

Virtual Endoscopy

IIIT-Delhi Technical Report

PRONOMA BANERJEE and KRUPA BHAYANI

ABSTRACT

One of the first concerns in any deep learning based research is to make sure the dataset is correct and usable. While working with a dataset initially promised to be aligned by the providers, we detected some major errors in the dataset. In our report, we have described our work on volumetric registration and segmentation to fix the errors found in the the Visible Korean Human dataset.

Additional Key Words and Phrases: image registration, segmentation, deformable, rigid, mutual information, multi-modal.

ACM Reference Format:

Pronoma Banerjee and Krupa Bhayani. 2021. Virtual Endoscopy: IIIT-Delhi Technical Report. 1, 1 (August 2021), 25 pages. <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

Authors' address: Pronoma Banerjee, f20190690@goa.bits-pilani.ac.in; Krupa Bhayani, f20180844@pilani.bits-pilani.ac.in.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2021 Association for Computing Machinery.

XXXX-XXXX/2021/8-ART \$15.00

<https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

1 INTRODUCTION

Any task in deep learning is most often heavily dependent on the dataset. In our project, we were worked with the Visible Korean Dataset. However, the dataset had several problems in alignment as well as segmentation. This report covers a detailed literature review on these concepts, as well as a description of our attempts to fix these errors.

2 LITERATURE REVIEW

2.1 Visible Korean Human dataset

The Visible Korean Human (VKH) project was started in 2001 at Ajou University, Suwon, Republic of Korea and was the second experiment that compiled data from magnetic resonance imaging (MRI), computed tomography (CT) and comprehensive anatomic images of a male cadaver; the other similar projects being the Visible Human Project (1994-1995) and Chinese Visible Human Project (2002-2003). This was the first successful attempt at producing color anatomic images of a human body. The study also produced segmented images of 11 structures derived from the anatomic images [21].

This research was preceded by the Visible Human Project (VHP), which was first tested on males in 1994 and on females in 1995 by the National Library of Medicine [7]. The experiment brought about a major breakthrough in medical imaging worldwide, leading to the foundation of three-dimensional software for virtual dissection, virtual endoscopy and virtual operation. However, the VHP data was incomplete in some aspects of data acquisition and had some major drawbacks. First, the data did not include complete MR and CT images of the entire body. Trunk and limbs were absent in the MR images and the upper limbs' lateral parts were cut off on the CT images. Second, the anatomical images of the VHP data did not include anatomical structures that were smaller than 0.2 mm since both the interval and pixel size of the images were 0.33 mm or greater. Third, the color of the anatomical images of the VHP data were not similar to a living body because formalin was perfused into the cadaver prior to serial sectioning. Last, the VHP did not publish the segmented images, which are essential to produce 3D images of each anatomical organ and helpful in making the virtual dissection software [7, 20].

The Visible Korean Project was carried out from March 2001 to August 2003 to generate MR, CT, anatomic and segmented images. A male cadaver, who died of Leukemia, was chosen for the experiment. He was 33 years old, with an average Korean body size (length: 1718mm (feet plantarflexed), weight: 55kg). The body was placed supine and immobilized using Mev-Green. Using body coil, horizontal MR images of the entire body were acquired at 1 mm interval; with inter slice gap: 0 mm; field of view: 480 mm × 480 mm; resolution, 512 × 512. The repetition time was fixed at 1,000 ms and echo time was fixed at 8 sec for increasing the signal/noise ratio. The CT images of the entire body were also acquired at same interval, field of view and resolution; using standard algorithm; voltage, 120 kV; electric current time, 280 mAs. The MR and CT images, each of voxel size 0.9375 mm × 0.9375 mm × 1 mm, were saved as tag image file format (TIFF) files (bit depth, 8 bits gray). 1718 MR and CT images were produced. Excessive margins of the MR and CT images, which did not include the body images, were cropped. [7][21]. Each cropped MR and CT image had 494 × 281 resolution, 8 bit gray and 144 KB file size.

The cadaver was then gradually embedded using small quantity of blue embedding agent at each step [20] and frozen to -70°C in the freezer, until the embedding box was fully filled with the embedding agent. This step-wise embedding was necessary to prevent the freezing embedding agent from pressing the cadaver and the alignment rods. The container was then placed longitudinally on the cryomacrotome and moved through the rotating cutting blade to produce serially sectioned images of thickness 0.2mm. 600 mm (width) × 400 mm (height) sized sectioned surfaces were digitally captured in both color and gray scale positioned using a mounted digital camera (DSC560, Kodak; resolution: 3,040 × 2,008; bit depth: 24 bits colour), capturing the cadaver, embedding agent, alignment rods, gray scale, and color patch; facing the center of the sectioned surfaces. A total of 8,590 anatomic images were downloaded and saved as TIFF files [20]. Each final anatomical image used had 2,468 × 1,407 resolution, 24 bits coloured, 10 MB file size.

Eleven important structures (skin, bones, liver, lungs, kidneys, urinary bladder, digestive tract, respiratory tract, arteries, brain, and heart) from the anatomic images were selected to be segmented. Contours of the anatomic structures were semiautomatically drawn on all 8590 anatomical images using the magnetic lasso tool of Adobe Photoshop. Each structure was colored to produce 8,590 segmented images [21]. The anatomical and segmented images were

stacked and volume-reconstructed to produce 3-D images. All 8590 anatomical images (3040 × 2008 pixels) were stacked at 0.2-mm intervals and subsequently volume-reconstructed to produce a 3-D image, which consisted of 3040 × 2008 × 8590 voxels. The segmented images were made into a 3-D image in a likewise manner [20].

Anatomic structures	Components
Skin	
Bones	
Liver	
Lungs	
Kidneys	
Urinary bladder	
Brain	Cerebrum*, cerebellum*, brain stem*, amygdaloid bodies*, thalamus*, pituitary gland*, lateral ventricles*, third ventricle*, mesencephalic aqueduct*, fourth ventricle*, dural venous sinus*
Digestive tract*	Oral cavity, pharynx, esophagus, stomach, small intestine, large intestine
Respiratory tract*	Nasal cavity, pharynx, larynx, trachea, bronchi, lobar bronchi, segmental bronchia
Arteries*	Ascending aorta, aortic arch, brachiocephalic trunk, common carotid arteries, external carotid arteries, subclavian arteries, axillary arteries, brachial arteries, radial arteries, ulnar arteries, thoracic aorta, abdominal aorta, celiac trunk, renal arteries, common iliac arteries, internal iliac arteries, external iliac arteries, femoral arteries, popliteal arteries, anterior tibial arteries, posterior tibial arteries

Anatomic structures	Components
Heart	Heart, right atrium*, left atrium*, right ventricle*, left ventricle*, right coronary artery*, left coronary artery*, tricuspid valve‘, mitral valve‘, pulmonary valve‘, aortic valve‘

*Luminal contours are segmented.

‘Fourteen brain components and 10 heart components are further segmented.

2.2 Image Registration

Image registration is defined as a process that overlays two or more images from various imaging equipment or sensors taken at different times and angles, or from the same scene, by finding corresponding structures, to geometrically align the images for analysis. One image is normally used as reference image (also known as the fixed, target or source images) and other deformed images (known as subject or moving images) are transformed and aligned with the reference image.

2.2.1 *Types of image registration:*

Image registration can be classified into several categories based on the transformation model that they use, method of acquiring feature points, manner of image acquisition and techniques.

- Based on transformation model:

Image registration can be categorized on the basis of rigid, affine, projective and non-rigid transformations. In the first three cases, the entire moving image can be mapped to the target image using a set of matrix operations including translation, rotation and scaling. In non-rigid registration, the reference and target images are locally misaligned. The transformation of pixels may be different depending on regions in the image, and are typically quantified and visualized through vector fields. Non-rigid registration is also known as deformable image registration or non-linear registration. The time complexity of the transforms in each transformation are characterized by their degrees of freedom.

Rigid body transformations have six degrees of freedom in three dimensions, i.e. three translations and three rotations. Non-rigid mapping requires more degrees of freedom to accurately describe the local differences in orientation.

- Based on the method of acquiring feature points:

The transformation algorithms can be further classified as feature based and intensity based. Feature-based methods establish a correspondence between distinct points, lines or contours. Intensity based methods compare intensity patterns in images via similarity metrics.

- Based on the manner of image acquisition:

Image registration can also be classified as multi-view (when images of the same object are captured from multiple view points), multi-temporal (when images of the same object are captured at different times, under different conditions) and multi-modal analysis (when different sensors are used to capture images of the same object). We can distinguish registration methods using similarity criteria according to the type of information they exploit. There are several approaches to do this for both mono-modal and multi-modal registration. In the mono-modal case, the use of standard similarity criteria (e.g., SSD or SAD) involving either intensities or multi-channel data extracted from the image through the application of filters is well-accepted by the community. In the multi-modal case, the use of information theoretic measures has become the prevalent solution.

- Based on technique:

Image registration techniques can be classified based on dimensionalities, domain of transformation, type of transformation, registration quality, parameters of registration, subject of registration, object of registration, nature of registration basis, interaction and the modalities involved [19].

Multi-modal deformable registration is the primary registration method used for our dataset. Most methods tried by us were intensity based methods.

2.2.2 Steps involved in Image registration

- . The process of image registration involves the following steps

- (1) Changing the voxel coordinates of the images to physical coordinates.
- (2) Feature detection to find common features in reference and the deformed image.

- (3) Feature matching to establish the correspondence between the non-aligned sensed image and the reference image.
- (4) Transform model assessment to determine the parameters of the mapping function (geometric transformation) using knowledge of the image acquisition process and correspondence between the feature points.
- (5) Alignment of the deformed image(s) to the reference image using the mapping function.

2.2.3 *Methods of registration:*

The various methods of image registration include the following [24]:

- (1) Extrinsic methods: Artificial objects which are easily detectable are attached to the patient's body, before capturing the images, which serve as external methods to be used for image matching.
- (2) Surface methods: Surfaces or boundaries or contours are generally distinct in medical images unlike landmarks. Surface-based approach is often employed for multi-modal rigid body registration. Heat and Hat algorithm [11], Iterative Closest Point Algorithm and Correspondence Matching Algorithm are successfully applied as registration algorithms for surface-based techniques.
- (3) Moments and principle axes methods: Used for registration of identical or very similar objects, by bringing their principal axes (orthogonal axes about which moment of inertia is minimized) into concurrence without employing any rigid or affine transformation.
- (4) Correlation based methods: Correlation based metrics such as correlation coefficient (CC) are used for analyzing the performance of mono-modal image registration. CC is 1 when the images are perfectly identical, CC is 0 when images are uncorrelated and CC is -1 when the images are anti-correlated. The equation is represented as:

$$CC = \frac{\sum_i(x_i-x_m)(y_i-y_m)}{\sqrt{\sum_i(x_i-x_m)^2}\sqrt{\sum_i(y_i-y_m)^2}}$$

where x_i, y_i are the intensities of i^{th} pixel in the reference and sensed image respectively, and x_m, y_m are the mean intensity of reference and sensed image respectively [19]. The features extracted from images are used to obtain the cross-correlation coefficients for image registration. Cross-correlation and Phase-correlation techniques based on Fourier domain, and normalized mutual information between the images adopting an Entropy

Correlation Coefficient (ECC)[16], have also been used for image registration. Correlation methods are, however, inefficient for multi-modal registration.

- (5) Mutual Information based methods: MI is another similarity metric using joint probability of intensities of comparable voxels. This method is widely used for voxel-based registration of multi-modal images. MI is maximum when both the images are perfectly aligned. The values of MI are always non-negative and symmetric, starting from zero and varying up to a high value. High mutual information depicts large reduction in uncertainty whereas zero MI value is clear indication that the two variables are independent. It is represented as:

$$MI = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log\left(\frac{p(x, y)}{p_1(x)p_2(y)}\right)$$

where $p(x, y)$ is the joint distribution function and $p_1(x), p_2(y)$ are the marginal distribution functions [19]. Gradient descent and various other optimization methods have been used to maximize mutual information.

- (6) Wavelet based methods: Wavelet-based feature extraction techniques along with normalized cross-correlation matching and relaxation-based image matching techniques are used for incorporating sufficient control points to reduce the local image degradations, for image registration. [24]
- (7) Soft computing methods: These methods have been developed recently and are based on Artificial Neural Networks, Fuzzy Sets and several Optimization Heuristics.

2.2.4 Components of a registration algorithm:

An image registration algorithm has 3 components:

- (1) A deformation model:

The choice of deformation models is important in the registration process because [24]

- It reflects the class of transformations that are desirable or acceptable, and therefore limits the solution to a large extent.
- The parameters that registration estimates through optimization strategy correspond to the degrees of freedom of the deformation model.
- It entails an important compromise between computational efficiency and richness of description, since increasing dimensionality and hence the descriptive power of the model might lead to increase in the model's complexity.

- It implies an assumption regarding the nature of the deformation to be recovered.

Deformation models can be classified as the following [24]:

(a) Geometric transformations derived from physical models:

- (i) Elastic body models: These are governed by the Navier Cauchy partial differential equation. In order to account for large deformations, non-linear elastic models have been proposed.
- (ii) Viscous fluid flow models: These models are governed by Navier-Stokes equation that is simplified by assuming a very low Reynold's number flow. D'Agostino et al. tackled multi-modal non-rigid registration using with the use of a viscous fluid model [9]. The algorithm is based on maximization of mutual information. The accuracy of the method was verified using simulated multi-modal MR images with known ground truth deformation. The results showed that the root mean square difference between the recovered and the ground truth deformation is smaller than 1 voxel.
- (iii) Diffusion models: In this case, the deformation is modelled by the diffusion equation $\Delta u + F = 0$. The Demon's algorithm for image registration, based on Maxwell's Demons, was first proposed by Thirion [28]. The proposed algorithm iterated between two steps: 1. Estimation of the demon forces for every demon (more precisely, the result of the application of a force during one iteration step, that is a displacement) 2. Update of the transformation based on the calculated forces. Depending on the way the demon positions are selected, the way the space of deformations is defined, the interpolation method that is used, and the way the demon forces are calculated, different variants can be obtained. The most suitable version for medical image analysis involved 1) selecting all image elements as demons, 2) calculating demon forces by considering the optical flow constraint, 3) assuming a non-parametric deformation model that was regularized by applying a Gaussian filter after each iteration, and 4) a trilinear interpolation scheme. The methods described by Thirion share the iterative approach, that is, iterating between estimating the displacements and regularizing to obtain the transformation. This iterative approach results in increased computational efficiency. Demon's algorithm has been applied to tackle several multi-modal registration problems [13][29][18]. The application of Demon's algorithm has also been extended to multi-channel images [22] and for different geometries [34].

- (iv) Curvature registration: These models are governed by the equilibrium equation:

$$\Delta^2 u + F = 0$$
- (v) Flows of diffeomorphism: In this case, the deformation is modeled by considering its velocity over time according to the Lagrange transport equation. This model has been used to develop the large deformation diffeomorphic metric mapping (LDDMM) framework.
- (b) Geometric transformations derived from interpolation theory: In interpolation theory, displacements, considered known in a restricted set of locations in the image, are interpolated for the rest of the image domain. In approximation theory, we assume that there is an error in the estimation of displacements. Thus, the transformation smoothly approximates the known displacements rather than taking the exact same values. These models are rich enough to describe the transformations that are present in image registration problems, while having low degrees of freedom and thus facilitating the inference of the parameters.
 - (i) Radial basis function (RBF): RBFs are one of the most important families of interpolation strategies, given by the equation $u(x) = \sum_{i=1}^N \omega_i \phi(\|x - p_i\|)$. This method has been widely used in image registration [24].
 - (ii) Elastic Body Splines: They were introduced by Davis *et al.* [10]. These splines are solutions of the Navier-Cauchy equilibrium equation for a homogeneous isotropic elastic body subjected to forces. When the force field that drives the registration based on the landmark correspondences is given as a radial symmetric function of the distance from the landmark, one can solve the equation analytically.
 - (iii) Free-form deformations: Free-form deformations (FFDs) is one of the most common types of transformation models in medical image registration. A rectangular grid $G = K_x \times K_y \times K_z$ is superimposed on the image (size $N_x \times N_y \times N_z$, $K_x \ll N_x, K_y \ll N_y, K_z \ll N_z$) that gets deformed under the influence of the control points. The dense deformation is given as a summation of tensor products of univariate splines. The displacement field is given by: $u(x) = \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 B_l(\mu_x) B_m(\mu_y) B_n(\mu_z) d_{i+l} d_{j+m} d_{k+n}$ where $i = \lfloor \frac{x}{N_x} \rfloor - 1$, $j = \lfloor \frac{y}{N_y} \rfloor - 1$, $k = \lfloor \frac{z}{N_z} \rfloor - 1$, $\mu_x = \frac{x}{N_x} - \lfloor \frac{x}{N_x} \rfloor$, $\mu_y = \frac{y}{N_y} - \lfloor \frac{y}{N_y} \rfloor$ and $\mu_z = \frac{z}{N_z} - \lfloor \frac{z}{N_z} \rfloor$. B_l denotes the l^{th} basis function of the B-spline and d denotes displacement. FFD is widely used because its transformation model is simple, it can

efficient provide smooth deformations and requires only a few degrees of freedom to describe local deformations.

- (iv) Basis functions from signal processing: Fourier and Wavelet analysis have been used for model transformations, since they can naturally provide a multi-resolution decomposition of the displacement field. This is a useful property for the coarse-to-fine schemes that are commonly applied in medical image registration to ease the computations and handle large deformations.
- (v) Locally affine models: Locally affine models parameterize the transformation by locally linear deformations.
- (c) Knowledge-based geometric transformations:
 - (i) Statistically-constrained geometric transformations: Statistical deformation models (SDMs) capture statistical information about deformation fields across a population of subjects. These methods are able to reduce the number of degrees of freedom, and consequently the computational demands of the problem, while achieving robust performance. Principal component analysis (PCA) has been used for studying shape variability on point correspondences [8]. Wouters et al. [33] used PCA to model the deformation and registration was performed by adjusting the coefficients of the principal components while maximizing mutual information (MI). PCA has also been used to learn an SDM to accelerate image alignment [26].
 - (ii) Geometric transformations inspired by biomechanical/biophysical models: finite element methods (FEMs) are employed to model the biomechanical/biophysical properties of the tissues under consideration. The main motivation behind using the methods of this category is the surmise that more informed priors regarding the biomechanical properties of the tissues will allow the reliable estimation of complex deformation fields with the use of few degrees of freedom. What is more, the limited search space results in improved efficiency when compared to the standard approaches.
- (d) Task-specific constraints:
 - (i) Topology preservation: The preservation of topology is equivalent to the invertibility of the deformation field. The Jacobian of the deformation field is very informative regarding the local properties of the deformation field. Attempts have been made to prevent singularities in the deformation field by tracking the Jacobian [5]. Another

way to preserve topology is by use of constraints on the Jacobian, as was done by Christensen *et al.* [4].

- (ii) Volume preservation: This is particularly important for images of incompressible anatomical structures like bones. FFD models have been used for volume preservation [27].
- (iii) Rigidity constraints: The presence of rigid anatomical structures in medical images motivates the incorporation of rigidity constraints in image registration. Non-rigid FFD registration methods have been locally constrained by penalizing deviations of the Jacobian from orthogonality by Loeckx *et al.* [15]. Other similar work have been done by Staring *et al* [25].

(2) An objective (mapping) function:

The image alignment can theoretically be performed by both backward mapping (target image voxels are mapped onto the source image) or forward mapping (source image coordinates are mapped back onto the target image). From implementation point of view, forward mapping is less advantageous [24].

The mapping (objective) function in steps (3) and (4) can be generalized as:

$$M(T, S \cdot W) + R(W)$$

The first term, M , termed interchangeably as the matching criterion, (dis)similarity criterion or distance measure, quantifies the level of alignment between a target image and a source image. The transformation W is applied to align the images. The second term, R , regularizes the transformation aiming to favor any specific properties in the solution that the user requires, and seeks to tackle the difficulty associated the image registration problem, which is by nature ill-posed [24].

(3) An optimization method:

For a non-convex and non-linear function, no closed-form general solution exists to estimate the registration parameters. The optimization methods locally maximize or minimize the objective function depending on how the matching term is chosen. The choice of matching term depends on anatomical properties of the organ, the nature (modality) of observations to be registered and the clinical setting in which registration is to be used.

Optimization methods for the objective function may be separated into two categories based on the nature of the variables that they try to infer:

- (a) Continuous Optimization: This class of methods solves optimization problems where the variables assume real values. They can be further classified into gradient descent methods, conjugate gradient methods, Powell's conjugate direction's method, Quasi-Newton (QN) methods, Gauss-Newton method, Levenberg-Marquardt algorithm, stochastic gradient descent methods. [24]. Registration is an inherently continuous problem. As a consequence, continuous optimization methods have been the main driving force behind registration algorithms.
- (b) Discrete Optimization: Methods in this class solve problem the variables take values from a discrete set. They can be further classified into graph-based methods, message passing methods and linear-programming (LP) approaches. Recently, discrete optimization techniques have been proposed to tackle deformable registration [24].

Evaluation of registration methods is a particularly difficult problem because of the lack of a “ground truth.” The absence of knowledge of correspondences between images makes the quantitative validation of the registration performance a challenging task. Moreover, because of the different requirements of the applications that are based on deformable registration, the notion of correspondence should vary according to application context, aiming to properly characterize error. Nonetheless, the increasing availability of annotated data sets (e.g., the LONI Probabilistic Brain Atlas [23], the Internet Brain Segmentation Repository - IBSR [12], the CUMC12 dataset2 acquired at the Columbia University Medical Center, the MGH10 dataset2 scanned at the MGH/MIT/HMS Athinoula A. Martinos Center for Biomedical Imaging) which has lead to several evaluation studies [14]. Furthermore, the development of evaluation projects for image registration (i.e., nonrigid image registration evaluation project—NIREP [6]) and the increasing understanding regarding the use of surrogate measures for the measurement of the accuracy of registration is further facilitating the comparison between different algorithms.

2.2.5 Toolkits for registration:

Recently several python toolkits are being used for image registration. Examples include image_registration (rigid), lcreg (rigid), imgreg2D (rigid), chi2_shift (rigid), PyElastix (non-rigid), dipy (non-rigid), ANTsPy (non-rigid), nireg (non-rigid), SimpleElastix(non-rigid), PIRT

(non-rigid), Python-register (non-rigid), Silx (non-rigid), GeFolki (non-rigid), pyStackReg (both), scikit-image (both), ITK (both) etc [2]. Most of these toolkits work perfectly for rigid registration as well as all kinds of brain registration. However full body registration or alignment of areas such as the abdomen still require a lot of further work.

3 EXPERIMENTATION

3.1 Segmentation

We used the segmentation provided with the dataset and gave unique labels to each segmented section. There were three levels of depth in each directory. Level one being the broad classification with classes like skeletal system, cardiovascular system, Skin, *etc*. Level two consists of specific segments. For example in skeletal system we have Cranium, Vertebral Column, Bones of Upper Limb, *etc*. Level three is the most deep classification such as for Cranium there are Cranium without mandible, Frontal bone, Parietal bone, *etc*. Apart from these classes we classified the remaining section of each 2 D image as background class with label 0. For all other classes we gave unique 16 bit labels where first bit is 0, next 4 bits are for level one, next 4 bits for level two and last 7 bits for level three. There were around 40 to 50 segmentation images which were different from the rest. Majority of the images were 2 D images with 0s and 1s but these images where 3 D and we had to manually verify which plane contain the useful information. In most of the cases all the planes where similar. Using the segmentation in this way we could remove the noise from cryo-sectioned images in the background. We created new cryo-sectioned volumes with clean background with 2 pixel dilation to ensure no information loss.

3.2 Image Registration

We experimented with several packages and libraries available, often tweaking some parameters for the code to work for 3D images.

For our experiment, we took a smaller subvolume of the abdomen of size 494x281x200 for CT and MRI volumes. We took 200 sectioned abdomen slices as our volume, resized it to one-fifth

of its original X and Y resolution (from 2468x1407 to 493x281) and then padded it to match the resolution of MRI and CT volumes (494x281).

We were informed by the providers of the dataset, that while the CT and cryo-sectioned images were acquired after freezing of the human body, the MRI was acquired before. Hence, the CT could be aligned to the sectioned volume by applying rigid registration only. However, the MRI volume had to be rigid registered as well as deformably registered to make it align with the cryo volumes.

3.3 Rigid registration

For registering CT volume to the sectioned volume, we tried multiple implementations in MATLAB and python and existing softwares, and finally got 3 results, using the libraries SimpleITK and Elastix in python, and using the MITK Software. While all the 3 results were good, the result from Elastix was the best. The results from the 3 methods are shown in Fig. 1.

The SimpleITK method used metric as Mattes Mutual information with gradient descent optimizer and linear interpolation. The Elastix method also used Mattes MI with adaptive stochastic gradient descent algorithm, and linear interpolator. MITK implemented the iterative closest point algorithm described in [3] and is the only point based registration that we will be mentioning in this section.

3.4 Deformable registration

For registering MRI volume to the sectioned volume, we first rigid registered MRI using SimpleITK/Elastix, then used the rigid registered MRI volume for deformable volumetric registration.

We tried over 40 methods for deformable registration. However very often we were unable to get the result, or a transformed array as a result at all. The methods for which we got results include the following from SimpleITK, Elastix and MITK:

- (1) Simple demons registration (SimpleITK)
- (2) Diffeomorphic demons registration (SimpleITK)
- (3) Symmetric forces and fast symmetric forces demons registration (SimpleITK)

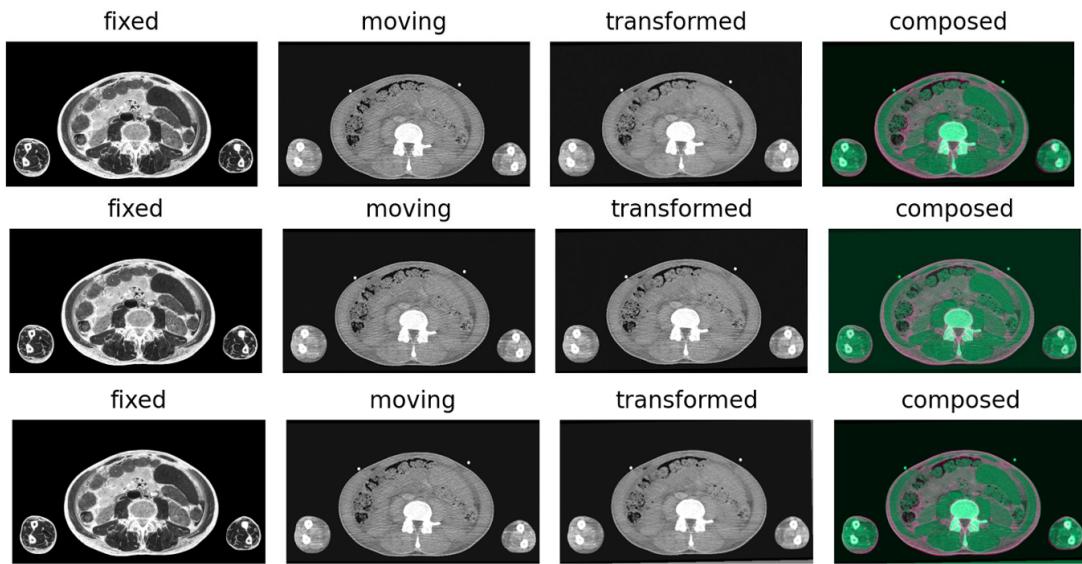


Fig. 1. Rigid registration of CT volume using SimpleITK, Elastix and MITK respectively. The reference is the gray-scale sectioned volume, moving volume is the CT volume, and the composition is of the reference and transformed volumes. The image shows the result for one slice.

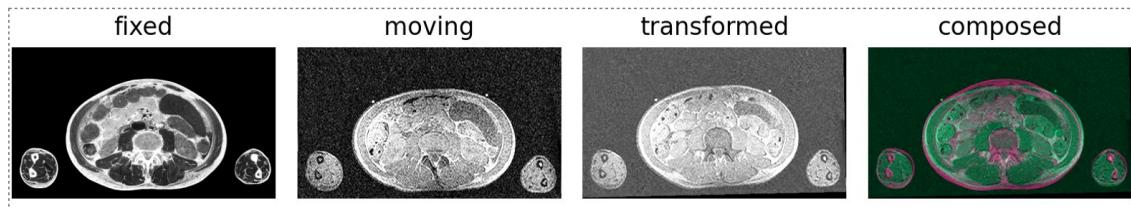


Fig. 2. Rigid registration of MRI volume using SimpleITK. The reference is the gray-scale sectioned volume, moving volume is the CT volume, and the composition is of the reference and transformed volumes, with red lines corresponding to fixed volume and green lines corresponding to the transformed volume. The image shows the result for one slice.

- (4) Bspline registration (SimpleITK)
- (5) Fast symmetric forces demons registration (MITK)
- (6) Level set motion registration (MITK)
- (7) BSpline registration (Elastix)

3.4.1 Demons registration (*SimpleITK*):

. SimpleITK has four demons registration filters, namely classic demons, diffeomorphic demons, symmetric forces demons and fast symmetric forces demons. All these algorithms use root mean square as their error metric. The underlying assumption of the demons framework is that the intensities of homologous points are equal. The code uses histogram matching to make the two images similar intensity wise prior to registration, which leads to the change of background colour of MRI volume after mapping to cryo. The classic demons algorithm is based on the algorithm proposed by Thirion [28]. The user supplied parameters include the number of iterations and the standard deviation of the Gaussian smoothing over the total displacement field. The algorithm uses Linear interpolator. The classic demons gave an RMS value of 0.207.

The diffeomorphic demons is based on the algorithm described in [31]. The algorithm extends the main idea of Demons, i.e. to develop an optimization procedure on the entire space of displacement fields, but in this case this theory is adapted to a space of diffeomorphic transformations, such that the resulting deformation field and its inverse are invertible. The RMS score in this method is 0.218.

Symmetric and fast symmetric forces demons algorithms produce a deformation field which results in symmetric forces on all ends of the pixels. The RMS scores of the two algorithms are 0.522 and 0.218, respectively.

The results obtained after applying each of the four filters is shown in Fig. 3.

3.4.2 BSpline registration (*SimpleITK*):

. This multi-modal registration algorithm uses BSplines for registration, with metric as Matthes mutual information, optimizer as gradient descent line search and linear interpolator. A deformation is defined on a sparse regular grid of control points and is varied by defining a deformation of each control point. The deformation at any point is obtained by using a B-spline interpolation kernel. A BSplineTransform usually has a large number of parameters which increases the complexity and duration of optimizing the deformation. The multi-resolution BSpline approach initially resamples the volumes to a lower resolution, and performs the registration at this low resolution with fewer parameters at the first level and then adapts or resamples the BSpline control points to a higher resolution at subsequent levels. The

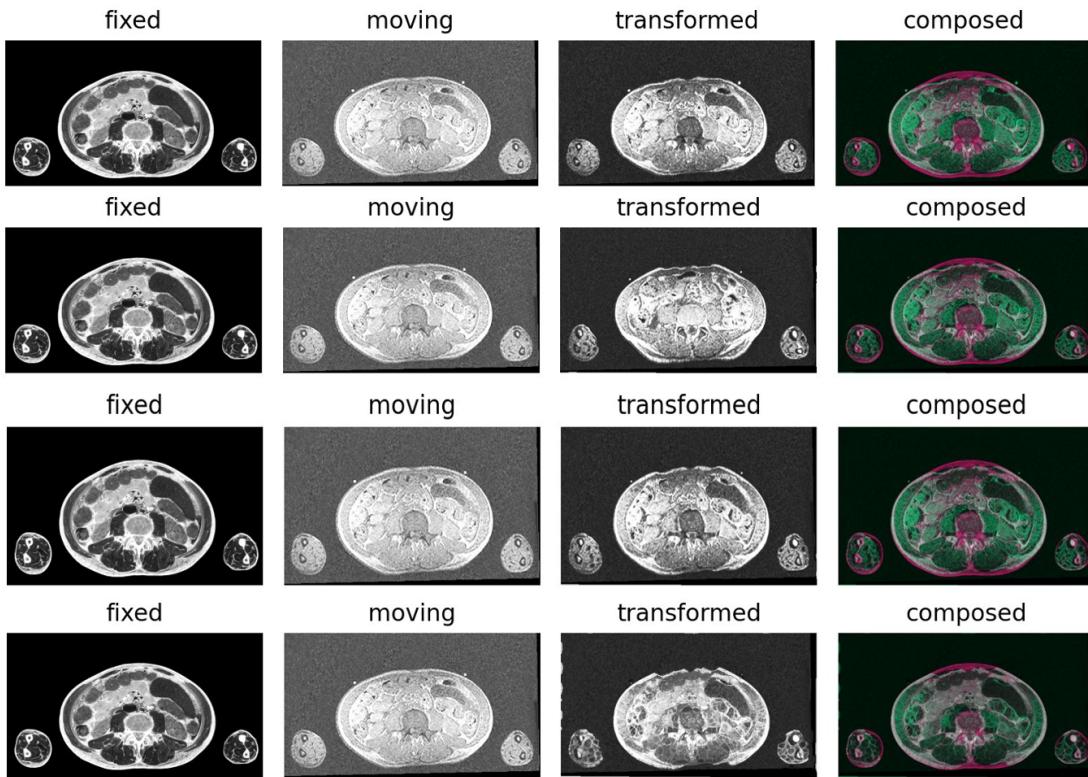


Fig. 3. Deformable registration result using SimpleITK using (1) simple demons algorithm, (2) diffeomorphic demons algorithm, (3) fast symmetric forces demons algorithm and (4)symmetric forces demons algorithm, respectively. The fixed volume is the gray-scale sectioned volume, moving volume is the rigid-registered MRI volume, and the composition is of the reference and transformed volumes, with red lines corresponding to reference volume and green lines corresponding to the transformed volume. The image shows the result for one slice.

multi-resolution framework uses two methods, SetShrinkFactorsPerLevel and SetSmoothingSigmasPerLevel. The former receives the shrink factors to apply when moving from one level of the pyramid to the next and the later receives the sigmas to use for smoothing when moving from level to level. The deformation field grid is defined by a user specified GridRegion, GridSpacing and GridOrigin. Each grid/control point has associated with it N deformation coefficients , representing the N directional components of the deformation. Deformation outside the grid plus support region for the BSpline interpolation is assumed to be zero. The

parameters for this transform is an $N \times N$ -D grid of spline coefficients. The user specifies the parameters as one flat array: each N-D grid is represented by an array in the same way an N-D image is represented in the buffer; the N arrays are then concatenated together to form a single array. Zero-order and third-order B-spline kernels are used to compute the probability density functions of the fixed and moving images, respectively. The result is shown in Fig. 4.

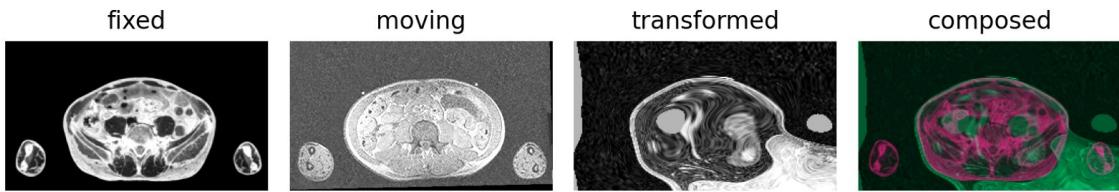


Fig. 4. Deformable BSpline registration result using SimpleITK. The reference is the gray-scale sectioned volume, moving volume is the rigid-registered MRI volume, and the composition is of the reference and transformed volumes, with red lines corresponding to reference volume and green lines corresponding to the transformed volume. The image shows the result for one slice.

3.4.3 Demons registration (MITK):

. This registration using the MITK software, used fast symmetric demons registration with mean square as error metric. It implements the algorithm described in [30]. Though was meant for mono-modal registration, gave results when MRI volume was mapped to the Elastix rigid-registered CT volume. The result is as given in Fig. 5.

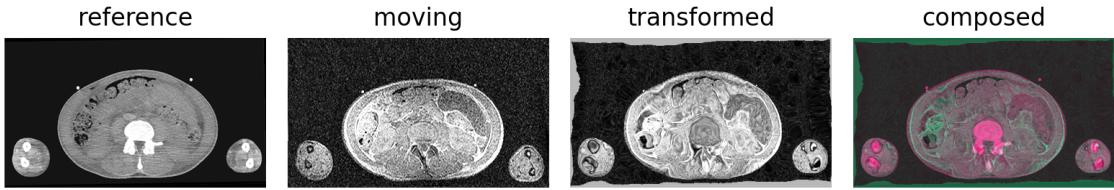


Fig. 5. Fast symmetric forces demons registration result using MITK. The reference is the rigid-registered CT volume, moving volume is the rigid-registered MRI volume, and the composition is of the reference and transformed volumes, with red lines corresponding to reference volume and green lines corresponding to the transformed volume. The image shows the result for one slice.

3.4.4 Level-set motion registration (MITK):

. The level set motion is an intensity morphing model using curve evolution theory, as implemented in [1]. The algorithm is meant for mono-modal registration with metric as mean square error as metric. However a result was obtained for our dataset when rigid-pre-registered MRI volume was mapped to the Elastix rigid-registered CT volume. The result is as given in Fig. 5.

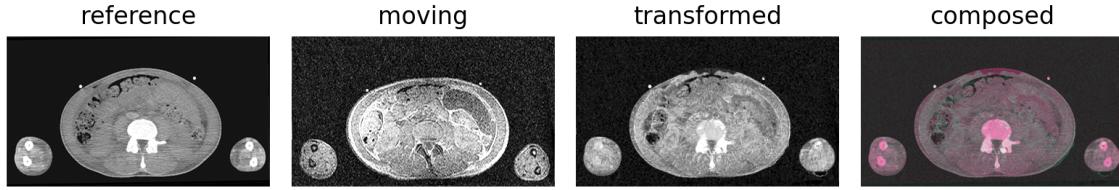


Fig. 6. Level-set motion registration result using MITK. The reference is the rigid-registered CT volume, moving volume is the rigid-registered MRI volume, and the composition is of the reference and transformed volumes, with red lines corresponding to reference volume and green lines corresponding to the transformed volume. The image shows the result for one slice.

3.4.5 BSpline registration (Elastix):

This multi-modal registration takes place in 2 steps. First there is a rigid registration using Mattes mutual information, using Eulers's transform using adaptive stochastic gradient descent algorithm, and using a linear interpolator. This is followed by free-form BSpline registration. The final physical grid spacing is set to 10mm along all dimensions. For the first result in Fig.7, a combination of mutual information and mean square error is used as metrics, along with a first order spline interpolator whereas for the second result a combination of mutual information and transform bending energy penalty were used as metrics, along with a third order (cubic) spline interpolator. These similarity metrics minimize the intensity variances over time. The algorithm performs constrained optimization using stochastic gradient descent with adaptive step-size estimation [17].

4 DISCUSSION AND CONCLUSION

While most of the algorithms described worked really well for rigid registration, none of them provided perfect results for deformable registration of our current volume. By far, the best result obtained was from the Elastix registration algorithms, however there was a significant

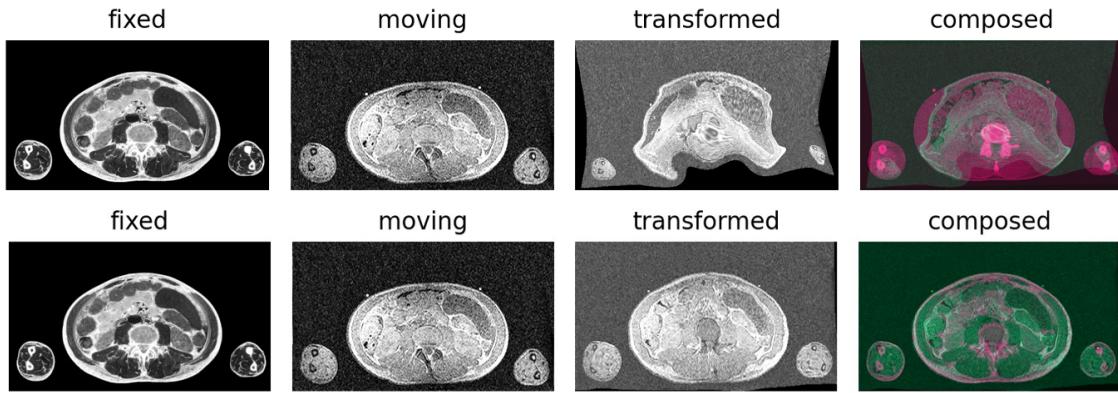


Fig. 7. 2 step registration using Elastix. The moving volume is first rigid-registered then BSpline registered, using a combination of Mutual Information and Mean Square, and Mutual Information and Transform Bending Energy Penalty, as metrics, respectively. The fixed volume is gray-scale sectioned volume, moving volume is MRI volume, and the composition is of the reference and transformed volumes, with red lines corresponding to fixed volume and green lines corresponding to the transformed volume. The image shows the result for one slice.

distortion of the bones of the MRI volume, which neither matched the original MRI nor the cryo volume to which it was mapped.

The most likely conclusion for above is that all the automatic deformable registration methods which gave results were intensity based. While intensity based methods have previously proven to more successful [32], the assumptions on intensities considered by most of the algorithms can cause errors in the registration process. The intensity patterns of cryo-sectioned images or CT and MRI images were vastly different, often being the opposite, for example at the bones which were white and high in intensity for sectioned and CT volumes, but was black and low in intensity in the MRI volumes. The method of acquiring the scans for MRI and CT are also different, with MRI focussing on the muscles, and CT focussing on the bones, which leads to a mismatch when intensity based registration is applied.

To handle the situation of dealing with multimodal registration, the metric used in most of the methods are based on the Mutual Information metric, to work out a statistical relationship between the two volumes to be registered. Mutual information is a very well established similarity measure since it manages different intensities between the modalities provided

that they are relatively consistent within each modality. However, the results confirm the drawbacks of MI based methods mentioned in [32]. The MI method of registration often fails when there are local intensity variations in the volume. Furthermore, MI only considers the statistical intensity relationships between both volumes and ignores the spatial and geometric information about the voxel. Hence MI methods, though often working relatively well (for example in case of Elastix), deviated significantly in terms in certain areas such as the bones.

A lot of deformable registration methods have been tried and tested on brain images. However, a lot of work is yet to be done for fixing large scale deformations of some other areas of the body, especially the abdomen.

This report has compiled a lot of information obtained from several papers as well as documentations. We haven't yet been able to find any existing method which works ideally in case of our dataset, but we have been able to get an understanding of why the algorithms failed, and what should have been different. We hope our report will be useful to anyone who tries to perform registration on full body volumetric registrations in future.

5 ACKNOWLEDGEMENT

We would like to thank Dr. Ojaswa Sharma for allowing us to work as interns in this project, and for guiding and supervising us throughout our internship. We would like to thank IIIT Delhi for providing us access to their machines and for several provisions which made us comfortable throughout the remote internship. Lastly, we would like to thank the other members of the Graphics Research Group for helping us with our work by sharing useful feedback and insights.

REFERENCES

- [1] [n.d.]. Image registration via level-set motion: Applications to atlas-based segmentation. 7, 1 ([n. d.]), 1–20. [https://doi.org/10.1016/S1361-8415\(02\)00063-4](https://doi.org/10.1016/S1361-8415(02)00063-4)
- [2] [n.d.]. *List of (non-rigid) image registration projects for Python.* <http://pyimreg.github.io/>
- [3] P.J. Besl and Neil D. McKay. [n.d.]. A method for registration of 3-D shapes. 14, 2 ([n. d.]), 239–256. <https://doi.org/10.1109/34.121791>
- [4] G.E. Christensen and H.J. Johnson. 2001. Consistent image registration. 20, 7 (2001), 568–582. <https://doi.org/10.1109/42.932742>

- [5] G.E. Christensen, R.D. Rabbitt, and M.I. Miller. 1996. Deformable templates using large deformation kinematics. 5, 10 (10 1996), 1435–1447. <https://doi.org/10.1109/83.536892>
- [6] Gary E. Christensen, Xiujuan Geng, Jon G. Kuhl, Joel Bruss, Thomas J. Grabowski, Imran A. Pirwani, Michael W. Vannier, John S. Allen, and Hanna Damasio. 2006. Introduction to the Non-rigid Image Registration Evaluation Project (NIREP). In *Biomedical Image Registration*, Josien P. W. Pluim, Boštjan Likar, and Frans A. Gerritsen (Eds.). Vol. 4057. Springer Berlin Heidelberg, 128–135. https://doi.org/10.1007/11784012_16
- [7] Min Suk Chung, Jin Yong Kim, Woo Sup Hwang, and Jin Seo Park. 2002. Visible Korean Human: Another trial for making serially sectioned images. *Proceedings of SPIE - The International Society for Optical Engineering* May (2002).
- [8] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham. 1995. Active Shape Models-Their Training and Application. 61, 1 (01 1995), 38–59. <https://doi.org/10.1006/cviu.1995.1004>
- [9] Emiliano D'Agostino, Frederik Maes, Dirk Vandermeulen, and Paul Suetens. 2003. A viscous fluid model for multimodal non-rigid image registration using mutual information. 7, 4 (2003), 565–575. [https://doi.org/10.1016/S1361-8415\(03\)00039-2](https://doi.org/10.1016/S1361-8415(03)00039-2)
- [10] M.H. Davis, A. Khotanzad, D.P. Flamig, and S.E. Harms. 1997. A physics-based coordinate transformation for 3-D image matching. 16, 3 (06 1997), 317–328. <https://doi.org/10.1109/42.585766>
- [11] J. Fitzpatrick, Derek Hill, and Calvin Maurer. 2021. *CHAPTER 8 Image Registration*. <https://cupdf.com/document/chapter-8-image-registration-semantic-scholar-450-image-registration-system.html>
- [12] Massachusetts General Hospital Center for Morphometric Analysis. [n.d.]. *Internet brain segmentation repository*. <http://www.cma.mgh.harvard.edu/ibsr/>
- [13] Alexandre Guimond, Alexis Roche, Nicholas Ayache, and Jean Meunier. 2001. Three-dimensional multimodal brain warping using the Demons algorithm and adaptive intensity corrections. *IEEE transactions on medical imaging* 20 (02 2001), 58–69. <https://doi.org/10.1109/42.906425>
- [14] Arno Klein, Jesper Andersson, Babak A. Ardekani, John Ashburner, Brian Avants, Ming-Chang Chiang, Gary E. Christensen, D. Louis Collins, James Gee, Pierre Hellier, Joo Hyun Song, Mark Jenkinson, Claude Lepage, Daniel Rueckert, Paul Thompson, Tom Vercauteren, Roger P. Woods, J. John Mann, and Ramin V. Parsey. 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. 46, 3 (07 2009), 786–802. <https://doi.org/10.1016/j.neuroimage.2008.12.037>
- [15] D. Loeckx, F. Maes, D. Vandermeulen, and P. Suetens. 2004. Nonrigid image registration using free-form deformations with a local rigidity constraint. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004: 7th International Conference, Saint-Malo, France, September 26-29, 2004. Proceedings, Part I (Lecture Notes in Computer Science, Vol. 3216)*. 639–646.
- [16] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. 1997. Multimodality image registration by maximization of mutual information. 16 (1997), 187–198. <https://doi.org/10.1109/42.563664>
- [17] C. T. Metz, S. Klein, M. Schaap, T. van Walsum, and W. J. Niessen. [n.d.]. Nonrigid registration of dynamic medical imaging data using nD + t B-splines and a groupwise optimization approach. 15, 2 ([n. d.]), 238–249. <https://doi.org/10.1016/j.media.2010.10.003>

- [18] Marc Modat, Tom Vercauteren, Gerard R. Ridgway, David J. Hawkes, Nick C. Fox, and Sébastien Ourselin. 2010. Diffeomorphic demons using normalized mutual information, evaluation on multimodal brain MR images, Benoit M. Dawant and David R. Haynor (Eds.). 76232K. <https://doi.org/10.1117/12.843962>
- [19] Sayan Nag. 2017. Image Registration Techniques: A Survey. (2017). <https://doi.org/10.17605/OSF.IO/RV65C> arXiv:1712.07540
- [20] Jin Seo Park, Min Suk Chung, Sung Bae Hwang, Yong Sook Lee, Dong-Hwan Har, and Hyung Seon Park. 2005. Visible Korean Human: Improved Serially Sectioned Images of the Entire Body. *IEEE Transactions on Medical Imaging* 24, Mar (2005), 352–360.
- [21] Jin Seo Park, Min Suk Chung, Sung Bae Hwang, Byeong-Seok Shin, and Hyung Seon Park. 2006. Visible Korean Human: Its Techniques and Applications. *Clinical Anatomy* 19 (2006), 216–224.
- [22] Jean-Marc Peyrat, Hervé Delingette, Maxime Sermesant, Chenyang Xu, and Nicholas Ayache. [n.d.]. Registration of 4D Cardiac CT Sequences Under Trajectory Constraints With Multichannel Diffeomorphic Demons. 29, 7 ([n. d.]), 1351–1368. <https://doi.org/10.1109/TMI.2009.2038908>
- [23] David W. Shattuck, Mubeena Mirza, Vitria Adisetiyo, Cornelius Hojatkashani, Georges Salamon, Katherine L. Narr, Russell A. Poldrack, Robert M. Bilder, and Arthur W. Toga. 2008. Construction of a 3D probabilistic atlas of human cortical structures. 39, 3 (02 2008), 1064–1080. <https://doi.org/10.1016/j.neuroimage.2007.09.031>
- [24] Aristeidis Sotiras, Christos Davatzikos, and Nikos Paragios. 2013. Deformable Medical Image Registration: A Survey. 32, Jul (2013), 1153–1190. <https://doi.org/10.1109/TMI.2013.2265603>
- [25] Marius Staring, Stefan Klein, and Josien P. W. Pluim. 2007. A rigidity penalty term for nonrigid registration: A rigidity penalty term for nonrigid registration. 34, 11 (10 2007), 4098–4108. <https://doi.org/10.1118/1.2776236>
- [26] Songyuan Tang, Yong Fan, Minjeong Kim, and Dinggang Shen. 2009. RABBIT: rapid alignment of brains by building intermediate templates, Josien P. W. Pluim and Benoit M. Dawant (Eds.). 725909. <https://doi.org/10.1117/12.811174>
- [27] Christine Tanner, Julia A. Schnabel, Daniel Chung, Matthew J. Clarkson, Daniel Rueckert, Derek L. G. Hill, and David J. Hawkes. 2000. Volume and Shape Preservation of Enhancing Lesions when Applying Non-rigid Registration to a Time Series of Contrast Enhancing MR Breast Images. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2000, Third International Conference, Pittsburgh, Pennsylvania, USA, October 11-14, 2000, Proceedings (Lecture Notes in Computer Science, Vol. 1935)*, Scott L. Delp, Anthony M. DiGioia, and Branislav Jaramaz (Eds.). Springer, 327–337. https://doi.org/10.1007/978-3-540-40899-4_33
- [28] J.-P. Thirion. 1998. Image matching as a diffusion process: an analogy with Maxwell's demons. 2 (1998), 243–260. [https://doi.org/10.1016/S1361-8415\(98\)80022-4](https://doi.org/10.1016/S1361-8415(98)80022-4)
- [29] A. Tristán-Vega, G. Vegas-Sánchez-Ferrero, and A Santiago. 2008. Local similarity measures for demons-like registration algorithms. *5th IEEE International Symposium on Biomedical Imaging* (2008), 1087–1090.
- [30] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, and Nicholas Ayache. [n.d.]. Diffeomorphic Demons Using ITK's Finite Difference Solver Hierarchy. <https://hal.inria.fr/inria-00616035>
- [31] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, and Nicholas Ayache. [n.d.]. Non-parametric Diffeomorphic Image Registration with the Demons Algorithm. In *Medical Image Computing and Computer-Assisted*

- Intervention – MICCAI 2007*, Nicholas Ayache, Sébastien Ourselin, and Anthony Maeder (Eds.). Vol. 4792. Springer Berlin Heidelberg, 319–326. https://doi.org/10.1007/978-3-540-75759-7_39
- [32] Jonghye Woo, Maureen Stone, and Jerry L. Prince. [n.d.]. Multimodal Registration via Mutual Information Incorporating Geometric and Spatial Context. 24, 2 ([n. d.]), 757–769. <https://doi.org/10.1109/TIP.2014.2387019>
- [33] Jeroen Wouters, Emiliano D’Agostino, Frederik Maes, Dirk Vandermeulen, and Paul Suetens. 2006. Non-rigid brain image registration using a statistical deformation model, Joseph M. Reinhardt and Josien P. W. Pluim (Eds.). 614411. <https://doi.org/10.1117/12.653081>
- [34] B.T.T. Yeo, M.R. Sabuncu, T. Vercauteren, N. Ayache, B. Fischl, and P. Golland. [n.d.]. Spherical Demons: Fast Diffeomorphic Landmark-Free Surface Registration. 29, 3 ([n. d.]), 650–668. <https://doi.org/10.1109/TMI.2009.2030797>