Revisiting Overlap Invariance in Medical Image Alignment
Nathan D. Cahill\textsuperscript{1,2}, Julia A. Schnabel\textsuperscript{1}, J. Alison Noble\textsuperscript{1}, and David J. Hawkes\textsuperscript{3}
\textsuperscript{1}Wolfson Medical Vision Laboratory, University of Oxford, Oxford OX1 3PJ, UK
\textsuperscript{2}Research & Innovation, Carestream Health, Inc., Rochester, NY 14608, USA
\textsuperscript{3}Centre for Medical Image Computing, University College London, London WC1E 6BT, UK
cahill@robots.ox.ac.uk

Abstract
In [8], Studholme et al. introduced normalized mutual information (NMI) as an overlap invariant generalization of mutual information (MI). Even though Studholme showed how NMI could be used effectively in multimodal medical image alignment, the overlap invariance was only established empirically on a few simple examples. In this paper, we illustrate a simple example in which NMI fails to be invariant to changes in overlap size, as do other standard similarity measures including MI, cross correlation (CCorr), correlation coefficient (CCoeff), correlation ratio (CR), and entropy correlation coefficient (ECC). We then derive modified forms of all of these similarity measures that are proven to be invariant to changes in overlap size. This is done by making certain assumptions about background statistics. Experiments on multimodal rigid registration of brain images\textsuperscript{1} show that 1) most of the modified similarity measures outperform their standard forms, and 2) the modified version of MI exhibits superior performance over any of the other similarity measures for both CT/MR and PET/MR registration.

1. Introduction

Soon after mutual information (MI) was introduced for use in multimodal image registration [2, 9], Studholme et al. [8] showed that MI suffers from the problem of overlap sensitivity. This problem arises over the course of registration because the size of the overlapping region between two images changes. For even the simple example of two images of a constant intensity disc centred in a black background, MI values fluctuate when one image is rotated about the centre of the disc and compared to the other. This contradicts the intuition that a useful image similarity measure should remain invariant to changes in the size of background regions.

In order to combat this problem, Studholme et al. proposed the normalized mutual information (NMI), which was shown empirically to be insensitive to changes in overlap size. NMI is closely related to Astola’s entropy correlation coefficient (ECC) [1], which was explored by Collignon [2] and Maes [5] for use in multimodal image registration.

Around the same time that NMI was proposed, another powerful similarity measure was introduced by Roche [7]. Roche showed that the correlation ratio (CR) is useful in multimodal image registration, and furthermore, it can be computed more efficiently than MI because it does not require the construction of a joint histogram. CR assumes only that some functional relationship exists between image code values, but it makes no assumption as to the form of that function. CR is a generalization of the correlation coefficient (CCoeff), which assumes a linear relationship between image code values. CCoeff and its non-normalized cousin cross correlation (CCorr) are commonly used in intramodal image registration [3, 4, 6].

In this paper, we show that CCorr, CCoeff and CR suffer from the same problem that Studholme recognized in MI: sensitivity to changes in overlap size. Furthermore, we illustrate a simple variant of the example shown in [8], where NMI (and ECC) fail to be overlap invariant. We then present modifications of CCorr, CCoeff, CR, MI, NMI, and ECC that are proven to be insensitive to changes in overlap size under certain assumptions on background statistics. Finally, we illustrate the performance of the modified similarity measures versus the standard similarity measures on a variety of rigid registration tasks.

2. Overlap Sensitivity of Standard Similarity Measures

There are many ways to gauge the degree of similarity between two images. In intramodal registration, the CCorr
and CCoeff describe the strength of the linear relationship between two images. If the images are thought of as random variables $A$ and $B$, the CCorr is simply the covariance $\text{Cov}(A, B)$, and the CCoeff is given by:

$$\rho_{A,B} = \frac{\text{Cov}(A, B)}{\sqrt{\text{Var}(A) \text{Var}(B)}} ,$$

(1)

where $\text{Var}(\bullet)$ indicates variance. Note that in image registration it is actually the square of these quantities that is maximized. The squared CCorr and squared CCoeff are used in order to be insensitive to the direction of the linear relationship between the images. Note that the contrast invariant distance defined in [6] is equivalent to $1 - \rho_{A,B}^2$.

Roche’s correlation ratio [7] is a generalization of the (squared) CCoeff, and it describes the strength of the functional relationship between two random variables. The CR is given by:

$$\eta_{B|A} = \frac{\text{Var}(E[B|A])}{\text{Var}(B)} ,$$

(2)

where $E[B|A]$ is the conditional expectation of $B$ given $A$.

Mutual information, NMI, and Astola’s ECC describe the strength of the statistical relationship between images. MI, NMI, and Astola’s ECC are given by:

$$\text{MI}(A, B) = H(A) + H(B) - H(A, B) ,$$

(3)

$$\text{NMI}(A, B) = \frac{H(A) + H(B)}{H(A, B)} ,$$

(4)

$$\text{ECC}(A, B) = \sqrt{2 - \frac{2H(A, B)}{H(A) + H(B)}} ,$$

(5)

where $H(\bullet)$ indicates entropy and $H(\bullet, \bullet)$ indicates joint entropy. In [5], it is actually a shifted version of the square of (5) that is maximized.

In order to illustrate the overlap sensitivity problem, consider the simple example shown in Fig. 1. The reference image (left) is of a disc centered on a dark background. The floating image (middle) is of a ring centered on a dark background, with the outer portion of the ring having the same diameter as the disc. Intuition would tell us that as the floating image rotates about the centre of the ring, any useful image similarity measure should remain constant.

In Fig. 1, if we assign the dark background a value of zero, and the lighter interior of the disc and ring a value of one, we can analytically compute values of the various similarity measures as functions of rotation angle $\theta$ and inner ring radius. For square images with side length 3, disc of radius 1, and ring with inner radius $r \in [0, 1]$, the squared CCorr, squared CCoeff, CR, MI, NMI, and squared ECC values are shown as functions of $\theta$ in Fig. 2.

For the case where $r = 0$, this example is the same as the disc example shown in [8], and the squared CCoeff, CR, NMI, and squared ECC values are all constant functions of $\theta$. However, for $r \in (0, 1)$, it is clear that none of the standard similarity measures are constant in $\theta$.

### 3. Image Statistics and Changing Overlap

In the example shown in Fig. 1, it would seem appropriate that any similarity measure that is used to compare the images be invariant to the rotation angle $\theta$. However, none of the standard similarity measures satisfy this notion of invariance, except for the special cases when $r = 0$ or $r = 1$. This is due to the fact that as the floating image rotates, the size of the overlapping region changes. As the overlap reduces in size, the proportion of background to foreground pixels shrinks, skewing the similarity measures.

In order to modify similarity measures to account for changes in overlap size, we first consider the relationship between image statistics and changing overlap. Let us define $\Omega_0$ to be the original overlap region between images $A$ and $B$. We consider that the overlap region changes as a function of time; i.e., $\Omega_t$ is the overlap region at time $t$, and we define $\omega_t$ to be the ratio of the change in overlap size to the original overlap size:

$$\omega_t = \frac{|\Omega_t| - |\Omega_0|}{|\Omega_0|} .$$

(6)

The relationship between $\Omega_0$ and $\Omega_t$ can be seen in the Venn diagram shown in Fig. 3.

Let us further define $\Omega_t' = \Omega_0 \cap \Omega_t^C$ to be the region lost and $\Omega_t = \Omega_t^C \cap \Omega_t$ to be the region gained. Now, we make the crucial assumption that $\Omega_t'$ and $\Omega_t$ share common statistics that are independent of $t$. This assumption holds, for example, when $\Omega_t'$ and $\Omega_t$ cover only background regions in both images.

In order to describe the statistics of images over a region, we introduce the following notation for the probability mass function (p.m.f.) and joint p.m.f.: $p_{\Phi}(x)$ denotes the probability that image $I$, when observed over region $\Phi$, takes on the value $x$, and $p_{I_1, I_2}(x, y)$ denotes the joint probability that images $I_1$ and $I_2$, when observed over region $\Phi$, take on the values $x$ and $y$, respectively.
The background p.m.f. of $\text{E}[B|A]$ is denoted by:

$$
\tilde{p}^\text{E}[B|A](x) := \frac{\text{E}[B|A](x)}{\tilde{p}_{\Omega_t}^A(x)} = \frac{p_{\Omega_t}^B(x)}{p_{\Omega_t}^A(x)}.
$$

(10)

For convenience, various first and second order statistics will be denoted by:

$$
\tilde{\mu}_A := \sum x \tilde{p}^A(x),
$$

(11)

$$
\tilde{\mu}_B := \sum x \tilde{p}^B(x),
$$

(12)

$$
\tilde{\sigma}^2_A := \sum (x - \tilde{\mu}_A)^2 \tilde{p}^A(x),
$$

(13)

$$
\tilde{\sigma}^2_B := \sum (x - \tilde{\mu}_B)^2 \tilde{p}^B(x),
$$

(14)

$$
\tilde{\gamma}^2_{A,B} := \sum (x - \tilde{\mu}_A)(y - \tilde{\mu}_B)p^A,B(x,y),
$$

(15)

$$
\tilde{\gamma}^2_{B|A} := \sum (x - \tilde{\mu}_B)^2 \text{E}[B|A](x).
$$

(16)

We will refer collectively to (7)–(16) as *background statistics*.

Under the common background statistics assumption, we can relate p.m.f.’s measured over $\Omega_t$ to those measured over $\Omega_0$. For example, using conditioning arguments, we can
The second order statistics are related by:

\[ p_{\Omega_0}^A(x) = \frac{\Omega_0^A}{\Omega_0} p_{\Omega_1}^A(x) + \frac{\Omega_0 \cap \Omega_1}{\Omega_0} p_{\Omega_0 \cap \Omega_1}^A(x) \]  

and

\[ p_{\Omega_0}^A(x) = \frac{\Omega_1^A}{\Omega_1} p_{\Omega_1}^A(x) + \frac{\Omega_0 \cap \Omega_1}{\Omega_1} p_{\Omega_0 \cap \Omega_1}^A(x) . \]

A small amount of algebraic manipulation, along with the substitution of (6) and (7) into (17) and (18), yields:

\[ p_{\Omega_1}^A(x) = (1 + \omega_t) p_{\Omega_1}^B(x) - \omega_t \bar{p}_{\Omega_1}^B(x) . \]

Using similar conditioning arguments, the following relationships between the other p.m.f.'s over \( \Omega_0 \) and \( \Omega_1 \) can be established:

\[ p_{\Omega_1}^B(x) = (1 + \omega_t) p_{\Omega_1}^B(x) - \omega_t \bar{p}_{\Omega_1}^B(x) , \]

\[ p_{\Omega_0}^{A,B}(x, y) = (1 + \omega_t) p_{\Omega_0}^{A,B}(x, y) - \omega_t \bar{p}_{\Omega_0}^{A,B}(x, y) \]

\[ E_{\Omega_1}[A] = (1 + \omega_t) E_{\Omega_1}[A] - \omega_t \bar{E}_{\Omega_1}[A] , \]

\[ E_{\Omega_0}[B] = (1 + \omega_t) E_{\Omega_0}[B] - \omega_t \bar{E}_{\Omega_0}[B] . \]

The second order statistics are related by:

\[ \text{Var}_{\Omega_0}(A) = (1 + \omega_t) \text{Var}_{\Omega_1}(A) - \omega_t \bar{\sigma}_A^2 \]

\[ -\omega_t (1 + \omega_t) (E_{\Omega_1}[A] - \bar{\mu}_A)^2 , \]

\[ \text{Var}_{\Omega_0}(B) = (1 + \omega_t) \text{Var}_{\Omega_1}(B) - \omega_t \bar{\sigma}_B^2 \]

\[ -\omega_t (1 + \omega_t) (E_{\Omega_1}[B] - \bar{\mu}_B)^2 , \]

\[ \text{Var}_{\Omega_0}(E[B|A]) = (1 + \omega_t) \text{Var}_{\Omega_1}(E[B|A]) - \omega_t \bar{\gamma}_{B|A}^2 \]

\[ -\omega_t (1 + \omega_t) (E_{\Omega_1}[B|A] - \bar{\mu}_{B|A})^2 , \]

\[ \text{Cov}_{\Omega_0}(A, B) = (1 + \omega_t) \text{Cov}_{\Omega_1}(A, B) - \omega_t \bar{\gamma}_{A,B} \]

\[ -\omega_t (1 + \omega_t) (E_{\Omega_1}[A] - \bar{\mu}_A) \cdot (E_{\Omega_1}[B] - \bar{\mu}_B) . \]

What is important to note in (19)–(28) is that the left hand side of each equation is independent of \( t \). This means that as \( t \) increases and the overlap region changes, the right hand side of each of these equations remains constant. It is this property that we will exploit in the next section to form overlap invariant versions of the similarity measures.

4. Modified Similarity Measures

If images \( A \) and \( B \) remain in a fixed position while the overlap region is allowed to change, the right hand sides of (19)–(28) remain constant as \( t \) increases. We can use this knowledge to construct new image similarity measure that are invariant to changes in overlap size.

First, we define the modified cross correlation (MCCorr) according to (28):

\[ \text{MCCorr}(A, B) = \frac{(1 + \omega) \text{Cov}(A, B) - \omega \bar{\gamma}_{A,B}}{-\omega (1 + \omega) (E[A] - \bar{\mu}_A) \cdot (E[B] - \bar{\mu}_B) .} \]

Note that we have dropped the subscript \( t \) from \( \omega \), and have dropped the subscripts from the expected values and covariance terms. This is done to simplify notation; it should be noted that \( E[A], E[B] \), and \( \text{Cov}(A, B) \) are to be computed over the current overlap region, and that an original overlap region must be assumed in order to compute \( \omega \).

A modified version of CCoeff can be constructed in a similar way. The modified correlation coefficient (MCCoeff) is given by:

\[ \text{MCCoeff}(A, B) = \frac{\text{MCCorr}(A, B)}{\sqrt{\text{MVar}(A) \sqrt{\text{MVar}(B)}} , \]

where

\[ \text{MVar}(I) = (1 + \omega) \text{Var}(I) - \omega \bar{\gamma}_I^2 \]

\[ -\omega (1 + \omega) (E[I] - \bar{\mu}_I)^2 . \]

As is the case with CCorr and CCoeff, the squares of the MCCorr and MCCoeff are the pertinent quantities to maximize for the purposes of image registration.

The correlation ratio can also be modified to be overlap invariant. The modified correlation ratio (MCR) is given by:

\[ \text{MCR}(B|A) = \frac{\text{MVar}(E[B|A])}{\text{MVar}(B)} , \]

where

\[ \text{MVar}(E[B|A]) = (1 + \omega) \text{Var}(E[B|A]) - \omega \bar{\gamma}_{B|A}^2 \]

\[ -\omega (1 + \omega) (E[B|A] - \bar{\mu}_{B|A})^2 . \]

To construct modified versions of MI, NMI, and ECC, we need to define modified entropy and modified joint entropy using the p.m.f.'s in (19)–(21):

\[ \text{MH}(I) = -\sum q^I(x) \log q^I(x) \]

and

\[ \text{MH}(A, B) = -\sum \sum q^{A,B}(x, y) \log q^{A,B}(x, y) . \]
where
\[ q^i(x) = (1 + \omega) p^i(x) - \omega \hat{p}^i(x) \] (36)
and
\[ q^{A,B}(x, y) = (1 + \omega) p^{A,B}(x, y) - \omega \hat{p}^{A,B}(x, y). \] (37)

The modified mutual information (MMI), modified normalized mutual information (MNMI), and modified entropy correlation coefficient (MECC) are then given by:

\[
\text{MMI}(A, B) = \text{MH}(A) + \text{MH}(B) - \text{MH}(A, B),
\]
\[
\text{MNMI}(A, B) = \frac{\text{MH}(A) + \text{MH}(B)}{\text{MH}(A, B)}, \quad \text{and}
\]
\[
\text{MECC}(A, B) = \sqrt{2 - \frac{2\text{MH}(A, B)}{\text{MH}(A) + \text{MH}(B)}}. \] (40)

As can be seen in Fig. 4, each of these modified similarity measures is constant with respect to changing overlap size when applied to the disc/ring example of Fig. 1.

5. Registration Experiment

In order to illustrate the behaviour of the modified similarity measures, we focus on the rigid registration case, and we use images from the Retrospective Image Registration Evaluation project. The RIRE project database contains CT, MR, and PET images for a variety of patients, and has a sequestered set of ground truth rigid body transformations that were computed from fiducial markers implanted in the skull. (The fiducial markers were removed from the images prior to retrospectively evaluating registration algorithms.) Results of the original RIRE study are given in [10].

In this paper, we used the images from nine patient datasets. Each patient dataset contains MR images from some or all of the following protocols: T1-weighted, T2-weighted, PD-weighted, and rectified versions of the T1, T2, and PD-weighted images. Five of the nine datasets contain both CT and PET images in addition to the MR images. Two of the datasets contain CT but not PET images, and the remaining two datasets contain PET but not CT images. The CT images have resolution 0.65 \times 0.65 \times 4.0 \text{ mm}^3, the MR images have approximate resolution 1.25 \times 1.25 \times 4.0 \text{ mm}^3, and the PET images has resolution 2.59 \times 2.59 \times 8.0 \text{ mm}^3. For ease of computation, we resampled each image to 6.0 \times 6.0 \times 6.0 \text{ mm}^3 isotropic resolution.

For each dataset, we rigidly registed the CT and/or PET image to all of the MR images. Axial views of the CT, MR-T1, and PET images from patient 5 are shown in Fig. 5. Rigid transformations were parameterized by three Euler angles and three translation parameters. Each of the similarity measures and their modified versions were optimized over the six-dimensional parameter space using the sequential quadratic programming (SQP) method implemented in

\begin{table}[h]
\centering
\begin{tabular}{|c|cc|}
\hline
Image & $\mu$ & $\sigma$ \\
\hline
CT & -998.66 & 4.91 \\
PET & 3.95 & 20.11 \\
MR T1 & 25.80 & 13.77 \\
MR T1-Rectified & 22.49 & 10.31 \\
MR T2 & 27.76 & 14.81 \\
MR T2-Rectified & 23.78 & 10.92 \\
MR PD & 29.66 & 15.80 \\
MR PD-Rectified & 25.54 & 11.74 \\
\hline
\end{tabular}
\caption{Estimated mean and standard deviation of background regions for the images from the patient 5 dataset.}
\end{table}

MATLAB’s Optimization Toolbox. Initial estimates of the solution were selected by translating the images to match their computed centroids.

For the CR, MI, NMI, ECC similarity measures and their modified versions, histograms (and joint histograms) were constructed with 64 (or 64 \times 64) equally spaced bins. Linear (or bilinear) interpolation was used to accumulate partial weights in neighboring bins.

5.1. Estimating Background Statistics

Each of the modified similarity measures requires knowledge of the background statistics of the reference and floating images. Probability mass functions, means, and standard deviations of the background regions were collected by manually selecting background regions from each volume. The estimated background mean and standard deviation values for patient 5 are displayed in Table 1.

Computing estimates of the joint p.m.f., covariance, and variance of the conditional mean of the background regions seems problematic, as it would require registered background regions for believable estimates. However, we will make the assumption for CT/MR and PET/MR registration that the background values of the reference and floating images are independent. This allows us to estimate the joint p.m.f. as the outer product of the background p.m.f.’s of the reference and floating image. Furthermore, this yields trivial values for the covariance and variance of the conditional mean of the background regions.

5.2. Results

The performance of similarity measures on the various registration tasks is measured via Target Registration Error (TRE). For the RIRE project [10], a number of anatomically meaningful volumes of interest (VOI) were annotated. TRE is computed as the average Euclidean distance (in mm) between VOI centroids in the reference image and their predicted positions after registration.

Tables 2 and 3 report the mean, median, and standard deviation of the TRE values measured across the VOI’s in
Figure 4. Values of modified similarity measures as functions of rotation angle $\theta$ and inner ring radius $r$ for the example in Fig. 1. Original overlap is considered to be the case when $\theta = 0$.

Figure 5. Axial slices of CT, MR (T1-weighted) and PET images from the patient 5 dataset.

A number of different observations can be made from these results. First, the entropy-based similarity measures (MI, NMI, ECC, and their modified versions) generally outperform the other similarity measures. This is to be expected because of the multimodal nature of this registration experiment. The utter failure of CR and MCR on the CT/MR registrations was initially baffling, leading us to believe we had an erroneous implementation. However, the improved performance on the PET/MR data (and especially the improvement of MCR over CR on this data) suggests that some other aspect of the CT/MR image combination may be causing trouble with CR and MCR.

Next, the comparison of MCCorr over CCorr is interesting. Even though average TRE errors near 10mm are not great, the MCCorr appears to provide a significant improve-
Perhaps the most interesting observations relate to the performance of MMI versus the other entropy based similarity measures. In CT/MR registration, MMI appears to outperform both NMI and ECC (and their modified versions). In addition, even though MMI has larger mean and median TRE than MI, the extent of outliers (or the number of misregistrations) is smaller. In PET/MR registration, the MMI TRE values appear to hover at the same levels as the NMI and ECC values. Compared with MI for PET/MR registration, however, MMI performs much better when comparing the extent of outliers. MMI does not appear to yield any misregistrations beyond TRE of approximately 10mm.

6. Conclusions and Future Work

In this paper, we showed that standard forms of a variety of similarity measures commonly used in image registration fail to be invariant to changes in overlap size. Using statistics describing the distribution of background intensities, we derived modified forms of the similarity measures that are invariant to changes in overlap size, provided that the common background statistics assumption holds. Publicly available data from the Retrospective Image Registration Evaluation project was used to illustrate performance differences between the standard similarity measures and their modified forms.

A number of areas are planned for future work. First, broader experiments should be performed to further analyze the robustness of image registration with the modified similarity measures. This can be done by expanding the experiment in this paper to test registration from a variety of random starting positions.

Second, the common background statistics assumption needs to be further explored. It may be that in practice, the modified similarity measures presented here may fail when this assumption is violated. If so, it would be prudent to determine where the effectiveness of the modified similarity measures wanes, and to derive alternate modified forms if possible.

Finally, these modified similarity measures should be evaluated in the context of nonrigid registration. Soon, images will be available from the Non-Rigid Image Registration Evaluation Project (NIREP), run by Gary Christensen at the University of Iowa. Much like the RIRE project, NIREP will enable a rigorous evaluation of registration techniques in a nonrigid setting.

References


<table>
<thead>
<tr>
<th>Modality</th>
<th>Statistic</th>
<th>CCorr</th>
<th>CCoeff</th>
<th>CR</th>
<th>MI</th>
<th>NMI</th>
<th>ECC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT/MR</td>
<td>Mean TRE</td>
<td>25.5205</td>
<td>9.4043</td>
<td>30.9385</td>
<td>3.1086</td>
<td>4.7849</td>
<td>4.4211</td>
</tr>
<tr>
<td></td>
<td>Median TRE</td>
<td>22.6913</td>
<td>8.2007</td>
<td>17.9066</td>
<td>2.4232</td>
<td>3.3555</td>
<td>3.2658</td>
</tr>
<tr>
<td></td>
<td>Std. dev. TRE</td>
<td>16.3848</td>
<td>6.4494</td>
<td>26.0851</td>
<td>2.8308</td>
<td>3.9574</td>
<td>3.3922</td>
</tr>
<tr>
<td>PET/MR</td>
<td>Mean TRE</td>
<td>23.1190</td>
<td>9.4246</td>
<td>11.8752</td>
<td>5.5307</td>
<td>4.2041</td>
<td>4.1255</td>
</tr>
<tr>
<td></td>
<td>Median TRE</td>
<td>19.6602</td>
<td>8.0162</td>
<td>7.7618</td>
<td>3.4900</td>
<td>3.5508</td>
<td>3.4714</td>
</tr>
<tr>
<td></td>
<td>Std. dev. TRE</td>
<td>15.7111</td>
<td>6.8273</td>
<td>11.0481</td>
<td>5.9444</td>
<td>2.4900</td>
<td>2.4271</td>
</tr>
</tbody>
</table>

Table 2. Statistics of VOI TRE aggregated across all patients for each of the standard similarity measures.

<table>
<thead>
<tr>
<th>Modality</th>
<th>Statistic</th>
<th>MCCorr</th>
<th>MCCoeff</th>
<th>MCR</th>
<th>MMI</th>
<th>MNMI</th>
<th>MECC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median TRE</td>
<td>9.3198</td>
<td>7.8501</td>
<td>55.4469</td>
<td>2.7208</td>
<td>3.5850</td>
<td>3.4791</td>
</tr>
<tr>
<td></td>
<td>Std. dev. TRE</td>
<td>3.4032</td>
<td>8.0407</td>
<td>29.7331</td>
<td>2.4503</td>
<td>4.0025</td>
<td>3.6606</td>
</tr>
<tr>
<td></td>
<td>Std. dev. TRE</td>
<td>6.6432</td>
<td>7.6677</td>
<td>4.9346</td>
<td>2.4197</td>
<td>2.6271</td>
<td>2.6592</td>
</tr>
</tbody>
</table>

Table 3. Statistics of VOI TRE aggregated across all patients for each of the modified similarity measures.

Figure 6. Box and whisker plots of TRE (in mm) after rigid registration of RIRE data using each similarity measure. Top and bottoms of each box are the quartiles; red line in the middle of each box is the median; whiskers are the extent of the data; red +’s are outliers.