Efficient and Robust Non-Rigid Least-Squares Rectification of Medical Images

A. Wong
Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario, Canada

J. Orchard
David R. Cheriton School of Computer Science
University of Waterloo
Waterloo, Ontario, Canada

Abstract - *One of the major problems facing medical imaging is the presence of geometric distortions inherent in an imaging technique. Image registration techniques are often used to correct for such geometric perturbations. Recently, it was proposed that the SSD cost function can be evaluated efficiently using the Fast Fourier Transform (FFT) to determine the optimal translation between two images. However, spatial distortions in medical images can be highly non-rigid in nature. This paper extends this efficient approach to allow for non-rigid alignment between two images through the use of patch correspondence and robust statistical model estimation techniques. This feature-based method is designed to be highly robust, making it suitable for aligning medical images with various forms of geometric distortions. Experimental results demonstrate high overall image rectification performance.*

Keywords: image rectification, registration, medical imaging, robust statistics, least-squares.

1 Introduction

Advances in medical imaging techniques have revolutionized the field of medical diagnostics. Different medical imaging techniques such as x-ray, computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET) give new insight on the human body that can be used in the diagnosis of a disease. However, a number of issues hinder the potential of medical imaging techniques. One such problem is the presence of geometric distortions in medical images. These geometric distortions can be a result of the imaging, as in the case of MRI. MRI is becoming the imaging technique of choice in clinical examinations due to its ability to distinguish between tissue types. However, a number of hardware-related and tissue-related issues cause geometric distortions and result in poor spatial localization when compared to CT [1]. Another source of geometric distortions is the actual operating conditions of the imaging device. It is particularly true for x-ray imaging devices, where the quality of the image is highly dependent on the technician and the positioning of the patient. This situation is compounded by the fact that medical images of the same object are taken at different times and/or by different technicians, each introducing a slightly different geometric distortion. An example of typical geometric distortions on medical images can be seen in Figure 1, which shows a slice from a cranial T1-weighted MRI scan and a distorted version of the slice. It can be observed that the distorted image differs from the original reference image in terms of location, orientation, and scale. This makes it difficult to compare medical images acquired at different times and/or different modalities for medical diagnostics of a patient. Therefore, the geometric rectification of distorted images allows for better consistency with previously acquired images as well as improved spatial localization.

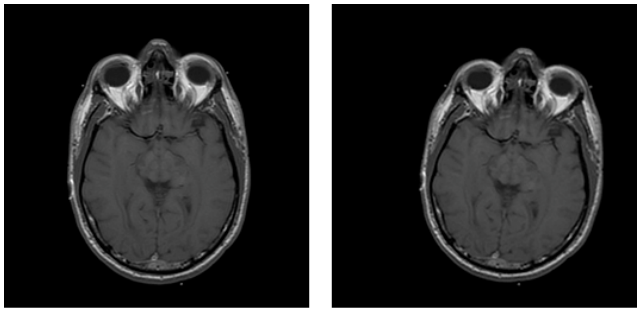

Figure 1: Left: T1-weighted MRI cranial slice
Right: Same slice with geometric distortions

One way of rectifying medical images is to have a medical expert manipulate the images manually with respect to a gold standard image. However, this is very laborious and time-consuming given the large volume of medical images acquired on a daily basis. Therefore, image registration techniques are often used to perform such geometric corrections by aligning an image to a gold standard image in an automated fashion. A number of different techniques have been proposed for the purpose of image registration. These techniques are often classified based on: 1) the feature space used as the basis of image comparison, and 2) the similarity metric used to perform the comparison. Some feature spaces include intensity, edges [2], gradient orientation [3], corners [4], and regions [5], and contours [6]. A number of different similarity metrics have been used for the purpose of image registration. Some of the more popular metrics include cross-correlation [7]-[9], phase correlation [10]-[13], sum of squared differences (SSD) [14]-[16], and mutual information [17]-[19].

Recently, it was proposed that a weighted SSD cost function may be evaluated efficiently through the use of the FFT [14][15]. This novel approach was demonstrated to perform 60 to 500 times faster than the direct approach of evaluating the SSD cost function [15]. However, this technique is limited to images where the difference between the two images is a translation and therefore is not suitable for the purpose of non-rigid image. Therefore, it is highly desired to take advantage of the efficient SSD cost evaluation technique while making it flexible enough to handle non-rigid distortions. The goal of this paper is to extend this efficient alignment framework to address this issue by making use of patch correspondence and robust statistical model estimation techniques.

The main contribution of this paper is an efficient feature-based SSD-type image alignment algorithm that can perform non-rigid alignment of images. This method is highly robust to distortions in position, orientation, and scale and therefore suitable for performing rectification on medical images with various geometric distortions. In this paper, we briefly outline the theory behind the FFT-based evaluation of the weighted SSD cost function, as well as the theory underlying patch correspondence and robust transformation model estimation. Section 2 also outlines proposed algorithm. The testing methods and test data are outlined in Section 3. Finally, experimental results performed to evaluate the rectification performance of the algorithm are discussed in Section 4, and conclusions are drawn based on the results in Section 5.

2 Theory

Before outlining the proposed rectification algorithm, it is important to discuss the theory behind the key concepts of the algorithm.

2.1 Weighted SSD Cost Function Evaluation using FFT

The cost function evaluation technique used in the algorithm is based on the technique introduced in [14] and [15]. Given the 2-D images *f* and *g*, the similarity between the two images within a region of interest (ROI) of g can be found by evaluating the weighted SSD cost between the images as

$$SSD = \sum_x \left(f(x) - g(x)\right)^2 w(x), \qquad (1)$$

where $w(x)$ is a weighting function over $g(x)$ in the range [0,1] (in the case of a ROI, 0 indicates the regions outside the ROI and 1 indicates, and x indicates the coordinate of a pixel. Using this cost function, a lower SSD cost indicates a higher image similarity. Likewise, the SSD cost between a shifted image f and an image g is

$$SSD(\delta) = \sum_x \left(f(x-\delta) - g(x)\right)^2 w(x), \qquad (2)$$

Where δ represents the shift. Therefore, to find the shift of f that provides the global optimal alignment of image f and g, it is necessary to find the shift δ that results in the lowest SSD cost function,

$$\delta = \arg\min_{\delta} \left[\sum_x \left(f(x-\delta) - g(x)\right)^2 w(x) \right]. \qquad (3)$$

This is commonly known as the least squares problem. If done in a direct fashion, each shifted image of f needs to be compared to g using the SSD cost function. This evaluation process is computationally expensive and so a more efficient way to determine the shift associated with the minimum SSD cost is desired. As described in [8] and [9], Equation 2 can be expanded and expressed as

$$\begin{aligned} SSD(\delta) = & \sum_x f^2(x-\delta)w(x) \\ & - 2\sum_x f(x-\delta)g(x)w(x) + \sum_x g^2(x)w(x). \end{aligned} \qquad (4)$$

The last term of Equation 4 is irrelevant to the optimization problem and can be ignored since it is independent of δ. The first two terms can be turned into convolutions. Therefore, the final weighted SSD cost function $PSSD(\delta)$ is expressed as

$$\begin{aligned} PSSD(\delta) = & \left\{ \bar{f}^2(x) * w(x) \right\}_\delta \\ & - 2\left\{ \bar{f}(x) * \left(g(x)w(x)\right) \right\}_\delta \end{aligned} \qquad (5)$$

where $\bar{f}(x) = f(-x)$ and $*$ indicates convolution. The bottleneck of this final cost function is the evaluation of the convolution operations. These can be evaluated very efficiently in the frequency domain due to the fact that convolutions in the spatial domain become multiplications in the frequency domain. Therefore, the terms in Equation 5 can be evaluated for all possible values of δ by taking the Fourier transform of the convolution terms, performing multiplication in the frequency domain, and then taking the inverse Fourier transform of the result, which can be expressed as

$$\begin{aligned} PSSD(\delta) = & F^{-1}\left\{ F\left(\bar{f}^2(x)\right) F\left(w(x)\right) \right\}(\delta) \\ & - 2F^{-1}\left\{ F\left(\bar{f}(x)\right) F\left(g(x)w(x)\right) \right\}(\delta). \end{aligned} \qquad (6)$$

Therefore, the performance bottleneck becomes the computation of the FFTs and the inverse FFTs (IFFTs). For large image sizes, significant performance

improvements can be achieved by evaluating the cost function in this manner since the FFT and the IFFT can be evaluated efficiently.

2.2 Patch Correspondence

The main drawback of using the weighted SSD cost evaluation technique described in Section 2.1 is that fact that it is only suitable for situations where the difference between two images is a translation. Since distortions in medical images include rotation, scale, and skew, it is not possible to apply this algorithm directly for the purposes of non-rigid image rectification. One way of getting around this issue is through the concept of patch correspondence.

Let us treat the content of an image as bounded by a rectangular box. It is true that non-rigid transformations on a rectangular box cannot be modeled using a single translation transformation. However, this is in direct contrast with the case for a single point in an image, where the only transformation is a single translation. Taking this concept one step further, it can be said that a small rectangular patch of neighboring pixels around a single point can be treated in a similar manner, as the geometric distortions in the patch can be treated as minimal relative to that of the overall image. As a result, it is possible to extract patches from the gold standard image *g* and find its corresponding patch in the distorted image *f* and represent the transformation as a simple translation. Finally, given enough pairs of matching patches (depending on the type of transformations supported), the estimated non-rigid transformation between the two images can be derived based on the patch centroids and used to rectify geometric distortions in the distorted image.

There are some issues that need to be considered for the use of patch correspondence in non-rigid image alignment. First, an automated method is needed for selecting patches from the gold standard image. These patches can then be used for finding corresponding patches in the distorted image. Each patch selected from the gold standard image should ideally be chosen so that its content is unique, and no other patch in the image is a close match. A method of selecting such patches is to detect points on the image that have high feature significance and to use them as the centroids of the patches. This can be accomplished by selecting feature points using the modified Harris corner detector outlined in [20]. An example of patch centroids selected from a T1-weighted MRI gold standard slice using this method is shown in Figure 2.

The second issue that needs to be considered is with regards to how patch correspondence can be evaluated using the efficient SSD cost function evaluation technique described in Section 2.1. A patch of the gold standard image can be regarded as an ROI, within which the distorted image is aligned using the weighted SSD cost function. The optimal shift between the two images within the ROI can then be treated as the translation vector that brings the distorted image in alignment with the gold standard image. As such, this vector can then be used to determine the position of the corresponding patch in the distorted image. This can be illustrated using a simple example. Imagine that a patch was selected with a centroid at $(x_1, x_2) = (50,50)$ in the gold standard image. To perform patch correspondence evaluation for this patch, the weighting function $w(x)$ defined in Section 2.1 is set up such that the weight for each pixel in the patch is 1, while all other weights are set to 0. When the distorted image *f* and the gold standard image *g* are aligned within the window specified by *w*, let the resulting optimal shift be $\delta = (20,20)$. From this, it can be determined that the corresponding patch in *f* has a centroid at $(x_1', x_2') = (x_1 - \delta_1, x_2 - \delta_2) = (30,30)$. Once all matching pairs of patches has been found, the centroids of matching patches are extracted to form a set of candidate control points for transformation model estimation. An example of matching patches found using this approach is shown in Figure 3.

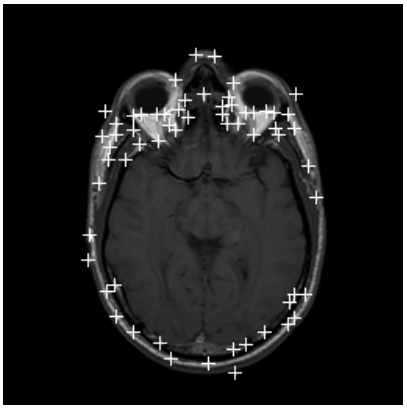

Figure 2: Patch centroids selected in T1-weighted MRI cranial slice used as gold standard

This approach to finding a corresponding patch based on patch-to-image comparisons is considerably different than traditional approaches, where specific patches are detected from both images and are matched on a patch-by-patch basis. One of the main reasons that patch-to-patch approaches are used is to reduce the search space, as a patch-to-image approach would be computationally expensive if computed in a direct manner. However, the

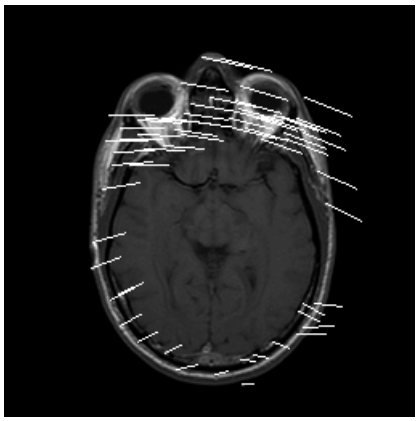

Figure 3: The translation vectors between matched patches shown relative to gold standard T1-weighted MRI slice

SSD cost evaluation method described in Section 2.1 is able to perform similarity evaluation for all possible shifts in a very efficient manner and therefore making it feasible to take the patch-to-image approach. There are two main advantages to the proposed method of finding matching patches. First, unlike the patch-to-patch approach, the SDD cost evaluation method finds the global optimum. Second, the proposed approach only needs to detect feature points from the gold standard image. Therefore, all subsequent alignments with a particular gold standard image would not require any form of feature detection and thus reduce computational complexity in the long run. Another performance optimization that can be performed is to precompute and store the Fourier transforms of $\bar{f}(x)$ and $\bar{f}^2(x)$, as they are used in the SSD cost function evaluation of each patch.

2.3 Robust Transformation Model Estimation

The transformation model that aligns the distorted image to the gold standard image needs to be determined based on the set of candidate control points found in Section 2.2. There are two main issues that need to be considered to determine the final transformation model. First, it is necessary to determine the type of transformation model needed for proper alignment between the distorted image and the gold standard image. The selection of a transformation model is dependent on the types of geometric distortions that exist in the distorted image. In the case of medical images, typical geometric distortions include translation, rotation, scale, and shear. A projective model can be used to express all of the above geometric distortions. The projective transformation model can be expressed as an invertible 3×3 matrix

$$T = \begin{bmatrix} t_0 & t_1 & t_2 \\ t_3 & t_4 & t_5 \\ t_6 & t_7 & t_8 \end{bmatrix}, \tag{7}$$

$$\text{where } x_1' = \frac{t_0 x_1 + t_1 x_2 + t_2}{t_6 x_1 + t_7 x_2 + t_8}, \tag{8}$$

$$\text{and } x_2' = \frac{t_3 x_1 + t_4 x_2 + t_5}{t_6 x_1 + t_7 x_2 + t_8}. \tag{9}$$

The second issue that needs to be considered regards how the parameters of the selected transformation model are estimated. One method of parameter estimation for a projective transformation matrix is the normalized direct linear transformation (DLT) algorithm described in [21]. One problem with using this parameter estimation algorithm directly on the set of candidate control points is the fact the DLT algorithm is a least-squares estimation method and therefore performs poorly in the presence of outliers. Since the error is squared, far outliers can have an enormous effect on the estimated transformation model. Once the set of candidate control points are selected using an automated algorithm as opposed to manually selected, the possibility of mismatched patches is relatively high. One method of addressing this issue is to prune outliers from the set of candidate control points through the use of the Random Sample Consensus (RANSAC) algorithm [22]. In the RANSAC algorithm, a number of data points are selected at random from the complete sample set, which in this case is the set of control points. The selected data points are then used to estimate a model, which in this case is the 3×3 projective transformation matrix. The number of outliers is then determined based on the estimated model. For this proposed algorithm, the number of outliers is determined by calculating: 1) the estimated transformed points $\underline{x'}$ of each control point x in the distorted image using the estimated model T_{est}, and 2) the estimated transformed points $\underline{x}$ of each control point x' in the gold standard image using the inverse mode T_{est}^{-1}. The SSD cost between $\underline{x'}$ and x', and between $\underline{x}$ and x is computed for each matching pair of control points, and the pairs whose cost is greater than or equal to a pre-defined threshold t are counted as outliers. This process is repeated for another K iterations, where

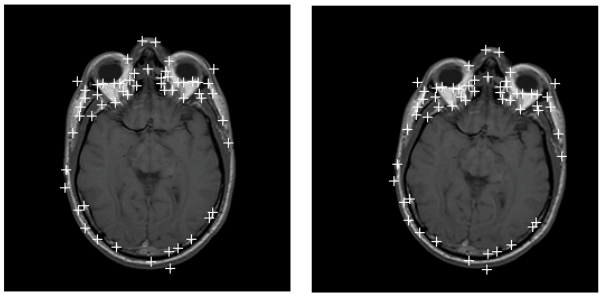

Figure 4: Left: Final control points in gold standard image Right: Final control points in distorted image

$$K = \frac{\log(1-p)}{\log\left(1-\left(n_{\text{inliers}}/N\right)^s\right)}, \tag{10}$$

and p is the desired estimated probability that all available data points fit the estimated model, n_{inliers} is the number of data points that fit the current estimated model, N is the total number of data points, and s is the minimum number of samples required to fit a model, which in the case of projective transformation model estimation is 4. At the end of the RANSAC algorithm, a final set of control points are determined by finding the set of inliers from the estimated model with the highest consensus. An example of the final set of control points for the distorted image and the gold standard image is shown in Figure 4. This set of control points, now free of outliers, is then used to estimate the final transformation model using the DLT algorithm.

2.4 Rectification Algorithm

Based on the theory presented, the alignment algorithm can be outlined as follows:

1. Given a distorted image f and a gold standard image g:
2. Extract a set of feature points from g using the modified Harris corner detector. Each of these feature points represents the centroid of a patch.
3. For each patch p detected from g, perform weighted SSD cost evaluation using FFT as described in Section 2.1 within a ROI bounded by p_g to determine the optimal shift between f and g. Use this shift as a translation vector to determine the location of the corresponding patch p_f in f.
4. Extract the centroids of matching patches to form a set of candidate control points.
5. Apply the RANSAC algorithm on the set of candidate control points using the projective transformation model to determine a final set of control points.
6. Use the DLT algorithm to estimate a final projective transformation model from the final set of control points.
7. Use the final projective transformation model to transform f into a geometrically corrected image f'.

3 Testing Methods

The proposed algorithm was implemented in MATLAB and was tested using three sets of images derived from the Visible Male dataset of the National Library of Medicine's Visible Human Project. All test images are 8-bit grayscale images. Each test set consists of a gold standard image and 5 distorted images generated from that image with random perturbations in location, orientation, scale, and projective skew. A description of each test set is given below.

THORAX: A 310×310 PD-weighted MRI coronal slice through the thorax.

HEAD: A 256×256 T1-weighted MRI horizontal slice through the head.

PELVIS: A 310×310 T1-weighted MRI coronal slice through the pelvis and upper thigh.

Each distorted image in a test set is rectified with the gold standard image in the same test set. To judge the rectification performance of the proposed algorithm, the mean squared error (MSE) of pixel intensity is computed between the rectified image and the gold standard image. The MSE between the original distorted image and the gold standard image is used as a reference.

Since we know the true transformation for each of our test cases, we can calculate how far each pixel is from its true position after rectification. We compute the average pixel displacement for all pixels, excluding background pixels. If the rectification is successful, the average pixel displacement should be small.

Table 1: MEAN SQUARED ERROR[1]

Test Set	Distorted Image	Rectified Image
THORAX	2097.00	12.00
HEAD	1519.40	8.44
PELVIS	2791.00	12.63
Overall	**2135.80**	**11.02**

1. Averaged over 5 test images in each test set.

Table 2: AVERAGE PIXEL DISPLACEMENT[2]

Test Set	Distorted Image	Rectified Image
THORAX	25.72	1.13
HEAD	22.78	1.69
PELVIS	35.75	2.69
Overall	**28.08**	**1.84**

2. Averaged over 5 test images in each test set.

4 Experimental Results

The MSE results are shown in Table 1. The proposed rectification method shows substantial improvements in MSE compared to the original distorted image in all test cases. Table 2 shows the average pixel displacements for the distorted and rectified images. Before rectification, the average pixel displacement is 28.08 pixels. After rectification, the average pixel displacement is reduced to only 1.84 pixels.

Examples of the rectification results achieved for each test set are shown in Figures 5, 6, and 7. In each case, the distorted image is correctly rectified with respect to the gold standard image. These results illustrate the effectiveness of the proposed image rectification algorithm for non-rigid distortions.

5 Conclusions and Future Work

In this paper, we have introduced a new method for efficient and robust least squares image rectification based on the concepts of patch correspondence and robust statistical model estimation. Experimental results show that overall rectification performance is relatively high. It is our belief that this method can be successfully implemented for medical image rectification purposes.

Future work includes investigating the effectiveness of quasi-orientation maps as a feature space for multi-modal non-rigid image registration.

Acknowledgments

The authors would like to thank the National Library of Medicine for the Visual Human Project data.

References

[1] D. Wang and D. Doddrell, "Geometric distortion in structural magnetic resonance imaging," in *Current Medical Imaging Reviews*, 2005, Vol. 1, No. 1, pp. 49-60.

[2] W. Kerwin and C. Yuan, "Active edge maps for medical image registration," in *Proc. SPIE*, 2001, Vol. 4322, pp. 516-526.

[3] A. Fitch, A. Kadyrov, W. Christmas, and J. Kittler, "Orientation Correlation," in *British Machine Vision Conference*, 2002, Vol. 1, pp. 133-142.

[4] P. Sand and S. Teller, "Video Matching," in *ACM Transactions on Graphics*, 2004, Vol. 22, No. 3, pp. 592-599.

[5] M. Ali and D. Clausi, "Automatic registration of SAR and visible band remote sensing images," in *IEEE International Geoscience and Remote Sensing Symposium*, 2002, Vol. 3, pp. 1331-1333.

[6] H. Li, B. Manjunath, and S. Mitra, "A contour-based approach to multisensor image registration," in *IEEE Transactions on Image Processing*, 1995, Vol. 4, No. 3, pp. 320-334.

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

[8] M. Solaiyappan and S. Gupta, "Predictive registration of cardiac MR perfusion images Using geometric invariants," in *Proc. ISMRM*, 2000, Vol. 1, p. 37.

[9] J. Le Moigne and R. Cromp, "Wavelets for remote sensing image registration and fusion," in *Proc. SPIE*, 1996, Vol. 2762, pp. 535-544.

[10] E. Castro, and C. Morandi, "Registration of translated and rotated images using finite Fourier transforms," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1987, Vol. 9, No. 5, pp. 700–703.

[11] B. Reddy and B. Chatterji, "An FFT-based technique for translation, rotation and scale invariant image registration," in *IEEE Transactions on Image Processing*, 1996, Vol. 5, No. 8, pp. 1266–1271.

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

[13] A. Averbuch and Y. Keller, "FFT based image registration," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2002, Vol. 4, pp. 3608-3611.

[14] A. Fitch, A. Kadyrov, W. Christmas, and J. Kittler, "Fast robust correlation," in *IEEE Trans. Image Processing*, 2005, Vol. 14, No. 8, pp. 1063–1073.

[15] J. Orchard, "Efficient global weighted least-squares translation registration in the frequency domain," in *Image Analysis and Recognition (ICIAR)*, 2005, pp. 116–124.

[16] B. Lucas and T. Kanade, "An interative image registration technique with an application to stereo vision," in *Proc. Imaging Understanding Workshop*, 1981, pp. 121-130.

[17] W. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volumn registration by maximization of mutual information", in *Medical Image Analysis*, Oxford University Press, Vol. 1, No. 1, pp. 35-51.

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

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

[20] J. Noble, "Descriptions of image surfaces," D.Phil. thesis, Robotics Research Group, Department of Engineering Science, Oxford University, p. 45, 1989.

[21] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, 2001, Cambridge University Press.

[22] M. Fischler and R. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," in *Comm. of the ACM*, 1981, Vol. 24, pp. 381-395.

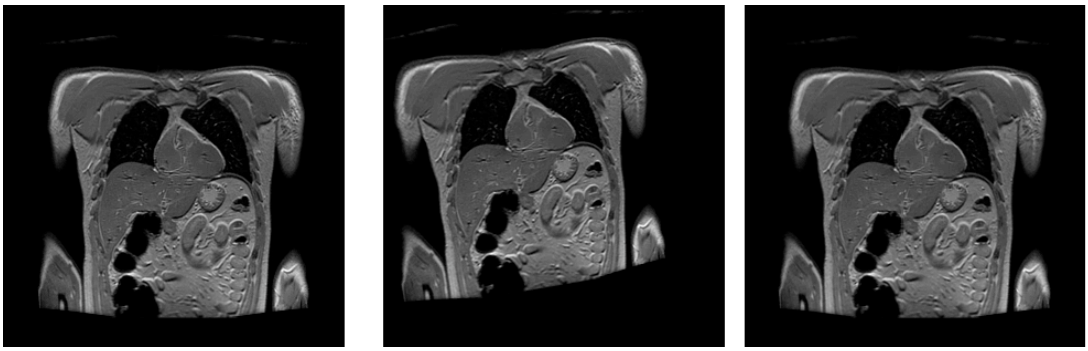

Figure 5: Rectification of PD-weighted MRI from THORAX test set
Left: Gold Standard; Center: Distorted; Right: Rectified

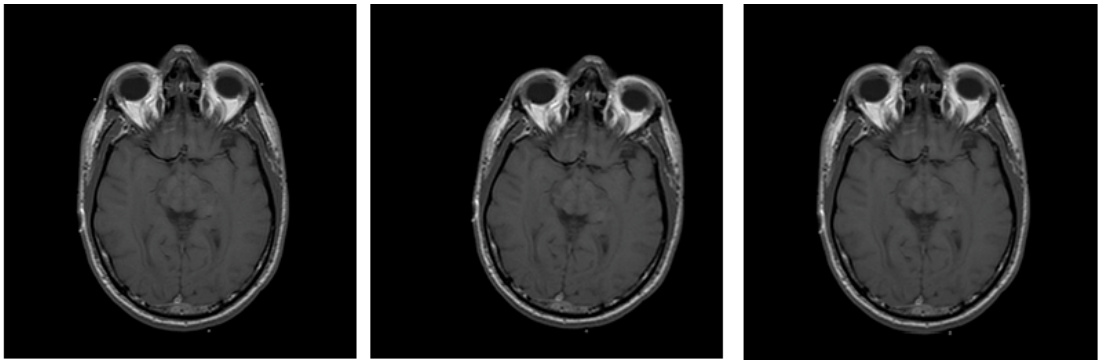

Figure 6: Rectification of T1-weighted MRI from HEAD test set
Left: Gold Standard; Center: Distorted; Right: Rectified

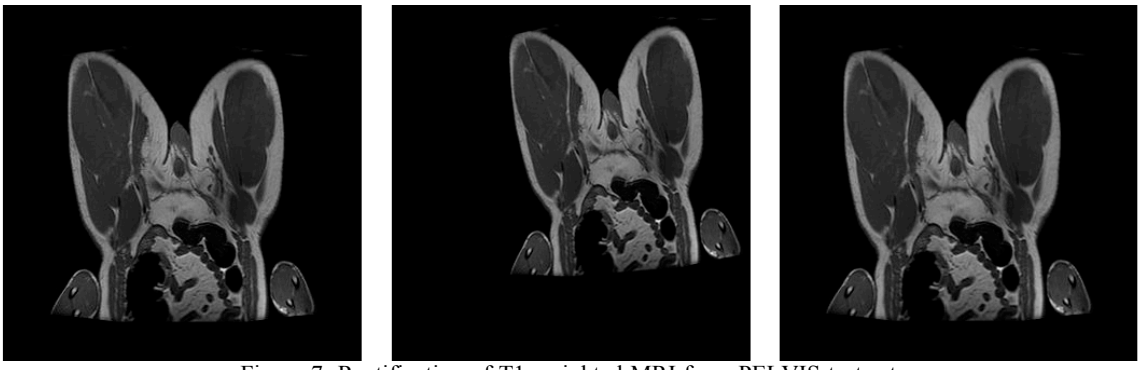

Figure 7: Rectification of T1-weighted MRI from PELVIS test set
Left: Gold Standard; Center: Distorted; Right: Rectified