CHAPTER 3: A MOVING LEAST SQUARE BASED FRAMEWORK FOR THORACIC CT IMAGE REGISTRATION

This chapter proposes an automatic registration process for tracing of the deformity path of the thoracic region based on feature detector cum descriptor Speeded Up Robust Feature (SURF) and Moving Least Squares (MLS). Corresponding control point pairs or landmarks in image pairs or groups can be used to define the deformation with respect to time, point of view or modality. The manual definition of the number of control points in an image such that it is enough to define all kinds of deformations in that respective image is a tedious task. Hence, an automatic definition of control points has been the way forward taken in the proposed work. The control point cloud on the images is defined by its feature set which is obtained by using the SURF detector/descriptor, which serves as an input for MLS algorithm to trace the deformations of the image (thoracic image in this case).

The proposed strategy begins with having a pair of images of same dimensions to be registered, obtained at different timestamps of a temporal image sequence. The corresponding anatomical landmark points are identified on both these images. These landmark points are inputs to a point cloud assisted continuous surface regeneration algorithm. This technique analytically solves a number of least squares problems to find the local elastic transformations. Applying these local transformations on the image pair creates the deformations throughout the considered sequence of images. The relative deformations between the image pair are accounted for by the surface reconstruction algorithm and are

adjusted using respective control point sets for both images thus registering one with respect to another. In our experiments, Target Registration Error (E_{TR}) is used as the quantitative measure for the evaluation of performance. The E_{TR} obtained for the dataset was found to be considerably lower than more traditional and prevalent transforms such as affine, thin-plate splines or finite element based approach. Therefore, the MLS-based method was found to be more suitable for the real-time applications of Image Guided Interventions (IGI) demanding higher speed and more accurate image registrations.

3.1 Introduction

Deformable images have been a constant focus of study and research in the field of image registration. There have been extensive intra and inter-patient studies and projects in this particular direction which have resulted in a variety of registration algorithms. Deformable registration is considered as an ill-posed problem because there is generally no unique solution to a registration problem. Usually it is formulated as an optimization problem. In the case of study of thoracic images from inhale to exhale phases (or vice versa), finding a finally registered image is a problem for which many algorithms and methods have been proposed, tested and compared over the time.

Apart from the peripheral deformations happening in the thoracic region, there are also local deformations of the internal organs (within the periphery) going on throughout the breathing process. The definitions of the deformity patterns governing those motions are quite unclear yet and much research hasn't happened towards this direction. It is a problem of bigger proportions for oncology researchers and radiologists alike to define and describe the inner deformations in their studies. This study's clinical relevance cannot be stressed

upon more. Since overall respiratory motion is related to lung function, it has a diagnostic value to itself. Any organ motion pertaining to breathing can lead to image artefact and position uncertainties during image guided clinical interventions. A particular case for such image guided interventions (IGI) can be the radiotherapy planning of thoracic and abdominal tumours; the respiratory motion causes important uncertainties and is a significant source of error [Keall, P. J. *et al.*, 2006]. Image registration has recently started playing an important role in this scenario; it helps in the estimation of any motion caused due to breathing during acquisition and the description of the temporal change in position and shape of the structures of interest by establishing the correspondence between images acquired at different phases of the breathing cycle [Ehrhardt, J. *et al.*, 2011]. The present study intends to shed light on the deformity paths of local deformations in the thoracic periphery which can be helpful as a prerequisite for radiation therapy (based on their dosimetric evaluations), tumour growth progression with time and also towards making deformable subject specific motion models more precise and accurate.

3.2 Background

The background study of this chapter initially includes a study of few most prominent proposed algorithms in the direction of study of the moving least squares and its applications. Then the proposed methods relating to image registration of thoracic CT images are discussed. The propositions are categorically discussed keeping in mind their acute relevance and their year of occurrence. Propositions occurring at a later instant in timeline are given higher priority in discussion in comparison to earlier works to establish better context. These methods are compared in a tabular format in table B.1 in Appendix B.

The concept of least square methodologies such as weighted least squares and moving least squares were first proposed in 1981 by P. Lancaster and K. Salkauskas [Lancaster, P. & Salkauskas, K., 1981]. They presented "an analysis of least squares methods for smoothing and interpolating scattered data/points". They proposed a non-interpolating least squares method as an alternate representation of the local approximation based on the choice of weight functions. This became the basis of more recent moving least squares, giving it its characteristic dynamic weight function choice option to project smoother surfaces for all data points coming into consideration in real time. In particular, they proved theorems concerning the smoothness of interpolants and the description of MLS processes as projection methods. The differences between interpolating and non-interpolating MLS method as projection methods were pointed and singled out. Effects of choice of weight functions and asymptotic behaviour of such single variable and multivariate functions have been studied in detail.

In the earlier discussed method, interpolating and non-interpolating non-linear least square methods have been discussed as projection of given scattered data/points. These scattered points were not necessarily representations/emulations of real world objects around us in multiple dimensions. One of the most prominent works in the direction of emulating/projecting real world objects from scattered point-sets/data using moving least squares came from M. Alexa & associates [Alexa, M. *et al.*, 2003]. This work stressed upon the use of point sets to represent shapes. It set its goal in defining surfaces from a set of points close to an original surface, this is approximated using MLS. A projection procedure has been defined which would project any point near the point set onto the surface. Then, the MLS surface is defined as the points

projecting onto themselves. The smoothness conjecture is motivated and respective projection is computed. The proposed model was tested on 'the Stanford bunny' (a computer graphics 3D test model developed by Greg Turk and Marc Levoy in 1994) along with other models. The proposed approach showed smoother silhouettes and more accurate highlights in comparison to more traditional methods like Splatting and Gouraud-shaded mesh model. Problem with both Splatting and Gouraud-shaded mesh was found to be that these models were not sampled densely enough exhibiting relatively inaccurate highlights. A parameter *h* connected to feature size is used such that features with radius smaller than *h* are smoothed out. Actual timings and memory requirements for the projection procedure depended heavily on the feature size *h*. As long as *h* was small with respect to the total diameter of the model, the use of main memory of the projection procedure was found to be negligible. For the Stanford bunny dataset, total 36,000 points of the bunny were projected on to a surface definition in roughly 10-30 seconds. It was found to be possible to provide a point set representation that conforms to a specified tolerance and the use of a point set (without connectivity) as a representation of shapes.

While the earlier works included using MLS to project known multivariate functions and established computer graphics 3D test models like the Stanford bunny and the Aphrodite statue. This work [Schaefer, S. *et al.*, 2006] by S. Schaeffer and associates implemented the concept of MLS to define deformations in rigid images of real world objects. They proposed an image deformation method based on Moving Least Squares using various classes of linear functions including affine, similarity and rigid transformations. These deformations were realistic and gave the user an impression of manipulating real-world objects.

Image deformations were built based on collections of points with which the user could control the deformation. A deformation function was constructed satisfying the three properties of Interpolation, Smoothness & Identity using MLS. The proposed method was applied for Affine, Similarity, Rigid & Elastic deformations on a set of images. It was found to perform deformations faster than the contemporary methods. Deformations were constructed such as to minimize the amount of local scaling and shear and restricting the classes of transformations used in Moving Least Squares to similarity and rigid-body transformations. This method, using MLS completely avoided the use of input image triangulation unlike the method proposed by T. Igarashi and associates [Igarashi, T. *et al.*, 2005], thus producing globally smooth deformations. It was showed how solutions could be computed directly from the closed-form deformation using similarity transformations thereby bypassing the non-linear minimization (contrary to Igarashi *et al.* 2005). The method is generalized enough to accommodate different distance metrics dependent on the topology of the shape rather than the simple, Euclidean distance used as weight factor.

The methods discussed till now deal with known multivariate function projections into point sets or data/points projections of established 3D models or definition of everyday objects' rigid image deformities using MLS. Richard Castillo and associates [Castillo, R. *et al.*, 2009] suggested a framework for deformable image registration of two images using MLS for corresponding sets of feature landmark point pairs in both images and evaluation of its spatial accuracy. They used an in-house developed Matlab® based software interface called *APRIL* (*A*ssisted *P*oint *R*egistration of *I*nternal *L*andmarks) to facilitate manual selection of landmark feature pairs between image volumes. This point set of the pair when

subjected to MLS, registered the source landmark point set to the corresponding target point set. The image registration error was calculated in terms of fiducial error or spatial errors. The uncertainty of spatial error U_{SE} estimates was found to be inversely proportional to the square root of the number of landmark point pairs and directly proportional to the standard deviation of spatial errors i.e. $U_{SE} \propto 1/(L_{pp})^{1/2}$ & $U_{SE} \propto SD_{SE}$. Cumulative distribution functions (CDFs) were generated from the corresponding set of error measurements for each case. To simulate the spatial error information derived from validation point sets of different size, uniform samples of the individual CDFs were obtained for sample sizes ranging from 10 to 5000. For each sample size, 100,000 independent sample sets were obtained. At each sample size increment, an independent calculation of the mean spatial error was performed for each of the 100,000 error samples. The feasibility of generating large (>1100) validation landmark sets has been demonstrated on five component phase pairs from clinically acquired treatment planning 4D CT data. The results demonstrate that large landmark point sets provide an effective means for objective evaluation of DIR with a narrow uncertainty range, and suggest a practical strategy for qualitative analysis of DIR spatial accuracy on a routine clinical basis. No proposition on the estimation of deformity between the registered image pairs were made though.

The EMPIRE10 challenge conducted by K. Murphy and associates [Murphy, K. *et al.*, 2011] was a study of Evaluation of Registration Methods on Thoracic CT. EMPIRE10 (Evaluation of Methods for Pulmonary Image REgistration 2010) is a public platform for fair and meaningful comparison and evaluation of non-rigid registration algorithms and techniques which are applied to a database of intra-patient thoracic CT image pairs. Evaluation of nonrigid

registration techniques is a nontrivial task. This is compounded by the fact that researchers typically test only on their own data, which varies widely. For this reason, reliable assessment and comparison of different registration algorithms has been virtually impossible in the past. The result of this study comprised of a comprehensive evaluation and comparison of 20 individual algorithms from leading academic and industrial research groups. All algorithms were applied to the same set of 30 thoracic CT pairs. Algorithm settings and parameters were chosen by researchers' expert in the configuration of their own method and the evaluation is independent, using the same criteria for all participants. Some methods up for comparison were: Asclepios1, Asclepios2, CMS, DIKU, DROP, elastix, IMI Lubeck Diffeomorph, Lyon FFD, MGH, Nifty Reggers, OFDP, picsl exp, picsl gsyn, Robust TreeReg Leuven, Spline MIRIT Leuven etc. All methods were fully automatic with the exception of MGH. It was found that generic registration algorithms performed better than data specific methods. It might still be the case that combining aspects of both could improve performance even further, particularly on more difficult scan pairs. The EMPIRE10 challenge enabled detailed, independent and fair evaluation of non-rigid registration algorithms.

Edward Castillo and associates [Castillo, E. *et al.*, 2014] proposed a moving least squares approach for computing spatially accurate transformations that satisfy strict physiologic constraints. It involved computation of physiologically realistic spatial transformation from a sparse point cloud of displacement estimates using MLS and any combination of upper bound, lower bound, or equality constraints placed on the Jacobian. MLS defined a spatial transformation from a sparse point cloud of estimated displacements and provided

simple analytic derivative estimates for all voxel locations; given displacement estimates from the automated block. Five publicly available (cases 6-10 from www.dir-lab.com) inhale/exhale thoracic CT image pairs each with 300 landmarks (for DIR validation) were registered by first obtaining a sparse point cloud of displacement estimates via block matching. Two MLS transformations were then computed, one with no Jacobian constraints and the other with strict contraction Jacobian constraints. Both MLS fields achieved similarly low average millimeter error on all five cases (between 1.16 – 1.26). However, the constrained MLS yielded a strict contraction (all Jacobian values between 0 and 1) while the unconstrained MLS resulted in regions of expansion (Jacobian values larger than 1) despite registering from inhale to exhale. The proposed MLS approach was found capable of producing Jacobian constrained transformations without any degradation in the spatial accuracy. Though applied to block match estimates, the approach can be employed in conjunction with displacement estimates from any DIR algorithm.

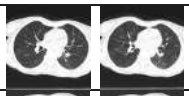
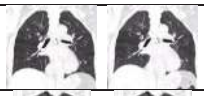
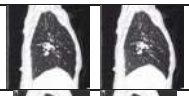
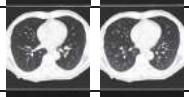
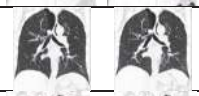
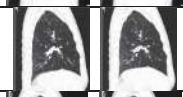
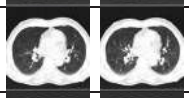
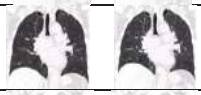
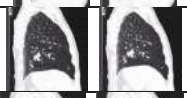
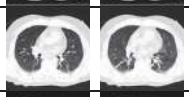
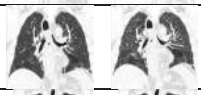
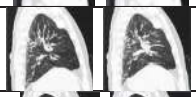
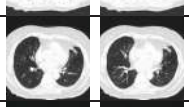
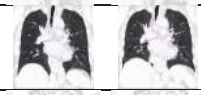
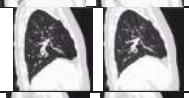
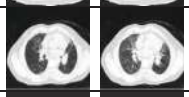
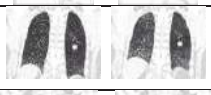
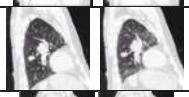
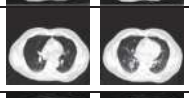
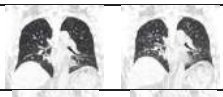
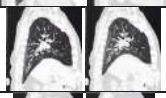
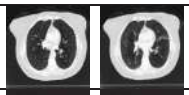
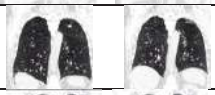
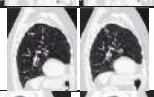
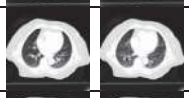
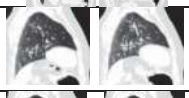
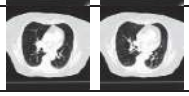
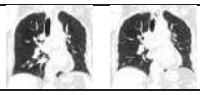
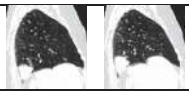
3.3 Method

3.3.1 Preparation

The dataset used comprised of a total (3×10)×10 i.e. 300 CT images across 10 subjects ranging from 396×396 to 432×400 pixels. There were 10 frames each for every anatomical plane i.e. Axial (supine), Coronal and Sagittal for all the 10 subjects acquired simultaneously with a gap of 0.1 second starting from time t = 0 to 1 seconds, thus called the 4DCT image dataset. All images were identified as $I_N^{AP}(x, y, t)$ where $\{N, t \in \mathbb{R}^+ | 1 \leq N \leq 10, 0.1 \leq t \leq 1\}$, (x, y) are the x & y coordinates in the Cartesian plane and *AP* signifies the three anatomical planes of view i.e. Axial (a), Coronal (c) and Sagittal (s). So, the sixth subject's Coronal CT

image acquired at t=0.6 sec will be identified as $\boldsymbol{I_6^c(x, y, 0.6)}$. Sample of images used from all viewpoints and all subjects at timestamps 0.1 and 0.6 seconds are summarized in Table 3.1.

Table 3.1: All three anatomical viewpoints for all the 10 subjects at time t=0.1 & 0.6 sec

	ANATOMICAL PLANES (at t=0.1 & 0.6 sec)		
	Axial	Coronal	Sagittal
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

3.3.2 Proposed Methodology

The methodology is as such that two images at different timestamps are taken; one with the earlier timestamp i.e. t=0.1 sec is considered as base/reference image and the one with the later timestamp i.e. t=0.6 sec deformed image (real time

image). These are fed into the SURF feature detector and a common corresponding feature set in the form of a Cartesian point cloud is obtained. This feature set then serves as input to the MLS algorithm as control point cloud corresponding to both reference and the deformed image. The MLS algorithm then traces the deformation in the real time image with respect to the corresponding base image. The overall process can be referred to in Figure 3.1.

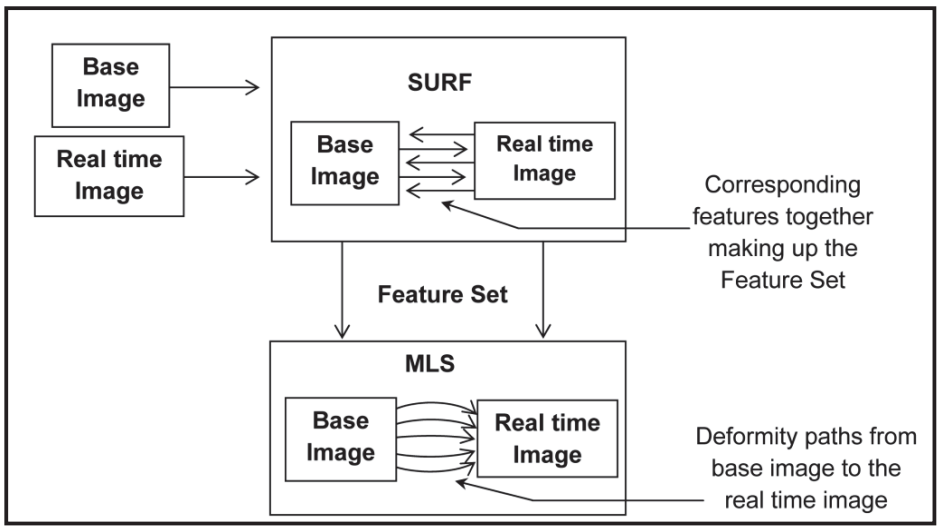


Figure 3.1: The proposed model

The proposed model uses the Speeded up robust feature detector (SURF) [Bay, H. *et al.*, 2006 & 2008] to obtain a feature set of the deformed image as well as the reference image. It detects and describes the feature set irrespective of any scaling and/or rotation in the corresponding images. SURF gives better results than previously proposed schemes with respect to repeatability, distinctiveness, and robustness, yet can be computed and compared much faster than any other state of the art feature detector [Pang, Y. *et al.*, 2012; Yoon, H. *et al.*, 2009]. This was achieved by relying on integral images for image convolutions; by building on the strengths of the leading existing detectors and descriptors (specifically, using a Hessian matrix-based measure for the detector, and a distribution-based

descriptor); and by simplifying these methods to the essential. This leads to a combination of novel detection, description, and matching steps [Bay, H. *et al.*, 2008] as can be seen in Figure 3.2. An implementation of the algorithm over the inhale and exhale frames for the first subject at t=0.1 and 0.6 sec respectively i.e. $I_1^A(x, y, (0.1, 0.6))$ is shown in Figures 3.3, 3.4 & 3.5.

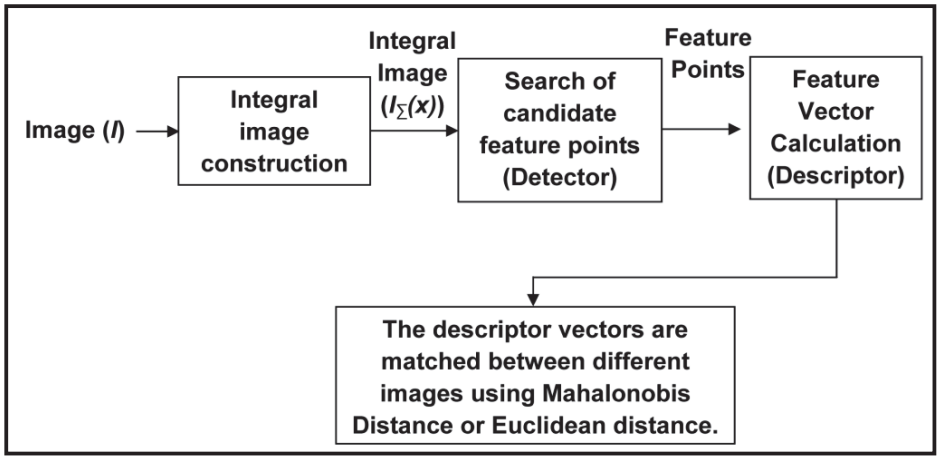


Figure 3.2: The working model for SURF

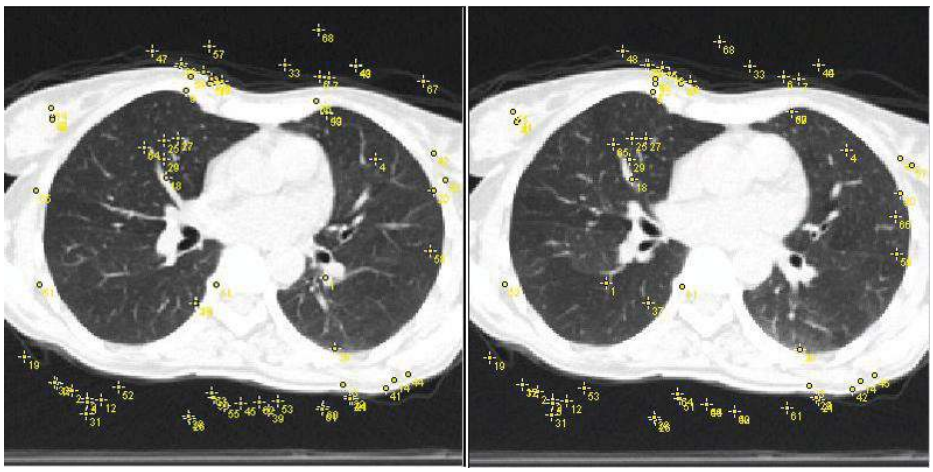


Figure 3.3: Corresponding feature points at their respective positions in the inhale (left) and the exhale (right) frame recorded at t= 0.1 and 0.6 seconds respectively

MLS has been successfully applied to surface reconstruction from points/point clouds and other point set surface definitions [Alexa, M. *et al.*, 2003,

Schaefer, S. *et al.*, 2006]. Given a set of control point pairs on the source and the target images, the MLS technique determines the transformation $T_v(x)$ that best minimizes the least square error expression:

$$\sum_i |T_v(p_i) - q_i|^2 \tag{3.1}$$

where p_i and q_i are the i[th] source and target control point pair respectively obtained by SURF algorithm. Since a single affine transformation is obtained as a result of the above transformation, there is no control over the scaling/shearing of the image. This issue is fixed by including a weighing function w_i to the least square error function which results in a different transformation function for each point of evaluation of the image.

$$\sum_i w_i |T_v(p_i) - q_i|^2 \tag{3.2}$$

where the weighing function w_i is of the form:

$$w_i = {}^{1}/_{\|p_i - v\|^{2\alpha}} \tag{3.3}$$

where v is the point of evaluation in eqn. (3.1-3.3), α is a parameter of the weighing function. This parameter changes values depending on the changing point of evaluation thus changing the weighing function (eqn. 3.3) and in turn changing the transformation for each point of the image. It performs better than its contemporaries while tracing deformations that are realistic and guides the user in manipulation of real-world objects. It also allows the user to specify the deformations using either sets of points or line segments, the later useful for controlling curves and profiles present in the image. For each of these techniques, it provides simple closed-form solutions that yield fast deformations, which can

be performed in real-time as is shown in Figure 3.5 as a result of using point-set features from the inhale and the exhale frames in Figure 3.3 & 3.4.

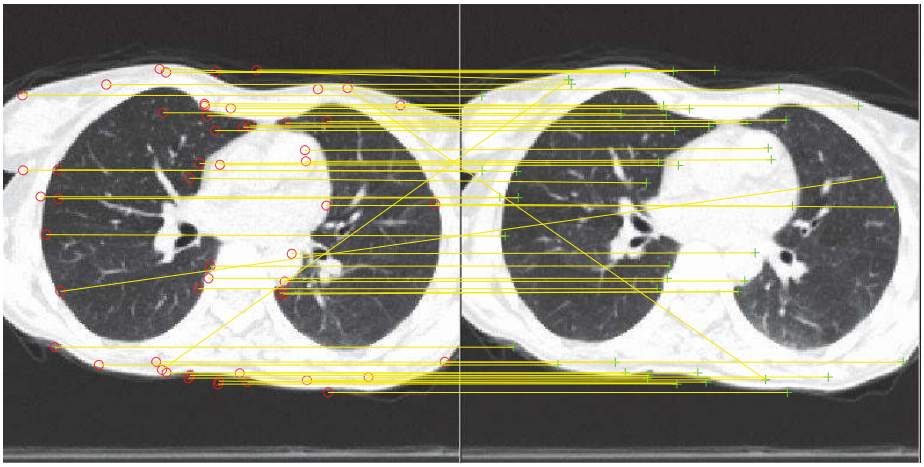


Figure 3.4: Corresponding feature points matched in the inhale (left) and the exhale (right) frame recorded at t= 0.1 and 0.6 seconds respectively

The combined implementation the SURF feature detector and MLS algorithm as a two-stage process attributes this methodology with faster processing speeds in feature detection/description and efficient tracking of the interest points (obtained from SURF) during their transition through frames/slices.

Figure 3.5: Registered axial image for the first subject from inhale frame at t=0.1 sec to exhale frame at t=0.6 sec

As we can see in the figure 3.3, an implementation of the SURF algorithm gave out matching features with their respective coordinate values in the inhale (t=0.1 sec) and the exhale (t=0.6 sec) frames for the axial AP CTs of subject 'case 1'. Similar process was employed for all subjects throughout all the three APs. This provided us with the respective coordinate point values for the mutual interest points between them. These frames along with the interest point values are used by the MLS algorithm to provide a registered image for every set of AP for a subject. Result of this process being the registered images for every subject through all three APs for all 10 subjects. Along with the registered images as output, the average translation of the interest points is also obtained in x and y Cartesian directions in 'pixel' units for all 10 subjects under observation from all three APs. The average translation is basically the collective deviation for all the interest points from an initial stage to the final stage in terms of Euclidean distance in pixel units in their own respective cases.

3.4 Results and Discussion

These average translations were compiled for all the 10 subjects with respect to the common denominator i.e. the number of frames/slices. Since, the number of frames taken into account was 6; there were total five translation gaps in between them. The translation data was comparatively classified for all 10 subjects into different graphical representations in terms of it being in x or y direction and the AP it belongs to. A widespread consensus now exists that it would be useful to use prior knowledge of respiratory organ motion and its variability to improve radiotherapy planning and treatment delivery [Blackall, J. M. *et al.*, 2006]. The

estimated deformation for a particular subject when computed using the proposed methodology can be compared and analysed against a corresponding standard atlas to assess the extent of abnormality. The figures 3.6 & 3.7 show the average x and y translation values for all subjects in axial AP. Looking into 'case 10' y-translation plot, the average deformations estimated over inter-frame durations of a single inhale-exhale process (stated above) was 0.278±0.11 pixels with a maximum value of 0.408 pixels while transitioning from frame 4 to 5. The variations in the deformations exhibited by the subject 'case 10' were significantly larger than the considered population average of 0.039±0.13 pixels. Similarly figures 3.8, 3.9 and 3.10, 3.11 show the average x, y translation for all interest points in the coronal and sagittal APs respectively. The maximum translation obtained for a case/subject during transition from one slice to another signifies the maximum variations in breathing pattern than the considered population and atlas data in general, which in this study happens to be subject 'case 10'. Similarly the minimum deviation obtained from either of the APs points out slice/frame transitions with no apparent anatomical deformation in the thoracic periphery of the subjects.

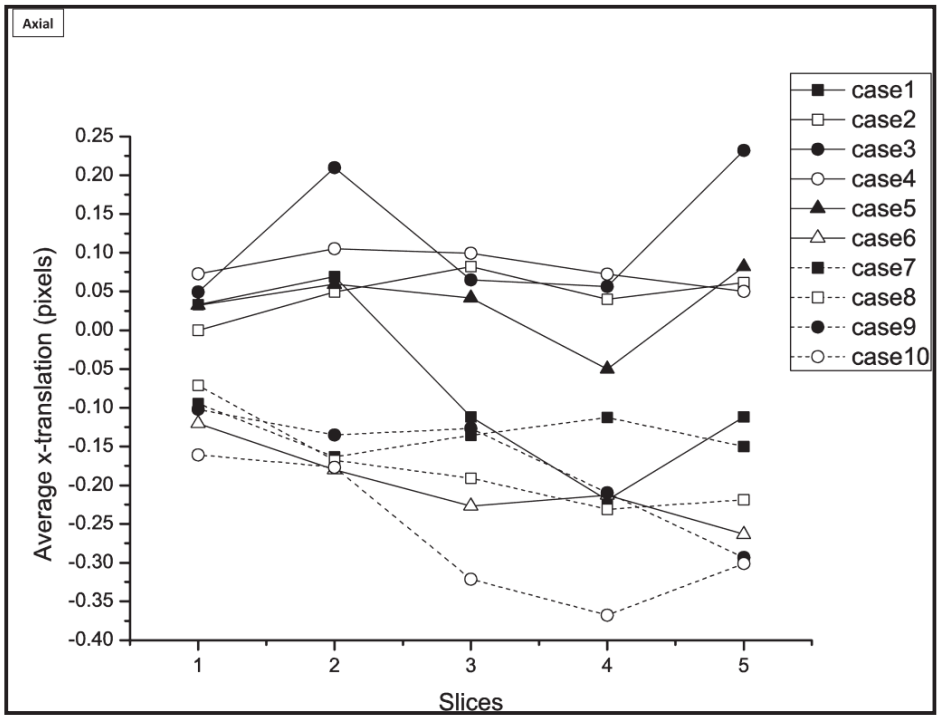


Figure 3.6: Average x-translations for axial AP

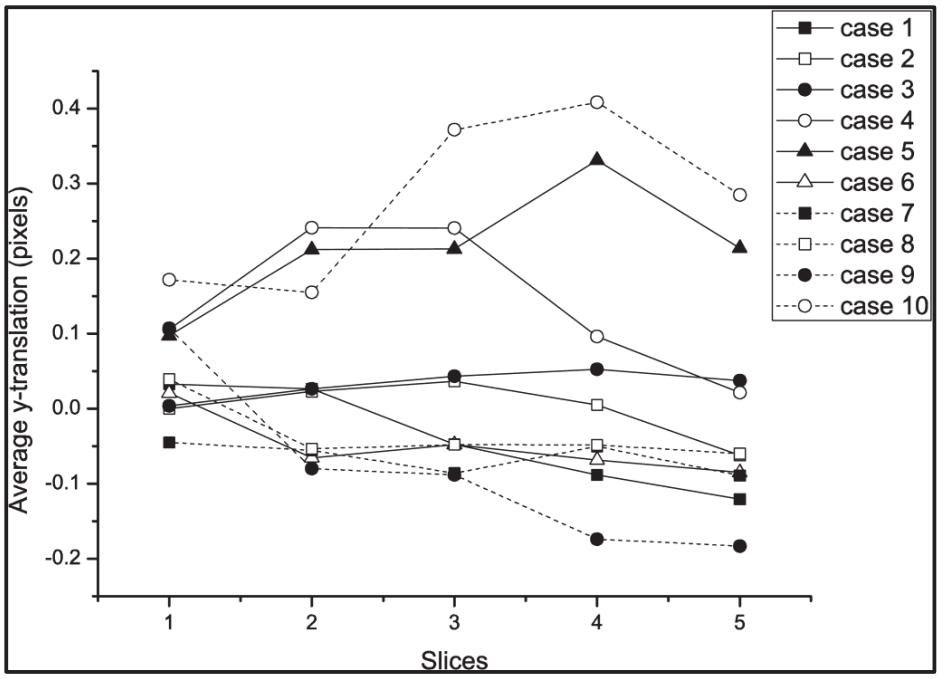


Figure 3.7: Average y-translations for axial AP

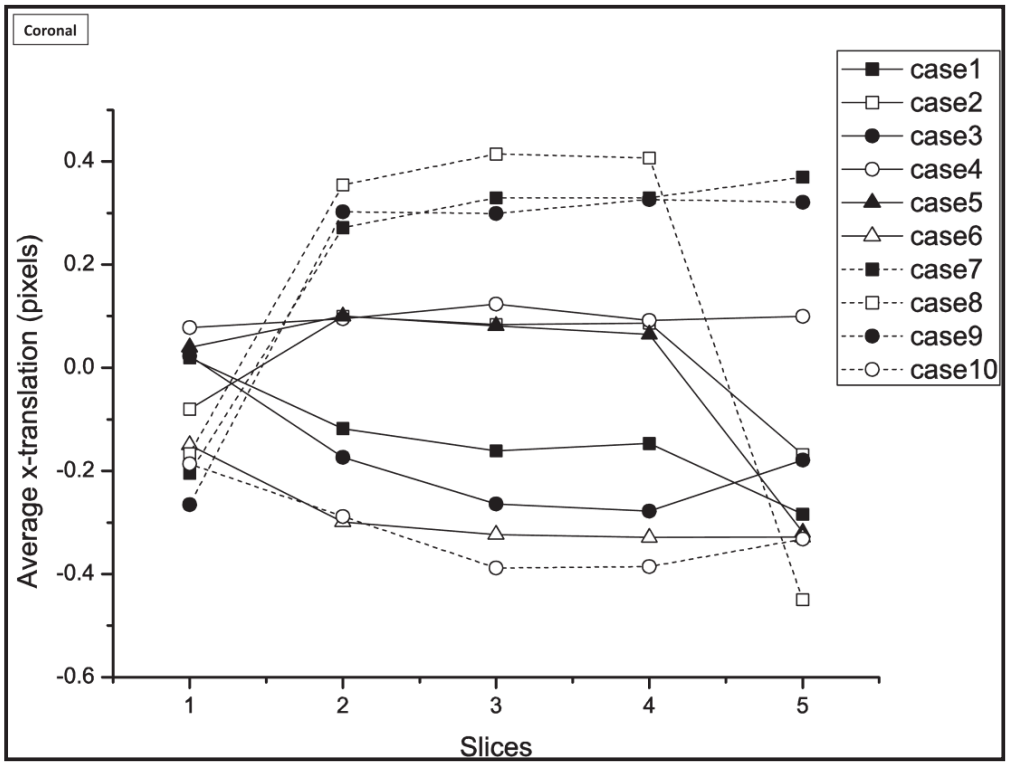


Figure 3.8: Average x-translations for coronal AP

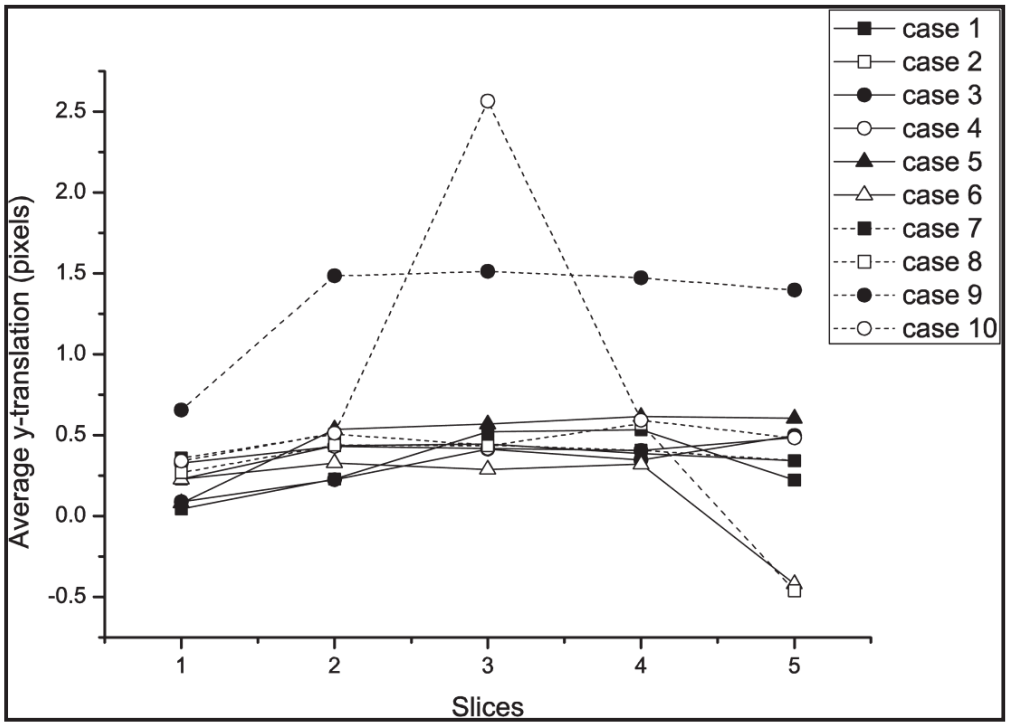


Figure 3.9: Average y-translations for coronal AP

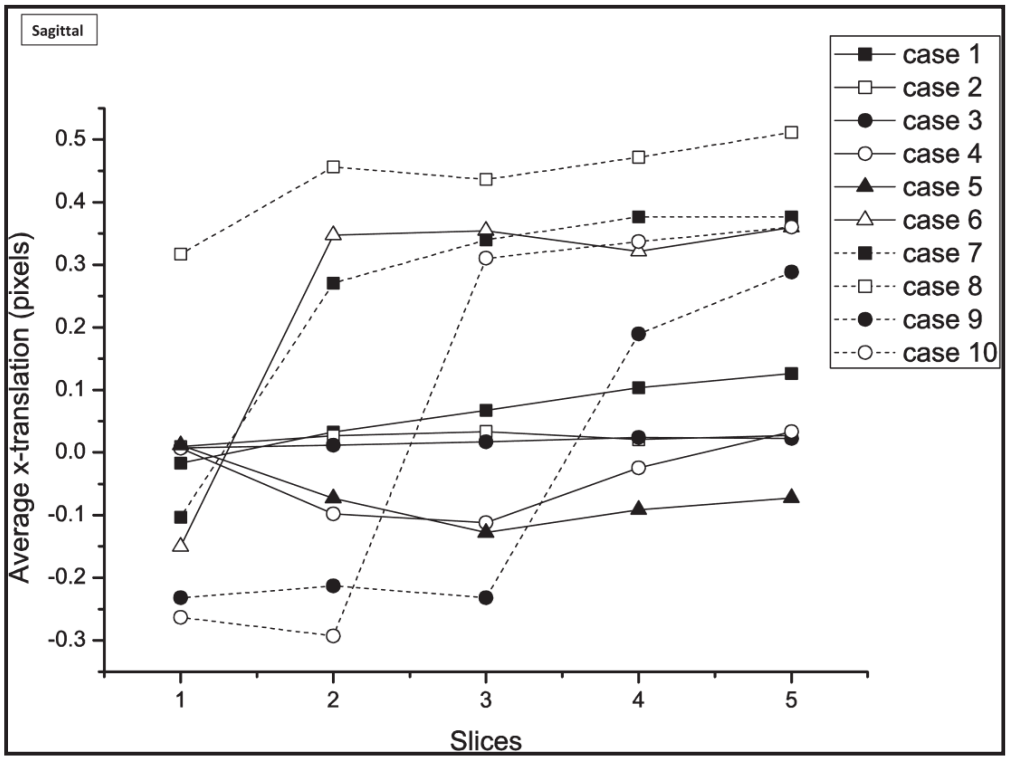


Figure 3.10: Average x-translations for sagittal AP

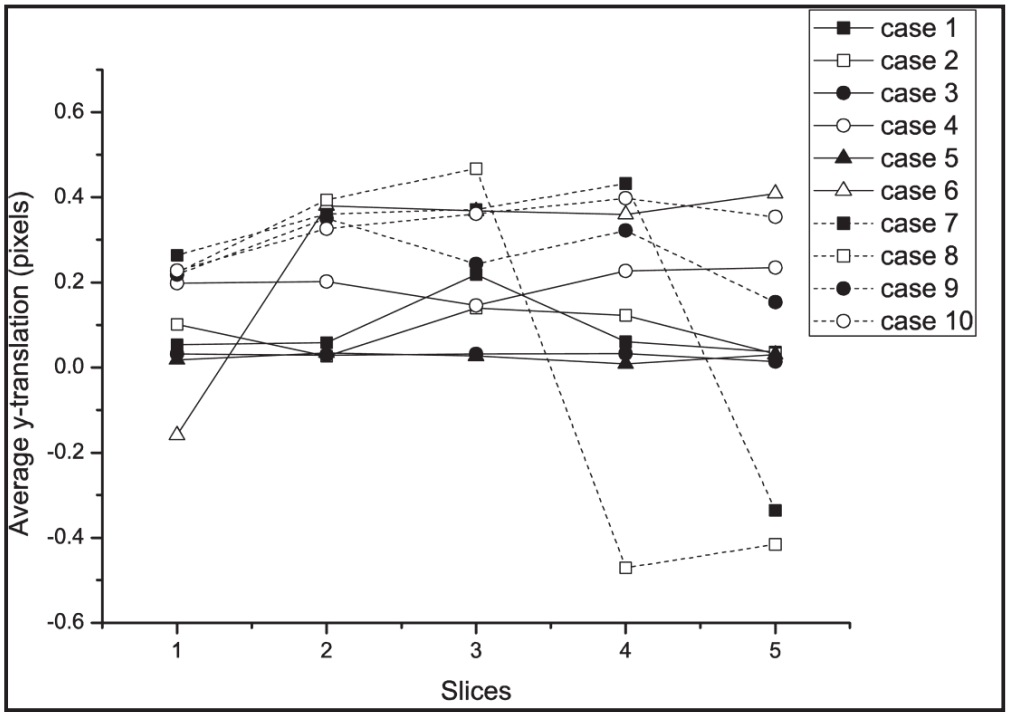


Figure 3.11: Average y-translations for sagittal AP

Table 3.2: Target registration error comparison through all APs for all subjects (*our method)

Subject (Case No.)	SIFT+MLS (Scale Invariant feature Transform)			SURF+TPS (Thin-plate Splines)			SURF+MLS*		
APs	Axial	Coronal	Sagittal	Axial	Coronal	Sagittal	Axial	Coronal	Sagittal
1	0.428	0.883	0.451	0.498	1.123	0.541	0.128	0.353	0.121
2	0.344	0.932	0.42	0.464	1.052	0.51	0.054	0.382	0.09
3	0.52	0.987	0.464	0.61	1.217	0.514	0.13	0.367	0.034
4	0.845	1.344	0.704	0.785	1.404	0.824	0.165	0.424	0.214
5	0.94	1.551	0.931	0.86	1.541	0.771	0.22	0.501	0.081
6	0.589	1.168	0.904	0.699	1.368	0.994	0.209	0.428	0.454
7	0.637	1.182	0.953	0.647	1.432	0.983	0.147	0.512	0.463
8	0.673	1.594	1.041	0.673	1.494	1.131	0.183	0.564	0.591
9	0.655	1.96	0.843	0.735	2.19	0.933	0.215	1.34	0.353
10	0.725	1.842	0.908	0.795	1.882	0.998	0.385	0.972	0.458

3.5 Conclusion

A methodology has been presented showing the how a feature point set generated by SURF can be used through MLS for deformable image transformations in medical images such as the thoracic pectus excavatum exhale and inhale frames used in this work. The accuracy of the deformable registration performed using the proposed methodology is assessed in terms of the target registration error (E_{TR}) and is compared with two other prevalent methods for all three APs (refer Table 2) for the same database. The error values (pixel values converted to mm using the resolution of images) for the proposed methodology are considerably lower than the other methods for almost all subjects considered. The use of SURF can be explained by its better performance in terms of smaller time complexities

(stated earlier) and common feature points than existing state of the art feature detector and descriptors. MLS however was the best available choice for Deformable Image Registration (DIR) based on landmark points. Although the proposed methodology provides with a fast and accurate way of DIR for medical images and thus an account of deformity in the thoracic periphery, there is much scope for improvement in the overall process. One way this can be achieved in future is by modifying the SURF and/or MLS procedures themselves involved in the process i.e. bringing newer versions of the existing methods better suited with the application. Another way is to improve and enhance the quality as well as the quantity of the database used. Also, the aforementioned procedure can provide better results if applied for a different human anatomy altogether.