

# BOOSTED METRIC LEARNING FOR 3D MULTI-MODAL DEFORMABLE REGISTRATION

Fabrice Michel<sup>(1,2)</sup> Michael Bronstein<sup>(3)</sup> Alex Bronstein<sup>(4)</sup> Nikos Paragios<sup>(1,2)</sup>

<sup>(1)</sup> Laboratoire MAS, Ecole Centrale Paris, Châtenay-Malabry, France

<sup>(2)</sup> Equipe GALEN, INRIA Saclay-Ile-de-France, Orsay, France

<sup>(3)</sup> Institute of Computational Science, Università della Svizzera Italiana, Switzerland

<sup>(4)</sup> Department of Electrical Engineering, Tel Aviv University, Israel

## ABSTRACT

Defining a suitable metric is one of the biggest challenges in deformable image fusion from different modalities. In this paper, we propose a novel approach for multi-modal metric learning in the deformable registration framework that consists of embedding data from both modalities into a common metric space whose metric is used to parametrize the similarity. Specifically, we use image representation in the Fourier/Gabor space which introduces invariance to the local pose parameters, and the Hamming metric as the target embedding space, which allows constructing the embedding using boosted learning algorithms. The resulting metric is incorporated into a discrete optimization framework. Very promising results demonstrate the potential of the proposed method.

*Index Terms*— Metric Learning, Multi-Modal Registration, 3D Deformable Registration, Gabor Feature Descriptor

## 1. INTRODUCTION

Deformable image registration is a notoriously difficult problem in medical image analysis, mostly due to the number of parameters to be estimated from the data. This is alleviated somehow when combining suitable registration metrics with regularization constraints imposing certain geometric properties on the deformation field. The definition of a metric can be either iconic or statistic. In the first case, one seeks for apparent similarities on an appropriate feature space (such as Sum of Absolute Differences (SAD), Sum of Squared Differences (SSD), Corelation Ratio (CR), etc.), while in the second, one seeks to optimize the measurement of statistical significance. The first class of methods is more suited to intra-modality, while the second one is often considered to establish correspondences between different nature of observations.

An *inter-modal* (multiple modality) deformable alignment setting, in which two images come from different imaging modalities, has important clinical applications such as fusion of anatomic and functional data. Compared to *intra-modal* (single modality) registration, the multi-modal case is much more challenging. In the single modality setting,

despite certain variations in image appearance, one can compare in a reasonable manner observations between images, thus making the optimization procedure employed for alignment the main component of the problem solution. Recent advances either in continuous [1] and discrete optimization [2] were able to a certain extent to provide good near-optimal solutions for intra-modal registration.

This is not the case for inter-modality, because defining a universal metric able to account for variations of images often acquired under very different conditions and having different underlying physical and anatomical properties is very challenging. Statistical methods exploiting joint densities are among the most common metrics used in multi-modal alignment. Mutual information [3], Kullback-Leibler- [4] or Jensenn-Rényi divergence [5] are the most prominent metrics used in this type of problems. These metrics work relatively well when the modalities have similar or correlated statistical properties, an assumption rarely satisfied in practice. On the other hand, these methods suffer from several limitations, including the lack of statistical interpretation when comparing modalities with different underlying imaging physics, the curse of dimensionality and the lack of sufficient support to build joint statistics, the lack of smoothness of the objective function with respect to the registration parameters, and its non-convexity.

In order to overcome this limitation, recently the idea of *metric learning* has been considered. Support vector machine (SVM)-based regression was used to construct the metric in [6]. These methods inherit the discontinuity of the SVM space while at the same time suffer from not being pose invariant. The use of direct regression between image modalities is another way of establishing correspondences as suggested in [7, 8]. The idea is to find a non-linear mapping between the two modalities such that one can predict/synthesize one modality from the second and convert the problem into intra-modal registration. The strength of such an approach is the visual representation of regression, however the problem of mapping a poor information space into a richer one is ill-posed and cannot be solved in the most general case.

Here, we follow our previous work [9], in which we introduced *multimodal similarity sensitive hashing*. Considered in the context of metric learning, the main idea of this approach is to embed the data of different modalities into a common metric space, whose metric is used to parametrize the multimodal metric. Multi-modal metric learning is thus posed as optimization over the data embeddings. Considering in particular the embedding into the space of binary strings with the Hamming metric, the problem can be formulated as boosted classification and solved employing boosting algorithms.

In this paper, we study the application of multi-modal metric learning to the problem of inter-modality alignment, and in particular cope with real 3D data, in which we have to take into account the very noisy nature of medical images, as well as the common illumination changes within the same modality. This setting forces us to have a learning technique that can be adapted to those conditions on one hand, and use invariant Gabor features on the other.

The rest of this paper is organized as follows: In Section 2, we formulate the problem of multi-modal alignment and a generic framework for its solution and we show how to learn such a metric from examples. Section 3 describes the implementation used in this paper and presents the results. Finally, Section 4 concludes the paper.

## 2. PROBLEM FORMULATION

### 2.1. Deformable Image Fusion

In the problem of non-rigid alignment, we are given a source and target images  $f$  and  $g$  (for simplicity, scalar-valued), defined on a domain  $\Omega$  ( $\Omega \subset \mathbb{R}^3$  in 3D alignment shown in our experiments here). In general, the images are related by a complicated relation,

$$g(x) = \mathcal{X} \circ f(\mathcal{T}(x)), \quad (1)$$

for all  $x$  in  $\Omega$ . This relation involves a geometric deformation  $\mathcal{T}$  and a non-linearity  $\mathcal{X}$  explaining the changes of appearance between corresponding points. Obviously,  $\mathcal{T}$  and  $\mathcal{X}$  are typically unknown.

State-of-the-art registration methods [2] attempt at estimating the deformation  $\mathcal{T}$  on a sparse grid  $\Omega' \subset \Omega$  ( $|\Omega'| \ll |\Omega|$ ) of *control points*,

$$\mathcal{T}(x) = x + \sum_{p \in \Omega'} \rho(\|x - x_p\|) \Delta_p, \quad (2)$$

where  $\Delta_p$  is the displacement vector of the control point  $x_p$ .

Moving a control point results in a local deformation of the image around it; the weighting function  $\rho$  measures the contribution of a control point in  $\Omega'$  to the displacement of a point in  $\Omega$ . The deformation field is found by maximizing the criterion of point-wise similarity between the target and deformed source images.

To tackle the problem of  $\mathcal{X}$ , it is usual to use some local similarity function  $s$ ,

$$E(\mathcal{T}) = \frac{1}{|\Omega'|} \sum_{p \in \Omega'} \int_{\Omega} \hat{\rho}(\|x - x_p\|) s(g(x), f(\mathcal{T}x)) dx \quad (3)$$

where  $\hat{\rho}$  is a normalized version of  $\rho$ . The choice of  $s$  is the critical part of the process. In order to avoid folding on the deformation grid, a smoothness term on  $\mathcal{T}$  is added.

For a practical and efficient numerical solution, problem (3) is posed as an assignment problem in the following way [2]: Let  $\mathcal{L} = \{u^1, \dots, u^k\}$  be a discrete set of labels corresponding to a quantized version of the deformation space  $\Theta = \{\Delta^1, \dots, \Delta^k\}$ . A label assignment  $u_p \in \mathcal{L}$  to a grid node  $x_p \in \Omega'$  is associated with displacing the node by the corresponding vector  $\Delta^{u_p}$ . The deformation field associated with a certain discrete labelling  $u$  is  $\mathcal{T}_u(x) = x + \sum_{p \in \Omega'} \rho(\|x - x_p\|) \Delta^{u_p}$ . Problem 3 can thus be posed as discrete Markov random field (MRF) optimization with respect to the labelling,

$$\begin{aligned} E(u) &= \frac{1}{|\Omega'|} \sum_{p \in \Omega'} \int_{\Omega} \hat{\rho}(\|x - x_p\|) s(g(x), f(\mathcal{T}_u x)) dx \\ &\approx \frac{1}{|\Omega'|} \sum_{p \in \Omega'} V_p(u_p), \end{aligned} \quad (4)$$

where  $V_p$  is a *singleton potential function* representing a local similarity measure. Such a formulation allows to plug in any similarity function without modifying the scheme itself.

When the source and the target images arise from different imaging modalities, we land at the problem of computing a cross-modality similarity. Modelling such similarity explicitly can be very difficult, but learning is possible given examples of aligned images, as described in the next section.

### 2.2. Cross-modality metric learning

Let  $X \subseteq \mathbb{R}^m$  and  $Y \subseteq \mathbb{R}^{m'}$  denote the space of local representation of the images. For example, these can be local patches or descriptors computed on the images in two modalities. Different representations can be used depending on the modality, so in general  $X$  and  $Y$  could dramatically differ in their structure and even dimensionality ( $m \neq m'$ ).

Furthermore, let  $s : X \times Y \rightarrow \{\pm 1\}$  denote an (unknown) similarity function between points in two images, which we assume for simplicity to be binary: corresponding points are similar and non-corresponding points are dissimilar. The similarity  $s$  partitions the set  $X \times Y$  of all pairs of data points into *positives*  $\mathcal{P} = \{(x, y) : s(x, y) = +1\}$  (e.g. descriptors computed at corresponding points in two images) and *negatives*  $\mathcal{N} = \{(x, y) : s(x, y) = -1\}$ . Given a training set of positive and negative examples, the goal of *similarity learning* is

to construct another binary similarity function  $\hat{s}$  that approximates  $s$  as faithfully as possible, in the sense that  $\hat{s} = +1$  for most pairs in  $\mathcal{P}$  and  $\hat{s} = -1$  for most pairs in  $\mathcal{N}$ .

Since  $X$  and  $Y$  are often incommensurable, it is impossible to directly compare points from  $X$  and  $Y$ . To cope with this problem, we embed the data points into some metric space  $(\mathbb{Z}, d_{\mathbb{Z}})$  by means of two maps  $\xi : X \rightarrow \mathbb{Z}$  and  $\eta : Y \rightarrow \mathbb{Z}$ . The distance  $d_{\mathbb{Z}}$  represents the dissimilarity of the embedded points, in the sense that the lower is  $d_{\mathbb{Z}}(\xi(x), \eta(y))$ , the higher is the probability that  $s(x, y) = +1$ .

A particular selection of the  $n$ -dimensional Hamming space  $\mathbb{H}^n$  as the embedding space  $\mathbb{Z}$  allows encoding each data point as a weighted  $n$ -bit binary string.  $h_i$  and  $\alpha_i$  encode the  $i^{\text{th}}$  bit and weight respectively.

$$d_{\mathbb{H}^n}(\xi(x), \eta(y)) = \frac{1}{2} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \alpha_i h_i(x, y) \quad (5)$$

The correlation between positive similarity of a pair of points and small Hamming distance between their corresponding codes implies that positives are likely to be mapped to the same code (here,  $h_i(x, y) = (2\xi_i(x) - 1)(2\eta_i(y) - 1)$  and  $\xi_i$  and  $\eta_i$  denote the  $i^{\text{th}}$  coordinate of the embeddings). This fact allows to interpret the Hamming embedding as *Similarity-Sensitive Hashing* [10], under which positive pairs have high collision probability, while negative pairs are unlikely to collide.

In [9] the Hamming metric is sequentially constructed as a superposition of weak binary classifiers  $h_i(x, y)$  by means of a boosted learning algorithm, generalizing this way the approach of [10] to multi-modal case. In this paper, we no longer set a constant  $\alpha_i$ , but use the standard optimization in AdaBoost (eq. 7) which yields a smoother metric and better training error rates. The boosted cross-modality similarity learning algorithm, whose structure follows the standard AdaBoost procedure [11], is summarized in Algorithm 1.

The maximization of  $r_i$  in (6) varies depending on the form of  $\xi$  and  $\eta$ . Using affine projections of the form  $\xi_i(x) = \text{sign}(p_i^T x + a_i)$  and  $\eta_i(y) = \text{sign}(q_i^T y + b_i)$ , where  $p_i$  and  $q_i$  are, respectively,  $m$ - and  $m'$ -dimensional unit vectors, and  $a_i$  and  $b_i$  are scalars, optimization is performed with respect to  $p_i, q_i, a_i, b_i$ . Vectors  $p_i, q_i$  are searched in the subspace spanned by the  $M$  (chosen as 5 in our experiments) largest left and right singular vectors of the matrix  $C = \sum_{k=1}^K w_i(k) s_k \bar{x}_k \bar{y}_k^T$  ( $\bar{x}_k$  and  $\bar{y}_k$  are  $x_k$  and  $y_k$  centered by their weighted means). Scalars  $a_i, b_i$  are found by an equivalent of a line search procedure.

### 3. MULTIMODAL ALIGNMENT

In the experiments in this paper, we deal with the alignment of multi-contrast MR image. In order to apply multi-modal metric learning to our alignment problem, we need to define the data space  $X$  and  $Y$  on which the similarity is defined, and

**Input:**  $K$  pairs  $(x_k, y_k)$  labelled by

$$s_k = s(x_k, y_k) \in \{\pm 1\}.$$

**Output:** maps  $\xi_i : X \rightarrow \{0, 1\}$  and  $\eta_i : Y \rightarrow \{0, 1\}$ , and scalars  $\alpha_i, i = 1, \dots, n$ .

Initialize weights  $w_1(k) = 1/K$ .

**for**  $i = 1, \dots, n$  **do**

    Select  $\xi_i$  and  $\eta_i$  such that  $h_i$  maximizes

$$r_i = \sum_{k=1}^K w_i(k) s_k h_i(x_k, y_k). \quad (6)$$

    Set  $\alpha_i$

$$\alpha_i = \frac{1}{2} \log(1 + r_i) - \frac{1}{2} \log(1 - r_i) \quad (7)$$

    Update weights according to

$$w_{i+1}(k) = w_i(k) e^{-s_k \alpha_i h_i(x_k, y_k)} \quad (8)$$

    and normalize by sum.

**end**

**Algorithm 1:** Boosted cross-modal similarity-sensitive embedding

the strategy to construct the training set on which the metric is learned. These aspects are addressed in this section.

#### 3.1. Gabor features

Since robustness to noise and intensity shifts is key in medical imaging applications, we used Gabor feature vectors that have such properties by definition. Such features are known to be robust to noise and generally applicable to many medical imaging modalities [12]. Given an image in each modality, we first extract anisotropic Gabor features at every pixel. To save on the computation time that would be required for the extraction of 3D Gabor features, we follow the works in [13] and use two orthogonal 2D Gabor filters :

$$\mathbf{g}(x, y) = \left( \frac{1}{2\pi\sigma_x\sigma_y} \right) e^{\left\{ -\frac{1}{2} \left( \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2j\pi W y \right\}}$$

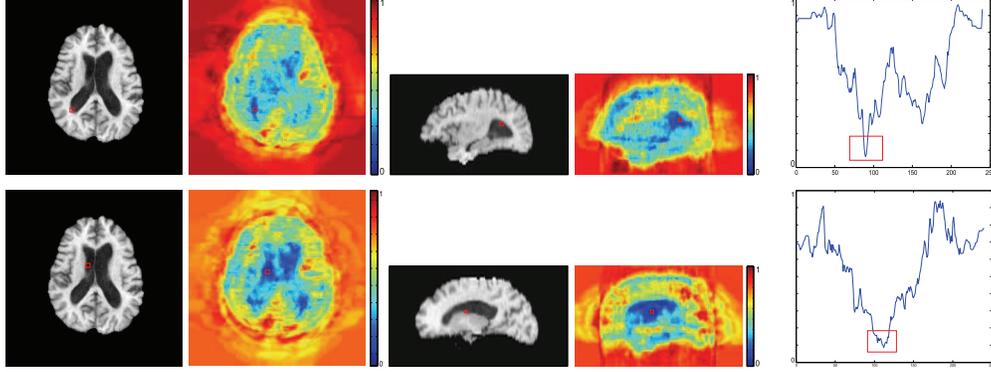
$$\mathbf{h}(y, z) = \left( \frac{1}{2\pi\sigma_y\sigma_z} \right) e^{\left\{ -\frac{1}{2} \left( \frac{y^2}{\sigma_y^2} + \frac{z^2}{\sigma_z^2} \right) + 2j\pi W y \right\}}$$

The Gabor filter bank is obtained through the generating functions with:

$$g_{m,n}(x, y) = a^{-m} g(x', y') \quad \text{and} \quad h_{m,n}(y, z) = a^{-m} h(y'', z')$$

$$x' = a^{-m} (x \cos \theta + y \sin \theta), \quad y' = a^{-m} (-x \sin \theta + y \cos \theta)$$

$$y'' = a^{-m} (y \cos \phi + z \sin \phi), \quad z' = a^{-m} (-y \sin \phi + z \cos \phi)$$



**Fig. 1.** Distance map: plot of the learned distance taken between the feature vector extracted in the red square position on the left on the T1-MRI and all of the feature vectors extracted on the corresponding co-registered T2-MRI image. Far right is a profile extracted on the same line as the reference position. Bottom row presents the less distinctive case, notice the 15 voxels neighbourhood around the extraction position.

$$\theta = \frac{n\pi}{N} \quad \text{and} \quad \phi = \frac{m\pi}{M}$$

where  $N$  the range of  $n$ , is the number of orientations and  $M$ , the range of  $m$  is the number of scales. In order to set  $\sigma_x, \sigma_y, a$  and  $W$  we use the values given in [14]:

$$a = \left( \frac{U_h}{U_l} \right)^{\frac{1}{M-1}} \quad \text{and} \quad W = U_h$$

$$\sigma_x = \frac{1}{2\pi\sigma_u} \quad \text{and} \quad \sigma_u = \frac{(a-1)U_h}{(a+1)\sqrt{2\ln 2}}$$

$$\sigma_y = \frac{1}{2\pi\sigma_v} \quad \text{and} \quad \sigma_v = \tan\left(\frac{\pi}{2N}\right) \sqrt{\frac{U_h^2}{2\ln 2} - \sigma_u^2}$$

We found that  $U_h = 0.2$  and  $U_l = 0.05$  yield the best results.

The zero-mean version of the Gabor filter bank was used in order to have invariance to intensity shifts. Since deformable registration at the fine scale performs local rotations and scaling, scale- and rotation-invariance are crucial for a good alignment. Reorganizing each feature vector into a 2-dimensional array w.r.t. scale and orientation, one can see that rotating or rescaling the image would amount to a translation in the first and the second dimension of the array, respectively. Following [15], we use the Fourier transform magnitude (FTM) to remove the dependence on rotation and scaling: the Fourier transform converts translation into a complex phase, which is removed by taking the magnitude. The symmetry of FTM representation allows keeping only half of the coefficients that are now scale- and rotation-invariant and also insensitive to intensity shift.

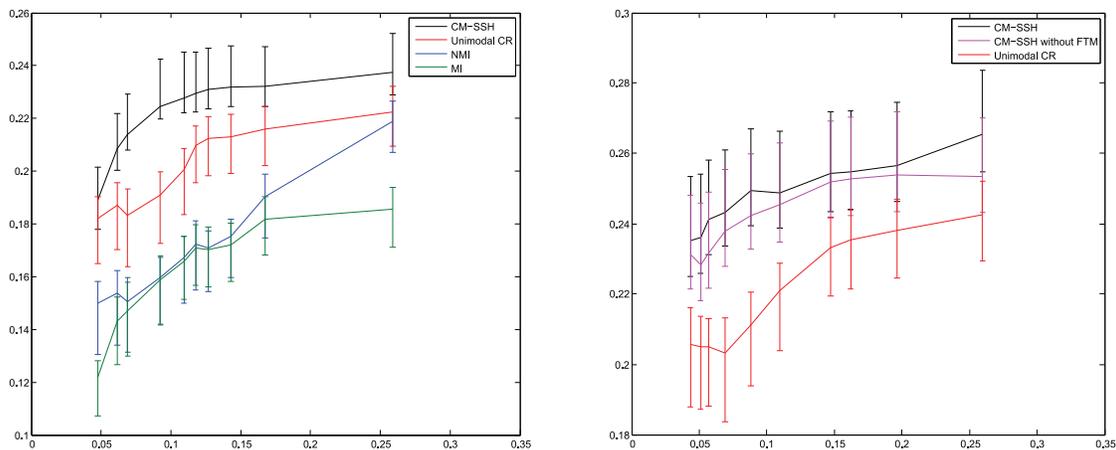
In our framework, the use of Gabor features can be considered as a pre-processing aimed at alleviating the learning burden: first, by decreasing the dimensionality of our data spaces  $X, Y$  (using the Gabor filter responses rather than raw patches), and second, by introducing rotation, scale, and intensity shift invariance by construction rather than by learning. We used  $M = 4$  and  $N = 12$ , which amounts to 96

Gabor coefficients, or 48 coefficients using the FTM. In order to have a fast implementation, we used the idea of [16], where infinite impulse response filter banks for isotropic Gabor filters are generated and applied it to the anisotropic case using the fast anisotropic Gaussian filter implementation described in [17]. Computation time for a  $256 \times 256 \times 48$  image on a Xeon X5560 is 12 seconds. The process can be highly parallelized, which has not been done here.

### 3.2. Training set construction

For our experiments, we used the MR brain images of ten patients. For each patient, perfectly co-registered T1-, T2-weighted and Proton-Density (PD) images were available. Four pairs of images were used for training; the rest was used for testing. The training dataset was designed using the groundtruth correspondence between the multi-modal images: feature vectors at corresponding location in two different modalities were considered similar, while two feature vectors extracted at a location distant 14 to 16 pixels from the groundtruth correspondence location were considered dissimilar. For the training set, we randomly picked features vectors in the four image pairs, with  $|\mathcal{P}| = 20 \times 10^3$  positive and  $|\mathcal{N}| = 200 \times 10^3$  negative pairs.

To visually assess the validity of the learned measure, we plot the learned metric in a point in the image in one modality to all the points in the image in second modality in Fig. 1 (since the data are 3D, two 2D slices are shown). It is interesting to observe that for some very distinctive points in the image, the distance is close to 0 in a very limited area around the point position (first row of Fig. 1), while in less distinctive image areas, the distance profile is more shallow around the point. We can note that the size of the valley around the point of interest in the latter case is around 15 voxels in radius, which is consistent with the training set creation.



**Fig. 2.** Increase in the DICE coefficient before and after alignment as a function of the harmonic energy. The solid line represents the average curve, and the whiskers around the line represent the maxima and minima for the method for the specific harmonic energy, CM-SSH represents our method. Left T2-MRI to T1-MRI registration, right T1-MRI to PD-MRI registration

### 3.3. Experimental Validation

The validation of our method was performed by a quantitative measure of the alignment of five of the test images to one given image (target). None of the images used for training were part of the test set. Alignment was performed in two steps: First, affine alignment of all the 5 images to the target, thus removing all effects due to affine alignment. Second, non-rigid registration between each of the five T2-weighted images and the T1-weighted target image (we also repeated the experiment with the couple PD images/T1 target).

The validation is done by comparison to a manual delineation of the ventricles. We used the deformation field obtained from the alignment to warp a delineation of the ventricle and compared it to the target using the DICE coefficient measuring the proportion of overlap between two segmentations. The smoothing term we introduced in Section 2, which is common to most of the registration algorithms usually prevents from having reliable measures. Indeed, across images and similarity measures, for the same value of the smoothing term parameter, the smoothness of the warp might appear completely different. To quantify the smoothness of the warp we use the harmonic energy defined in [18] as the average over all voxels of the squared Frobenius norm of the Jacobian of the displacement field. The lower the harmonic energy, the more rigid and smooth the transformation.

In Fig. 2, we show the evolution of the increase in DICE coefficient before and after registration with respect to the harmonic energy. We compare our method (with invariance and without) to the most commonly used metrics in multi-modal medical image alignment in the case of T1-T2 alignment, and for reference, provide the results in the uni-modal case (T1-T1) with Correlation Ratio (CR) metric. Each curve represents a different method, for each method we tested the registration from 5 patients to one. None of the patients were

part of the training set. Each curve was computed on 20 points, which yields 100 experiments for one single curve, only a dozen points per curve are displayed. The solid line represents the average curve, and the whiskers around the line represent the maxima and minima for the method for the specific harmonic energy. In figure 3, we show an example of registration.

In these results we clearly see the superiority of our learned metric over most commonly used similarity measures. Most importantly, we see that our method performs better than the ideal case of the uni-modal Correlation Ratio.

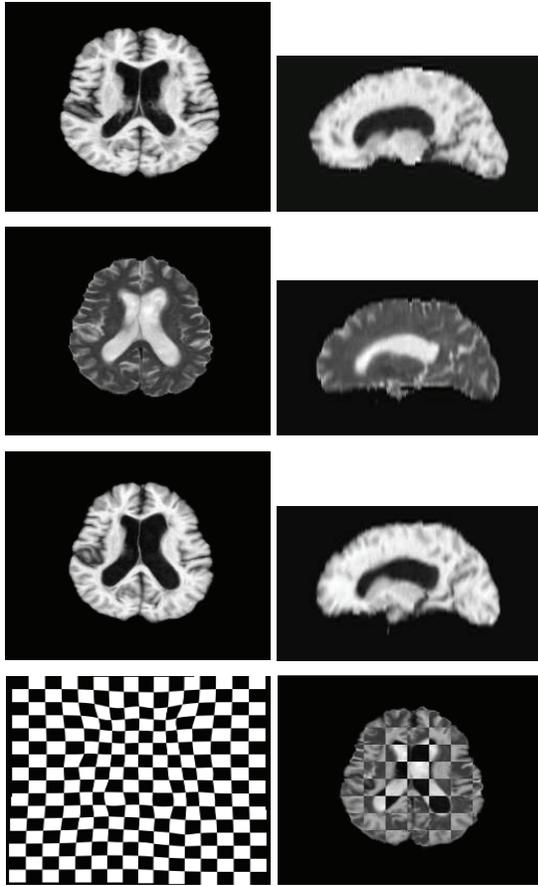
## 4. CONCLUSION

In this paper, we used the recent advances in metric learning to learn an optimal similarity measure for image registration. The extreme flexibility of the learning algorithm allowed us to design robust feature vectors. The use of zero-mean Gabor filters rendered invariant with their FTM, makes this learning insensitive to intensity shifts, robust to noise and adequate for image registration. We showed in our experiments on two types of modalities that in all we have results superior to all commonly used statistical metrics, and that we even have better results than one would have in a common uni-modal setting.

Future work consists in looking into continuous embedding in order to introduce some desired continuity properties on the metric space. Furthermore, looking into non-linear regression and kernel-based mappings could improve the performance of the method when aiming to fuse images with non-linearly related appearance variability.

## 5. ACKNOWLEDGMENTS

The work was partially supported from the STEREO+ grant of the Medicen Competitive Cluster of Ile-de-France region.



**Fig. 3.** Top row: Source Image T1-MRI image. Second Row: target T2-MRI image. Third Row: deformed image after multi-modal deformable registration. Bottom Row: left, deformation field of the registration, right, checker-board image between the target and the deformed source.

## 6. REFERENCES

- [1] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Non-parametric diffeomorphic image registration with the demons algorithm," *LNCS*, vol. 4792, pp. 319, 2007.
- [2] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, and N. Paragios, "Dense image registration through mrfs and efficient linear programming," *MedIA*, vol. 12, no. 6, pp. 731–741, 2008.
- [3] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE-TMI*, vol. 16, no. 2, pp. 187–198, 1997.
- [4] A. C. S. Chung, W. M. Wells, A. Norbash, and W.E.L. Grimson, "Multi-modal image registration by minimizing kullback-leibler distance," in *MICCAI*, 2002, pp. 525–532.
- [5] Y. He, A.B. Hamza, H. Krim, et al., "A generalized divergence measure for robust image registration," *IEEE-TSP*, vol. 51, no. 5, pp. 1211–1220, 2003.
- [6] D. Lee, M. Hofmann, F. Steinke, Y. Altun, N.D. Cahill, and B. Schölkopf, "Learning the similarity measure for multi-modal 3d image registration," in *IEEE-CVPR*, 2009.
- [7] G. Charpiat, M. Hofmann, and B. Schölkopf, "Kernel methods in medical imaging," in *Handbook of Biomedical Imaging*, N. Paragios, J. Duncan, and N. Ayache, Eds. Springer, Berlin, Germany, 12 2008.
- [8] F. Michel and N. Paragios, "Image transport regression using mixture of experts and discrete Markov Random Fields," in *IEEE-ISBI*, 2010, pp. 1229–1232.
- [9] M.M. Bronstein, A.M. Bronstein, F. Michel, and N. Paragios, "Data fusion through cross-modality metric learning using similarity-sensitive hashing," in *IEEE-CVPR*, 2010, pp. 3594–3601.
- [10] G. Shakhnarovich, *Learning task-specific similarity*, Ph.D. thesis, MIT, 2005.
- [11] Y. Freund and R.E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," in *Proc. European Conf. Computational Learning Theory*, 1995.
- [12] Y. Ou and C. Davatzikos, "DRAMMS: deformable registration via attribute matching and mutual-saliency weighting," *IPMI*, 2009.
- [13] Y. Zhan and D. Shen, "Automated segmentation of 3D US prostate images using statistical texture-based matching method," *MICCAI*, pp. 688–696, 2003.
- [14] BS Manjunath and WY Ma, "Texture features for browsing and retrieval of image data," *IEEE-PAMI*, vol. 18, no. 8, pp. 837–842, 2002.
- [15] I. Kokkinos and A. Yuille, "Scale Invariance without Scale Selection," in *CVPR*, 2008.
- [16] A. Bernardino and J. Santos-Victor, "Fast IIR isotropic 2-D complex Gabor filters with boundary initialization," *IEEE-TIP*, vol. 15, no. 11, pp. 3338, 2006.
- [17] JM Geusebroek, AWM Smeulders, and J. Weijer, "Fast Anisotropic Gauss Filters," *IEEE-T Im. Proc.*, vol. 13, no. 8, pp. 938–943, 2003.
- [18] B.T.T. Yeo, T. Vercauteren, P. Fillard, J.M. Peyrat, X. Pennec, P. Golland, N. Ayache, and O. Clatz, "DT-REFinD: Diffusion tensor registration with exact finite-strain differential," *Medical Imaging, IEEE TMI*, vol. 28, no. 12, pp. 1914–1928, 2009.