Towards quantitative follow-up for anatomical and functional whole-body MRI

Jakub Ceranka¹²*, Frederic Lecouvet³, and Jef Vandemeulebroucke¹²

¹ Vrije Universiteit Brussel (VUB), Department of Electronics and Informatics (ETRO), Pleinlaan 2, 1050 Brussels, Belgium;
² imec, Kapeldreef 75, B-3001 Leuven, Belgium;
³ Institut de Recherche Exprimentale et Clinique (IREC), Universite Catholique de Louvain, Belgium
* jceranka@etrovub.be

Introduction: The value of whole-body MRI is constantly growing and is currently employed in several pathologies, including soft tissue tumour detection, diagnosis and prognosis of multiple myeloma, and evaluation of treatment response assessment in bone metastases. Anatomical MRI can provide excellent definition of anatomical structures and underlying pathologies, however, due to its non-quantitative nature, intensities cannot be compared from one acquisition to another. Intra-patient follow-up MR images acquired in the same scanner do not only suffer from spatial misalignment but also intensity inhomogeneities, making the absolute MR intensity values inherently non-comparable. Such inter-scan intensity inhomogeneities make it difficult to derive reproducible measurements and limits the use of treatment response maps only to quantitative MR modalities, e.g. apparent diffusion coefficient (ADC). In this work, we have investigated and compared the performance of several intensity standardization algorithms for whole-body MRI images based on the registration of image histograms.

Methods: The investigated method consists of two steps. Firstly, the follow-up whole-body image is spatially registered to a baseline image using B-spline deformable registration. Secondly, an intensity standardization algorithm based on a histogram registration is applied to the follow-up image. Four different intensity standardization strategies aimed at matching the intensity profiles of a baseline and follow-up image were investigated and evaluated. This included linear histogram scaling in the range up to the 99.8% intensity percentile; linear piecewise histogram image registration and deformable image registration performed between a reference and target image histograms. Additionally, an extension to deformable histogram registration was evaluated: an equidistant partitioning method along the cranio-caudal direction of the whole-body image smoothed by B-spline interpolation at the junctions of the partitions. Experiments were performed on a total of 10 whole-body anatomical and diffusion-weighted image pairs (baseline/follow-up) of patients with metastatic bone disease and healthy volunteers. Intensity profile similarity was assessed with a mean absolute intensity difference and Kullback-Leibler histogram similarity metric.

Results: A deformable equidistant partitioning method showed a superior performance in comparison to the other evaluated intensity standardization methods and significantly improved both mean absolute difference and Kullback-Leibler similarity with respect to non-standardized image.

Conclusion: This work compares several methods of histogram registration based intensity standardization methods that can help to overcome the non-quantitative nature of whole-body MRI images, allowing for extraction of important imaging parameters, visualization of whole-body MR treatment response maps and assessment of severity of pathology based on its MR intensity profile.
The Effect of Motion on Elasticity Calculation in MR Elastography

Emily M. Chan1*, Daniel Fovargue1, David Nordsletten1, Vicky Goh1, Matt Kelly2, Ralph Sinkus1 and Julia A. Schnabel1
1 School of Biomedical Engineering & Imaging Sciences, King’s College London, UK;
2 Perspectum Diagnostics, Oxford, UK;
* emily.m.chan@kcl.ac.uk

Introduction: Magnetic Resonance Elastography (MRE) has shown promise in assessing the stage of liver fibrosis [1], by quantitatively measuring tissue stiffness. Determining the stiffness involves imaging over three breath-hold acquisitions, with a fourth reference breath-hold, in order to capture the propagation of shear waves through the tissue in 3D. However, motion between these breath-holds (Fig. 1) is not taken into account in the elasticity calculation; therefore, this work aims to investigate motion as a possible source of error.

Methods: An MRE dataset [2] containing 57 patients was used, where each patient had four breath-hold images acquired of 80×80×8 voxels, with a resolution of 4×4×4mm. The motion across the dataset was estimated by registering each breath-hold image to the reference using B-spline free-form deformations [3] and calculating the mean deformation magnitude for each patient. The results from 5 cases were not used, as registration failed due to poor imaging quality (N=52). The deformations acquired from the registration were then used in a modified MRE reconstruction algorithm [4] to recalculate the elasticity, which was compared to the original elasticity before motion correction.

Results: An average deformation magnitude of 3.87±0.918mm (almost the size of one voxel) over the liver was observed in the dataset (Fig. 2), with a range of 1.97–5.98mm. All cases exhibited a change in elasticity after motion correction, as shown in Fig. 3, with an average change of 360±136Pa over the 52 cases. A Pearson coefficient of 0.422 (p=0.039) indicated a weak association between the mean deformation and the mean change in elasticity (Fig. 4).

Conclusion: The results show the extent of motion likely to be present between breath-holds in MRE, as well as a preliminary indication of the effect this may have on the elasticity, which will be further investigated through simulations.


Acknowledgements: This work was supported by the Medical Research Council [grant number MR/N018028/1] and Perspectum Diagnostics
Multivariable prediction of DIR accuracy in 4D CT of thoracic cancer patients

Celia Juan de la Cruz1, Martin F. Fast1, and Jan-Jakob Sonke1*
1 Department of Radiation Oncology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX, Amsterdam, The Netherlands;
* j.sonke@nki.nl

Introduction: Deformable image registration (DIR) plays a key role in respiratory motion management in radiotherapy. Inaccurate deformations may lead to erroneous patient representations and thus inaccurate treatment planning or patient setup. Landmark-based metrics are frequently used for DIR validation; however, they are time-consuming and not always available. Other quantitative approaches can automatically be derived from the registration, but their relationships with DIR errors are unclear. The purpose of this work was to evaluate the ability of different DIR-derived quality metrics to predict DIR errors of three different DIR algorithms.

Methods: The inhale phase of 4D CTs of 16 thoracic cancer patients with 300 (Castillo (2009, 2010), Phys. Med. Biol. 54, 55) and 100 (Vandemeulebrouke (2011), Med. Phys. 38) manually identified landmarks was deformed to the exhale phase using a feature-based dense displacement algorithm (ADMIRE 2.0, Elekta AB, Stockholm, Sweden), a phase-based optical flow algorithm (WIMP, in-house software) and a multi B-splines registration method with sliding constraint (Elastix 4.7, Delmon (2013), Phys. Med. Biol. 58). The ground-truth error (Target Registration Error (TRE)) was calculated as the Euclidean 3D distances between the true and the deformed landmark position. Landmarks were also spatially labelled based on 8 different regions defined inside the lung.

Four DIR-derived quality metrics were evaluated for each algorithm in every landmark: Inverse Consistency Error (ICE), Transitivity Error (TE), a modified version of the Distance Discordance (DD) (Vickress (2017), Phys. Med. Biol. 62), and Spatial Smoothness (SS) (Schreibmann (2012), J. Appl. Clin. Med. Phys.13). Univariate, multivariate and mixed-effect (with patient and landmark location as random effects) linear models were implemented to predict TRE using the DIR-derived metrics. Landmarks with TRE <1 mm were removed from the analysis to minimize the impact of observers’ uncertainty. A log-transformation was performed on the model parameters to normalize fit residuals.

Results: The TRE for the three algorithms are shown in Table 1. The DDM correlates best with the TRE for all three algorithms (WIMP (R²=0.54), ADMIRE (R²=0.09), Elastix (R²=0.21)). Goodness of fit of each model is denoted by the R² in Table 2.

Conclusion: DDM had the highest association with landmark derived DIR errors for all algorithms evaluated. Algorithms with higher accuracy had lower model goodness of fit. Using multivariate or mixed-effects model, slightly improved the prediction of the results.

<table>
<thead>
<tr>
<th>Table 1: Median TRE values and percentile range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median TRE (mm)</td>
</tr>
<tr>
<td>WIMP</td>
</tr>
<tr>
<td>ADMIRE</td>
</tr>
<tr>
<td>Elastix</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2: Models' goodness of fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>R² (TRE=DDM)</td>
</tr>
<tr>
<td>R² (TRE=ICE)</td>
</tr>
<tr>
<td>R² (TRE=SS)</td>
</tr>
<tr>
<td>R² (TRE=TE)</td>
</tr>
<tr>
<td>Adjusted R² Multivariate</td>
</tr>
<tr>
<td>Marginal R² Mixed-Effects</td>
</tr>
</tbody>
</table>
Training deep convolutional neural networks for deformable image registration of pulmonary CT images.
Koen A.J. Eppenhof*, Maxime W. Lafarge, and Josien P.W. Pluim

1Eindhoven University of Technology, The Netherlands; 2University Medical Center Utrecht, The Netherlands; * k.a.j.eppenhof@tue.nl

Introduction: Deformable image registration can be time consuming and often needs extensive parameterization to perform well on a specific application. We present a deformable registration method based on a three-dimensional convolutional neural network, together with a framework for training such a network. In one forward pass, the network directly estimates the deformation field from two input images, which requires less than a second.

Methods: We use the well-known U-net neural network architecture. The network outputs the deformation field between two 3D images as three maps for the x-, y-, and z-components of the displacement. The network is trained on synthetically transformed images, which means that no manually annotated data is required. The synthetic transformations consist of a sequence of three transformations that operate at different scales. This multi-scale approach aids in generating larger displacements while preventing folding or tearing of the image. To cope with larger transformations, we propose to train the network by building up the U-net during training, starting with the lower resolution layers and slowly introducing higher resolution layers and higher resolution images and deformation fields.

Results: We present results on registration of affinely registered pulmonary CT inhale-exhale images, validated by target registration errors on corresponding landmark annotations. The network was trained on the CREATIS data set and validated on the DIRLAB data set to show generalization to different data. The TRE values are compared to running the Elastix registration software on the same images using pipelines with and without lung masks. We show that the performance of the network approaches the performance of Elastix without lung masks (TRE 2.66 ± 1.64 mm for the network versus 2.17 ± 3.22 mm for Elastix). The network is able to make deformation field estimates in 0.6 seconds, which is substantially faster than other methods.

Conclusion: We have trained and validated a convolutional neural network for deformable image registration and show that the results are approaching that of Elastix at vastly shorter registration times, while being more robust to outliers. Other advantages of the method include the lack of explicit parameterization of similarity metrics and optimization methods, and the fact that no ground truth labels are needed to train the network.
Landmark propagation over cardiac 4D-CT using non-rigid registration

D. Farotto1,2*, P. Vanden Berghe1, R. Wirix-Speetjens1, J Vander Sloten2
1 Materialise, Belgium; 2 KU Leuven, Belgium;
*Dario.farotto@materialise.be

Introduction: Transcatheter mitral valve replacement has recently emerged as a therapeutic alternative for patients with mitral valve regurgitation [1]. Various devices are being evaluated, showing that careful design and sizing, adapted to the target patient population, are crucial to their success. Therefore, detailed knowledge about the deforming heart tissue over the full cardiac cycle is necessary. Cardiac 4D-CT captures the full cardiac cycle, allowing annotations of landmarks for the detection of the heart tissue trajectories. However, manual detection of these landmarks is a time consuming and error-prone process. In this study, three non-rigid image registration schemes were used to automate the landmark annotation process, propagating landmark locations from a single phase to all phases of the cardiac cycle.

Methods: For eight aortic 4D-CT scans (11 phases), 8 relevant landmarks on the left ventricle were manually indicated by three cardiac experts. Ground truth landmarks were defined as the average of these three indications, allowing to measure the inter-user error. The manual indication was performed using Mimics 20 Research (Materialise, Leuven). The ground truth from one template phase was then propagated to the remaining phases using two different strategies: a pairwise strategy, where the template image always refers to the phase preceding the target image (pw_1), and a pairwise strategy, where the template image is kept constant (pw_2). The registrations were performed with the registration package elastix. The bias related to the choice of the reference phase was accounted for by repeating the process starting from all phases.

Results: Figure 1, left, depicts the boxplots of the errors for all landmarks and patients. The right plot shows the error as a function of the phase, averaged over all the patients, for a crucial landmark on the mitral valve, the lateral trigone. The accuracy clearly degrades as the distance traveled by the landmark increases.

Discussion and Conclusion: The obtained results are encouraging: the propagation accuracy of the pairwise registration schemes is close to the accuracy of manual indication. The high accuracy indicates that non-rigid registration could indeed be used to propagate landmarks and potentially also segmentation masks and respective 3D-objects over the cardiac cycle, making the efficient analysis of the deforming heart tissue more attainable.

References
Determine accuracy of registration algorithms in lung subtraction CT with the LUMIC Challenge

Dagmar Grob*1, James Meakin1, Luuk Oostveen1, Bram van Ginneken1, Monique Brink1 and Ioannis Sechopoulos1
1 Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Nijmegen, The Netherlands
* Dagmar.Grob@radboudumc.nl

Introduction: Subtraction computed tomography (CT) subtracts an unenhanced, pre-contrast injection, CT scan from an enhanced, post-contrast injection, CT scan to obtain pulmonary iodine maps. To align the geometry of the two scans, the unenhanced scan is registered with the enhanced scan. Usefulness of subtraction CT requires a high accuracy of the registration algorithm. This challenge aims to evaluate the performance of these algorithms by using an anthropomorphic digital phantom to simulate a chest CT scan and hence determine the accuracy of algorithms in registering the unenhanced and enhanced CT images.

Methods: The LUMIC challenge is aimed at comparing the accuracy and performance of algorithms in registering simulated, phantom-based unenhanced and enhanced CT images of the XCAT phantom (Duke University, Durham, North Carolina, U.S.A.). The simulated scans are available with voxel size 0.6x0.6x1.0 mm³. The unenhanced CT images are generated with different diaphragm locations compared to those in the enhanced images: 3 mm (small difference), 8 mm (average clinical difference), and 20 mm (large difference), all without pulmonary pathology. A training set is available with diaphragm level differences of 2, 6, and 17 mm. Since the voxel-by-voxel displacement between the two scans is known, the residual error between the estimated displacement fields and the known truth can be determined for all lung voxels.

Results: The results submitted in response to this challenge will be calculated as the median, 75th and 90th percentile of the residual values for each of the different diaphragm levels. Per residual value, a ranking will be determined by comparing the values of each submission, and finally the average ranking will be calculated. Three algorithms have already been evaluated, resulting in the following residual errors for the large diaphragm-difference with the anisotropic voxel data [median (75th – 90th percentile)]: 0.93 mm (1.58 mm – 2.56 mm), 0.92 mm (1.57 mm – 2.58 mm), 0.93 mm (1.60 mm – 2.74 mm). For all registration algorithms, the largest errors were seen in the areas around the heart and close to the diaphragm.

Conclusion: This challenge is a useful way to determine and compare the performance of registration algorithms of the chest. In addition, this challenge will provide the location-specific residual error information by showing several axial slices of the residual error, enabling the participants to identify problematic areas for their registration algorithms. The challenge is available on https://lumic.grand-challenge.org.
Enabling manual intervention for otherwise automated registration of large image series

Roman Grothausmann*, and Christian Mühlfeld

1 Institute of Functional and Applied Anatomy, Hannover Medical School, Hannover, Germany;
* grothausmann.roman@mh-hannover.de

Introduction: Aligning thousands of images from serial imaging techniques can be a cumbersome task. Methods and programs for automation exist but often need case-specific tuning of many meta-parameters (e.g., mask, pyramid-scales, denoise, transformtype, method/metric, optimizer and its parameters). Other programs that apparently depend on few parameters often hide many of the remaining ones (initialized with defaults) often cannot handle challenging cases satisfactory.

Methods: Instead of spending much time on the search for suitable meta-parameter that yield a usable result for the complete image series, the described approach allows to provide specific meta-parameters for problematic image pairs or to intervene by manually aligning image pairs. The meta-parameters or manually found transform is then used by the automatic alignment as initial transformation that is then optimized as in the pure automatic cases. Therefore the manual alignment does not have to be precise. This way the worst case time consumption is limited and can be estimated (manual alignment of the whole series) in contrast to tuning of meta-parameters of pure auto-alignment of complete series which can hardly be guessed.

This introduces a trade-off concerning the use of given time: either on tuning meta parameters in order to improve unsupervised auto alignment or on simple repetitive initial manual alignment of failed cases of the auto alignment.

Results: “Register Virtual Stack Slices” (Fiji) is able to align more images of a sequence of lung tissue (see images below) without intervention (~100) but registering a series of ~2500 images in total did not succeed within a week of trying to find suitable meta-parameters whereas the use of the developed script (http://github.com/romangrothausmann/elastix_scripts/tree/reg2transformed_01 using SimpleElastix, in conjunction with Makefile rules calling “midas” from IMOD for manually providing initial transforms) allows us to finish the task within less than a week.

Conclusion: The described proof of principle combines automated alignment with manual intervention such that ideally the automation does the whole work but also ensures that “in the worst case” at least the result of a pure manual alignment will be achieved.

This comes with need to balance the two time consuming tasks:
either tuning the meta-parameters for the automated alignment (total time can be unlimited)
or helping the automation with initial manual alignments (total time is limited).
While the implementation through Makefiles serves as a proof of principle, it can be further improved to a more intuitive and user-friendly program.

![Image 1](image1.png)
![Image 2](image2.png)
![Image 3](image3.png)
A Large Deformation Diffeomorphic Approach to Registration of CLARITY Images via Mutual Information

Kwame S. Kutten¹*, Nicolas Charon¹, Michael I. Miller¹, J. Tilak Ratnanather¹, Jordan Matelsky¹, Alexander D. Baden¹, Kunal Lillaney¹, Karl Deisseroth², Li Ye², and Joshua T. Vogelstein¹
¹ Johns Hopkins University, USA; ² Stanford University, USA;
* kkutten1@jhu.edu

Introduction: CLARITY is a method for converting biological tissues into translucent and porous hydrogel-tissue hybrids. This facilitates interrogation with light sheet microscopy and penetration of molecular probes while avoiding physical slicing. Deformable registration of acquired images to a standard atlas is an essential step in building brain maps. It enables the precise localization of synapses, axons and cell bodies.

Methods: We developed a pipeline for registering CLARIfied mouse brains, acquired using light-sheet microscopy, to the Allen Reference Atlas (ARA)'s Scanning Two-Photon (STP) template. Differences between these modalities make registration using absolute intensity values impractical. Thus we derived a Mutual Information based LDDMM (MI-LDDMM) algorithm from an optimal control standpoint. As acquired image volumes were over a terabyte in size, they were far too large for work on personal computers. Therefore the NeuroData computational infrastructure was deployed for multi-resolution storage and visualization of these images and annotations on the web. Source code was integrated into the NeuroData Registration (ndreg) Python module (https://github.com/neurodata/ndreg)

Results: Eleven CLARITY volumes were registered to the ARA using MI-LDDMM. Quantitative comparison using fiducial landmarks showed that MI-LDDMM outperformed traditional Sum of Squared Differences (SSD) LDDMM and was comparable to SyN ANTs. Furthermore we demonstrated how a cascaded multi-resolution approach can improve registration quality while reducing algorithm run time.

Conclusion: We developed and validated MI-LDDMM for registering CLARITY image volumes to a standard brain atlas. Future work may involve further code optimizations to reduce runtime.
Diffeomorphic Surface Registration for Measuring Cortical Thickness

Kwame S. Kutten1*, Sylvain Arguillère1, Laurent Younes1, J. Tilak Ratnanather1,
1 Johns Hopkins University, USA
* kkutten1@jhu.edu

Introduction: The cerebral cortex is a highly folded structure whose inner surface lies on the white matter and outer surface follows the pia matter on the brain’s exterior. Its vertical structure consists of columns with ends orthogonal to these surfaces. Horizontally it has six layers whose thicknesses can vary greatly from sulci to gyri. In order to accurately describe the depths and locations of cortical structures, a curved coordinate system must be developed.

Methods: In any brain region the inner and outer surfaces are directly related to each other due to cortical folding occurring during brain development. Thus using the inner and outer cortical surfaces as the template and target respectively, registration was done within the LDDMM framework. To reflect the columnar structure, time-varying velocity fields were constrained to be orthogonal to the deformed template at all times. The curved paths followed by inner surface vertices during deformation formed the basis of the coordinate system. Thickness was defined as the length of these paths. We tested our approach on auditory cortical surfaces reconstructed from MR scans of a congenitally deaf cat and a normally hearing cat.

Results: The figure shows the curves between surfaces for a normally hearing cat (left column) and a deaf cat (right column). Thicknesses are show in millimeters on the outer surface. Our results suggested that FreeSurfer underestimates thickness due to its dependence on Euclidean distance.

Conclusion: We introduced a technique for constructing a cortical coordinate system through diffeomorphic registration. Future work will involve validation through comparison of curves with cortical columns in histology.
The Continuous Registration Challenge

Kasper Marstal1*, Floris Berendsen2, Niels Dekker2, Marius Staring2 and Stefan Klein1

1 Biomedical Imaging Group Rotterdam Erasmus MC, University Medical Center Rotterdam Departments of Medical Informatics and Radiology, P.O. Box 2040, 3000 CA
2 Division of Image Processing (LKEB), Department of Radiology, Leiden University Medical Center, P.O. Box 9600, 2300 RC Leiden, the Netherlands

* k.marstal@erasmusmc.nl

Introduction: Image registration is a common preprocessing step in many different medical image analysis applications. However, registration methods are implemented in a large variety of toolboxes, and the methods are rarely compared on the same datasets. The lack of standardization makes it challenging for end-users to select the right registration algorithm and hyperparameters for their application.

To standardize comparison, registration methods can enter a Grand Challenge (GC). GCs are competitions with standardized data sets, evaluation methods, and experimental setups that focus on specific research topics. The experiments are run by third parties which ensures fair, independent evaluations. However, most modern GCs are static, one-time events, that allow closed source contributions. This hampers collaboration and reproducibility.

Methods: To address these limitations, and inspired by modern software development practices, we proposed the Continuous Registration Challenge (CRC). For this challenge, we developed a fully automatic platform for benchmarking registration methods on many different data sets. The platform consists of a C++ API for running registrations, a Python framework and a Continuous Integration system for running experiments, a compute backend, and a website with public leaderboards. The platform and all submissions are open source. The C++ API, SuperElastix, was designed with a role-base architecture that allows many different registration paradigms to co-exist in the same framework. The challenge focuses on pairwise registration of lungs and brains, two problems frequently encountered in clinical settings.

Results: The system described above is used for the Continuous Registration Challenge. The system allows researchers to test their methods in a standardized way, using a fully automatic experimental setup. The experiments are running every week, and participants can follow the leaderboards, which are updated every weekend (https://continuousregistration.grand-challenge.org/leaderboard/). The leaderboards will continue to be updated and the repository will be open for contributions even after the challenge ends. The results will be presented and discussed at the Workshop On Biomedical Image Registration (WBIR 2018, https://wbir2018.nl/). All participants are invited to collaborate on a paper which we plan to submit to a leading journal in the field.

Conclusion: We present an open source framework for the continuous and automated benchmarking of image registration algorithms.
High-Performance, Open-Source Microscopy Image Montaging

Matthew McCormick1, Dženan Zukić1, François Budin1, Michael Jackson2, Dennis Dimiduk2, David Rowenhorst1, Sean Donegan4, and Michael Groeber4

Kitware, Inc, USA; 2 BlueQuartz Software, USA; 3 Navy Research Labs, USA; 4 Air Force Research Labs, USA;

* matt.mccormick@kitware.com

Introduction: Large image montaging is an image registration problem encountered in scientific microscopy. The fusion of 2D microscopy image tiles into a continuous, whole 2D image or 3D volume is critical to understand large structures and the relationship between structures at multiple length scales. While excellent open source tools are available in Fiji / ImageJ to perform montaging, we describe the development of additional open source tools to perform high-performance, algorithmically advanced montaging on scientific images. These tools are based on a C++ software architecture.

Methods: Phase correlation-based montaging methods are being actively developed to address the microscopy image stitching problem. The re-usable open source methods are developed as a module for the Insight Toolkit library for developers and exposed as a plugin in the DREAM.3D applications for end users. Users or developers can also script batch processing jobs through Python bindings.

Results: Precision, accuracy, and computational throughput are evaluated on ground truth tiles sampled from a large image and the NIST Stitching Challenge dataset. Translation between two tiles based on FFT-based phase-correlation is improved through: a) pre-processing to remove exposure bias, b) multiple FFT-padding methods, c) multiple sub-pixel phase-correlation peak interpolation methods in the real or complex phase correlation image, d) discarding local maxima based on the translation magnitude and the distribution of phase correlation magnitudes, and e) regularization from the translations of nearby tiles. To facilitate flexible resampling operations, visualization of intermediate results, and mitigate interpolation errors, the resulting output translation transformations can be serialized to and deserialized from file.

Conclusion: Microscopy image tile stitching is required to understand microstructures at multiple scales. There are considerable practical challenges to this problem, and montaging methods are informed by approaches to register two images through their patches. In this work, we present our initial efforts to enable high-throughput, reproducible montaging for both research software developers and domain scientists.
Automatic Initialization for 3D ultrasound-CT Registration During Liver Tumor Ablations

Dirk Schut1*, Theo van Walsum2
1 TU Delft, The Netherlands; 2 Biomedical Imaging Group Rotterdam, The Netherlands; *dirk_schut@live.nl

Introduction: During liver tumor ablations, ultrasound(US) guidance is used to insert a needle into a tumor. However, the tumor is not always visible on the US images. Therefore image fusion with a pre-operative contrast enhanced CT scan may be used to visualize the tumor in the US images. This multimodal fusion requires an initial (approximate) alignment between CT and US images, and a subsequent continuous update (tracking) of this alignment. This abstract presents an automated way of initializing the alignment. Challenges include: the large search space of possible US probe orientations, the difference in appearance between US and CT images and the fact that the appearance of US scans is orientation dependent.

Methods: Prior to the intervention, in the CT scan segmentations are obtained of the skin, of the boundary of the liver and of the blood vessels inside of the liver. The skin segmentation is based on thresholds and region growing, the vessel segmentation is based on vesselness filtering and for the liver boundary, manual segmentations were already available for the dataset. During the intervention, blood vessels and the liver boundary are automatically segmented in the US scan, using a neural network with a 3D U-net architecture [1]. The dataset consists of only 40 scans so data augmentation with small affine deformations is used during training.

To find a rigid registration, a cost function is minimized that combines the distance between closest points in both segmentations and between the US scanner and the skin. To provide robustness against segmentation errors a trimmed distance function [2] is used that discards the biggest distances. Because of the big search space the SaDE (self-adaptive differential evolution) optimizer [3] is used.

Results: The 3D U-net segmentation has an average Dice score of 0.518 on the liver boundary and 0.624 on the blood vessels, both using 4-fold cross validation. Using these segmentations for registration, 18 out of 35 initializations were considered good, and 7 were considered acceptable by the author. When using the manual ground truth segmentation instead, 24 out of 35 were considered good and 4 acceptable.

Conclusions: Segmentations of the liver boundary, blood vessels and the skin can be used for registration between CT and US images. Moreover, automatic US segmentations can be obtained using a 3D U-net. However, further improvements are needed to apply fully automatic US-CT image fusion in practice.

References:
Towards A Robust Deep CT-Ultrasound Registration Method for Image Guidance in Liver Tumor Ablation

Yuanyuan Sun1*, Wiro Niessen1, Adriaan Moelker1 and Theo van Walsum1
1 Erasmus University Medical Center, Rotterdam, the Netherlands; 2 Department of Imaging Physics, Faculty of Applied Sciences, Delft University of Technology, Delft, the Netherlands
*y.sun@erasusmc.nl

Introduction: Real-time overlay of CT in ultrasound (US) image guidance is ideal for liver tumor ablation, which helps the clinicians to accurately visualize the target. To this end, we have been investigating a deep learning method for single- and multi-modality alignment of CT and US volumes. In this work, we present a network that directly estimates the displacement vector (DV) from a pair of 2D image patches, without implementing an explicit similarity metric like conventional methods.

Methods: The proposed CNN architecture DVNet takes a pair of 2D patches as input, which is similar to the one in [1]. The output of DVNet is a two-channel patch which resembles as the displacement of the center of the patch. To generate the patch pairs, a patch of $73 \times 73$ pixels (the pixel size ranges between 0.37 and 0.75 mm) was first sampled from the fixed image. Then, the other patch of the same size was generated from the moving image (aligned with the fixed image) with a synthetic DV (random integer between $\pm 15$ pixels in x, y).

Experiments & Results: We have applied the method to various registration scenarios, from monomodal to (simulated) multi-modal with inaccurate ground truth:
1. Mono-modality: The fixed and the moving image are the same (either CT or US) here except that there is Gaussian noise ($\sigma = 5$) on the moving image. The final error on the test data is less than 1 mm after 100 epochs;
2. Multi-modality (simulated): Then, towards CT - US registration, we first did experiments on CT and simulated US (a linear combination of CT and its directional gradient magnitude: $\alpha CT + (1 - \alpha)|\nabla CT|$ [2]). The final error on the test data is less than 1 mm after 200 epochs (for $\alpha = 0.5$ or random $\alpha$);
3. Inaccurate ground truth: As it is very difficult to obtain accurate ground truth for CT - US registration, we also investigate the robustness of DVNet on inaccurate ground truth. With the disturbance (a random value $\in [-4, 4]$ or $[-8, 8]$) pixels) on all the target displacement vectors, DVNet could still learn the right displacements, for both single-modality and (simulated) multi-modality.
4. Full CT-US registration: this is work-in-progress.

Conclusion: In this work, first, we have investigated a method to estimate the translation between a pair of image patches of single modality. Second, we did several experiments towards a method for CT-US registration. In the future, we are going to investigate if DVNet could work on US – US simulation from CT or vice versa.

References:
Estimating inter-patient registration uncertainty for image-based data mining

E. Vasquez Osorio\(^1\)*, A. McWilliam\(^1\), and M. van Herk\(^1\)

\(^1\) The University of Manchester, UK; The Christie NHS Trust, UK

* eliana.vasquezosorio@manchester.ac.uk

**Introduction:** Data mining allows generating new information by examining large pre-existing datasets. In the context of radiotherapy, dose delivered to anatomical locations and its relation to outcomes, e.g., survival or toxicity, is of interest. In order to analyze these relationships, image-based data mining can be used e.g. we identified the base of the heart as a dose sensitive region affecting survival in lung cancer patients [McWilliam et.al. Eur J Cancer. 2017]. In image-base data mining, voxel-wise statistical analysis is performed after dosimetric data from hundreds of patients are mapped to a common reference anatomy using non-rigid registration. Uncertainties related to these registrations should be quantified and accounted for to ensure robust data analysis. We propose a methodology to quantify these uncertainties.

**Methods:** Manually identified points of interest (POIs) were used to estimate uncertainties. Considering that perfect registrations should map corresponding landmarks to the same geometrical location, variations on these mapped POIs can be used to estimate the registration uncertainty. We identified the random (standard deviation of the mapped POIs) and systematic (error dependent on the reference anatomy) components of the registration error. We applied this methodology to two sites: thorax (focused in the lung/heart) and pelvis, using 2 open-source non-rigid registration packages: NiftyReg and Elastix.

**Results:** Registrations were performed for 753 lung cancer patients and 265 anal cancer patients using 5 reference anatomies (patient CTs). Two POIs in the heart were identified for all lung cancer patients and 9 POIs in the pelvic bone for all anal cancer patients. The following figure summarizes the residuals of these points after registration as well as the random and systematic components of the registration errors:

**Conclusion:** Inter-patient registration aimed at image-based data mining has considerable uncertainties, regardless of the registration algorithm used, and should be accounted for [McWilliam et.al. Eur J Cancer. 2017]. Different registration algorithms are more suitable for different sites: in our application, soft tissues in the thorax were better registered using NiftyReg while the bony anatomy of the pelvis was better registered using Elastix. Special care selecting the reference anatomy is essential for robust results as demonstrated by the different distributions of the residuals. We are currently carrying out an inter-observer study to differentiate POI identification and registration uncertainties.