Wavelet Energy Map: A Robust Support for Multi-modal Registration of Medical Images *

Olivier Pauly[1], Nicolas Padoy[1,2], Holger Poppert[3], Lorena Esposito[3] and Nassir Navab[1]
[1]Computed Aided Medical Procedures, Technische Universität München, Germany
[2]LORIA-INRIA Lorraine, Nancy, France
[3]Neuro-Kopf Zentrum, Klinikum Rechts der Isar, Technische Universität München, Germany
{$pauly, padoy, navab$}@cs.tum.edu

Abstract

Multi-modal registration is the task of aligning images from an object acquired with different imaging systems, sensors or parameters. The current gold standard for medical images is the maximization of mutual information by computing the joint intensity distribution. However intensities are highly sensitive to various kinds of noise and denoising is a very challenging task often involving a-priori knowledge and parameter tuning. We propose to perform registration on a novel robust information support: the wavelet energy map, giving a measure of local energy for each pixel. This spatial feature is derived from local spectral components computed with a redundant wavelet transform. The multi-frequential aspect of our method is particularly adapted to robust registration of images showing ambiguities such as tissues, complex textures and multiple interfaces. We show the benefits of the wavelet energy map approach in comparison to the classical framework in 2D and 3D rigid registration experiments on synthetic and real data.

1. Introduction

Image registration is a crucial preprocessing step in all image analysis tasks in which information from various imaging sources needs to be combined. These sources of information can be acquisitions from different viewpoints of an object, at different times or with different sensors [20]. Establishing the correspondences between images acquired with different medical imaging modalities is a challenging task known as multi-modal registration. To identify the geometric transformation that maps the coordinate system of one modality to the other [21], objective functions that evaluate the quality of alignment known as similarity measures are optimized. The choice of the appropriate measure is not straightforward, because it implicitly models the relationship between the different images to register. Indeed, this measure quantifies how well images are registered according to the transformation parameters [16]. As modeling the physical relationship between different imaging modalities is very difficult, statistical measures have become more and more popular.

Since its introduction by Viola and Wells [19] and Collignon et al [3], mutual information remains the state of the art of multi-modal registration of medical images. Several other entropy-based measures have also been introduced: for example, the normalized version of mutual information proposed by Studholme et al. [18] or the Kullback-Leibler distance introduced by Chung et al. [2]. Furthermore, a quantitative-qualitative measure of mutual information has been presented by Luan et al. [12] to take the saliency of each image voxel into account. All these different entropy definitions use the same support of information: the intensity distribution. But image intensities are very inclined to be corrupted by noise due to different phenomena that can occur during the acquisition procedure. Indeed, medical images such as magnetic resonance suffer very often from different types of noise due to interferences between electronic devices, which can dramatically influence registration results. Image denoising is however a very challenging task, because the type of noise has to be known or modeled to perform an efficient filtering.

To gain in robustness, Gan and Chung [7] introduced a novel spatial feature named maximum distance-gradient-magnitude (MDGM) for rigid registration of medical images. Each pixel is characterized by the most dominant local variation and its intensity value. Again, taking into account intensity values can lead to misregistration in presence of noise. Because of its ability to extract features characterizing local frequency components, the Discrete Wavelet

*This research is partially funded by the Munich Center for Advanced Photonics and by Prof. Eckstein's Gefäßchirurgie Department, Klinikum Rechts der Isar.

Transform (DWT), whose main application is data compression, has been recently introduced in the field of image registration. Le Moigne et al. [14] and Sharman et al. [17] proposed to perform the registration on a feature space formed by the dominant local variations. In a coarse to fine strategy, wavelet coefficients are selected with a magnitude above a certain threshold. This selection method is also used by Hongli et al. [9] on approximation coefficients computed with a slightly different wavelet transform scheme. Using a Complex Wavelet Transform (CWT), Oubel et al. [15] present a 2 steps registration framework, in which a first alignment is done on low frequency and refinement on high frequency coefficients from the first decomposition level. But relying on these high frequency components is not a safe strategy, especially in the case of high frequency noise. Li et al. [11] propose an energy feature based on the coefficients of the first decomposition level computed with a Discrete Frame Wavelet Transform. Each pixel being only characterized by the highest part of the frequency spectrum, again, these features are not reliable in the presence of noise.

To take fully advantage of this kind of transforms, we combine the information contained in *all* sub-bands of the frequency spectrum. In this paper, we propose to perform the registration on a novel feature map we name *wavelet energy map* (WEM), whose computation is parameter-free and which is very robust to the noise present in the original images. The WEM measures the local signal energy at each pixel and is computed from local spectral components in its neighborhood. These spectral components are obtained with the redundant wavelet transform [4], which gives the best approximation of a space-frequency representation. For registration tasks, the energy probability distribution of the WEM is used as input for mutual information. The method does not require any additional a-priori information or parameter, but only a slightly increased initial computation. The multi-frequential aspect of this approach is especially adapted to registration of medical images presenting ambiguities such as tissues with complex textures or multiple interfaces. We demonstrate its value on a wide range of experiments on synthetic and real images. In the remaining of this paper, section 2 will define the wavelet energy map and the registration framework. Section 3 will present experiments demonstrating three properties of the WEM:

1. ***correctness***: energy and intensity maps give the same global maximum for mutual information
2. ***robustness***: registration on local energies is robust to noise
3. ***efficiency***: mutual information computed on wavelet energy maps outperforms the classical approach in terms of robustness for an equivalent accuracy

2. Methods

2.1. Problem statement

The goal of multi-modal image registration is to identify the geometric transformation that maps the coordinate system of one modality to the other. Let us consider two 2D images defined on the domains Ω_1 and Ω_2 with intensity functions $I_1 : \Omega_1 \subset R^2 \rightarrow R$ and $I_2 : \Omega_2 \subset R^2 \rightarrow R$. The registration task can be defined as a maximization problem, in which we want to estimate the best transform T according to a chosen similarity measure S computed on the discrete overlap domain $\Omega = \Omega_1 \cap T(\Omega_2)$:

$$T = \underset{T}{\operatorname{argmax}}\ S_\Omega(I_1, T(I_2)) \tag{1}$$

Since intensities are highly sensitive to noise, we propose to evaluate the similarity in a more robust feature space. We introduce a novel spatial feature map named *wavelet energy map* (WEM) giving a measure of local energy around each pixel of the original images. In the following parts, we define the concept of local energy and its computation from local spectral components extracted with a redundant wavelet transform. We will discuss the 2D case for better readability, the extension to three dimensions being straightforward. In that case, T is the composition of a translation and a rotation.

2.2. Energy vs. Intensity

In signal processing, the energy of a signal $x(t)$ is defined as:

$$E = \int_t |x(t)|^2 \tag{2}$$

and in the discrete case, for an image with intensity function I:

$$E = \sum_{i,j} |I(i,j)|^2 \tag{3}$$

Interpreting a zero image as a flat surface, energy can be understood as the work capacity accumulated during its deformation to the final relief. In the previous equation, energy is expressed in terms of intensities. Because they do not provide any contextual information, intensity values are not a safe support of information. Their variation frequencies, in contrary, offer a safer support by involving spatial context. Parseval's theorem guarantees that in a Hilbert space (we assume intensity functions being elements of a Hilbert space, for example $L^2(R^2)$) the energy of a signal x can also be determined from its frequency spectrum:

$$\int_t |x(t)|^2\, dt = \int_f |X(f)|^2\, df \tag{4}$$

with X being the Fourier transform of x. In the discrete case we obtain:

$$\sum_{i,j} |I(i,j)|^2 = \sum_f |F_I(f)|^2 \tag{5}$$

where F_I is the 2D Fourier transform of the image. The major drawback of the Fourier transform is its poor resolution in space domain. To define a *local* energy as a spatial feature, we need a frequency-space representation to know which spectral components exist at any given position in the image. A relatively new method introduced by Grossmann and Morlet [8] known as wavelet transform provides the best approximation of this space-frequency representation.

2.3. Extraction of local spectral components

2.3.1 The redundant wavelet transform

The traditional discrete wavelet transform (DWT) projects a signal onto an orthogonal wavelets basis. Its principle is to extract iteratively the information contained in each sub bands of the frequency spectrum [13]. In practice, the DWT is performed by passing the image through a cascade of orthogonal high pass (H) and low pass (L) filters to select each sub bands and analyze their content. The resulting decomposition coefficients are then down-sampled according to the Nyquist-Shannon sampling theorem as represented on Fig. 1. The original image is decomposed in 4 components: HH corresponds to the application of high pass filters in x and y directions, HL to high pass in x and low pass in y direction, LH to the contrary and LL to low pass filters in both directions. The LL component is then redecomposed in 4 components and the process is repeated. HH, HL and LH components are called details and LL approximation coefficients.

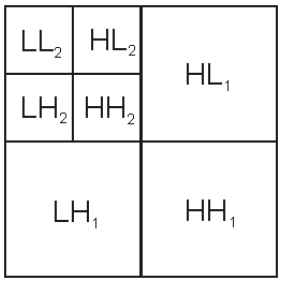

Figure 1. Discrete Wavelet Transform of a 2D image.

Orthogonality is a crucial property ensuring the most exact conservation of the spectral information. DWT does not only provide orthogonality between each sub bands, but also between their components. Its major drawback is the down-sampling operation that results in a loss of position information. Hence, we use another transform known as redundant discrete wavelet transform (RDWT) that basically removes the down-sampling operation. The RDWT also known as "Algorithme à trous" produces an over complete representation of the image and is considered as a better approximation of the continuous wavelet transform [4]. It is implemented by using a bank of filters (refer to the filter cascade on Fig. 2). In the one-dimensional case, the signal is filtered by a low l and a high pass h as shown below:

$$v_{j+1}[n] = \sum_{k=1}^{p} v_j[k] l_j[n - 2^j k] \quad (6)$$

$$w_{j+1}[n] = \sum_{k=1}^{p} v_j[k] h_j[n - 2^j k] \quad (7)$$

with v_{j+1} being the approximation and w_{j+1} the detail component at the decomposition level $j+1$ and p the size of the filter. This is analog to a classical filtering of the signal by iteratively inserting zeros, or in other words "holes" ("trous" in french) between all coefficients of the filters. In

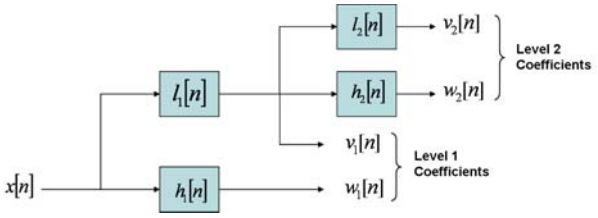

Figure 2. Filter Bank for Redundant Discrete Wavelet Transform of a 1D signal.

the two-dimensional case, each row and column of the original image are treated like a one dimensional signal. By introducing redundant information, the RDWT is not orthogonal but projects the signal onto a frame. A frame can be a stable and redundant representation of signals if its basis verifies the Heisenberg-Weyl condition [5]. The whole frequency axis is then covered by this representation and it can be considered according to Daubechies [5] as a quasi-orthogonal expansion. Such a representation helps to characterize textures of an image and increases the robustness to additive noise [6]. This redundancy has the main advantage to permit a better localization of each spectral components in the image: indeed, to each pixel corresponds a set of coefficients characterizing the local spectrum.

2.3.2 Choice of the wavelet basis

Since the RDWT removes the down-sampling operation, the spatial sampling rate is fixed across all scales. This gives to this transform a translational invariance property in contrary to the traditional DWT. Unfortunately the RDWT is not rotational invariant. To reduce the impact of this rotational non invariance, we focus on two types of wavelets that have a compact support: *orthogonal* wavelets from the Daubechies family and the *biorthogonal* Cohen-Daubechies-Feauveau 9/7 wavelets.

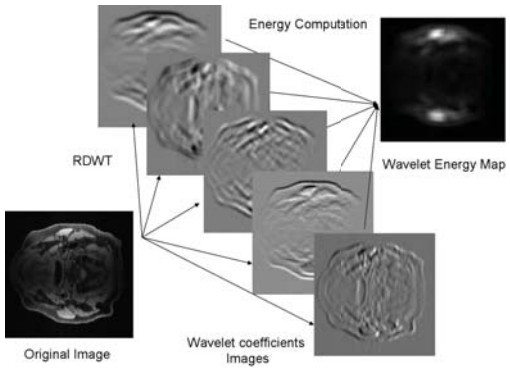

Figure 3. The wavelet energy map computation.

2.4. Local energy formulation

After decomposing an image I with a RDWT in m subbands with m being the maximum number of possible levels, we obtain for each pixel (i, j) a set of $3m + 1$ coefficients (3 detail components per level and the approximation from the last level) providing information on local frequency components. Let $w(i, j)$ be the coefficient vector:

$$w(i,j) = (w_1(i,j), w_2(i,j), \ldots, w_{3m+1}(i,j)) \quad (8)$$

Using Parseval's theorem, we can define a local energy $\mathcal{W}(i, j)$ computed from the local spectral components:

$$\mathcal{W}(i,j) = \|w(i,j)\|^2 = \sum_{k=1}^{3m+1} |w_k(i,j)|^2 \quad (9)$$

We name *wavelet energy map* (WEM) the array containing all values $\mathcal{W}(i, j)$. Its computation is summarized by Fig. 3.

2.5. Energy based registration framework

In terms of WEMs, equation 1 becomes:

$$T = \underset{T}{\operatorname{argmax}}\ S_\Omega(\mathcal{W}_1, T(\mathcal{W}_2)) \quad (10)$$

with $\mathcal{W}_1$ and $\mathcal{W}_2$ being the WEMs computed from both images to align. Since different imaging systems emphasize different characteristics of an object, the resulting WEMs will highlight different structures. As presented in the introduction, statistical similarity measures are the current standard in multi-modal registration. Thus, we propose to use the mutual information (MI) to evaluate the statistical relationship between both wavelet energy maps. First, we normalize them between 0 and 1:

$$\overline{\mathcal{W}}(i,j) = \frac{\mathcal{W}(i,j) - \min_{i,j}(\mathcal{W}(i,j))}{\max_{i,j}(\mathcal{W}(i,j))} \quad (11)$$

By dividing the domain $[0, 1]$ in N bins $\mathcal{B}_1, \ldots, \mathcal{B}_N$, we can determine the probability of a pixel x to fall in the bin $\mathcal{B}_k$:

$$p(x \in \mathcal{B}_k) = \frac{\#\left\{(i,j),\ \overline{\mathcal{W}}(i,j) \in \left[\frac{k}{N}\ \frac{k+1}{N}\right[\right\}}{\#\Omega} \quad (12)$$

where $\#\cdot$ is the cardinality operator. It is then possible to compare the shared amount of information in the energy maps of both images by using the classical definition of MI:

$$\begin{aligned} &MI(\overline{\mathcal{W}}_1, T(\overline{\mathcal{W}}_2)) = \\ &H(\overline{\mathcal{W}}_1) + H(T(\overline{\mathcal{W}}_2)) - H(\overline{\mathcal{W}}_1, T(\overline{\mathcal{W}}_2)) \end{aligned} \quad (13)$$

with $H(\overline{\mathcal{W}})$ being Shannon's definition of information entropy:

$$H(\overline{\mathcal{W}}) = -\sum_{k=1}^{N} p(x \in \mathcal{B}_k) log_2(p(x \in \mathcal{B}_k)) \quad (14)$$

The joint entropy $H(\overline{\mathcal{W}}_1, T(\overline{\mathcal{W}}_2))$ is defined by using the probability of the pixel x to respectively fall into the bins $\mathcal{B}_k$ and $\mathcal{B}_l$ in the maps $\overline{\mathcal{W}}_1$ and $T(\overline{\mathcal{W}}_2)$.

3. Experiments and Results

First, experiments on synthetic datasets show that energy and intensity maps give the same global maximum for mutual information. Further tests reveal the robustness of our approach to gaussian noise. Finally, 2D and 3D experiments on real medical datasets illustrate the benefits of a WEM based registration framework. The different wavelets transforms are based on the Matlab implementation by Gabriel Peyré and the Rice wavelet toolbox. To compute a consistent WEM, three conditions have to be fulfilled:

1. Images must have the same pixel size,
2. the domains where the RDWT is computed must have a size which is a factor of 2,
3. both images have to be decomposed in the same number of levels.

Convergence to the right solution depends much more on the topography of the search space offered by the similarity measure than on the optimizer. Hence, we can choose a Downhill-Simplex optimizer, that does not require any gradient information, to solve our measure maximization tasks. The quality of registration will be evaluated by using the *target registration error* (TRE). The TRE is computed by comparing the positions of a set of points $\{p_i,\ 1 \leq i \leq M\}$ after being mapped by the estimated transform T and by the ground truth transform G:

$$TRE = \frac{1}{M} \sum_{i=1}^{i=M} \|T(p_i) - G(p_i)\| \quad (15)$$

In the following, mutual information computed on wavelet energy map will be denoted by MEI (Mutual Energy Information), while the classical approach on intensity maps by MII (Mutual Intensity Information). Different MEI

based on Haar, Daubechies 4 (D4) and Cohen-Daubechies-Feauveau 9/7 (CDF) wavelet bases have been evaluated. All experiments were performed with MATLAB 7.5.0 on a Intel Core 2 Duo CPU 2.40 GHz.

3.1. Correctness

The goal of these experiments is to show that registration performed with MEI leads to the same global maximum than by using MII. Therefore, we use synthetic images and plot both measures to compare their global maxima and smoothness. For a better understanding and visualization of the results, we analyze separately rotation and translation. The images contain ambiguities resulting in several local maxima to demonstrate the superiority of our approach in terms of smoothness of the search space.

Experiment 1:

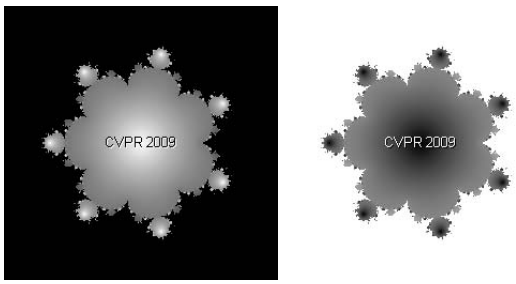

Figure 4. Experiment 1: the Mandelbrot fractal image and its inverse used for visualizing the similarity measure as function of the rotation parameter.

We use a Mandelbrot fractal with equation $f(z) = z^8 + c$ which has interesting multi-frequential characteristics. Indeed, it shows fine structures at arbitrary small scales, as observed in medical images presenting tissues with complex textures. This fractal shows an orientation ambiguity: we can distinguish 7 different global rotation maxima. Thus a "*CVPR 2009*" detail is added to give an orientation to the whole image (see Fig. 4). We compare MEI and MII similarity measures for this image and its inverse, both having a resolution of 256x256. Similarity values are plotted for a rotation angle varying between -90 and $+90$ degrees. As shown on Fig. 5 (left), the global maxima perfectly correspond for both approaches. Besides, the WEM emphasizes the right solution by smoothing other local maxima contrary to the intensity map.

Experiment 2:

In our second experiment, we use the following sum of cosinus to simulate multiple interfaces such as those observed in many medical images: $f(x) = a_1 cos(2\pi f_1 x) + a_2 cos(2\pi f_2 x) + a_3 cos(2\pi f_3 x)$. The resulting image shows ambiguities in both x and y directions. As before, we add the "*CVPR 2009*" detail giving a unique solution (see Fig. 6). We compare MEI and MII for this image and its inverse, both having a resolution of 128x128. Similarity values are plotted for translation parameters varying between -20 and $+20$ pixels in both directions. As shown on Fig. 7 the global maximum perfectly corresponds for both approaches. This experiment also reveals the abilities of the WEM to both emphasize the global maximum and offer a smoother search space, which are very valuable for optimization purposes.

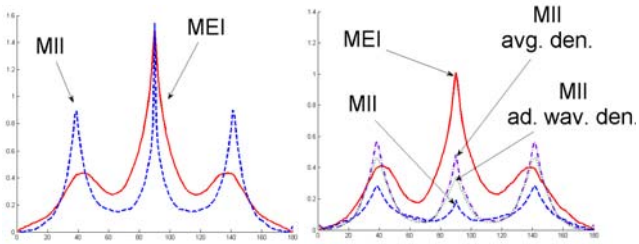

Figure 5. Experiment 1 (left) and Experiment 3 (right): plot of the similarity measures for rotation angles between -90 and $+90$ degrees. On this figure, D4 wavelet has been used to compute the WEM.

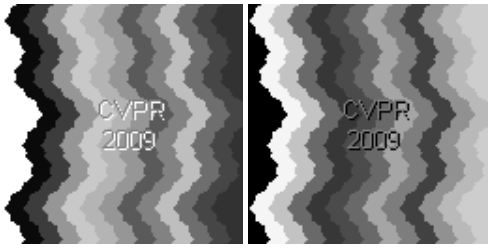

Figure 6. Experiment 2: the multiple interfaces image and its inverse used for visualizing the similarity measure as function of the translation parameters.

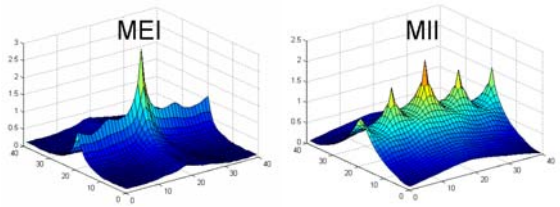

Figure 7. Experiment 2: Measures plotted for variations in translation. On this figure, D4 wavelet was used to compute the WEM.

3.2. Robustness to noise

The goal of these experiments is to argue for the superiority of MEI in terms of robustness in comparison to the classical approach, even when a denoising step has been performed prior to the similarity computation. Two denoising methods are used: averaging and adaptive wavelet denoising [1]. The later method is parameter-free and based on soft-thresholding of the wavelet coefficients. It was chosen for fair comparison with wavelet energy maps. Denoising being a challenging task usually involving a priori knowledge on the type of noise, a "denoising-free" measure such as MEI is very valuable for robust registration. We will use for the following experiments a gaussian

noise model, that affects independently all pixels of the images and thus highly corrupts the information content. We show in experiment 3 that MEI preserves its search space from distortions on the same synthetic images corrupted by gaussian noise. Experiment 4 shows its robustness to different levels of noise and is performed on an image without any ambiguity. A real registration framework is used for evaluation of the search space formed by the three transform parameters.

Experiment 3

The impact of gaussian noise on the smoothness of the search space of MEI, MII and MII preceded by a denoising step is analyzed. We use the same figures and setup than in the previous section, and add to all images an additive gaussian noise with $\sigma = 20\%$ of the maximum intensity value. Figures 5 (right) and 8 show that our approach

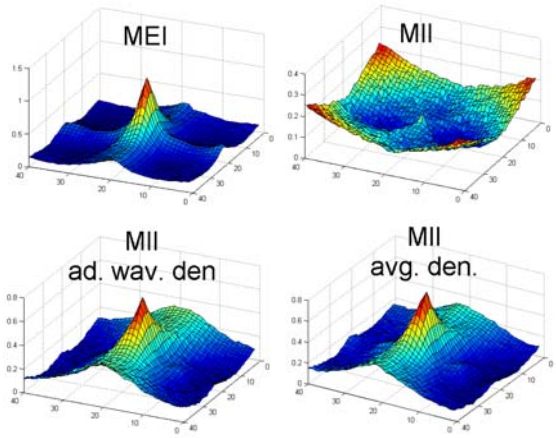

Figure 8. Experiment 3: From left to right, top to bottom: MEI, MII without denoising, MII with adaptative wavelet denoising and MII with averaging denoising for variation in translation. In this figure, D4 wavelet was used to compute the WEM.

preserves the global maximum in presence of gaussian noise unlike the others. Even when a denoising step is applied prior to the computation of the similarity measure, MEI offers a smoother search space with a meaningful global maximum. The rotational case even reveals that MII looses the right solution.

Experiment 4

In this experiment, we evaluate the robustness to an increasing amount of gaussian noise. MEI is compared to MII and MII preceded by a denoising step on an image without any ambiguity. We use the portrait of Lena with a resolution of 128x128 pixels. Registration to its inverse image is performed with a Downhill Simplex optimizer by starting from 10 random initial positions within the range -3 to $+3$ pixels translation and -3 to $+3$ degrees rotation. Fig. 9 shows the mean target registration error in function of the percentage of noise. For the classical computation of mutual information, the TRE increases dramatically in contrast to our approach that always presents the smallest error. Even when a denoising step is performed prior to the registration, results illustrate the benefits of MEI in terms of robustness.

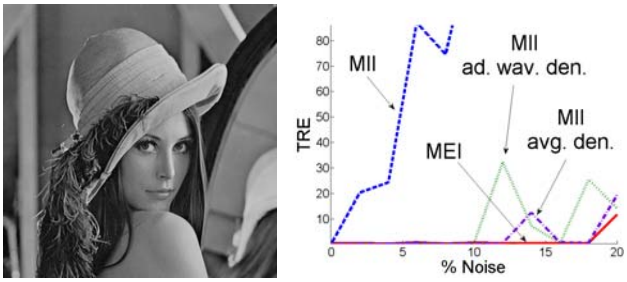

Figure 9. Registration performed on Lena image and its inverse to evaluate the robusness to noise of each method.

3.3. Efficiency on medical images

In this part, we evaluate the efficiency of our approach in 2D and 3D multi-modal registration experiments. First, 2D experiments are conducted on magnetic resonance (MR) images from sequences acquired with different system parameters, namely T1, T2, PD and TOF acquisitions. Then, 3D experiments are performed on MR and SPECT volumes with an increasing amount of noise. MEI computed with Haar, D4 and CDF wavelet bases are compared to the classical MII. To evaluate the registration efficiency of each method, we distinguish between *success rate* and *accuracy*. A registration is considered as successful when the final target registration error is inferior to a given threshold t_e. Otherwise, the approach did not converge in the neighborhood of the right solution. The accuracy is evaluated as the mean target registration error computed on the cases where all methods have converged under t_e.

The chosen multi-modal datasets contain ambiguities which can lead classical approaches to misregistration. Experimental results illustrate the ability of our method to cope with such ambiguities by emphasizing the right global maximum and smoothing other local extrema.

3.3.1 2D registration experiments: Real Magnetic Resonance datasets

While T1, T2, PD sequences provide different information related to tissue characteristics (refer to Fig. 10), TOF gives dynamic information related to the blood flow in arteries. The circular shape of the neck makes the registration task ambiguous in 2D. Indeed, to recover the right rotation parameter, registration can only rely on corresponding tissues or interfaces appearing in each modalities.

In the following experimental setup, 2D registration tests are conducted on all possible combinations of T1, T2, PD and TOF images taken from 8 patients. They have a resolu-

tion of of 128 x 128 with a pixel size of 1.25mm x 1.25mm. Knowing the ground truth position for each dataset, we give a random initial perturbation within the TRE range of 20mm. The threshold t_e is set to 10mm which corresponds to 50% of the initial TRE range. Results presented in table 1 reveal the benefits of our approach on real medical datasets: MEI shows the best success rates for an equivalent accuracy. In hard tasks such as registrations to TOF images, both success rate and accuracy are better. Even though MII is an accurate measure, it shows more local extrema than MEI when images present ambiguities. These local extrema trap the optimizer, leading thereby to more misregistration errors. In contrast, by smoothing these local extrema, our approach offers a better success rate.

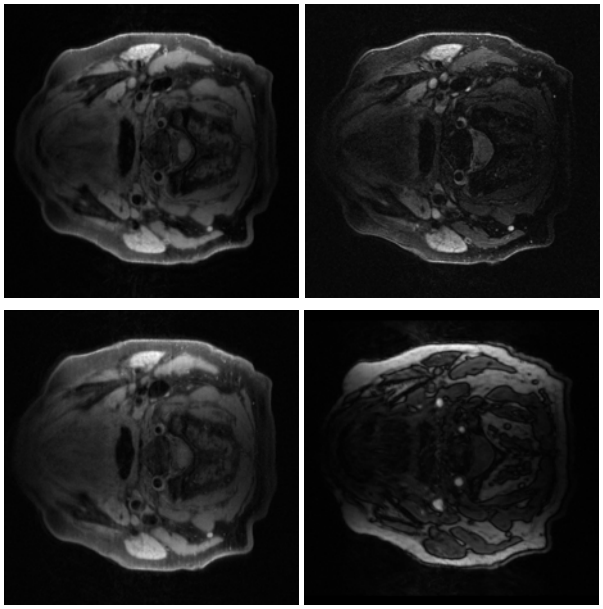

Figure 10. From left to right, and top to bottom: T1, T2, proton density (PD) spin echo sequences and Time of Flight (TOF) MR Angiography gradient echo sequence of the neck of the same patient.

3.4. 3D registration experiments: T1 Magnetic Resonance and SPECT-Tc volume

Single photon emission computed tomography (SPECT) is a nuclear medicine tomographic imaging technique used to provide information related to the blood flow. In 3D, SPECT volumes present a blurry cloud aspect with smooth intensity variations that do not correspond to any structure visible in the MR volume. This makes the recovery of transformation parameters for such a task challenging.

In the following experimental setup, 3D registration tests are conducted on MR and SPECT volumes of 4 patients taken from the Whole Brain Atlas online database [10]. They have an in-plane resolution of 128 x 128 with a voxel size of 1.67mm x 1.67mm x 1mm. When the size in z is not

Success rate in %				
	Haar MEI	D4 MEI	CDF MEI	MII
T1/TOF	87.50%	87.50%	87.50%	84.38%
T1/T2	100%	100%	100%	98.96%
T1/PD	100%	100%	100%	98.96%
T2/TOF	86.46%	87.50%	87.50%	85.42%
T2/PD	100%	100%	100%	98.96%
PD/TOF	87.50%	87.50%	87.50%	82.29%

Target Registration Error in mm					
		Haar MEI	D4 MEI	CDF MEI	MII
T1/TOF	*mean*	2.84	3.18	2.80	3.75
	std dev	1.69	1.96	0.92	2.16
T1/T2	*mean*	0.91	0.90	0.99	0.64
	std dev	0.64	0.49	0.59	0.55
T1/PD	*mean*	0.95	1.04	1.05	0.83
	std dev	0.55	0.55	0.52	0.58
T2/TOF	*mean*	2.99	3.25	3.19	3.26
	std dev	1.78	2.03	2.02	2.32
T2/PD	*mean*	1.14	1.02	1.25	0.60
	std dev	0.61	0.56	0.73	0.44
PD/TOF	*mean*	2.68	3.03	2.79	3.21
	std dev	1.82	1.91	1.61	2.32

Table 1. 2D registration: **success rate** and **final TRE** on T1, T2, PD and TOF images.

a power of 2, a zero-padding is performed at the boundaries of the volume before the RDWT. Knowing the ground truth for each dataset, an initial perturbation is applied within a range of 14mm of initial TRE. By using 20 initializations for each patient, and this for an increasing amount of noise, we can investigate the ability of each measure to converge towards the right solution and thereby assess their robustness. The threshold t_e is set to 7mm which corresponds to 50% of the initial TRE range. Results presented in Fig. 11 show that MEI offers better success rate for a better accuracy. With an increasing amount of noise, even though an averaging denoising step is performed prior to registration, our approach leads to better results. Since noise is localized in a small part of the frequency spectrum, its impact is minimized by the computation of the WEM. As in the 2D experiments, our method copes better with the ambiguities caused by the aspect of the SPECT volumes.

4. Conclusion

In this paper, we proposed to perform the registration on a feature map called *wavelet energy map* (WEM) instead of using the original image. We empirically showed that mutual information performed on the WEM leads to the same solution than the classical approach on intensity maps. Moreover, its multi-frequential aspect permits to emphasize the global maximum in ambiguous cases containing multiple local extrema, offering thereby a smoother search space, even in presence of noise. 2D and 3D registration experiments on real medical datasets

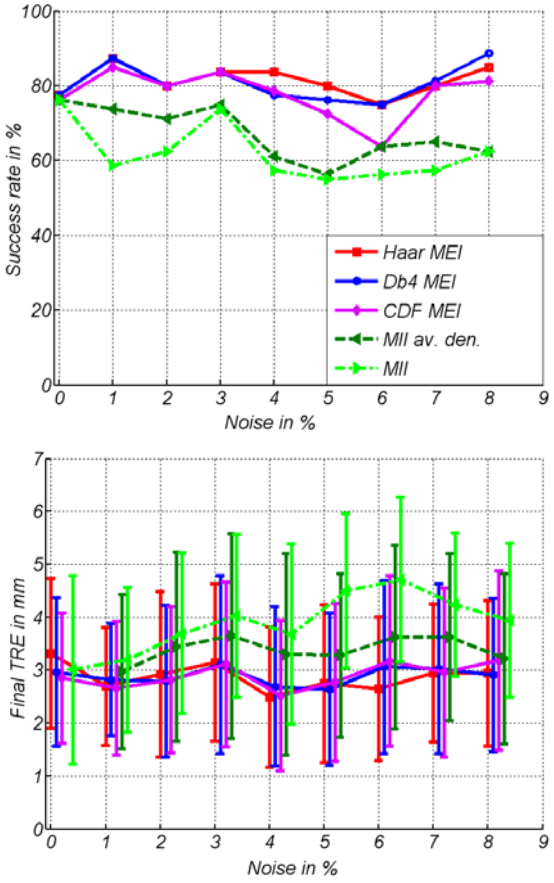

Figure 11. 3D registration: Plot of **success rate** and **final TRE** according to an increasing amount of noise for MR-SPECT volumes.

illustrated the efficiency of our approach in comparison to the classical framework which is more sensitive to noise and image ambiguities. In future work, we plan to address the rotational non-invariance issue of the redundant wavelet transform, for instance by using filters computed on more orientations.

Acknowledgments: We would like to thank Max Baust, Jose Gardiazabal, Pierre Georgel, Martin Groher and Diana Mateus for their valuable comments.

References

[1] S. G. Chang, B. Yu, and M. Vetterli. Adaptive wavelet thresholding for image denoising and compression. *IEEE Trans. on Image Processing*, 9(9):1532–1546, 2000. 5

[2] A. Chung, W. Wells, A. Norbash, and W. Grimson. Multi-modal image registration by minimising kullback-leibler distance. *Proc. of MICCAI Conf.*, pp:525–532, 2002. 1

[3] A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal. 3d multi-modality medical image registration using feature space clustering. *Proc. of Computer Vision, Virtual Reality and Robotics in Medicine Conf.*, pp:195–204, 1995. 1

[4] S. Cui and Y. Wang. Redundant wavelet transform in video signal processing. *Proc. of Image Processing, Computer Vision, & Pattern Recognition Conf.*, pp:191–196, 2006. 2, 3

[5] I. Daubechies, A. Grossman, and Y. Meyer. Painless non orthogonal expansions. *J. Math. Phys.*, 27:1271–1283, 1986. 3

[6] J. E. Fowler. The redundant discrete wavelet transform and additive noise. *IEEE Signal Processing Letters*, 12(9):629–632, 2005. 3

[7] R. Gan and A. Chung. Multi-dimensional mutual information based robust image registration using maximum distance-gradient-magnitude. *Proc. of IPMI Conf.*, pp:210–221, 2005. 1

[8] A. Grossman and J. Morlet. Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM J. Math.*, 15:723736, 1984. 3

[9] S. Hongli and H. Bo. Image registration using a new scheme of wavelet decomposition. *IEEE Proc. of Instrumentation and Measurement Technology Conf.*, pp:235–239, 2008. 2

[10] K. A. Johnson and J. A. Becker. The whole brain atlas. *http://www.med.harvard.edu/AANLIB/home.html*. 7

[11] S. Li, J. Peng, J. T. Kwok, and J. Zhang. Multimodal registration using the discrete wavelet frame transform. *Proc. of ICPR Conf.*, pp:877–880, 2006. 2

[12] H. Luan, F. Qi, and D. Shen. Multi-modal image registration by quantitative-qualitative measure of mutual information (q-mi). *Proc. of CVBIA Conf.*, pp:378–387, 2005. 1

[13] S. G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(7):674–693, 1989. 3

[14] J. L. Moigne, W. J. Campbell, and R. F. Cromp. An automated parallel image registration technique based on the correlation of wavelet features. *IEEE Trans. On Geoscience and Remote Sensing*, 40(8):1849–1864, 2002. 2

[15] E. Oubel, A. F. Frangi, and A. O. Hero. Complex wavelets for registration of tagged mri sequences. *Proc. of ISBI*, pp:622–625, 2006. 2

[16] A. Roche, G. Malandain, N. Ayache, and S. Prima. Towards a better comprehension of similarity measures used in medical image registration. *Proc. of MICCAI Conf.*, pp:555–567, 1999. 1

[17] R. Sharman. A fast and accurate way to register medical images using wavelet modulus maxima. *Pattern Recognition Letters*, 21(6):447–462, 2000. 2

[18] C. Studholme, D. L. G. Hill, and D. J. Hawkes. Automated 3-d registration of mr and ct images of the head. *Medical Image Analysis*, 1(2):163–175, 1996. 1

[19] P. Viola and W. Wells. Alignment by maximization of mutual information. *IJCV*, 24(2):137–154, 1997. 1

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

[21] L. Zollei, J. Fisher, and W. Wells. A unified statistical and information theoretic framework for multi-modal image registration. *Proc. of IPMI Conf.*, pp:366–377, 2003. 1