Medical Image Analysis 15 (2011) 125–132

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

A curvelet-based patient-specific prior for accurate multi-modal brain image rigid registration

M. Freiman *, M. Werman, L. Joskowicz

School of Engineering and Computer Science, The Hebrew University of Jerusalem, Israel

ARTICLE INFO

Article history:
Received 7 September 2009
Received in revised form 15 June 2010
Accepted 26 August 2010
Available online 25 September 2010

Keywords:
Multi-modal image registration
Mutual information
Non-uniform sampling
Curvelet transform

ABSTRACT

We present a new non-uniform sampling method for the accurate estimation of mutual information in multi-modal brain image rigid registration. Most existing density estimators used for mutual information computation incorrectly assume that the intensity of each voxel is independent from its neighborhood. Our method uses the 3D Fast Discrete Curvelet Transform to reduce the sampled voxels' interdependency by sampling voxels that are less dependent on their neighborhood, and thus provide a more accurate estimation of the mutual information and a more accurate registration. The main advantages of our method over other non-uniform sampling schemes are that: (1) it provides more accurate estimation of the image statistics with fewer samples; (2) it is less sensitive to the variability of anatomical structures shapes, orientations, and sizes, and; (3) it yields more accurate registration results. Extensive evaluation on 1000 synthetic registrations between T1 and T2-weighted clinical MRI images and 20 real clinical registrations of 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. Our method has the lowest mean registration errors recorded to date for CT-MR image registration in the RIRE website for methods tested on more than five datasets. These results indicate that our sampling scheme can be used to achieve more accurate multi-modal registration required for image guided therapy and surgery.

1. Introduction

Multi-modal rigid image registration is a key step in medical image analysis for diagnosis, Image Guided Therapy (IGT) (Webster et al., 2009) and Image Guided Surgery (IGS) (Gering et al., 2001; Joskowicz et al., 2006). It is required to align image datasets from different modalities into a common coordinate frame to provide an informative, quantitative view of the clinical situation. For IGS and IGT, the Targets Registration Error (TRE) distribution is of importance due to patient safety issues. Although many registration methods have been proposed in the past decade, accuracy improvement is still possible and of importance for both diagnosis, IGT and IGS (Greve and Fischl, 2009). For reviews of existing image-based registration methods, see (Maintz and Viergever, 1998; Hajnal et al., 2001; Zitova and Flusser, 2003; Modersitzki, 2004).

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 similarity measures, such as intensity difference or intensity correlation measures, are of limited use for multi-modal image registration. Information-theoretic similarity measures, such as Mutual Information (MI) (Collignon et al., 1995; Wells et al., 1996), have proved to be superior since they do not assume a linear relation between the images intensities. Subsequent refinements, including Normalized Mutual Information (NMI) (Studholme et al., 1999) and Mattes' formulation (Mattes et al., 2003) increase the registration robustness and convergence range. For a survey of MI-based registration, see (Pluim et al., 2003).

MI-based registration methods estimate image entropies by modeling the image voxels as independent and identically distributed (i.i.d) random variables. However, as shown in Rueckert et al. (2000), the occurrence of a voxel intensity value is dependent on its neighboring voxels intensity values, so the voxels should not be modeled as i.i.d.

To overcome these limitations, we have developed a non-uniform sampling scheme based on the 3D curvelet transform (Candes et al., 2005a; Ying et al., 2005). The curvelet transform is a generalization of the wavelet transform that provides a compact representation of the image in the curvelet domain. The curvelet transform was designed to provide a nearly optimal sparse representation of images with non axis-aligned edges (Candes and Donoho, 2004). It is less sensitive to noise and is not biased to

* Corresponding author. Address: School of Engineering and Computer Science, The Hebrew University of Jerusalem, Givat Ram campus, Jerusalem 91904, Israel. Tel.: +972 2 6585371; fax: +972 2 6585727.
E-mail address: freiman@cs.huji.ac.il (M. Freiman).
URL: http://www.cs.huji.ac.il/~freiman (M. Freiman).

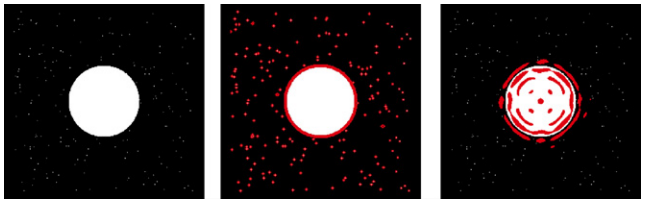
1361-8415/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.media.2010.08.004

axis-aligned structures as other sampling masks (Sabuncu and Ramadge, 2004; Bhagalia et al., 2009; Sundar et al., 2007; Luan et al., 2008).

Fig. 1 illustrates the properties of a mask generated by our method compared to a mask generated by a gradient-based method on a 2D synthetic example. The synthetic image consists of a centered sphere and a noisy background. The gradient-based mask selects the pixels with the largest gradient magnitude without considering their edge scale and orientation. Therefore, most of the mask is related to the noise inside the image. In contrast, our curvelets-based mask considers also the edge orientation and scale of each pixel, and therefore selects regions inside the circle (object) to the mask rather than selecting regions with noisy pixels. This suggests that curvelet-based non-uniform sampling can provide a good estimation of the image entropies with a small number of samples.

We demonstrate the effectiveness of our method for the registration of clinical brain T1 to T2-weighted MRI brain images and for the registration of clinical CT images to Proton Density (PD) and T1 and T2-weighted MRI images. All images were taken from the public RIRE database (West et al., 1997). For MRI T2-MRI T1 registration, 1000 synthetic transformations were used to evaluate the performance of our method on real clinical images with noise and intensity inhomogeneities. Our method yields a mean and maximum TRE of 1.6 mm (std = 1.6 mm) and 4.5 mm respectively, with respect to the ground truth. The mean TRE improvement is 1.5 mm, and maximal TRE improvement is 1.1 mm with respect to existing sampling methods (Wells et al., 1996; Sabuncu and Ramadge, 2004; Bhagalia et al., 2009). For CT-MRI registration, 20 real clinical situation transformations were computed. Our method yields a mean and maximum target registration error of 0.73 mm (std = 0.3 mm) and 1.6 mm, respectively, with respect to the ground truth. The mean TRE improvement is 0.3 mm and the maximal TRE improvement is 1 mm with respect to existing sampling methods (Wells et al., 1996; Sabuncu and Ramadge, 2004; Bhagalia et al., 2009). Our method yields the lowest mean registration errors recorded to date for CT-MR image registration in the RIRE website for methods tested on at least five datasets, and more accurate results compared to recently published results on this database (Hahn et al., 2010) and to other databases with similar image properties (Wong et al., 2010).

The rest of this paper is organized as follows. Section 2 summarizes previous works that addressed the voxels inter-dependency in MI estimation and discusses their drawbacks. Section 3 describes the theoretical motivation of curvelet-based sampling for MI-based registration. Section 4 presents a brief summary of the curvelet transform. Section 5 describes our new method for the generation of the image sampling mask with the curvelet transform. Section 6 describes our two-step curvelet-based registration method. Section 7 presents the experimental setup and results of a comparative registration accuracy evaluation. Section 8 concludes the paper.



(a) original synthetic image (b) gradient-based mask (c) Our curvelets-based mask

Fig. 1. Illustrative example: (a) original synthetic example. The object is the circle, with random noise added to it; (b) gradient-based mask, and; (c) our curvelets-based mask. The masks are overlayed on the original image in red. The gradient-based mask is sensitive to the noisy pixels while the curvelets-based mask captures the object edges, and ignores the noisy pixels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2. Previous work

Recent research addresses the inter-voxels dependency by incorporating spatial information into the similarity measure. We distinguish between three main approaches: (1) hybrid similarity measures; (2) higher-order MI, and; (3) non-uniform sampling.

The hybrid methods use a two-term based similarity measure in which the first term is the intensity-based MI and the second term incorporates spatial information. Pluim et al. (2000) use a gradient-based term as the spatial term in their hybrid similarity measure. Rui and Yen-Wei (2007) use multi-resolution wavelets to compute the spatial term. Gan et al. (2008) use the Maximum Distance-Gradient (MDG) vector field for the spatial term. These measures have several drawbacks. First, they often yield less accurate registrations, although they are more robust than those obtained with standard MI measures. 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 depend on the registration domain and require extensive fine-tuning.

Higher-order MI based methods consider higher-order image properties as the random variables instead of the voxels' intensities in the MI computation. The higher-order information can include image gradients (Butz and Thiran, 2001), voxel co-occurrence matrices, or voxel neighborhood regions (Rueckert et al., 2000; Russakoff et al., 2004; Bardera et al., 2006), Gaussian scale space derivatives (Holden et al., 2004; Legg et al., 2009), gradient vector fields (Yujun and Cheng-Chang, 2006), gradient intensities (Shams et al., 2007), Gibbs random fields (Zheng, 2008), spatial locations (Sabuncu, 2006; Staring et al., 2009), or wavelet features (Pauly et al., 2009). While these measures improve the registration accuracy, they require complex and time-consuming computations.

Non-uniform image sampling is an indirect method for reducing the inter-voxels' dependency in MI-based registration. The adaptive sampling replaces the spatial terms and the high dimensional variables that are used in the similarity measure with adaptive sampling of voxels that are less dependent on their neighboring voxels. Voxels along the edges are then used for the image IPDF's estimation. Sabuncu and Ramadge (2004) and Bhagalia et al. (2009) estimate the image IPDF by sampling more densely regions with higher gradient image magnitudes. While this speeds up the registration, the image gradients are computed locally for each voxel, and thus cannot detect edge-like structures. Thus, they are highly sensitive to noisy voxels such as those shown in Fig. 1. Luan et al. (2008) introduce a qualitative mutual information measure where the weight of each voxel is computed using its spherical neighborhood entropy. However, sphere-like regions are not common anatomical structures. Also, the registration is biased by noisy regions since they yield higher entropy and thus have a higher contribution to the MI value. Sundar et al. (2007) use a multi-resolution image sampling technique with octrees. An octree 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 to axis-aligned structures, which are uncommon in medical images.

3. Non-uniform sampling for MI estimation

We first define and motivate the use of curvelet-based sampling for the MI computations for registration. Let $f(x)$ and $g(x)$ be the

two images where x represents the image spatial domain. The MI between the images is defined as:

$$MI(f(x),g(T(x))) = H(f(x)) + H(g(T(x))) - H(f(x),g(T(x))) \quad (1)$$

where $H(f(x))$ and $H(g(T(x)))$ are the marginal entropies, and $H(f(x),g(T(x)))$ is the joint entropy of $f(x)$ and $g(T(x))$. The registration goal is to find a transformation T that maximizes the MI based metric between the images:

$$\arg\max_T MI(f(x),g(T(x))) \quad (2)$$

An image $f(x)$ is an n-dimensional random variable; its marginal entropy is defined as:

$$H(f(x)) = H(f(x_1),f(x_2),f(x_3),\ldots,f(x_n)) \quad (3)$$

When all the voxels $f(x_i)$ in the image are independent and identically distributed (i.i.d) random variables, the image entropy is:

$$H(f(x)) = \sum_x H(f(x)) \quad (4)$$

Thus, the image marginal entropy can defined as:

$$H(f(x)) = -\sum_{i=1}^{N} p(i)\log(p(i)) \quad (5)$$

where N is the number of possible intensity values, i is the image intensity value, and $p(i)$ is its probability. Similarly, the joint entropy of images $f(x)$ and $g(T(x))$ is defined as:

$$H(f(x),g(T(x))) = -\sum_{i=1}^{N}\sum_{j=1}^{M} p(i,j)\log(p(i,j)) \quad (6)$$

where N and M are the number of possible intensity values in images f and g, respectively, i and j are the indices of the intensity values in the images, and $p(i,j)$, is the joint intensity probability. Standard density estimators widely used in medical image registration are the histogram-based and the Parzen window based estimators (Pluim et al., 2003). To simplify the density estimation, these estimators assume that the voxel intensities are independent.

However, most of the voxel intensities are not independent from their neighborhood, and thus the real marginal entropy of the image $f(x)$ should be defined as:

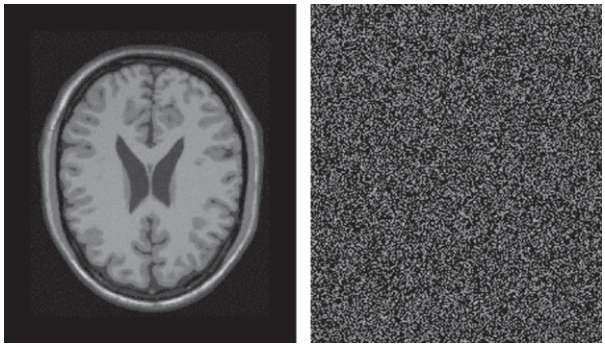
$$H(f(x)) = -\sum_{i=1}^{N} p(i|S)\log(p(i|S)) \quad (7)$$

where S is the dependency neighborhood of the voxels with intensity value i and $p(i|S)$ is the probability that intensity value i will appear in the image given its dependency neighborhood S. Clearly:

$$-\sum_{i=1}^{N} p(i|S)\log(p(i|S)) \leqslant -\sum_{i=1}^{N} p(i)\log(p(i)) \quad (8)$$

To illustrate this idea, consider a MRI T1 slice of a brain and the synthetic image created by randomly shuffling the voxel locations in the original image (Fig. 2). The entropy values of both images computed with Eq. (5) are identical (0.48). However, the original image compression rate using lossless compression is much higher (0.28) than that of the synthetic image (0.46) and thus its entropy value should be much lower when the inter-voxels dependencies are taken into account.

This bias occurs both in the estimation of the marginal entropies of f and g and in the estimation of the joint entropy between them. The marginal entropies bias can be assumed to be independent of the image pair alignment and thus does not affect the registration accuracy. However, the joint entropy computation incorrectly assumes that the intensity of voxel x_i in the image f depends only on the intensity of voxel $T(x_i)$ in the image g. In practice, it also depends on the intensity values of the neighboring voxels $T(x_i)$ in the image g (Sabuncu, 2006).



(a) MRIT1slice (b) After random shuffling

Fig. 2. Illustration of image entropy estimation: (a) slice of MRI T1 image from the Brainweb database (Collins et al., 1998), and; (b) synthetic image generated by randomly shuffling the voxel locations in (a).

Unfortunately, an explicit model for the inter-voxels dependencies is not available. However, for the entropy estimation, we can reduce this bias by selecting only voxels that are less dependent on their neighborhood. Formally, our goal is to design a sampling function s that minimizes the error between the estimated entropy $\widehat{H_s}(f(x))$ using the sampling mask s and the real image entropy $H(f(x))$:

$$\arg\min_s \|\widehat{H_s}(f(x)) - H(f(x))\|^2 \quad (9)$$

Note that the real entropy H cannot be directly computed, as the inter-voxels dependencies are unknown. The sampling function s should select voxels that are less dependent on their neighborhood.

4. The curvelet transform

Wavelets have proven to be very useful tool in signal processing, as they provide sets of sub-band decompositions of one-dimensional (1D) piecewise smooth functions (Mallat, 1999). Sub-band analysis captures most of the signal samples interdependency (Davis and Nosratinia, 1998). Thus, wavelets can be used for accurate density estimation from both noisy and dependent samples (Donoho et al., 1996; Masry, 1994).

However, these wavelets properties do not extend directly to 2D and 3D piecewise smooth functions with discontinuities, as those present in images. Several works describe wavelet extensions to 2D and 3D (Candes et al., 2005a; Freeman and Adelson, 1991; Do and Vetterli, 2005). Among these transforms, the Curvelet transform (Candes et al., 2005a; Ying et al., 2005) is a nearly optimal sparse representation of typical objects f that are C^2 except for discontinuities along piecewise C^2 non axis-aligned edges. In addition, the error bound between the original function f and the reconstructed function f_n^c from n curvelet coefficients is (Candes and Donoho, 2004):

$$\|f - f_n^c\|_{L_2}^2 \leqslant K \cdot n^{-2} \cdot (\log n)^3, n \to \infty \quad (10)$$

where K is a constant and n is the number of reconstruction coefficients. This approximation is within a poly-log factor of the optimal number of samples, which is asymptotically n^{-2} (Mallat, 1999). The curvelet-based representation is far more sparse than the image Fourier decomposition convergence $\left(n^{-\frac{1}{2}}\right)$, and of that of the wavelet decomposition convergence (n^{-1}). This indicates that the curvelet-transform coefficients better identify the voxels that are less dependent on their neighbor image regions. Consequently, the curvelet

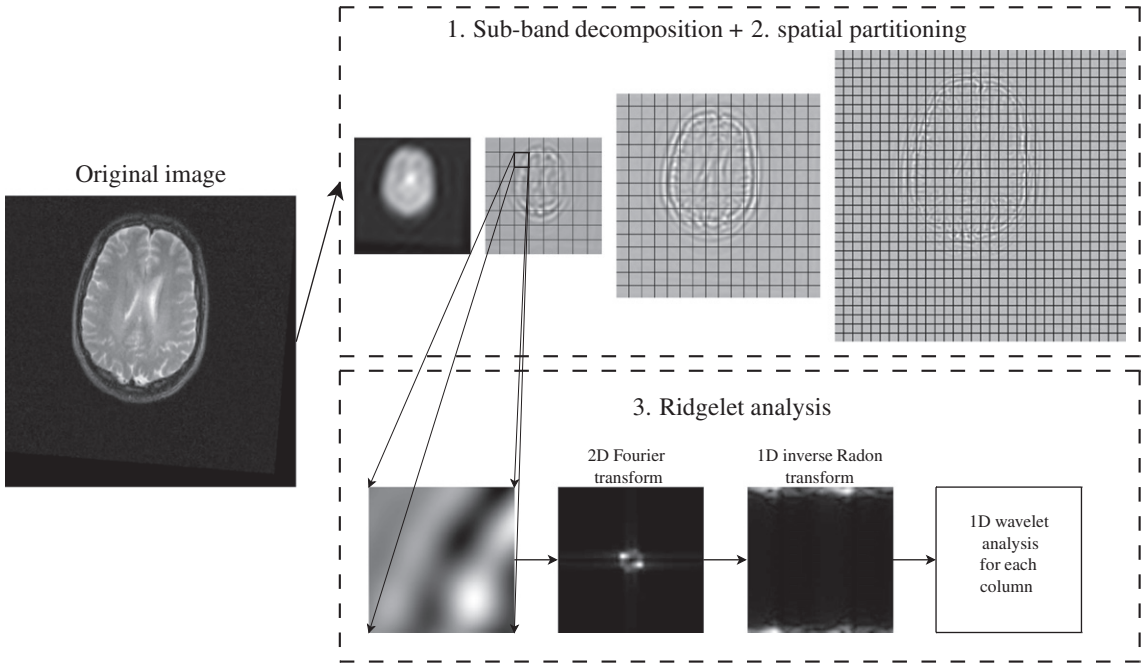


Fig. 3. The curvelet transform computation. The original image (left) is filtered using a high-pass filter to enhance the images edges. Then, the high-pass filtered image is decomposed into sub-bands (1) to allow the analysis of the edges at different scales. Each sub-band is smoothly partitioned into square regions (2) so that each square is analyzed separately. The ridgelet transform is then applied to each region (3). It uses the 2D Fourier transform followed by the 1D inverse Radon transform to generate a set of 1D representations of edges in different orientations. Each resulting 1D signal is then analyzed with 1D wavelet-based analysis to detect the edge. This scheme allows the detection of non-axis aligned edges in different orientations and scales.

coefficient supporting regions should yield a better estimation of the image entropy and therefore a more accurate registration.

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 $C(\ell,w,k)$ representing 2D discontinuities, where ℓ is the scale index, w represents the related orientation, and $k = [k_1,k_2]$ is the spatial location in the image. The curvelet decomposition of an image is constructed in three steps (Fig. 3):

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. *Spatial partitioning:* Each sub-band is smoothly windowed into square regions at each 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 2D discontinuities of scale ℓ and orientation w at spatial location k. The coefficients are computed by applying the FDCT to the image:

$$C(\ell,w,k) = \mathrm{FDCT}[f(x)] \tag{11}$$

where x are the spatial coordinates of the image f. This representation directly extends to the 3D spatial domain, where $k = [k_1,k_2,k_3]$ (Ying et al., 2005).

5. Curvelet-based sampling mask computation

We now describe how to compute a curvelet-based sampling function that minimizes the error between the estimated and the real entropy.

The goal of the sampling function is to identify the most informative regions at multiple scales based on the image curvelet decomposition. The sampling function computation consists of three steps: (1) image curvelet decomposition; (2) discontinuities coefficients enhancement, and; (3) informative regions identification at each scale.

First, we apply the curvelet transform (Eq. (11)) to the image to obtain a set of curvelet coefficients C at different scales ℓ, orientations w, and spatial locations k. The resulting coefficients are then enhanced by applying a soft-thresholding function to the curvelet coefficients:

$$F(\alpha) = a[S(c(\alpha-b)) - S(-c(\alpha+b))] \tag{12}$$

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

$$a = 1/[S(c(1-b)) - S(-c(1+b))]$$

which scales the curvelet coefficients to the $[-1,1]$ range. Coefficients with small absolute values are considered noise and are thus set close to zero. Those with large absolute values correspond to edges and thus are set close to ±1.

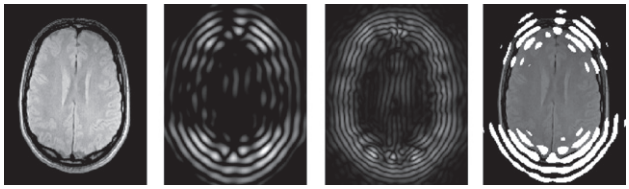
To identify the informative regions at scale m, we retain the curvelet coefficients of this scale and set all others to zero. The resulting curvelet coefficients $\widetilde{C}_m$ are:

$$\widetilde{C}_m(\ell,w,k) = \begin{cases} F(C(\ell,w,k)) & \text{for } \ell = m \\ 0 & \text{for } \ell \neq m \end{cases} \tag{13}$$

Next, we apply the inverse curvelet transform IFDCT to the coefficients:

$$\tilde{s}_m(x) = \mathrm{IFDCT}[\widetilde{C}_m(\ell,w,k)] \tag{14}$$

The resulting function $\tilde{s}_m$ describes a new image in which prominent discontinuities are represented by voxels with large absolute values.



(a) original image (MR-PD) (b) discontinuities at coarse scale (c) discontinuities at fine scale (d) final coarse curvelet mask

Fig. 4. Illustration of the curvelet-based sampling mask computation: (a) original MR-PD image; (b) discontinuities at a coarse scale generated from the thresholded curvelets features; (c) discontinuities at a finer scale, and; (d) resulting curvelet-based mask at the coarse scale, overlayed on a darkened version of the original image. The mask is shown in bright white.

We generate a binary sampling mask $M_{m,p}$ from $\tilde{s}_m$ at each scale from the resulting images by thresholding. We select the desired percentage of voxels p with the largest absolute values:

$$M_{m,p} = \{x : |\tilde{s}_m(x)| > P_p[|\tilde{s}_m(x')|]\} \tag{15}$$

where $P_p[|\tilde{s}_m(x)|]$ is the pth percentile of the intensities appearing in image $|\tilde{s}_m(x')|$, sorted in descending order. Note that taking the voxels with the highest absolute value causes the mask to include the edge voxels and their neighborhood voxels. It thus better approximates the i.i.d assumption in the entropy computations. Fig. 4 illustrates the mask computation process.

6. Curvelet-based registration method

Our 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 sampling mask may reduce the overlap between the images significantly, and thus yield an erroneous registration. Therefore the mask can be used only when the images are coarsely aligned. e.g. when their mean TRE is ±3 mm.

We use the Normalized Mutual Information (NMI):

$$NMI(f(x), g(T(x))) = \frac{H(f(x)) + H(g(T(x)))}{H(f(x), g(T(x)))} \tag{16}$$

for both steps, where H is the entropy between the source image $f(x)$ and the target image $g(T(x))$ with respect to rigid transformation T. The target image intensities $g(T(x))$ are estimated by linear interpolation. The NMI is computed with Parzen-windowed histogram-based estimators of the image intensity probability distributions with histogram smoothing (Mattes et al., 2003; Thévenaz and Unser, 2000).

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 a 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 step.

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

$$\widehat{P}(f(x) = i_t, g(T(x)) = i_s) = \frac{1}{N} \sum_{x \in M_{1,p}} K(f(x) - i_t, g(T(x)) - i_s) \tag{17}$$

where N is the total number of voxels in the mask $M_{1,p}$, and $K(\cdot,\cdot)$ is a cubic spline-based Parzen windowing kernel (Mattes et al., 2003). The optimization algorithm, parameter settings, and halting conditions are the same for both steps.

7. Experimental results

To evaluate the effectiveness of our sampling and registration method, we conducted two experiments on clinical multi-modal images. The first one quantifies the performance of curvelet-based sampling rigid registration between T1 and T2-weighted MRI images and compares it to uniform (Wells et al., 1996) and gradient-based (Sabuncu and Ramadge, 2004) sampling. The second quantifies the performance of curvelet-based sampling rigid registration between CT and MRI images from different protocols and compares it to uniform and gradient-based sampling registration.

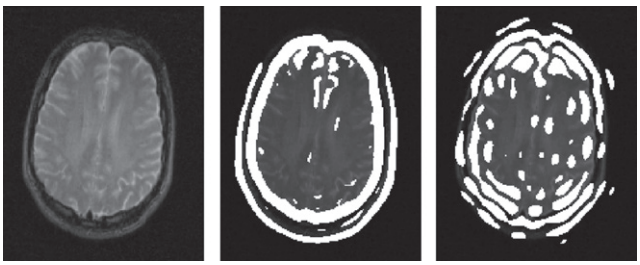
7.1. Images

We use the multi-modal Retrospective Image Registration Evaluation (RIRE) Project database, commonly known as the Vanderbilt Database (West et al., 1997). The RIRE project is a public domain platform that is used to compare retrospective multi-modal registration techniques developed around the world. The database 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). Since the images are real and not synthetic ones, they include acquisition noise and intensity inhomogeneity, which make the registration task more challenging.

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.

Although the RIRE database is relatively old, it is, to the best of our knowledge, the only database for which high-accuracy ground-truth transformations from attached fiducial markers are available. Ground-truth transformations from other publicly available datasets are much less accurate and include the inherent human operator or registration algorithm error used to generate them.

Since the image voxels in the RIRE datasets are anisotropic ($x = y = 1.27$ mm, $z = 4.11$ mm), we preprocess them to create isotropic versions for the sampling mask computation only. We apply



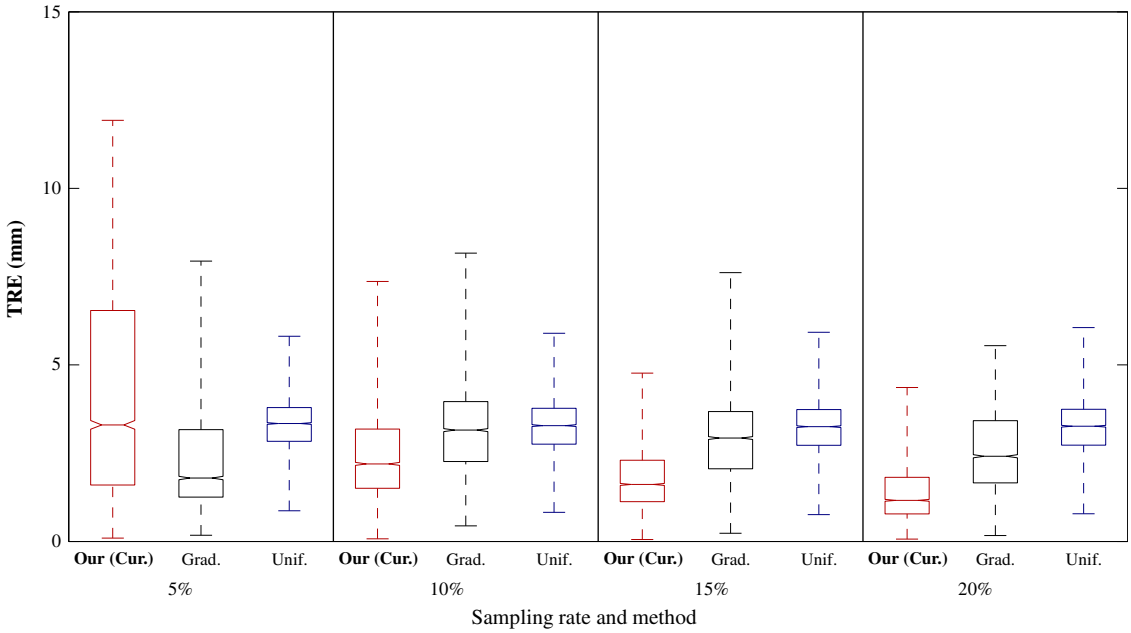
(a) original image (b) gradient mask (c) curvelet mask

Fig. 5. Illustrative example: (a) original MRI-T2 slice; (b) gradient-based mask, and; (c) curvelets-based mask. All masks are overlayed on the original image in bright white. The gradient-based mask captures mostly the skull due to its strong gradient-magnitude, while our curvelets-based mask captures also less strong edges inside the brain.

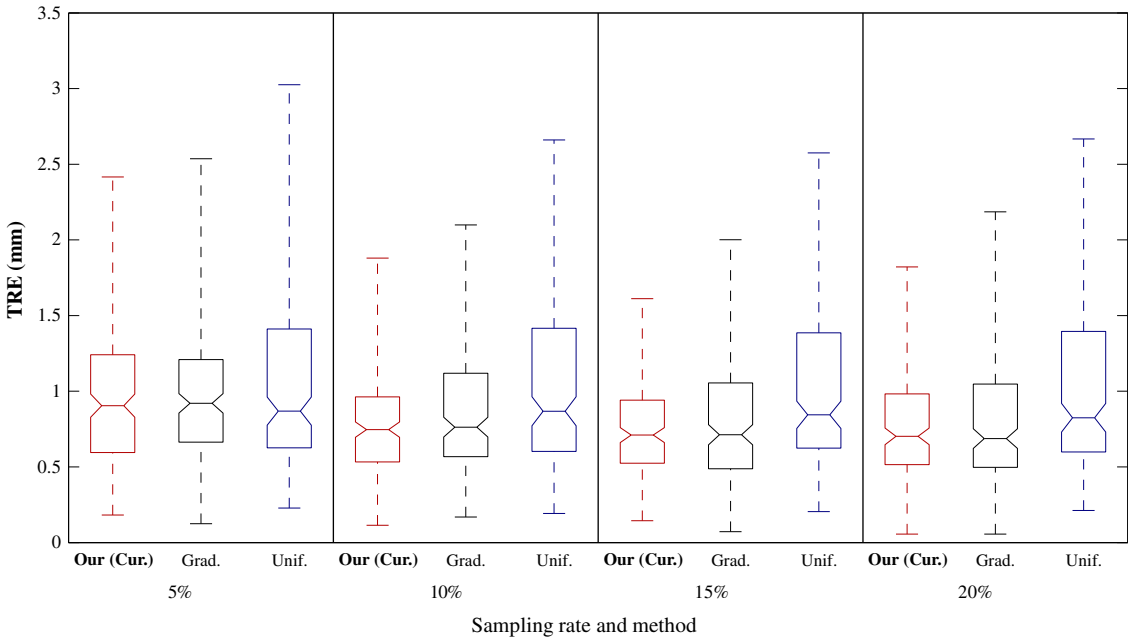
a 3D curvelet transform on the entire volume and generate from it the corresponding 3D regions representation. We then select the voxels with the largest absolute values to generate the image sampling mask. The soft-threshold function parameters were set experimentally to $b = 0.1$ and $c = 40$ for all images. For the 3D gradient-based sampling (Sabuncu and Ramadge, 2004), we compute 3D gradient magnitudes by filtering the entire volume in the x, y, and z directions with the first-order derivative of a Gaussian kernel ($\sigma = 3$, voxels = 3.8 mm). The gradient mask is then created by selecting the voxels with the largest absolute values.

Fig. 5 shows a representative example of the mask generated by our method and by the gradient-based method on a T2 slice from the RIRE database. Note that although the image includes a relatively large amount of noise and intensity inhomogeneity, our mask successfully identifies the informative regions inside the image.

We implemented our curvelet-based rigid registration method with the curvelet package (Candes et al., 2005b) and with the Elastix software library (Klein et al., 2010). In all cases, we used the NMI measure from Elastix with 32 histogram bins and a Parzen-window



(a) MRI T1-MRI T2 TRE distribution box-plots



(b) CT-MRI TRE distribution box-plots

Fig. 6. Box-plots of the registration results. The results of our method (red) and those of the gradient-based (black) and uniform sampling (blue) with respect to the sampling percentile used in the masks are shown. On each box-plot the central mark is the median, the edges of the box are the 25th and 75th percentiles, and the whiskers extend to the most extreme TRE measured on the targets. (a) results for 1000 MRI T2-T1 registrations, and; (b) results for 20 CT-MRI {T1,T2,PD} registrations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

estimator with a first-order B-spline kernel. A regular-step gradient descent optimizer with initial step length of 1 and maximum of 200 steps were used for all registrations.

For the coarse registration step, we used stochastic uniform sampling with 25% of the 3D image voxels. For the fine registration step, we generated several masks with different sampling percentiles (5–20%) of the voxels with each sampling method and used them in the registration. The number of scales used for masks generation depends on the image size. For the images used in this experiment, only the coarsest scale curvelet-based mask was used for registration. We found that using finer scale masks induced additional noise to the registration process and yielded less accurate results. Note that since the images are clinical images with both noise and intensity inhomogeneity, the results reflect our method accuracy in actual clinical conditions.

7.2. MRI T2-T1 rigid registration

For MRI T2-MR T1 registration, we use the training set of images from the RIRE database. For this set, the ground-truth transformations between the CT and the MR images are available as described above. Initially, we computed the MRI T2-T1 ground-truth transformation using matrix multiplication. Then, we applied 1000 synthetic transformations to the T1 image in the range of $[-23,23]$ degrees for the rotation and in the range of $[-160,160]$ mm for the translation about each axis. Finally, the images were registered using our method and compared to the uniform and gradient-based sampling registration results.

The TRE was measured on 10 clinically relevant points defined by an expert on the T2 image. Fig. 6a summarizes the results of our method and those of the gradient-based and uniform sampling with respect to the sampling percentile used in the masks.

For uniform sampling, the number of samples does not improve the registration accuracy as suggested in Klein et al. (2007). However, for both gradient-based and curvelet-based sampling, the number of samples does improve the accuracy. For the 5% sampling percentile, curvelet-based sampling performs worse than uniform and gradient-based sampling since the curvelet mask did not include enough information and the gradients mask includes most of the skull surface which helps it to perform the registration. For the other sampling percentiles, curvelet-based sampling consistently yielded a smaller median TRE with a much tighter distribution compared to that of the gradient-based mask. The most significant improvement of the curvelet-based sampling was recorded in the 20% sampling percentile, where the mean TRE was improved by more than 1.5 mm, which is ~50% improvement, compared to the other sampling methods.

7.3. CT-MRI rigid registration

In this experiment, following the RIRE evaluation framework (West et al., 1997), the goal was to compute 20 registrations: 6 CT-MRI T1 registrations, 7 CT-MRI T2 registrations, and 7 CT-MRI PD registrations. The resulting transformations for each registration were then submitted to the RIRE website (West et al., 1997) for validation. The TRE was measured on 10 clinically relevant points defined by the database providers, and were blind to us.

Fig. 6b summarizes the registration results of our method (red) and those of the gradient-based and uniform sampling with respect to the sampling percentile used in the masks. As in the MRI T2-T1 experiment, the number of samples does not improve the registration accuracy using uniform sampling. However, for both gradient-based and curvelet-based sampling, the number of samples does improve the accuracy. Curvelet-based sampling consistently yielded a smaller median registration error for all sampling percentiles, with about 0.3 mm improvement for the 15% sampling percentile compared to the uniform sampling, and a tighter TRE distribution compared to the gradient-based sampling. The most significant improvement of the curvelet-based sampling was recorded on the maximal error, which was reduced from 2.57 mm for uniform sampling to 1.61 mm, an improvement of about 1 mm (~40%). Using our method with 15% sampling yields the following results: mean = 0.73 mm, std = 0.3 mm, median = 0.71 mm, max = 1.61 mm. These are the best results recorded on the RIRE website to date, compared to other published methods evaluated on more than 5 registrations, including methods that used physical coordinates based uniform sampling (Thévenaz et al., 2008) and parameters optimization (Hahn et al., 2010). Using more than 15% of the voxels with the curvelet-based sampling scheme reduced the overall accuracy although the reduction is negligible. This is explained by the fact that these voxels have low curvelet coefficients values, and are thus more related to noise.

The registration times were almost linearly dependent on the number of samples. Both gradient-based and curvelet-based sampling requires almost the same time for each sampling percentile (120 s for 15% sampling percentile), while uniform sampling required relatively less time (90 s for 15% sampling percentile), but yielded less accurate results. The average mask computation time was on average 10.7 s (std = 0.07 s) for each registration on a dual-processor 2 GHz PC with 2 GB RAM running Linux.

8. Conclusions

This paper presents a new non-uniform sampling method for the accurate estimation of the mutual information between two images. The method uses the 3D Fast Discrete Curvelet Transform to identify the edges inside the image and therefore to reduce the voxels' interdependency for more accurate estimation of the image entropy. 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 uniform and gradient-based sampling for both MR T2-MR T1 and CT-MR multi-modal rigid brain image registration.

We are planning to apply our method to other multi-modal linear image registration tasks, and with larger span of modalities such as fMRI.

Acknowledgments

This research was partially funded by the European FP7 ROBOCAST Project No. 215190. Moti Freiman is also supported in part by the Hebrew University Hoffman Leadership and Responsibility Fellowship Program. 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

Bardera, A., Feixas, M., Boada, I., Sbert, M., 2006. High-dimensional normalized mutual information for image registration using random lines. In: WBIR'06. LNCS, vol. 4057. Springer, pp. 264–271.

Bhagalia, R., Fessler, J., Kim, B., 2009. Accelerated nonrigid intensity-based image registration using importance sampling. IEEE Trans. Med. Imag. 28 (8), 1208–1216.

Butz, T., Thiran, J., 2001. Affine registration with feature space mutual information. In: MICCAI'01. LNCS, vol. 2208. Springer, pp. 549–556.

Candes, E., Donoho, D., 2004. New tight frames of curvelets and optimal representations of objects with piecewise c^2 singularities. Commun. Pure Appl. Math. 57 (2), 219–266.

Candes, E., Demanet, L., Donoho, D., Ying, L., 2005a. Fast discrete curvelet transforms. Multiscale Model. Simul. 5, 861–899.
Candes, E., Demanet, L., Donoho, D., Ying, L., 2005b. The curvelet transform software package. <http://www.curvelet.org>.
Collignon, A., Vandermeulen, D., Suetens, P., Marchal, G., 1995. Automated multimodality medical image registration using information theory. In: Proc. of the 14th Int. Conf. on Inf. Processing in Med. Imaging (IPMI'95), vol. 3, pp. 263–274.
Collins, D., Zijdenbos, A., Kollokian, V., Sled, J., Kabani, N., Holmes, C.A.C.E., 1998. Design and construction of a realistic digital brain phantom. IEEE Trans. Med. Imag. 17 (3), 463–468.
Davis, G., Nosratinia, A., 1998. Wavelet-based image coding: an overview. Appl. Comput. Control Signals Circ. 1 (1).
Do, M., Vetterli, M., 2005. The contourlet transform: an efficient directional multiresolution image representation. IEEE Trans. Image Process. 14 (12), 2091–2106.
Donoho, D.L., Johnstone, I.M., Kerkyacharian, G., Picard, D., 1996. Density estimation by wavelet thresholding. Ann. Stat. 24 (2), 508–539.
Freeman, W., Adelson, E., 1991. The design and use of steerable filters. IEEE Trans. Patt. Anal. Mach. Intell. 13 (9), 891–906.
Gan, R., Chung, A., Liao, S., 2008. Maximum distance-gradient for robust image registration. Med. Image Anal. 12 (4), 452–468.
Gering, D., Nabavi, A., Kikinis, R., Hata, N., O'Donnell, L., Grimson, W.F.A.J., Black, P., Wells, W.I., 2001. An integrated visualization system for surgical planning and guidance using image fusion and an open MR. J. Magn. Reson. Imag. 13, 967–975.
Greve, D.N., Fischl, B., 2009. Accurate and robust brain image alignment using boundary-based registration. NeuroImage 48 (1), 63–72.
Hahn, D.A., Daum, V., Hornegger, J., 2010. Automatic parameter selection for multimodal image registration. IEEE Trans. Med. Imaging 29 (5), 1140–1155.
Hajnal, J., Hill, D., Hawkes, D. (Eds.), 2001. Medical Image Registration. CRC Press. ISBN 0849300649.
Holden, M., Griffin, L., Saeed, N., Hill, D., 2004. Multi-channel mutual information using scale space. In: MICCAI'04. LNCS, vol. 3216. Springer, pp. 797–804.
Joskowicz, L., Shamir, R., Freiman, M., Shoham, M., Zehavi, E., Umansky, F., Shoshan, Y., 2006. Image-guided system with miniature robot for precise positioning and targeting in keyhole neurosurgery. Comput. Aid. Surg. 11 (4), 181–193.
Klein, S., Staring, M., Pluim, J., 2007. Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines. IEEE Trans. Image Process. 16 (12), 2879–2890.
Klein, S., Staring, M., Murphy, K., Viergever, M., Pluim, J., 2010. elastix: a toolbox for intensity-based medical image registration. IEEE Trans. Med. Imag. 29 (1), 196–205. <http://elastix.isi.uu.nl>.
Legg, P.A., Rosin, P., Marshall, A., Morgan, J., 2009. A robust solution to multi-modal image registration by combining mutual information with multi-scale derivatives. In: MICCAI'09. LNCS, vol. 5761. Springer, pp. 616–623.
Luan, H., Qi, F., Xue, Z., Chen, L., Shen, D., 2008. Multimodality image registration by maximization of quantitative-qualitative measure of mutual information. Pattern Recogn. 41 (1), 285–298.
Maintz, J., Viergever, M., 1998. A survey of medical image registration. Med. Image Anal. 2 (1), 1–36.
Mallat, S., 1999. A Wavelet Tour of Signal Processing. Academic Press.
Masry, E., 1994. Probability density estimation from dependent observations using wavelets orthonormal bases. Stat. Probabil. Lett. 21 (3), 181–194.
Mattes, D., Haynor, D., Vesselle, H., Lewellen, T., Eubank, W., 2003. PET-CT image registration in the chest using free-form deformations. IEEE Trans. Med. Imag. 22 (1), 120–128.
Modersitzki, J., 2004. Numerical Methods for Image Registration. Oxford University Press.
Pauly, O., Padoy, N., Poppert, H., Esposito, L., Navab, N., 2009. Wavelet energy map: a robust support for multi-modal registration of medical images. In: IEEE Conf. Comp. Vis. Patt. Rec. (CVPR'09).
Pluim, J., Maintz, J., Viergever, M., 2000. Image registration by maximization of combined mutual information and gradient information. IEEE Trans. Med. Imag. 19 (8), 809–814.
Pluim, J., Maintz, J., Viergever, M., 2003. Mutual-information-based registration of medical images: a survey. IEEE Trans. Med. Imag. 22 (8), 986–1004.
Rueckert, D., Clarkson, M., Hill, D., Hawkes, D., 2000. Non-rigid registration using higher-order mutual information. In: Proc. SPIE Medical Imaging 2000: Image Processing, pp. 438–447.
Rui, X., Yen-Wei, C., 2007. Wavelet-based multiresolution medical image registration strategy combining mutual information with spatial information. Int. J. Innov. Comput. Inform. Control (IJICIC) 3 (2), 285–296.
Russakoff, D., Tomasi, C., Rohlfing, T., Maurer, C.R.J., 2004. Image similarity using mutual information of regions. In: ECCV'04. LNCS, vol. 3023. Springer, pp. 596–607.
Sabuncu, M., 2006. Entropy-based image registration. Ph.D. thesis, Dept. of Electrical Engineering, Princeton University, Princeton, NJ.
Sabuncu, M., Ramadge, P., 2004. Gradient based non-uniform subsampling for information-theoretic alignment methods. In: Proc. of the 26th Int. Conf. of the IEEE Eng. in Medicine and Biology Society (IEMBS'04), vol. 1, pp. 1683–1686.
Shams, R., Kennedy, R., Sadeghi, P., Hartley, R., 2007. Gradient intensity-based registration of multi-modal images of the brain. In: IEEE Int. Conf. on Comp. Vision (ICCV'07).
Staring, M., van der Heide, U.A., Klein, S., Viergever, M.A., Pluim, J., 2009. Registration of cervical MRI using multifeature mutual information. IEEE Trans. Med. Imaging 28 (9), 1412–1421.
Studholme, C., Hill, D., Hawkes, D., 1999. An overlap invariant entropy measure of 3D medical image alignment. Pattern Recogn. 32 (1), 71–86.
Sundar, H., Shen, D., Biros, G., Xu, C., Davatzikos, C., 2007. Robust computation of mutual information using spatially adaptive meshes. In: MICCAI'07. LNCS, vol. 4791. Springer, pp. 950–958.
Thévenaz, P., Unser, M., 2000. Optimization of mutual information for multiresolution image registration. IEEE Trans. Image Process. 9 (12), 2083–2099.
Thévenaz, P., Bierlaire, M., Unser, M., 2008. Halton sampling for image registration based on mutual information. Sampl. Theory Signal Image Process. 7 (2), 141–171.
Webster, G., Kilgallon, J., Ho, K., Rowbottom, C., Slevin, N., Mackay, R., 2009. A novel imaging technique for fusion of high-quality immobilised MR images of the head and neck with CT scans for radiotherapy target delineation. Br J. Radiol. 82 (978), 497–503.
Wells, W., Viola, P., Atsumi, H., Nakajima, S., Kikinis, R., 1996. Multi-modal volume registration by maximization of mutual information. Med. Image Anal. 1 (1), 35–51.
West, J., Fitzpatrick, J., Wang, M., et al., 1997. Comparison and evaluation of retrospective intermodality brain image registration techniques. J. Comput. Assist. Tomogr. 21 (4), 554–566. <http://www.insight-journal.org/rire>.
Wong, A., Clausi, D., Fieguth, P., 2010. CPOL: complex phase order likelihood as a similarity measure for MR-CT registration. Med. Image Anal. 14 (1), 50–57.
Ying, L., Demanet, L., Candes, E., 2005. 3D discrete curvelet transform. In: Proc. of the 11th Int. Conf. on Wavelets.
Yujun, G., Cheng-Chang, L., 2006. Multi-modality image registration using mutual information based on gradient vector flow. In: Proc. of the 18th Int. Conf. on Pattern Recognition (ICPR'06), pp. 697–700.
Zheng, G., 2008. Effective incorporation of spatial information in a mutual information based 3D-2D registration of a CT volume to X-ray images. In: MICCAI'08. LNCS, vol. 5242. Springer, pp. 922–929.
Zitova, B., Flusser, J., 2003. Image registration methods: a survey. Image Vision Comput. 21 (11), 977–1000.