-NOVEM): "A novel approach to myocardial regeneration" to H.M.M. van Beusekom et al. under grant nr. ISO43050;
- A. Leemans for providing the DTI-BRAIN data;
- R. Klaassen, B. F. Coolen and A. J. Nederveen for providing the DCE-ABDOMEN data;
- B. F. Coolen and A. J. Nederveen for providing the T1VFA-CAROTID data;
- N. M. deSouza, L. Bernardin and N. Douglas, Institute of Cancer Research, London, UK, for providing the ADC-ABDOMEN data. The ADC-ABDOMEN data were acquired in the context of the QuIC-ConCePT project.

References

- [1] N. Ali Ahmed and D. V. Gokhale. Entropy expressions and their estimators for multivariate distributions. *IEEE Trans Inf Theory*, 35(3):688–692, 1989.
- [2] S. Balci, P. Golland, M. Shenton, and M. Wells. Free-form B-spline deformation model for groupwise registration. In *MICCAI Workshop*, pages 23–30, Brisbane, Australia, 2007.
- [3] A. J. Bell. Co-information lattice. In *4th International Symposium on Independent Component Analysis and Blind Source Separation*, pages 921–926, 2003.
- [4] R. Bellman. *Adaptive control processes: a guided tour*. Princeton University Press, New Jersey, 1961.
- [5] B. F. Coolen, D. H. J. Poot, M. I. Liem, L. P. Smits, S. Gao, G. Kotek, S. Klein, and A. J. Nederveen. Three-dimensional quantitative T1 and T2 mapping of the carotid artery: Sequence design and in vivo feasibility. *Magn Reson Med*, 75(3):1–10, 2015.
- [6] N. De Geeter, G. Crevecoeur, L. Dupré, W. Van Hecke, and A. Leemans. A DTI-based model for TMS using the independent impedance method with frequency-dependent tissue parameters. *Phys Med Biol*, 57(8):2169–2188, 2012.
- [7] D. J. Galas, N. A. Sakhnenko, A. Skupin, and T. Ignac. Describing the complexity of systems: multivariable set complexity and the information basis of systems biology. *J Comput Biol*, 21(2):118–140, 2014.
- [8] X. Geng, G. E. Christensen, H. Gu, T. J. Ross, and Y. Yang. Implicit reference-based group-wise image registration and its application to structural and functional MRI. *NeuroImage*, 47(4):1341–1351, 2009.
- [9] J.-M. Guyader, L. Bernardin, N. H. Douglas, D. H. Poot, W. J. Niessen, and S. Klein. Influence of image registration on apparent diffusion coefficient images computed from free-breathing diffusion MR images of the abdomen. *J Magn Reson Im*, 42(2):315–330, 2015.
- [10] W. Huizinga, D. H. J. Poot, J. Guyader, R. Klaassen, B. F. Coolen, and M. V. Kranenburg. PCA-based groupwise image registration for quantitative MRI. *Med Image Anal*, 29:65–78, 2016.
- [11] R. Klaassen, O. Gurney-Champion, E. ter Voert, A. Heerschap, M. Bijlsma, M. Besselink, G. van Tienhoven, J. Nio, P. C., J. Wilmink, H. van Laarhoven, and A. Nederveen. Motion correction of high temporal 3T dynamic contrast enhanced MRI of pancreatic cancer - preliminary results. In *Proceedings of the 22st Annual Meeting International Society for Magnetic Resonance in Medicine*, page 3667, 2014.
- [12] S. Klein, J. Pluim, M. Staring, and M. Viergever. Adaptive stochastic gradient descent optimisation for image registration. *Int J Comput Vision*, 81(3):227–239, 2009.
- [13] S. Klein, M. Staring, K. Murphy, M. Viergever, and J. Pluim. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imag*, 29(1):196–205, 2010.
- [14] S. Korkmaz, D. Goksuluk, and G. Zararsiz. MVN: An R package for assessing multivariate normality. *The R Journal*, 6(December):151–162, 2014.
- [15] A. Leemans, J. Sijbers, S. De Backer, E. Vandervliet, and P. Parizel. Multiscale white matter fiber tract

- coregistration: A new feature-based approach to align diffusion tensor data. *Magn Reson Med*, 55(6):1414–1423, 2006.
- [16] W. J. McGill. Multivariate information transmission. *Psychometrika*, 19(2):317–325, 1954.
- [17] J. Pluim, J. Maintz, and M. Viergever. Mutual information based registration of medical images: a survey. *IEEE Trans Med Imag*, 22(8):1–21, 2003.
- [18] D. H. J. Poot and S. Klein. Detecting statistically significant differences in quantitative MRI experiments, applied to diffusion tensor imaging. *IEEE Trans Med Imag*, 34(5):1164–1176, 2015.
- [19] Y. D. Reijmer, A. Leemans, S. M. Heringa, I. Wielaard, B. Jeurissen, H. L. Koek, and G. J. Biesels. Improved sensitivity to cerebral white matter abnormalities in Alzheimer’s disease with spherical deconvolution based tractography. *PLoS ONE*, 7(8):1–8, 2012.
- [20] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. Nonrigid Registration Using Free-Form Deformations : Application to Breast MR Images. *IEEE Trans Med Imag*, 18(8):712–721, 1999.
- [21] D. Seghers, E. D. Agostino, F. Maes, and D. Vandermeulen. Construction of a brain template from MR images using state-of-the-art registration and segmentation techniques. In *MICCAI*, pages 696–703, Rennes, Brittany, 2004.
- [22] C. E. Shannon. A mathematical theory of communication. *Bell Syst Tech J*, 27:379–423, 1948.
- [23] N. H. Timm. *Applied multivariate analysis*. Springer, New-York, 2002.
- [24] N. Timme, W. Alford, B. Flecker, and J. M. Beggs. Synergy, redundancy, and multivariate information measures: an experimentalist’s perspective. *J Comput Neurosci*, 36:119–140, 2014.
- [25] N. van der Aa, A. Leemans, F. J. Northington, H. L. Van Straaten, I. C. Van Haastert, F. Groenendaal, M. J. N. L. Benders, and L. S. De Vries. Does diffusion tensor imaging-based tractography at 3 months of age contribute to the prediction of motor outcome after perinatal arterial ischemic stroke? *Stroke*, 42(12):3410–3414, 2011.
- [26] P. Viola and I. Wells, W.M. Alignment by maximization of mutual information. *Proceedings of IEEE International Conference on Computer Vision*, pages 16–23, 1995.
- [27] C. Wachinger and N. Navab. Simultaneous registration of multiple images: similarity metrics and efficient optimization. *IEEE Trans Pattern Anal*, 35(5):1221–1233, 2013.
- [28] H.-C. Wang, J.-L. Hsu, and A. Leemans. Diffusion tensor imaging of vascular parkinsonism: structural changes in cerebral white matter and the association with clinical severity. *Arch Neurol*, 69(10):1340–1348, 2012.
- [29] S. Watanabe. Information theoretical analysis of multivariate correlation. *IBM J Res Dev*, 4(1):66–82, 1960.