Curvelet-Based Sampling for Accurate and Efficient Multi-Modal Image Registration

M.N. Safran[a] and M. Freiman[b] and M. Werman[b] and L. Joskowicz[b]

[a]Interdisciplinary Center for Neural Computation, The Hebrew University of Jerusalem, Israel
[b]School of Engineering and Computer Science, The Hebrew University of Jerusalem, Israel.

ABSTRACT

We present a new non-uniform adaptive sampling method for the estimation of mutual information in multi-modal image registration. The method uses the Fast Discrete Curvelet Transform to identify regions along anatomical curves on which the mutual information is computed. Its main advantages of over other non-uniform sampling schemes are that it captures the most informative regions, that it is invariant to feature shapes, orientations, and sizes, that it is efficient, and that it yields accurate results. Extensive evaluation on 20 validated clinical brain CT images to Proton Density (PD) and T1 and T2-weighted MRI images from the public RIRE database show the effectiveness of our method. Rigid registration accuracy measured at 10 clinical targets and compared to ground truth measurements yield a mean target registration error of 0.68mm(std=0.4mm) for CT-PD and 0.82mm(std=0.43mm) for CT-T2. This is 0.3mm (1mm) more accurate in the average (worst) case than five existing sampling methods. Our method has the lowest registration errors recorded to date for the registration of CT-PD and CT-T2 images in the RIRE website when compared to methods that were tested on at least three patient datasets.

Keywords: Registration, Multiresolution and wavelets

1. INTRODUCTION

Multimodal image registration is a key step in medical image analysis for diagnostic and surgical purposes. It is required to align image datasets from different modalities into a common coordinate frame, and thus provide a more informative, quantitative view of the clinical situation. For a general review of registration methods, see.[1–4]

One of the most popular methods for image registration compares the intensity values in both images and finds the transformation that maximizes the similarity between them. Due to the non-linear intensity dependency between different imaging modalities, standard intensity-based similarity measures, such as intensity-difference based or correlation-based measures, have limited use for multimodal image registration. Information-theoretic similarity measures, such as Mutual Information (MI),[5,6] have been more successful since they do not assume a linear relation between the image intensity values in both images.

Subsequent refinements, including Normalized Mutual Information (NMI)[7] and the Mattes formulation[8] increase the robustness and convergence range of the registration. However, MI-based registration methods have three key limitations: 1) they do not incorporate spatial information that can significantly improve the registration accuracy and robustness, 2) they are sensitive to interpolation artifacts,[9–12] and 3) their computation is time-consuming. For a detailed survey of MI-based registration see.[13]

Recent research proposes methods to improve the registration by adding a spatial term to the similarity measure. Pluim et al.[14] include a gradient-based term with both spatial and intensity-based mutual information. Xu and Chen[15] use multi-resolution wavelets to compute the gradient-based spatial term. Gun et al[16] use instead

Further author information: (Send correspondence to M. Freiman)
M.N. Safran: E-mail: moshe.safran@mail.huji.ac.il, Telephone: +972-2-6585371
M. Freiman: E-mail: freiman@cs.huji.ac.il, Telephone: +972-2-6585371
M. Werman: E-mail: werman@cs.huji.ac.il, Telephone: +972-2-6584931
L. Joskowicz: E-mail: josko@cs.huji.ac.il, Telephone: +972-2-6586299

Medical Imaging 2009: Image Processing, edited by Josien P. W. Pluim, Benoit M. Dawant,
Proc. of SPIE Vol. 7259, 72590M · © 2009 SPIE
CCC code: 1605-7422/09/$18 · doi: 10.1117/12.810695

the Maximum Distance-Gradient (MDG) vector field for the gradient term. These combined similarity measures suffer from several drawbacks. First, although these methods yield more robust registration results than those obtained with standard MI measures, their accuracy is often lower, so a second registration refinement step is usually necessary. Second, the spatial term significantly increases the computation time. Third, the weight coefficients used to combine the spatial and MI terms in the similarity measure require extensive fine-tuning and depend on the registration domain.

Alternative approaches incorporate spatial information into the MI-based similarity measure based on higher-order image properties. These include voxel co-occurrence matrices or voxel neighborhood regions,[17–19] gradient vector fields,[20] gradients intensities,[21] and Gibbs random fields.[22] While these measures do improve the registration accuracy, they require complex and time-consuming computations. Luan et al.[23] introduce a qualitative mutual information measure that incorporates anatomical information about the voxels which yields a more robust similarity measure. However, this method requires prior information about the anatomical structures, and cannot be applied in a generic, anatomy and protocol independent registration framework.

Adaptive image sampling has been proposed as an indirect method for incorporating spatial information into MI-based registration. The idea is to treat each voxel as a variable in a Markov random field instead of as an independent variable. Thus, the contribution (weight) of each voxel to the mutual information also incorporates the voxel neighborhood information. The weight is the probability of the voxel to be sampled during the estimation of the image intensity Probability Distribution Functions (PDF). Adaptive sampling replaces the spatial terms and the variables that are added to the similarity measure. Sabuncu et al.[24] estimate the image PDF by proportionally sampling more densely regions with higher gradient image magnitudes. While this yields a faster registration, the image gradients are highly sensitive to image quality and noise, thus affecting the registration convergence and quality. Sundar et al.[25] propose to use octree image partitioning as a multi-resolution sampling technique. An octree is built from the original image is split into spatially adaptive homogeneous regions of different sizes according to their content. The main drawbacks of octree sampling are its high sensitivity to noise and its bias for axis-aligned structures, which are uncommon in medical images.

To overcome these limitations, we have developed a novel adaptive sampling scheme based on the curvelet transform.[26,27] The curvelet transform is a higher dimensional generalization of the wavelet transform. It is designed to represent piecewise smooth functions at different scales and angles. In this scheme, curved edges are approximated with very few coefficients. We use the curvelet transform to define the informative regions in the image and to estimate the image intensity PDF for MI computation in these regions. We propose a two-step registration scheme consisting of a coarse registration with a standard MI-based method, followed by a fine registration refinement with a curvelet-based MI estimation registration.

Extensive evaluation on 20 validated clinical brain CT images to Proton Density (PD) and T1 and T2-weighted MRI images from the public RIRE database show the effectiveness of our method. Rigid registration accuracy measured at 10 clinical targets and compared to ground truth measurements yield a mean target registration error of 0.68mm(std=0.4mm) for CT-PD and 0.82mm(std=0.43mm) for CT-T2. This is 0.3mm (1mm) more accurate in the average (worst) case than exhaustive sampling and four other adaptive sampling schemes.[24,25] Our method has the lowest registration errors recorded to date for the registration of CT-PD and CT-T2 images in the RIRE website when compared to methods that were tested on at least three patient datasets. Our method is more efficient than existing methods and reduces the interpolation artefacts common in MI-based metrics.[12]

2. THE CURVELET TRANSFORM

Wavelets are having a growing impact on signal processing over the last decade. Their key advantages are that they provide a good approximation for one-dimensional (1D) piecewise smooth functions[28] and are effective for identifying point discontinuities. Unfortunately, these properties do not extend to 2D and 3D piecewise smooth functions with 1D and 2D discontinuities which are used to model images. Recent works describe wavelet extensions to 2D and 3D.[26,29,30] In particular, the curvelet transform[26] has been successfully used for image denoising[31] and has a natural extension to 3D.[27] We briefly describe the 2D curvelet transform next.

The Fast Discrete Curvelet Transform (FDCT) is an invertible map from the space of 2D images to the set of coefficients $C(\ell, w, k_1, k_2)$ representing 2D curves, where ℓ is the scale index, w represents the related orientation, and k_1, k_2 are the spatial location indices in the image. The curvelet decomposition of an image is constructed in three steps:

1. **Sub-band decomposition:** The image f is filtered into several sub-bands representing the image response to a low-pass filter and to a series of band-pass filters at different scales ℓ.
2. **Smooth partitioning:** Each sub-band is smoothly windowed into square regions at an appropriate scale.
3. **Ridgelet analysis:** Each square region is then decomposed into a set of 1D directional signals by applying the 2D Fourier transform followed by the 1D inverse Radon transform. Each resulting 1D signal is then analyzed with wavelets.

The resulting coefficients sparsely describe the curves of scale ℓ and orientation w at spatial location (k_1, k_2). The coefficient set obtained by applying the FDCT to image $f(\vec{x})$ is:

$$C(\ell, w, k_1, k_2) = \text{FDCT}[f(\vec{x})] \tag{1}$$

3. MASK COMPUTATION

The sampling pattern is defined by an image mask, which is computed as follows. The goal is to represent the image regions that contain strong curves at multiple scales with the curvelet decomposition. The mask computation proceeds in four steps: 1) image curvelet decomposition; 2) curve trace enhancement; 3) image curves computation for each scale, and; 4) mask region selection

To decompose the image into curvelets, we apply the curvelet transform (Eq. 1) to the image. The result consists of a set of curvelet coefficients (C) at different scales (ℓ), orientations (w), and spatial locations (k_1, k_2).

To enhance the curves, we apply to the curvelet coefficients a soft-thresholding function:

$$F(\alpha) = a[S(c(\alpha - b)) - S(-c(\alpha + b))] \tag{2}$$

where $S(x) = 1/(1 - e^{-x})$ is the sigmoid function, b, c are predefined threshold parameters, and

$$a = 1/[S(c(1 - b)) - S(-c(1 + b))] \tag{3}$$

This function scales the curvelet coefficients in the $[-1, 1]$ range. Coefficients with small absolute values (noise) are set close to zero, while those with large absolute values (edges) are set close to ± 1.

To compute the image curves at scale m, retain the coefficients of this scale and set all others to zero. The resulting curvelet coefficients $\tilde{C}_m$ are:

$$\tilde{C}_m(\ell, w, k_1, k_2) = \begin{cases} F(C(\ell, w, k_1, k_2)) & , \quad \ell = m \\ 0 & , \quad \ell \neq m \end{cases} \tag{4}$$

To obtain the curves in the image at scale m, we apply the inverse curvelet transform IFDCT to the coefficients:

$$\tilde{f}_m(\vec{x}) = \text{IFDCT}\left[\tilde{C}_m(\ell, w, k_1, k_2)\right] \tag{5}$$

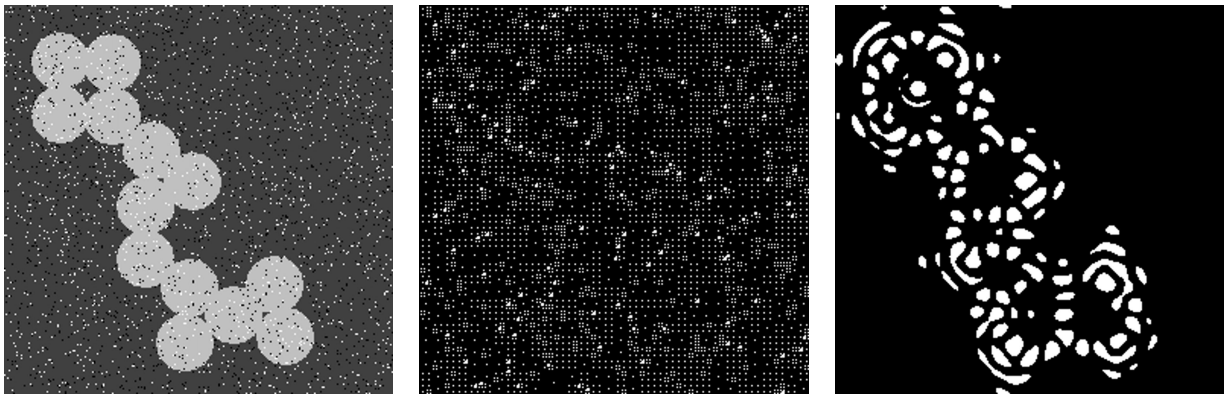
In the resulting image, prominent curves are represented by pixels with large absolute values.

Finally, we generate a mask containing the desired percentage of pixels p by selecting the pixels with the largest absolute values. The resulting mask is defined as:

$$M_{m,p} = \left\{\vec{x} : \left|\tilde{f}_m(\vec{x})\right| > P_p\left[\left|\tilde{f}_m\left(\vec{x'}\right)\right|\right]\right\} \tag{6}$$

where $P_p\left[\left|\tilde{f}_m(\vec{x})\right|\right]$ is the p-th percentile of the intensities appearing in image $\left|\tilde{f}_m\left(\vec{x'}\right)\right|$, sorted in descending order.

Fig. 1 illustrates the mask generated from a noisy synthetic image with the octree-based sampling mask and our curvelet-based method. Note how the octree-based mask is highly sensitive to noise. In contrast, our method robustly ignores the noise and captures the informative structures.



(a) original image (b) quadtree mask (c) curvelet mask

Figure 1. Mask examples on a synthetic image: (a) original image with added impulse noise; (b) quadtree-based mask, and (c) curvelet-based mask.

4. CURVELET-BASED REGISTRATION METHOD

The curvelet-based registration method consists of two steps. In the first step, a coarse registration is computed with an MI-based similarity measure using stochastic uniform sampling. In the second step, the registration is refined by deterministic curvelet-based sampling. The two steps are necessary because the spatial information term in the MI computation often leads to a narrow attraction range.[14]

In both steps, we use the Normalized Mutual Information (NMI):

$$NMI\left(f\left(\vec{x}\right), g\left(T\left(\vec{x}\right)\right)\right) = \frac{\left(H\left(f\left(\vec{x}\right)\right) + H\left(g\left(T\left(\vec{x}\right)\right)\right)\right)}{H\left(f\left(\vec{x}\right), g\left(T\left(\vec{x}\right)\right)\right)} \tag{7}$$

where H is the entropy between the source image $f\left(\vec{x}\right)$ and the target image $g\left(T\left(\vec{x}\right)\right)$ with respect to rigid transformation T. The points $\vec{x}$ are taken from the source image. The target image intensities $g\left(T\left(\vec{x}\right)\right)$ are estimated by linear interpolation. The NMI is computed with Parzen-windowed histogram-based estimators of the image intensity probability distributions with histogram smoothing.[8, 32]

In the first step, we compute the image histograms and the NMI with randomly sampled points selected from the source image with uniform probability. We maximize the NMI with standard regular-step gradient descent optimization; new random points are selected at each iteration. The resulting transformation is then used as the initial transformation for the second registration step.

In the second step, we calculate the image histograms for the NMI using all of the points included in the curvelet mask at the coarsest scale, $M_{1,p}$. The joint histogram estimator $\hat{P}(.,.)$ of the source and target images voxel intensities, i_t and i_S, is:

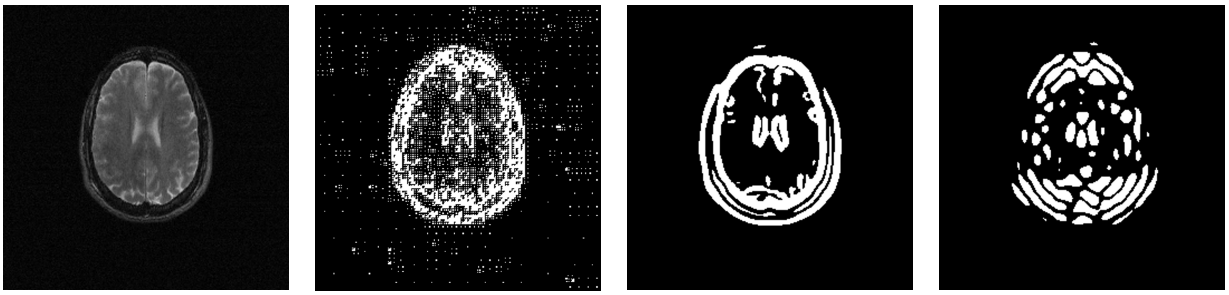
$$\hat{P}\left(f\left(\vec{x}\right) = i_t, g\left(T\left(\vec{x}\right)\right) = i_s\right) = \frac{1}{N} \sum_{\vec{x} \in M_{1,p}} K\left(f\left(\vec{x}\right) - i_t, g\left(T\left(\vec{x}\right)\right) - i_s\right) \tag{8}$$

where N is the total number of pixels in the mask $M_{1,p}$, and $K\left(\cdot,\cdot\right)$ is a cubic spline-based Parzen windowing kernel.[8] The optimization algorithm, parameter settings, and halting conditions are the same for both steps.

5. EXPERIMENT

We have implemented our registration method with the[33] software library and with the curvelet package.[34]

We use the multimodal Retrospective Image Registration Evaluation (RIRE) Project database, commonly known as the Vanderbilt Database,[35] in all our experiments. The RIRE project is public domain platform that is used to compare retrospective multimodal registration techniques developed around the world. The database



(a) original image (b) octree mask (c) gradient mask (d) curvelet mask

Figure 2. Illustrative example: (a) original MRI-T2 slice, (b) octree-based mask, (c) gradient-based mask, and; (d) curvelet-based mask.

consists of clinical brain images of seven patients (with and without tumors), including CT, MR Proton-Density (MR-PD), MR T1 time relaxation (MR-T1), and MR T2 time relaxation (MR-T2). We register 20 CT images (subject images) to their corresponding MRI images (template images): 7 MR-PD images, 6 MR-T1 images, and 7 MR-T2. In all cases, we use the rectified MR images, corrected for scanner-dependent geometric distortion.

The ground-truth registration transformations which constitute the gold standard were acquired prospectively with implanted fiducial markers. The transformations remain hidden from the public. Prior to disclosure, the markers were erased from the images. Researchers then perform a retrospective blind registration task and report back to the site their computed registration transformations. These transformations are then compared to the gold standard and ranked with respect to other algorithms according to their target registration error (TRE). The TRE is measured in millimeters at ten predefined clinically relevant target locations. For training purposes, an additional dataset with clinical images from all modalities is provided with the ground truth transform.

5.1 Methodology

To evaluate the effectiveness of our two-step registration method and compare it to that of the other four sampling-based methods, we individually registered all 20 CT-MR dataset pairs with all methods and recorded the results.

For the coarse registration step, we used the same image mask for all methods. The mask is created by randomly sampling 25% of the image pixels in each slice. For the fine registration step, we sampled 10% of the pixels with each sampling method and used them the image mask (for the no sampling method, we use all the image pixels). The resulting transformations for each registration were then submitted to the RIRE website[35] for validation.

Due to the anisotropy of the MR image voxels ($x = y = 1.27$mm, $z = 4.11$mm), we perform the sampling computations on individual image slices and then compose the results in 3D. For our curvelets-based sampling method, we apply a 2D curvelet transform on each image slice and generate from it the corresponding 2D curve representation. The 2D curves are combined into a 3D curve representation image, from which we select the pixels with the largest absolute values to generate the image sampling mask. The soft-threshold function parameters were set to $b = 0.1$ and $c = 40$ for all images.

For gradient-based sampling, we compute 2D gradient magnitudes by filtering each image slice in the x and y directions with the first-order derivative of a Gaussian kernel ($\sigma = 3$ pixels $= 3.8$ mm). The gradient mask is then created by selecting 10% of the pixels with the largest gradient magnitude values.

For octree-based sampling, we compute quadtrees for each image slice and create quadtree-based masks by sampling the center of each 2D block. Since the number of samples is not always the same for all images, we included a percentage close to it ($11.6\% \pm 2\%$).

Fig. 2 shows examples of the masks generated by the three adaptive sampling methods. Observe that the gradient and octree based masks induce a high fragmentation, while the curvelets masks capture the image curves. Note how our mask captures the informative structures while ignoring other, less informative regions.

Table 1. Registration results. Comparison between different sampling schemes. The target registration error is in mm. The last column shows the average (std) running time for each method.

METHOD	CT - PD			CT - T2			CT - T1		
	Mean	Max	σ	Mean	Max	σ	Mean	Max	σ
No-sampling	1.04	2.68	0.59	1.04	2.62	0.51	0.91	2.08	0.54
Uniform	1.05	2.51	0.6	1.03	2.63	0.55	0.98	2.17	0.56
Gradient	0.85	1.89	0.47	0.93	2.20	0.44	0.88	1.95	0.46
Octree	0.73	1.57	0.38	0.95	2.76	0.57	0.72	1.37	0.29
Curvelet	**0.68**	**1.59**	**0.40**	**0.82**	**1.71**	**0.43**	**0.92**	**2.90**	**0.56**

5.2 Results

Table 1 shows the registration results of our method and the other six methods.

For the CT-PD and CT-T2 registrations, curvelet-based sampling consistently outperformed all other methods, with both lower mean and lower median TREs. In addition, our method has the lowest registration errors recorded to date for the registration of CT to rectified PD and T2 images in the RIRE website when compared to methods that were tested on at least three patient datasets. Registration with our curvelet-based sampling method yields a mean accuracy of 0.68mm(std=0.4mm) for CT-PD and 0.82mm(std=0.43mm) for CT-T2. This is 0.3mm (1mm) more accurate in the average (worst) case than the other five sampling methods.

Note that the overall performance of all three adaptive sampling methods was better than that of the exhaustive sampling method. This improvement was not due to the stochastic sampling introduced in the first registration step, which on its own performed similarly to or worse than the exhaustive sampling scheme.

For the CT-T1 modality, octree-based sampling yield the smallest TREs. Our sampling method was less effective for this type of registration, most likely because the CT and T1 modalities are very similar, so their mutual information is spread in the entire image and not just along the curves.

All registrations were performed on a dual-processor 2GHz PC with 2GB RAM running Linux. The average running time of our curvelet-based sampling method was 87 ± 69secs, which is was faster on average than 103 ± 76secs for gradient-based sampling, 92 ± 35secs for octree-based sampling. The running time without sampling is 512 ± 127secs, which is significantly larger.

6. CONCLUSIONS

This paper presents a new adaptive non-uniform sampling method for the accurate estimation of the mutual information between two images. The method uses the discrete curvelet transform to detect the informative regions in the images, while eliminating non-informative regions and noise. We use a two-step registration scheme with uniform sampling for the MI estimation during the first step for coarse registration and the curvelet-based sampling mask during the second step for fine registration. Our experimental results show that our sampling method yields a significant improvement in registration accuracy compared to other adaptive sampling schemes, including exhaustive sampling, for CT-MR (PD,T2) multimodal image registration.

We are currently working in the application of our method to other multimodal image registration tasks such as PET-MR. We plan to apply our sampling method for non-rigid registration.

Acknowledgments

This research was partially funded by the European FP7 ROBOCAST project. The images for this research were provided by the Retrospective Image Registration Evaluation Project No. 8R01EB002124-03, National Institutes of Health, Principal Investigator, J. Michael Fitzpatrick, Vanderbilt University, Nashville, TN, United States.

REFERENCES

[1] Zitova, B. and Flusser, J., "Image registration methods: a survey.," *Image Vision Comput.* **21**(11), 977–1000 (2003).

[2] Maintz, J. B. A. and Viergever, M. A., "A survey of medical image registration," *Medical Image Analysis* **2**(1), 1–36 (1998).

[3] Hajnal, J. V., Hill, D. L. G., and Hawkes, D. J., eds., [*Medical Image Registration*], CRC Press (2001). ISBN 0849300649.

[4] Modersitzki, J., [*Numerical Methods for Image Registration*], Oxford University Press (2004).

[5] Collignon, A., Vandermeulen, D., Suetens, P., and Marchal, G., "Automated multimodality medical image registration using information theory," in [*Proc. of the 14th Int. Conf. on Information Processing in Medical Imaging (IPMI'95)*], **3**, 263–274, Kluwer Academic Publishers (June 1995).

[6] Wells, W., Viola, P., Atsumi, H., Nakajima, S., and Kikinis, R., "Multi-modal volume registration by maximization of mutual information," *Medical Image Analysis* **1**(1), 35–51 (1996).

[7] Studholme, C., Hill, D., and Hawkes, D., "An overlap invariant entropy measure of 3d medical image alignment," *Pattern Recognition* **32**(1), 71–86 (1999).

[8] Mattes, D., Haynor, D. R., Vesselle, H., Lewellen, T. K., and Eubank, W., "Pet-ct image registration in the chest using free-form deformations.," *IEEE Trans. Med. Imaging* **22**(1), 120–128 (2003).

[9] Tsao, J., "Interpolation artifacts in multimodality image registration based on maximization of mutual information," *Medical Imaging, IEEE Transactions on* **22**, 854–864 (July 2003).

[10] Pluim, J. P. W., Maintz, J. B. A., and Viergever, M. A., "Interpolation artefacts in mutual information-based image registration," *Computer Vision and Image Understanding* **77**(2), 211 – 232 (2000).

[11] Ji, J. X., Pan, H., and Liang, Z.-P., "Further analysis of interpolation effects in mutual information-based image registration," *Medical Imaging, IEEE Transactions on* **22**, 1131–1140 (Sept. 2003).

[12] Thévenaz, P., Bierlaire, M., and Unser, M., "Halton sampling for image registration based on mutual information," *Sampling Theory in Signal and Image Processing* **7**, 141–171 (May 2008).

[13] Pluim, J., Maintz, J., and Viergever, M., "Mutual-information-based registration of medical images: a survey," *IEEE Trans. Med. Imaging* **22**, 986–1004 (August 2003).

[14] Pluim, J., Maintz, J., and Viergever, M., "Image registration by maximization of combined mutual information and gradient information," *IEEE Trans. Med. Imaging* **19**, 809–814 (August 2000).

[15] Xu, R. and Chen, Y.-W., "Wavelet-based multiresolution medical image registration strategy combining mutual information with spatial information," *International Journal of Innovative Computing, Information and Control (IJICIC)* **3**(2), 285–296 (2007).

[16] Gan, R., Chung, A. C., and Liao, S., "Maximum distance-gradient for robust image registration," *Medical Image Analysis* **12**(4), 452–468 (2008).

[17] Rueckert, D., Clarkson, M. J., Hill, D., and Hawkes, D., "Non-rigid registration using higher-order mutual information," in [*Proc. SPIE Medical Imaging 2000: Image Processing*], 438–447 (2000).

[18] Russakoff, D., Tomasi, C., Rohlfing, T., and Maurer, C. R. J., "Image similarity using mutual information of regions," in [*Proc. of the 8th Eur. Conf. on Computer Vision (ECCV'04)*], *LNCS* **3023**, 596–607 (May 2004).

[19] Bardera, A., Feixas, M., Boada, I., and Sbert, M., "High-dimensional normalized mutual information for image registration using random lines.," in [*3rd Int. Workshop on Biomedical Image Registration (WBIR'06)*], Pluim, J. P. W., Likar, B., and Gerritsen, F. A., eds., *LNCS* **4057**, 264–271, Springer (2006).

[20] Yujun, G. and Cheng-Chang, L., "Multi-modality image registration using mutual information based on gradient vector flow," in [*Proc. of the 18th Int. Conf. on Pattern Recognition (ICPR'06)*], 697–700, IEEE Computer Society, Washington, DC, USA (2006).

[21] Shams, R., Sadeghi, P., and Kennedy, R. A., "Gradient intensity: A new mutual information-based registration method.," in [*CVPR*], IEEE Computer Society (2007).

[22] Zheng, G., "Effective incorporation of spatial information in a mutual information based 3d-2d registration of a ct volume to x-ray images.," in [*MICCAI (2)*], Metaxas, D. N., Axel, L., Fichtinger, G., and Szkely, G., eds., *LNCS* **5242**, 922–929, Springer (2008).

[23] Luan, H., Qi, F., Xue, Z., Chen, L., and Shen, D., "Multimodality image registration by maximization of quantitative-qualitative measure of mutual information," *Pattern Recogn.* **41**(1), 285–298 (2008).

[24] Sabuncu, M. and Ramadge, P., "Gradient based nonuniform subsampling for information-theoretic alignment methods," in [*Proc. of the 26th Int. Conf. of the IEEE Eng. in Medicine and Biology Society (IEMBS'04)*], **1**, 1683–1686 (2004).

[25] Sundar, H., Shen, D., Biros, G., Xu, C., and Davatzikos, C., "Robust computation of mutual information using spatially adaptive meshes," in [*Proc. of the 10th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI'07)*], Ayache, N., Ourselin, S., and Maeder, A., eds., *LNCS* **4791**, 950–958, Springer (2007).

[26] Candes, E. J., Demanet, L., L., D. D., and Ying, L., "Fast discrete curvelet transforms," *Multiscale Model. Simul.* **5**, 861–899 (2005).

[27] Ying, L., Demanet, L., and Candes, E., "3D discrete curvelet transform," in [*Proc. of the 11th Int. Conf. on Wavelets*], (July 2005).

[28] Mallat, S., [*A wavelet tour of signal processing*], Academic Press (1998).

[29] Freeman, W. and Adelson, E., "The design and use of steerable filters," *IEEE Trans. Pattern Analysis and Machine Intelligence* **13**(9), 891–906 (1991).

[30] Do, M. N. and Vetterli, M., "The contourlet transform: an efficient directional multiresolution image representation.," *IEEE Transactions on Image Processing* **14**(12), 2091–2106 (2005).

[31] Starck, J., Candes, E., and Donoho, D., "The curvelet transform for image denoising.," *IEEE Trans. on Image Processing* **11**(6), 670–684 (2002).

[32] Thevenaz, P. and Unser, M., "Optimization of mutual information for multiresolution image registration," *Image Processing, IEEE Transactions on* **9**, 2083–2099 (Dec 2000).

[33] `http://elastix.isi.uu.nl`.

[34] `http://www.curvelet.org`.

[35] West, J., Fitzpatrick, J., Wang, M., and et al, "Comparison and evaluation of retrospective intermodality brain image registration techniques," *Journal of Computer assisted Tomography* **21**, 554–566 (Jul-Aug 1997). `http://www.insight-journal.org/rire`.