A UNIFIED FEATURE-BASED REGISTRATION METHOD FOR MULTIMODALITY IMAGES

Jie Zhang, Anand Rangarajan

Dept. of CISE, University of Florida, Gainesville, FL

ABSTRACT

While mutual information-based methods have become popular for image registration, the question of what underlying feature to use is rarely discussed. Instead, it is implicitly assumed that intensity is the right feature to be matched. We depart from this tradition by first beginning with a set of feature images—the original intensity image and three directional derivative feature images. This "feature extraction" is performed on both images in a typical intermodality registration setup. Assuming the existence of a training set of registered images, we find the best projection onto a single feature image by maximizing the normalized mutual information (NMI) between the two images w.r.t. the projection weights. After discovering the best feature to match using normalized mutual information as the criterion, we use the same projection coefficients on new test images. We show that affine NMI-based registration of the test images using the new best "feature" is more noise resistant than using image intensity as the default feature. Since the assumption of a registered, training set of images is problematic, we extend the idea to the bootstrap case, wherein we use imperfectly registered images (obtained by using NMI on the original intensity pair) as a training set. The best feature combination is computed using the imperfectly registered pair of images. We show that subsequent NMI-based registration of the best feature image pair is able to improve upon the original imperfect registration. Results are shown on 2D coronal, axial and sagittal slices drawn from a 3D MRI volume of proton density (PD) and T2- weighted images.

1. INTRODUCTION

Why should the vanilla image intensity be used in mutual information-based registration? Our motivation for this paper begins with that question. While mutual information-based methods [1, 2] have been largely successful at alleviating the problem of matching two images drawn from different modalities, the question of why the intensity should be the preferred feature to match is almost never asked. Exceptions are [3] where mutual information and gradient information are combined in registration, [4] where optimal edge information was used with mutual information in registration and [5] where intensities and ICA-derived features are combined in a minimum spanning tree (MST) entropy-based registration algorithm. While these approaches begin with the assumption that the combination of intensities and features is better than intensity alone, they do not demonstrate, as we do, the existence of a single best feature image that outperforms the original image intensity.

We begin with the assumption that a set of filters is available and is capable of extracting useful, usually "edge-based" information from the original intensity images. Given two sets of feature images corresponding to the two input images that we seek to register, we seek the best projection of each set of feature images onto a single feature image. We use normalized mutual information (NMI)[6] as the criterion for determining the best projection of the set of feature images onto a single dimension. NMI is used rather than mutual information (MI)[[7]] because MI favors images with high entropy–higher the marginal entropy of an image, higher the mutual information. Since feature images have lower marginal entropy than the original intensity image, MI is not a good criterion for finding the best projection of the set of feature images.

Our approach can be run in both training set mode and bootstrap mode. In the training set mode, we assume the existence of a set of registered images. Given the set of registered images, we run a set of filters to get two sets of feature images. After we get the two sets of feature images, we determine the best projection onto a single feature dimension by maximizing NMI between the two sets of feature images w.r.t. the projection coefficients. Having determined the projection coefficients, we perform the same set of operations on the test set (where ground truth is unavailable) to get two "best" feature images. Subsequently, we run an NMI-based affine registration algorithm between the two feature images to obtain the best spatial mapping that brings the two feature images into register. The same spatial mapping can be used on the original image volumes as well. In the bootstrap mode, we run NMI on the original images to get an approximate registration. Using this in lieu of the training set, we then repeat the above procedure to find the best feature images and then rerun an affine NMI registration algorithm. This process can be repeated if desired.

2. FEATURE COMBINATION AND REGISTRATION

2.1. Feature Combination

The definitions of MI and NMI of two random variables X and Y are

$$MI(X,Y) = H(X) + H(Y) - H(X,Y) \tag{1}$$

and

$$NMI(X,Y) = \frac{H(X) + H(Y)}{H(X,Y)} \tag{2}$$

where $H(X)$ and $H(Y)$ are the marginal entropies of X and Y and $H(X,Y)$ is the joint entropy of X and Y.

Assume that we have two images $I^{(1)}$and $I^{(2)}$ which are already registered. Let $\{f_k^{(1)}\}$ and $\{f_k^{(2)}\}$,$k = 1, \ldots, K$ be K feature images corresponding to $I^{(1)}$ and $I^{(2)}$. The feature images correspond to a set of filters that are run on both images. There is

We acknowledge support from NSF 0196457 and 0307712

no *a priori* reason to use the same set of filters on both images. As mentioned above, given the two sets of feature images, we determine a set of projection coefficients that map each set of feature images onto a single "best" feature image. Normalized mutual information (NMI) [6] is used as the criterion to find the single, best dimension of feature projection. The objective function used is

$$E(W^{(1)}, W^{(2)}) = NMI((W^{(1)})^T F^{(1)}, (W^{(2)})^T F^{(2)}) \quad (3)$$

where $(W^{(r)})^T$ is the transpose of the column vector $W^{(r)} = \begin{bmatrix} w_1^{(r)} \\ \vdots \\ w_K^{(r)} \end{bmatrix}$ and $F^{(r)} = \begin{bmatrix} f_1^{(r)} \\ \vdots \\ f_K^{(r)} \end{bmatrix}$,$r = 1, 2$. We denote by $F^{(r)}$, the set of feature images of image $I^{(r)}$,$r = 1, 2$. The objective function in (3) is maximized to get the best projection coefficients $(W^{(1)}, W^{(2)})$. By maximizing (3), we obtain the best linear combination of features.

As mentioned in the introduction, we have found that MI is not a good criterion to determine the best feature combination. MI is biased toward images with higher entropy and usually favors the original intensity image or an even noisier image! NMI does not suffer from the same bias.

2.2. Affine image registration with the "best" feature images

Assuming that the best feature combination has been discovered, we then proceed with NMI-based affine registration. Let $T = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ e & f & 1 \end{bmatrix}$ be a 2D affine transformation, where $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ can be decomposed into shear, scale and rotations and $\begin{bmatrix} e & f \end{bmatrix}$ are the x- and y-translations. Assume that images $I^{(1)}$and $I^{(2)}$are two noisy, unregistered images and $F^{(1)}$and $F^{(2)}$ are two sets of feature images. Let $W^{(r)}$,$r \in 1, 2$ be feature combination coefficients achieved by maximizing the objective function (3) on a representative training set. Then we register images $I^{(1)}$ and $I^{(2)}$ by maximizing NMI between the best feature images $(W^{(1)})^T F^{(1)}$ and $(W^{(2)})^T F^{(2)}$. This is a standard affine registration step using NMI as the criterion.

$$T^* = \arg\max_{\mathbf{T}} NMI((W^{(1)})^T F^{(1)}, (W^{(2)})^T F^{(2)}(T)) \quad (4)$$

where $F^{(2)}(T)$ is the set of feature images of image $I^{(2)}(T)$, which is the affine transformed version of image $I^{(2)}$ with affine transformation T.

3. EXPERIMENTS RESULTS

Our experimental results use the powerful Brainweb simulated MRI volumes for a normal brain [8]. The simulations are based on an anatomical model of the normal brain with the main advantage being that the ground truth is known and can be used for validation.

3.1. Robustness of Feature Images to Noise

In our experiments, we use pairs of 2D slices from 3D PD and T2 MR brain images. The basic feature set includes image intensity I, x-direction derivative $\nabla_x I$, y-direction derivative $\nabla_y I$, cross derivative $\nabla_c I$ and their product $I * \nabla_x I$,$I * \nabla_y I$,$I * \nabla_c I$,$\nabla_x I * \nabla_y I$, $\nabla_x I * \nabla_c I$,$\nabla_y I * \nabla_c I$. The best projection onto a single dimension of the best feature combination of features $W^{(r)}$,$r = 1, 2$ is obtained by maximizing the objective function (3). The original intensity range of the PD and T2 images is in [0 1]. We add different levels of Gaussian noise with mean zero and varying standard deviations (0.1, 0.2 and 0.4) to the images. As a simple example, we rotate the T2 image within a range from -20 degrees to 20 degrees and compute the NMI between the single best pair of feature images. This is compared to the NMI obtained from the original intensity pair. The NMI plots shown in Figure 1 demonstrate that the single best feature found is more robust to noise than the original intensity.

	ground truth	results from intensity	results from feature
s	0.5	0.45	0.5
t	0.5	0.45	0.5
θ	10	12	10
ϕ	10	11.5	10
e	5	5	5
f	5	4	5

Table 1. Registration results of 2D sagittal PD and T2 slices using intensity and best feature pair

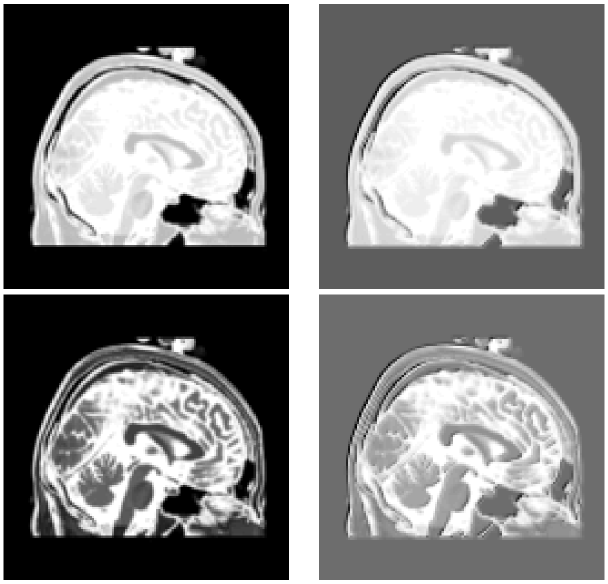
3.2. Affine Registration of PD and T2 2D image slices

We demonstrate results on 2D slices of PD and T2-weighted MR images. The affine transformation is decomposed into a product of shear, scale and rotations. Let $T = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ e & f & 1 \end{bmatrix}$ be an affine transformation.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 2^s & 0 \\ 0 & 2^{-s} \end{bmatrix} R(\theta) \begin{bmatrix} 2^t & 0 \\ 0 & 2^t \end{bmatrix} R(\phi)$$

where s and t are scale and shear parameters, $R(\theta)$ and $R(\phi)$ are two rotation matrices with rotation angle θ and ϕ. In the experiments, the range of the shear and scale parameters is [-1 1], the range of rotation parameters are [-45 45] degrees and the range of translations is [-10 10] pixels. We used 2D slices drawn at sagittal, coronal and axial orientations. The best linear combination is first estimated by maximizing the NMI on the feature sets of noiseless, registered PD and T2-weighted images, whose noised and transformed versions are used in following registration. We then transform the T2-weighted image with an affine transformation T. The shear and scale parameters are 0.5 and the two rotation parameters are 10 degrees, with the translations along the x and y directions being 5 pixels. Hence $T = \begin{bmatrix} 1.9095 & -0.5130 & 0 \\ 0.2565 & 0.4548 & 0 \\ 5 & 5 & 1 \end{bmatrix}$, $s = 0.5$,$t = 0.5$,$\theta = 10$,$\phi = 10$,$e = 5$ and $f = 5$. Gaussian noise with mean 0 and standard deviation 0.2 is added to PD and transformed T2 images. NMI is subsequently used as the criterion to register these two images. We used a coarse-to-fine search strategy to find the best T^* in (4). NMI is computed only on the overlap area of the two images with nearest neighbor interpolation used for the transformation of the image in all experiments.

Figure 2 shows 2D sagittal PD, transformed T2 slices and results recovered using the best feature images. Table 1 shows the registration results using intensity and the best features.Figure 3



(a) intensity image (b) feature image

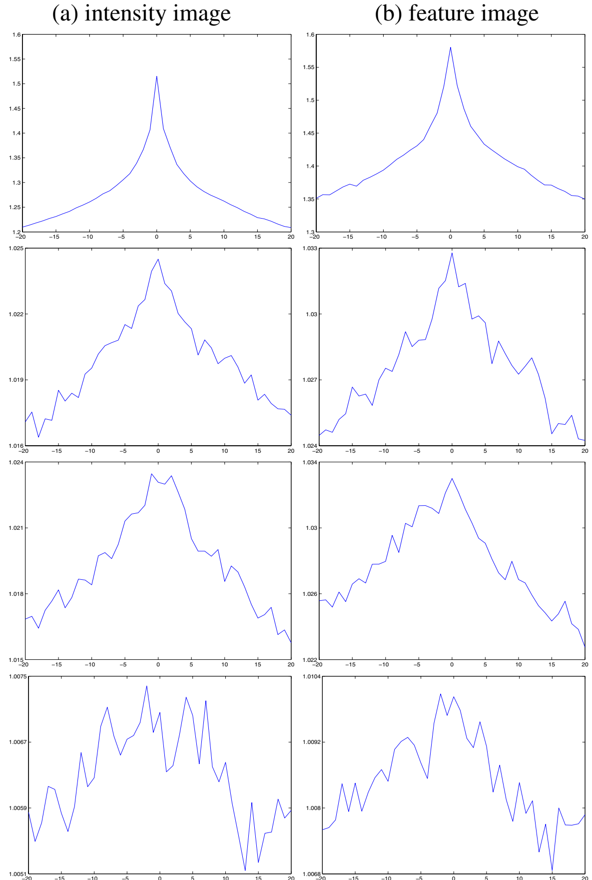
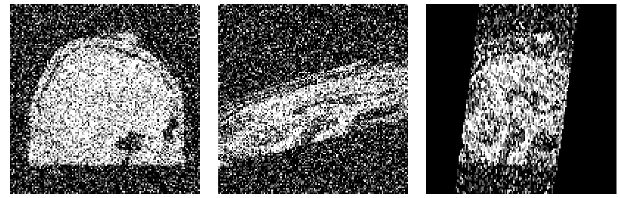


Fig. 1. Left: From top to bottom, NMI between the original intensity pair with rotation. Right: From top to bottom, NMI between the best feature image pair with rotation. The rotation range is from -20 degrees to 20 degrees. From top to bottom: Gaussian noise added with mean 0 and std. deviation 0, 0.1, 0.2 and 0.4.

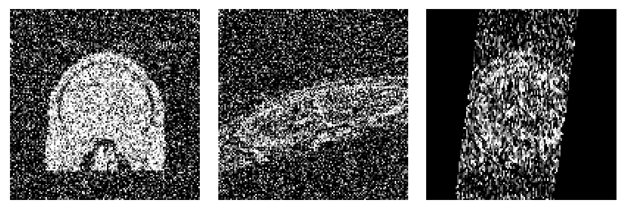


(a) PD (b) T2 (c) recovered T2

Fig. 2. 2D sagittal slices

	ground truth	results from intensity	results from feature
s	0.5	0.45	0.48
t	0.5	0.45	0.52
θ	10	11.5	10.6
ϕ	10	11	9.4
e	5	5	5
f	5	5	5

Table 2. Registration results of 2D coronal PD and T2 slices using intensity and best feature pair

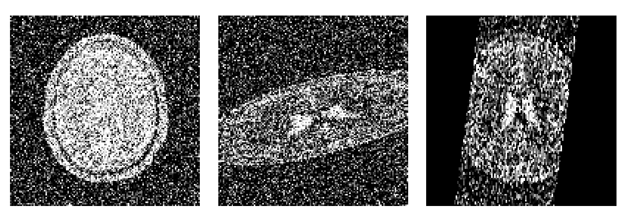


(a) PD (b) T2 (c) recovered T2

Fig. 3. 2D coronal slices

	ground truth	results from intensity	results from feature
s	0.5	0.4	0.5
t	0.5	0.4	0.5
θ	10	12	10
ϕ	10	11	10
e	5	5	5
f	5	5	5

Table 3. Registration results of 2D axial PD and T2 slices using intensity and best feature pair



(a) PD (b) T2 (c) recovered T2

Fig. 4. 2D axial slices

	ground truth	results before bootstrap	results after
s	0.5	0.45	0.5
t	0.5	0.45	0.5
θ	10	12	10.6
ϕ	10	11.5	10.2
e	5	5	5
f	5	4	5

Table 4. Registration results of 2D sagittal PD and T2 slices before and after bootstrap

shows 2D coronal PD, transformed T2 slices and results recovered using the best feature images. Table 2 shows the registration results using intensity and the best feature image pair. Figure 4 shows 2D axial PD, transformed T2 slices and results recovered using the best feature images. Table 3 shows the registration results using intensity and the best feature image pair.

From the registration results above, we observe that better results are achieved using the feature image pair than using the original intensity pair. The feature image sets were combined using projection coefficients obtained from noiseless, registered, training images. However, it would be much more interesting to see if the above approach extends to the case where we do not have *registered*, training images. We turn to this case, next.

3.3. Bootstrapping the feature combination with imperfect registration

We now describe a bootstrap approach wherein NMI is used to register two images using intensity as the feature after which we apply the above technique to get the best set of projection coefficients. These projection coefficients are used to get the best feature pair. NMI is then used on the best feature pair after which the process can be repeated (if necessary). In this experiment, we register images (same images used in section 3.2) using image intensity at the first step. Then we use these imperfectly registered images as training images to extract the best projection coefficients. After projecting the feature images onto the single best dimension, we use the new feature images as input to a new NMI-based registration to boost the registration results achieved by maximizing NMI on image intensity. The second columns of Table 4, 5 and 6 are registration results obtained by maximizing NMI between the original image intensity pairs. These are used to extract the best projection coefficients of features. The third columns of Table 4, 5 and 6 are registration results obtained by maximizing NMI between the best feature image pairs.We see that maximizing NMI between the best feature image pairs is able to improve upon the registration result obtained by using NMI on the original image intensities. Despite the fact that the best feature images were obtained from images which were somewhat mis-registered, we were able to improve upon the original registration result using the *imperfectly* registered images as a training set.

4. CONCLUSIONS

We have demonstrated a proof of concept of the basic idea that feature images can improve upon the original intensities when using NMI as a registration measure. The best feature combination of images in each modality is obtained by projecting the set of feature images onto the single best dimension, with NMI being the

	ground truth	results before bootstrap	results after
s	0.5	0.45	0.5
t	0.5	0.45	0.52
θ	10	11.5	9
ϕ	10	11	11
e	5	5	5
f	5	5	5

Table 5. Registration results of 2D coronal PD and T2 slices before and after bootstrap

	ground truth	results before bootstrap	results after
s	0.5	0.4	0.52
t	0.5	0.4	0.52
θ	10	12	9.4
ϕ	10	11	9.6
e	5	5	5
f	5	5	5

Table 6. Registration results of 2D axial PD and T2 slices before and after bootstrap

criterion once again. Results were shown both in cases where a training set existed and in the bootstrap case where the original image intensities were used to get a crude registration with subsequent refinement using the best feature combination. Obviously, a careful validation effort is required to go beyond the proof of concept stage. We also plan to extend the basic idea of best feature combination to non-rigid registration using NMI.

5. REFERENCES

[1] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," in *Fifth Intl. Conf. Computer Vision (ICCV)*. 1995, pp. 16–23, IEEE Press.

[2] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multi-modality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187–198, 1997.

[3] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Image registration by maximization of combined mutual information and gradient information," *IEEE Trans. on Medical Imaging*, vol. 19, no. 8, pp. 809–814, 2000.

[4] T. Butz and J-P. Thiran, "Affine registration with feature space mutual information," in *MICCAI 2001*, pp. 549–556. Utrecht, NL, 2001.

[5] H. Neemuchwala, A. Hero, P. Carson, and C. Meyer, "Local feature matching using entropic graphs," in *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2004.

[6] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3d medical image alignment," *Pattern Recognit.*, vol. 32, no. 1, pp. 71–86, 1999.

[7] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley and Sons, New York, NY, 1991.

[8] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. Med. Imag.*, vol. 17, no. 3, pp. 463–468, 1998.