Manifold Learning Based Registration Algorithms Applied to Multimodal Images

Mohammad Farid Azampour *1, Aboozar Ghaffari *2, Azam Hamidinekoo *3 and Emad Fatemizadeh *4

Abstract— **Manifold learning algorithms are proposed to be used in image processing based on their ability in preserving data structures while reducing the dimension and the exposure of data structure in lower dimension. Multi-modal images have the same structure and can be registered together as mono-modal images if only structural information is shown. As a result, manifold learning is able to transform multi-modal images to mono-modal ones and subsequently do the registration using mono-modal methods. Based on this application, in this paper novel similarity measures are proposed for multi-modal images in which Laplacian eigenmaps are employed as manifold learning algorithm and are tested against rigid registration of PET/MR images. Results show the feasibility of using manifold learning as a way of calculating the similarity between multi-modal images.**

Keywords: Image Registration, Manifold Learning, Similarity Measure, Laplacian Eigenmaps

I. INTRODUCTION

Multimodal image registration- defined as a spatial map between images- has a broad use in medical applications to obtain insights regarding accurate comparison between images like evaluating the evolution of a disease in a patient or an organ changes over different times [1] or with different angels [2] and also the information fusion obtained by various imaging modalities [3]. Therefore, every point in one image has a corresponding point on other images for their alignment. Moreover, registration algorithms can be divided into two categories by the type of information used [1]: feature based methods in which three essential steps as feature extraction, correspondence establishment and transformation estimation are fulfilled. Another group includes intensity based methods that use the intensities of the voxels and define a measure of similarity between different images and the registration parameters are optimized when maximizing the respective measure [4]. Consequently, some statistical measures like mutual information [5] have become popular which assumes that two pixels having the same intensity in the first image also have the same intensity in the second image. In contrary, feature based methods are found to be robust to intensity variations for key point extraction and description [6]. However, the multimodal image registration is affected by more substantial intensity variations that make such approaches not applicable enough [2]. Recently in [7], manifold learning was employed in order to extract structural information out of multi-modal images and then get use of the obtained information in registration. In other words, manifold learning algorithms are able to reveal structures by processing data in a high dimensional space and present it in a lower dimensionality which can reveal structural similarity between multimodal images. In this research, some novel similarity measures based on manifold learning of both reference and float images are proposed in which the principals of Laplacian eigenmaps algorithm [8] is used. At first, the difference of Laplacian matrix of two images are introduced as a similarity measure and then in another definition the difference of manifold learning output under an appropriate constraint is presented as similarity measure. In the latter approach two different constraints and problems are defined to find the similarity measure. The first problem tries to compute one image manifold subject to being as similar as possible to the other image and the second one calculates both image manifolds simultaneously subjected to finding manifolds with minimum differences in lower dimension. Finally the respective similarity measures are analyzed for the rigid registration of multi-modal brain MR and PET images.

Subsequently, in the following sections, common manifold learning techniques and their applications in image processing are briefly reviewed. Then a technique to transform multi-modal images to mono-modal ones is explained and some novel similarity measures based on manifold learning in multi-modal images are introduced. In the third section, the proposed transformation is tested and similarity measures are validated by using them in a rigid registration process of PET/MR images. Putting all together, the conclusion is presented.

II. BACKGROUND

Manifold learning techniques have become popular after the introduction of Isomap [9] and LLE [10] in 2000. Since then other learning approaches were introduced like Laplacian [8] and Hessian eigenmaps [11] and each tries to represent high dimension data structures in lower dimensions. Among these methods, Laplacian eigenmaps approach has attracted researchers attention more because of the best structure preservation with lower computational complexity [12]. Manifold learning in image registration can be categorized into two groups. In the first group, manifold learning is used in order to find correspondence between samples of two image sets. These methods were initially become known by Ham [13] and then a semi-supervised approach was introduced by Mahadevan and Wang [14].

* Biomedical Signal and Image Processing Laboratory (BiSIPL), Electrical Engineering Department, Sharif University of Technology, Tehran, Iran
1 mfazampour@ee.sharif.edu
2 a_ghaffari@sharif.edu
3 hamidinekoo@ee.sharif.ir
4 fatemizadeh@sharif.edu

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

Later Zha *et al*. [15] proposed an unsupervised method for image matching. This group of algorithms are known as *Manifold Alignment Algorithms*.

The second group is related to two image alignment by means of manifold learning. Among the methods suggested, Navab *et al*. [7] declared that manifold learning is capable of detecting structures and so gives the structural similarity between images taken from different modalities. They presented their solution for the multimodal image registration. In the proposed method, each pixel of the image is converted to 225-Dimensional vector based on a 15×15 neighboring region around that pixel and then the respective data are transformed into 1-Dimensional data using Laplacian eigenmaps algorithm which has got the least computational expense compared to other manifold learning algorithms. For N data this computational complexity is $O(N^2)$. Afterward, images are reconstructed after aligning the created manifolds that is the multi-modal images are converted to mono-modal images. After this step, common mono-modal image registration algorithms can be implemented.

III. METHOD

As mentioned, manifold learning has got various applications in multi-modal image processing and the main reason lies in the existence of similar structural information from different modalities and subsequently manifold learning capability in detecting such structures. Thus, similar results are expected from manifold learning of two images with similar structural information. In the following a technique to transform multi-modal images to mono-modal images is explained and then some novel similarity measures based on manifold learning in multi modal images are presented.

A. Multi-modal to mono-modal transformation

Major disadvantages of transforming multi-modal to mono-modal images are computational complexity and required memory capacity considering number of image pixels. In this research, a suitable solution for the reduction of computational expenses while preserving transformation accuracy is presented. In this method a set of pixels on edges via Canny [16] algorithm and those selected by the uniform mesh are considered as the representatives of image information. From this point forward, this way of selecting pixels is called *manifold pixel selection* (*MPS*). Thus by limiting the number of participating pixels in learning process, computational expense and the required memory would decrease significantly. In order to reduce complexity even more, some data in 225-Dimensional space that have a distance less than a defined threshold value are removed. After manifold learning from the opted data out of the image, decreased dimension data are aligned to [0, 1] interval which determine the intensity values of pixels in the new modality. In the next step, the reference image is reconstructed in accordance to new pixel values in the new modality. To this end a 5×5 neighborhood around each pixel is formed which is the minimum region around each pixel that contains a part of structural information and the maximum value conserving details. The distance of each pixel is calculated with labeled pixels considering the 25-Dimensional vector and therefore the intensity value of the nearest labeled pixel will be assigned to unlabeled pixels.

B. Multi-modal similarity measure based on manifold learning

In this section, a new multi-modal similarity measure is defined based on the characteristics of manifold learning algorithms that is the detection of data structure in low dimensional space. Images of an object obtained from different modalities do not have an obvious relation considering only their pixel intensities but these images have structural relations due to the common source. If a measure is capable of extracting such structures then fine registration would be possible. In this research proposed similarity measures are categorized in three groups: the first group investigates the similarity between neighborhood matrix and Laplacian matrix. In the second group, manifold learning for just one image is done while in the third group, simultaneous manifold learning of both images is performed. The latter two measures are found to be more accurate by exerting a constraint to the manifold learning problem.

C. Similarity measure based on Laplacian matrix

As stated in advance, by a 15×15 neighboring region around each pixel and forming the 225-Dimensional vector, structural information of an image patch can approximately be taken into consideration. Subsequently, by forming neighborhood graph in Laplacian eigenmaps algorithm, each datum is assigned to its respective nearest neighbor in high dimension space which equals connecting pixels with similar structures in image space. MPS method is employed to make calculations possible. If the selected pixel locations are just adopted from the source image, then the maximum similarity between its Laplacian matrix and neighborhood graph can be expected in the case images are aligned. Put it together the following measures are definable:

1) In the first measure the number of nonzero elements of the neighborhood graph matrix is defined as a similarity measure.

$$S_1(T) = \left\| G^R - G^F(T) \right\|_0 \tag{1}$$

where G^R is the neighborhood graph matrix of the reference image and G^F is the float image matrix. In (1) L1 or L2 norms can be used too but L0 norm leads to better results.

2) In the second measure, after forming Laplacian matrices of the reference and the float images, their difference is calculated neglecting values less than a defined threshold which shows minor differences in structure. Therefore, by putting threshold on the difference matrix, remained coefficients will be those with significant differences in reference and float image

matrices and the number of the respective nonzero elements are considered as the similarity measure.

$$L_T = L^R - L^F \quad (2)$$

$$L_T(i,j) < \varepsilon \rightarrow L_T(i,j) = 0 \ \forall i,j \quad (3)$$

$$S_2 = \|L_T\|_0 \quad (4)$$

D. Similarity measure based on manifold learning of one image

In this method manifold learning is applied to one of the images trying to find low dimension presentation as similar as possible to the other image. Again MPS method is used to make computations tractable. When two images are completely aligned together then the resulted vector from manifold learning of the first image is very similar to the vector obtained from the second image. Mathematically it can be written as (5):

$$I^* = \min f(I) = \min \sum_i (I_i - I_j)^2 w_{ij}^I + \alpha \sum_i (I_i - J_i)^2 \quad (5)$$

where the first term is the cost function in Laplacian eigenmaps and the second one expresses the constraint leading the algorithm to find a proper presentation in low dimension space which is closer to the respective values in the second image. w_{ij}^I represents coefficients calculated for i'th and j'th pixels of I in original Laplacian eigenmaps method. The cost function can be stated in the matrix form as (6):

$$f(I) = 2trace(I^T L^I I) + \alpha ||I - J||_2^2 \quad (6)$$

where L^I is the Laplacian matrix calculated for selected pixels of I. In order to find the minimum of (6), its derivative should be equal to zero as the following:

$$\begin{aligned} 2 \times L^I I^* + \alpha(I^* - J) = 0 \\ I^* = (2 \times L^I + \alpha)^{-1} \times (\alpha J) \end{aligned} \quad (7)$$

Using the resulted I^* and its difference with the second image vector is defined as the similarity measure as (8):

$$S = ||I^* - J||_2 \quad (8)$$

In the above calculations each of either the reference or the float image can be used as the first or the second image. Now both cases are investigated as following:

$$S_3(T) = ||(2 \times L^F(T) + \alpha)^{-1} \times (\alpha R) - R||_2 \quad (9)$$

$$S_4(T) = ||(2 \times L^R + \alpha)^{-1} \times (\alpha F(T)) - F(T)||_2 \quad (10)$$

S_3 tries to liken the low dimension representation of the float image to the reference image that is more consistent with the general registration definition compared to the second measure. In S_4 the Laplacian matrix of the reference image is fixed so it is enough to find the inverse of matrix $(2 \times L^R + \alpha)$ once and computational complexity is much less than S_3.

E. Similarity measure based on simultaneous manifold learning of both images

The last introduced similarity measure simultaneously performs the manifold leaning of both images and resemblance of calculated manifolds is defined as a similarity measure. For this aim, we define a new cost function consisting of two terms regarding the manifold learning problem of each images and a third term trying to make learnt manifolds similar. The final form of the cost function will be as (11):

$$\begin{aligned} S(T) = \min_{R,F(T)} \sum_{i,j} (r_i - r_j)^2 w_{i,j}^R + \sum_{i,j} (f_i - f_j)^2 w_{i,j}^F \\ + \sum_i \beta (r_i - f_j)^2 \\ \text{s.t.} \quad [RF]^T [RF] = I \end{aligned} \quad (11)$$

This cost function has a trivial completely zero answer as the cost function in Laplacian eigenmaps. Thus, in order to prevent from converging to this undesirable result, one constraint is added to the problem too as shown in (12) in matrix form:

$$\begin{aligned} S(T) = \min_{R,F(T)} tr(R^T L^R R) + tr(F^T L^F F) \\ + tr(\beta (F - R)^T (F - R)) \end{aligned} \quad (12)$$

$$L^R = D^R - W^R \quad (13)$$

$$L^F = D^F - W^F \quad (14)$$

$$D^R = \begin{cases} \sum_j w_{i,j}^R, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

$$D^F = \begin{cases} \sum_j w_{i,j}^F, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

In (13), W^R and W^F corresponds to coefficient matrices in Laplacian eigenmaps calculated for reference and float images, respectively. Using Lagrange coefficients and the derivative of the cost function, we will have:

$$\begin{bmatrix} L^R - \beta I & \beta I \\ \beta I & L^F - \beta I \end{bmatrix} \times \begin{bmatrix} R \\ F \end{bmatrix} = \lambda \begin{bmatrix} R \\ F \end{bmatrix} \quad (17)$$

$$L_T \times V_T = \lambda V_T \quad (18)$$

Like the main problem in Laplacian eigenmaps algorithm, the solution of the above problem leads to the eigenmap problem. When solving the above problem, two other similarity measures are definable. One is the magnitude of the low dimension presentation difference as in (18). The other one is the value of $S(T)$ so that when V_T is the eigenvectors of L_T then $S(T)$ will be equal to L_T eigenvalues as (19):

$$S_5(T) = \|R^* - F^*(T)\|_2 \quad (19)$$

$$S_6(T) = trace(V_T^T L_T V) = \lambda \quad (20)$$

IV. RESULTS

In this section, at first proposed multi-modal to mono-modal transformation is tested by applying to MR images obtained from Brain Web [17] then similarity measures are validated by using them in rigid registration of MR and PET images from RIRE database [18].
Original images taken from different modalities are shown in Fig. 1, pixels selected by MPS after manifold learning are presented in Fig. 2 and afterwards Fig. 3 shows reconstructed images in new modality. Non-rigid registration done using SSD as similarity measure is depicted in Fig. 4. Registration result in original modality is shown in Fig. 5.

(a) Reference

(b) Float

Fig. 1. Original images.

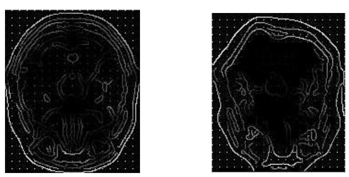


Fig. 2. pixels selected by MPS after manifold learning.

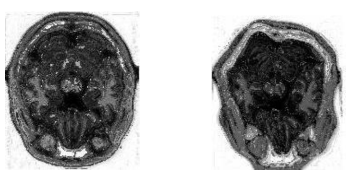


Fig. 3. Reconstructed images in new modality.

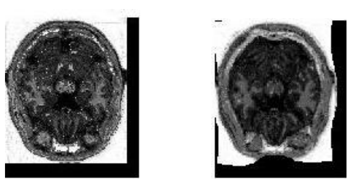
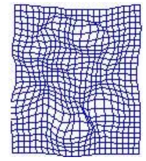
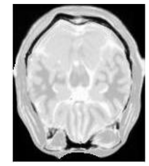


Fig. 4. Registered images in new modality.



(a) B-spline control points map



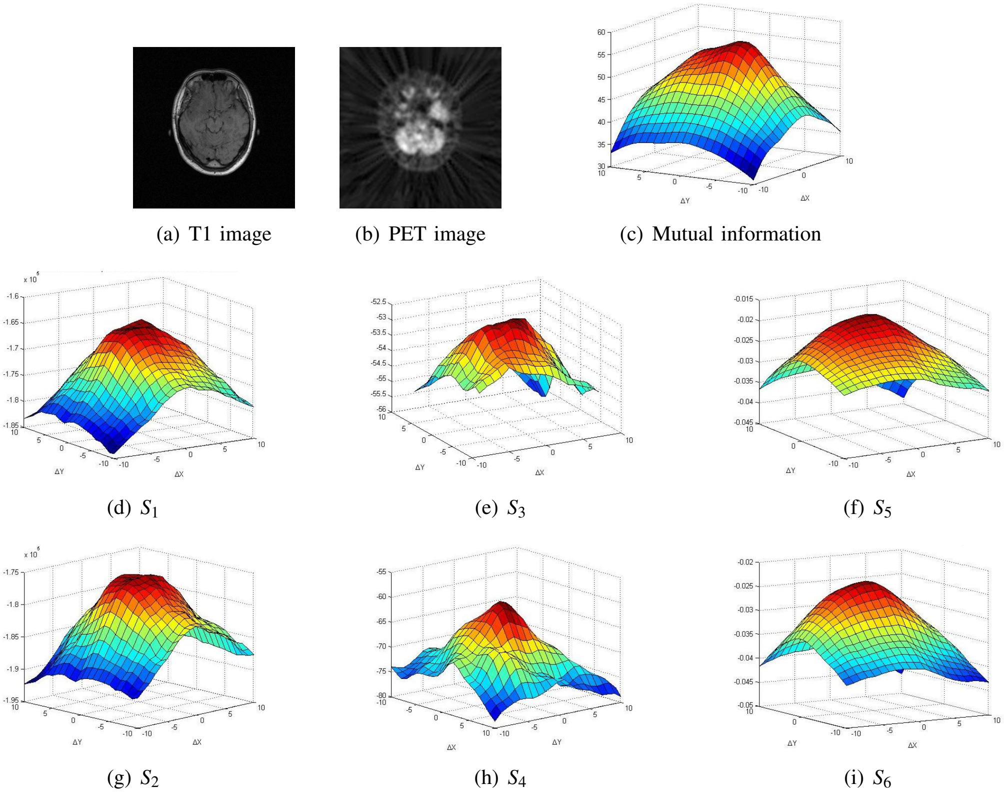
(b) Float image after registration in original modality

Fig. 5. Registration result in original modality

As shown above, proposed transformation changed multi-modal images to mono-modal with much less computation. For tested images of size 128×128 the complexity will approximately be 1.5% of original method.
To validate proposed measures, similarity of MR and PET images were calculated with respect to different translation in x and y directions. Images were chosen from different slices taken from one subject. A total of 23 image pairs were used and surface plots obtained are then compared to the same one created by Mutual Information. Results show capability of proposed measures in rigid registration of PET/MR. Surfaces for one of these pairs are shown in Fig. 6. It can be concluded from Fig. 6 that global extremum of proposed similarity measures happen in the same place as mutual information. For S_5 and S_6, surface plots shows less local extremum and a more distinctive maximum than others.Among proposed similarity measures, S_4 needs less computation and shows acceptable results in these tests. Furthermore, because of having linear relation to intensity of float image pixels, derivative base on this measure can be developed to be used in non-rigid registration.

V. DISCUSSION AND CONCLUSION

In this paper we introduced some similarity measures based on manifold learning features and application in image processing especially the capability of transforming multi-modal images to mono-modal ones. Manifold learning tries to reveal data structures; moreover, multi-modal images almost have the same structural information. As a result, equivalent manifolds for multi-modal images are expected. Proposed measures employed Laplacian eigenmaps to calculate image manifolds and can be categorized in three groups. The first group considers the neighborhood or Laplacian matrix difference as a measure of similarity. In the second group, the manifold of one image is primarily learnt and then its comparison with the second image is considered as a measure. Finally, the third group consists of methods computing both image manifolds simultaneously and then defining their difference as a similarity measure.To validate the proposed measures, a rather difficult task of registration was done; rigid registration of PET/MR images was tested. Results show acceptable performance where all measures display an extremum at the same place as mutual information. S_5 and S_6 similarity measures have more significant extrema with a smoother surface compared to other measures including mutual information.
Furthermore, among the proposed measures, only S_1 acts free from a regularization parameter which can be counted as an advantage. In S_2, the threshold should be given and in our tests it was in range [0.05 0.2]. Moreover, S_4 and S_3 have α and it was suggested to be in range [0.001 0.05]. In addition, S_5 and S_6 have β that according to our tests can be chosen



(a) T1 image (b) PET image (c) Mutual information

(d) S_1 (e) S_3 (f) S_5

(g) S_2 (h) S_4 (i) S_6

Fig. 6. Plot of similarity measures with respect to translation in x and y direction. For better visualization negative of measures are shown. Maxima indicate best alignment.

from [0.001 0.01]. Although performance of proposed methods do not vary significantly in suggested ranges but fine tuning can lead to a more accurate registration.

From computational complexity point of view, S_4 outperforms other methods since only matrix inverse should be computed once and a simple matrix multiplication is needed each time. Besides, because of a linear relationship between pixel intensities of the float image and S_4, gradient based approach for non-rigid registration of multi-modal images can be derived.

REFERENCES

[1] Barbara Zitova and Jan Flusser. Image registration methods: a survey. *Image and vision computing*, 21(11):977–1000, 2003.

[2] VRS Mani et al. Survey of medical image registration. *Journal of Biomedical Engineering and Technology*, 1(2):8–25, 2013.

[3] Derek LG Hill, Philipp G Batchelor, Mark Holden, and David J Hawkes. Medical image registration. *Physics in medicine and biology*, 46(3):R1, 2001.

[4] JB Maintz and Max A Viergever. A survey of medical image registration. *Medical image analysis*, 2(1):1–36, 1998.

[5] Frederik Maes, Andre Collignon, Dirk Vandermeulen, Guy Marchal, and Paul Suetens. Multimodality image registration by maximization of mutual information. *Medical Imaging, IEEE Transactions on*, 16(2):187–198, 1997.

[6] Tinne Tuytelaars and Krystian Mikolajczyk. Local invariant feature detectors: a survey. *Foundations and Trends® in Computer Graphics and Vision*, 3(3):177–280, 2008.

[7] Christian Wachinger and Nassir Navab. Manifold learning for multi-modal image registration. In *BMVC*, pages 1–12. Citeseer, 2010.

[8] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396, 2003.

[9] Joshua B Tenenbaum, Vin De Silva, and John C Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.

[10] Sam T Roweis and Lawrence K Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.

[11] David L Donoho and Carrie Grimes. Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. *Proceedings of the National Academy of Sciences*, 100(10):5591–5596, 2003.

[12] Paul Aljabar, Robin Wolz, and Daniel Rueckert. Manifold learning for medical image registration, segmentation, and classification. *Machine learning in computer-aided diagnosis: medical imaging intelligence and analysis. IGI Global*, 2012.

[13] Jihun Ham, Daniel Lee, and Lawrence Saul. Semisupervised alignment of manifolds. In *Proceedings of the Annual Conference on Uncertainty in Artificial Intelligence*, volume 10, pages 120–127, 2005.

[14] Chang Wang and Sridhar Mahadevan. Manifold alignment using procrustes analysis. In *Proceedings of the 25th international conference on Machine learning*, pages 1120–1127. ACM, 2008.

[15] Yuru Pei, Fengchun Huang, Fuhao Shi, and Hongbin Zha. Unsupervised image matching based on manifold alignment. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(8):1658–1664, 2012.

[16] John Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):679–698, 1986.

[17] Chris A Cocosco, Vasken Kollokian, Remi K-S Kwan, G Bruce Pike, and Alan C Evans. Brainweb: Online interface to a 3d mri simulated brain database. In *NeuroImage*. Citeseer, 1997.

[18] Jay West, J Michael Fitzpatrick, Matthew Y Wang, Benoit M Dawant, Calvin R Maurer Jr, Robert M Kessler, Robert J Maciunas, Christian Barillot, Didier Lemoine, Andre Collignon, et al. Comparison and evaluation of retrospective intermodality brain image registration techniques. *Journal of computer assisted tomography*, 21(4):554–568, 1997.