Indirect Knowledge-Based Approach to Non-Rigid Multi-Modal Registration of Medical Images

Alexander Wong
Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario, Canada, N2L 3G1
Email: a28wong@engmail.uwaterloo.ca

William Bishop
Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario, Canada, N2L 3G1
Email: wdbishop@uwaterloo.ca

Abstract— **Information acquired using different medical imaging techniques (e.g., MRI, PET, CT, etc.) can be combined to get a clear understanding of the overall condition of a patient for the purpose of diagnosis. Registering images from different modalities without a priori knowledge is difficult since the images may have very different intensity mappings and structural characteristics. This paper presents a novel approach to the multi-modal registration of medical images through the use of a priori knowledge to align medical images using an indirect mapping. The proposed algorithm uses stored information from successful alignment results to infer a relationship between the input images from different modalities. This relationship is then used to estimate the transformations needed to align the medical images together. Experimental results show that a high level of accuracy can be achieved using the proposed algorithm to align medical images from different modalities.**

I. INTRODUCTION

A variety of medical imaging techniques have been developed to acquire unique perspectives on the human body. These techniques include x-rays, computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET). It is often beneficial to combine image data collected using different techniques. This process is known as medical image data fusion. This process enables physicians to get a better understanding of the overall condition of a patient for the purpose of diagnosis. For example, PET scans are very useful for providing metabolic information regarding functional processes in the body. However, PET scans do not provide the clear anatomic information provided by MRI scans or CT scans. Therefore, it is very useful to combine PET scans with MRI scans or CT scans to provide a better understanding of a patient's overall condition.

One important technique used in medical image data fusion is image registration which aligns images to each other. Different techniques have been proposed for image registration. One common approach to categorizing image registration methods is based on the similarity metrics used to compare the images for the purpose of alignment. Using this approach, image registration techniques can be generally categorized as follows:

1) correlation [1]–[3]
2) sum of squared differences (SSD) [4]–[7]
3) mutual information [8]–[12]

For the purpose of medical image data fusion, multi-modal registration techniques can be utilized to align images acquired using different modalities. Current multi-modal registration techniques such as those based on SSD and mutual information attempt to find similarities between images obtained from different modalities in a direct fashion without a priori knowledge. This is often very difficult as images acquired from different modalities can have very different intensity mappings. An example of this is shown in Fig. 1 where a T1-weighted MRI cranial slice is compared to a T2-weighted MRI cranial slice. Alignment is further complicated by the fact that structural characteristics captured by one imaging technique may not be captured by another imaging technique. This is illustrated in Fig. 2 where a PD-weighted MRI cranial slice is compared to a CT cranial slice. The tissue information captured by the MR slice is absent in the CT cranial slice, making it difficult to find geometric relationships between the two images.
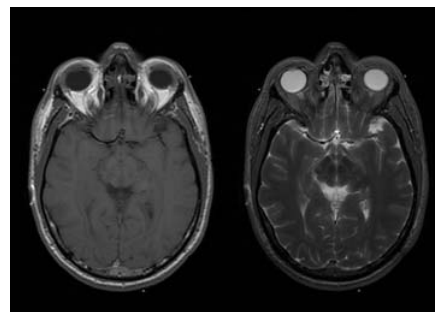

Fig. 1. A cranial slice (a) T1-weighted MRI (b) T2-weighted MRI

This paper describes a novel approach to the multi-modal registration of medical images using a priori knowledge to align medical images in an indirect fashion. The proposed algorithm stores information from successful alignment results into a knowledge learning system. The proposed algorithm compares a set of input medical images with information stored in the knowledge base to infer a relationship between the input images. This relationship is then used to estimate the transformation model needed to align the input images together, accounting for various geometric distortions. Exper-
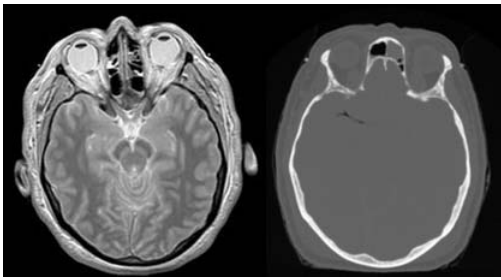
Fig. 2. A cranial slice (a) PD-weighted MRI (b) CT

imental results show that a high level of accuracy can be achieved using the proposed algorithm for aligning medical images from different modalities.

In this paper, the theory behind the proposed algorithm is described and explained in detail in Section II. The experimental results used to demonstrate the effectiveness of the proposed algorithm are presented in Section III. Finally, conclusions are drawn in Section IV.

II. THEORY

Prior to describing the proposed algorithm, it is important to first explain the underlying theory behind the algorithm.

A. Indirect Approach to Multi-Modal Registration

The proposed approach to multi-modal registration is motivated by two key observations. First, as discussed in the introduction, it is often very difficult to perform registration on images acquired using different modalities due to the structural and intensity differences between the images. Most multi-modal image registration techniques (with the exception of neural network approaches [13]) are naive in the sense that they attempt to align images from different modalities in a direct fashion without a priori information about the images. This is in contrast to human experts who are able to align images with very different structural and intensity characteristics with relative ease. One of the reasons that a human expert is able to do this is because he or she is able to leverage previous experience to infer relationships between images with similar underlying content but vastly different characteristics. By learning what is considered a correct alignment between image pairs, a human expert can use this knowledge to determine the relationship between image pairs that are similar to previously aligned pairs. This observation motivates the use of a priori knowledge from previous successful alignments to improve registration accuracy.

The second observation is that it is significantly easier to register two medical images from the same imaging modality than two images from different imaging modalities. Medical images acquired using the same imaging modality have similar pixel intensity characteristics and similar structural characteristics. These similarities make it much easier to find geometric relationships between such images. This observation motivates the use of an indirect approach to multi-modal registration,

where the geometric relationship between images from different modalities is inferred through the geometric relationships between the input images and well-defined images of the same modality from previous knowledge. This indirect relationship inference is illustrated in Fig. 3.
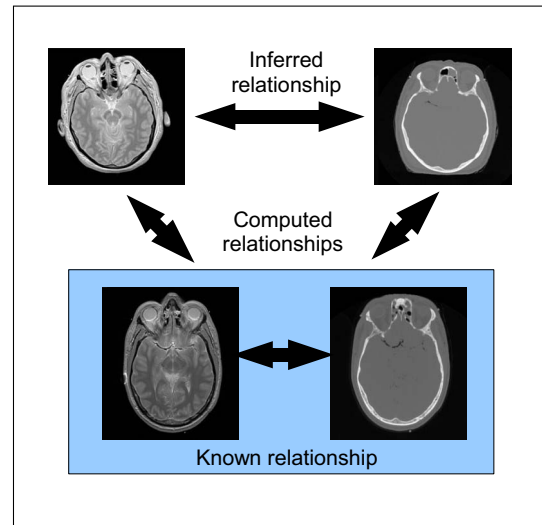

Fig. 3. Relationship inference

The proposed algorithm makes use of these key observations for the purpose of multi-modal image registration. First, the algorithm establishes a computed geometric relationship between each of the input medical images and images with the same modality in previous knowledge instances. Then, the algorithm uses this knowledge to infer a final geometric relationship between the input medical images for the purpose of alignment.

B. Knowledge Base

The proposed knowledge-based approach to multi-modal image registration requires the construction of a knowledge base containing knowledge about correct alignments between medical images acquired using different imaging modalities. This is synonymous to the memory of a human expert, which retains knowledge learned from previous successes to improve his or her performance in future tasks. The knowledge base used in the proposed system is built upon a hierarchical tree structure. Each non-leaf node signifies either a type of medical image (e.g., heart scan, brain scan, etc.) or a modality pairing of images (e.g., CT-MRI, PET-CT, etc.). Each leaf node contains a knowledge instance that represents the knowledge gained from a successful alignment between two medical images. The knowledge stored within each knowledge instance varies depending upon the type of features and/or method used in the relationship inferencing process. For testing purposes, a normalized cross-correlation similarity metric was used with corners as interest points. Therefore, each knowledge instance contains the coordinates for interest points for each image in the image pair as well as the pixel intensity information corresponding to an image patch around each interest point.

The pixel intensity information within that image patch is used as the local descriptor for that interest point. Most importantly, each knowledge instance contains information regarding the geometric relationship between the two medical images. This relationship can be represented as a transformation matrix $T_{knowledge}$ that maps one image to the other. This information can then be used to determine the alignment between input medical image pairs.

C. Relationship Inferencing

With a knowledge base in place, it is possible to apply this knowledge to draw a relationship between input medical images acquired using different imaging modalities. This is analogous to the cognition of a human expert, where inputs are analysed and compared with existing knowledge to draw associations between different inputs. For the proposed approach, the cognitive association process is broken down into two processes: i) similarity search, and ii) relationship estimation.

In the similarity search, input medical images are compared to a set of knowledge instances within the knowledge base to evaluate the similarity between the input images and images represented by a knowledge instance. Additional parameters such as image types (e.g., CT-MRI, heart scan, etc.) may also be supplied along with the input medical images. These additional parameters help narrow the search space to improve computational speed and registration accuracy. For the purpose of this paper, the input images are denoted as images A and B, and the images stored within the selected knowledge instance are denoted as images C and D. During each similarity comparison, the input images are compared to medical image pairs represented by a knowledge instance. Different feature spaces and similarity metrics may be used in the similarity search. For testing purposes, a similarity value S between each input image and each image from the image pair represented by the knowledge instance is computed by matching local descriptors using a normalized correlation metric. The final similarity score between the input images and the image pair within a knowledge instance (if modalities are unspecified for the input images) is calculated as follows:

$$Similarity = \max(S_{AC}, S_{AD}) + \max(S_{BC}, S_{BD}) \qquad (1)$$

where S_{ij} is the similarity value between input image i and image j from the knowledge instance. If modalities are known for the input images, then the similarity score will be based on similarity values between images of the same imaging modalities. The knowledge instance with the highest similarity score is then selected for the relationship estimation process.

In relationship estimation, the geometric relationship between the input images is inferred based on the geometric relationships between the input images and the knowledge instance. Consider the case where A is most similar to C and B is most similar to D. The geometric relationships between A and C and between B and D can be found using either uni-modal or multi-modal image registration techniques. For testing purposes, a standard cross-correlation based image

registration method was used with pixel intensities as a feature space. The geometric transformation between images was then estimated using the Direct Linear Transformation (DLT) method [14]. The use of this technique demonstrates that a uni-modal image registration technique can be used effectively by the proposed algorithm to perform multi-modal image registration. The geometric relationships between A and C and between B and D are represented by transformation matrices T_{AC} and T_{DB}, which map the coordinates from C to A, and the coordinates from B to D respectively. Since the geometric relationship between C and D are known in the knowledge instance, the geometric relationship between A and B can be estimated as:

$$T_{AB} = T_{AC}T_{knowledge}T_{DB} \qquad (2)$$

where $T_{knowledge}$ is the transformation matrix stored in the knowledge instance that maps the coordinates from D to C.

D. Transformation and Alignment

Using the transformation matrix determined in the relationship estimation process, B can then be aligned to A by transforming the coordinates of B as follows:

$$X_{B'} = T_{AB}X_B \qquad (3)$$

where X_B represents the original x-y coordinates of B and $X_{B'}$ represents the transformed x-y coordinates of B.

E. Proposed Algorithm

Based on the above theory, the proposed approach can be outlined in the following manner:

Given input images A and B:

1) Determine a set of knowledge instances from knowledge base for similarity comparison (based on additional parameters if available).
2) Compare images A and B with each knowledge instance from the selected set and compute the corresponding similarity score.
3) Select best matched knowledge instance based on the computed similarity scores. The image pair represented by the knowledge instance is denoted as C and D, where C is most similar to A and D is most similar to B.
4) Determine geometric relationship between A and C using image registration technique to yield transformation matrix T_{AC}.
5) Determine geometric relationship between B and D using image registration technique to yield transformation matrix T_{BD}.
6) Compute geometric relationship between A and B based on T_{AC}, T_{DB}, and $T_{knowledge}$ from the knowledge instance. This yields transformation matrix T_{AB}.
7) Align image B to image A using transformation matrix T_{AB}.

III. Experimental Results

To evaluate the effectiveness of the proposed algorithm, the algorithm was tested under two scenarios using medical imaging data from the Visible Male and Visible Female datasets from the National Library of Medicine's Visible Human Project. In the first scenario, alignment was performed between MRI-CT cranial slices from the Visible Female dataset without prior knowledge regarding the patient. The knowledge base was trained with image slices from the Visible Male dataset that were artificially distorted to simulate a larger set of patient data. In the second scenario, alignment was performed between MRI-PET cranial slices from the Whole Brain Atlas [15] with prior knowledge about the patient. Therefore, the knowledge base was trained with different artificially distorted image slices from the same patient. The root-mean squared error (RMSE) was calculated relative to gold standard control points and averaged over 10 trials for each scenario. The alignment accuracies for both scenarios are reported in Table I. It can be seen that the mean pixel error is relatively low for both cases. The MRI-CT case resulted in lower alignment accuracy than the MRI-PET case. The lower accuracy can be attributed to the lack of prior knowledge about the patient. An alignment between PET-CT cranial slices is illustrated in Fig. 4. It can be seen that the image slices were correctly aligned. This demonstrates the effectiveness of the proposed algorithm in situations where large intensity and structural differences exist between images from different imaging modalities.

TABLE I

ALIGNMENT ACCURACY

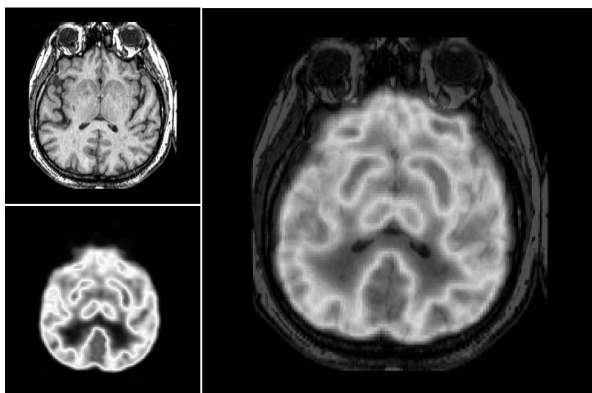| Scenario	RMSE (pixels)
MRI-CT Cranial	4.6827
MRI-PET Cranial	2.3220

Fig. 4. Alignment of MRI and PET images of a cranial slice

IV. Conclusions

This paper presents a novel approach to the multi-modal registration of medical images through the use of a priori knowledge to align images in an indirect manner. The algorithm utilizes knowledge from previous successful medical image alignments to infer a relationship between input images acquired using different modalities. Experimental results demonstrate the effectiveness of the proposed scheme at achieving a high level of registration accuracy. It is believed that this method can be used successfully for aligning medical images of different modalities for the purpose of medical image data fusion.

References

[1] A. Cideciyan, "Registration of ocular fundus images: an algorithm using cross-correlation of triple invariant image descriptors," *IEEE Engineering in Medicine and Biology Magazine*, vol. 14, no. 1, pp. 52-58, 1995.

[2] A. Averbuch and Y. Keller, "FFT based image registration," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 3608-3611, 2002.

[3] L. Heng, E. Hwa, H. Kok, "High accuracy registration of translated and rotated images using hierarchical method," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 6, pp. 2211-2214, 2000.

[4] J. Orchard, "Efficient global weighted least-squares translation registration in the frequency domain," *Proceedings of Image Analysis and Recognition*, pp. 116-124, 2005.

[5] A. Fitch, A. Kadyrov, W. Christmas, and J. Kittler, "Fast robust correlation," *IEEE Transactions on Image Processing*, vol. 14, no. 8, pp. 1063-1073, 2005.

[6] A. Wong and J. Orchard, "Efficient and robust non-rigid least-squares rectification of medical images," *Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition*, June 2006.

[7] J. Ashburner and K. Friston, "Nonlinear spatial normalization using basis functions," *Human Brain Mapping*, vol. 7, pp. 254-266, 1999.

[8] C. Studholme, J. Little, G. Penny, D. Hill, D. Hawkes, "Automated multimodality registration using the full affine transformation: applications to MR and CT guided skull base surgery," *Proceedings of Visualisation in Biomedical Computing, Lecture Notes in Computer Science 1131*, pp. 601-606, 1996.

[9] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, G., P. Suetens, "Multi-modality medical image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 16, no. 2, pp. 187-198, 1997.

[10] P. Viola and W. Wells, "Alignment by maximization of mutual information," *International Journal of Computer Vision*, vol. 24, no. 2, pp. 137-154, 1997.

[11] R. Shekhar and V. Zagrodsky, "Mutual information-based rigid and nonrigid registration of ultrasound volumes," *IEEE Transactions on Medical Imaging*, vol. 21, no. 1, pp. 9-22, 2002.

[12] S. Liao and A. Chung, "Multi-modal image registration using the generalized survival exponential entropy," *Proceedings of the 9th International Conference on Medical Image Computing and Computer Assisted Intervention, Lecture Notes in Computer Science 4191*, pp. 964-971, 2006.

[13] W. Wang, F. Ngo, J. Chen, R. Huang, K. Chou, and R. Liu, "PET-MRI image registration and fusion using artificial neural networks," *Biomedical Engineering - Applications, Basis & Communications*, vol 15, pp. 95-99, 2003.

[14] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2001.

[15] K. Johnson and J. Becker, "The whole brain atlas," Internet: http://www.med.harvard.edu/AANLIB/home.html, January 2007.