ACCURACY MEASUREMENT FOR MULTI-MODAL RIGID REGISTRATION USING FEATURE DESCRIPTORS

Hauler F.*, Jurisic M. Furtado H, Birkfellner W.

*frida.hauler@meduniwien.ac.at

Center of Medical Physics and Biomedical Engineering, Medical University Vienna

Software for the Use of Multi-Modality images in External Radiotherapy

1 - Introduction

In radiotherapy (RT), different image modalities like CT and MRI or CBCT help to define tumor structures before applying a high dose of ionizing radiation to tumor regions (called target volume). These images need to be registered, finding an optimal geometrical transformation which aligns one dataset (moving image) with corresponding areas into an other dataset (fixed image) taken at various point in time or by different scanners [1].

The accuracy of multi-modal image registration is crucial to spare the surrounding healthy tissues; therefore a reliable evaluation method for registration outcome is needed.

The gold standard validation methods are visual inspection by experts and fiducial-based evaluation[2]. However visual inspection is time consuming and prone to errors. The fiducial-based evaluation is an invasive method when fiducial markers are fixated to the bone or organs.

Therefore, a robust non-invasive automated method is needed for validating the registration accuracy in RT.

The aim of this study is to introduce and evaluate an automatic landmark-based accuracy measure using feature descriptors for multi-modal rigid registration.

2 - Methods and Materials

As evaluation of our method we used a multi modality dataset consisting of CT and as moving images CBCT, MR-T1, MR-T2 scans of a porcine head with seven fixed fiducials and known registration gold standard.

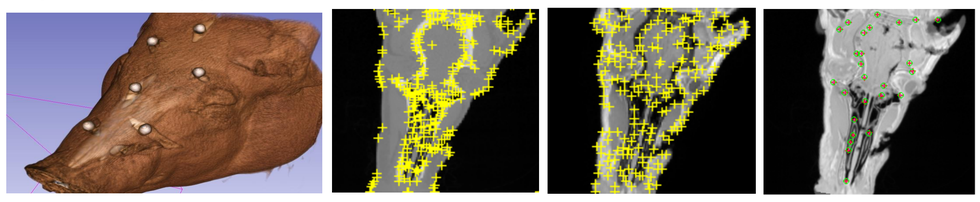

Fig. 1. a.) Porcine phantom with bone fiducials. Finding feature points on b.) CT and c.) MR-T1 coronal slices of the pig head phantom. d.) Overlayed CT and MR-T1 images with matched pair-points.

Furthermore, the robustness of evaluation method was tested on 10 brain and 25 lung dataset.

For registration we used a commercial software, Analyze 11.0. Consistently, the CT is the fixed image and depending of the anatomical region the moving image is CBCT or MR-T1, MR-T2, MR-T1 contrast enhanced with Gd.

After the rigid registration, features on both fixed and moving images are located by SURF descriptor[3],[4] on each slice. The interest points are distinctive locations like corners, blobs and T-junctions found by Hessian detector (Fig. 2a-2b).

The corresponding points from each image are matched by correlation(Fig. 2c). From the presumptive matches found by correlation, the outliers have been eliminated by RANdom SAmple Consensus (RANSAC) [5] (Fig. 2d-2e, Fig. 3b-3c).

The mean euclidean distance between the remained inlier pair matches (Fig. 2f, Fig. 3d) will give the accuracy measure of the rigid registration.

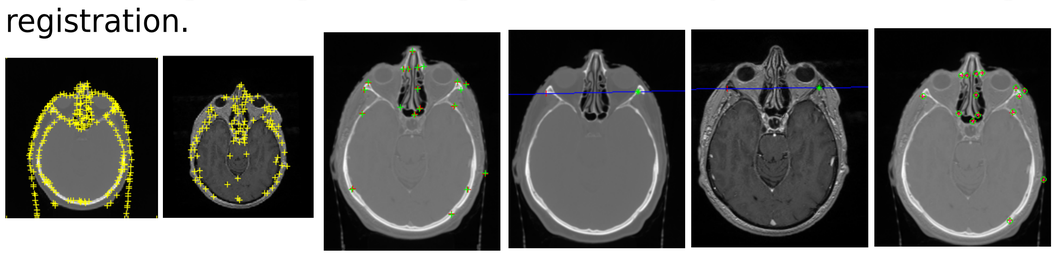

Fig. 2. Feature points found by SURF detector on CT, MR-T1 axial slice of a glioma patient (a,b). Presumptive matching of the pair-interest points by correlation (c), eliminating outliers by RANSAC (d,e), the resulting, overlayed landmark-pairs, after eliminating the outliers from the presumptive match. (f).

The accuracy measure based on feature descriptors was calculated for every slide from each three views.

In absence of fiducial marker data set in clinical cases, the accuracy measure have been compared by manual landmark based evaluation.

Representative slices have been chosen from all image modalities to contain features used by radiation oncologist during visual inspection.

Approximatively on each 10th slice same features were manually annotated from all three views. Finding the corresponding landmarks from the moving and fixed slices was made by the same algorithm as for the automated case.

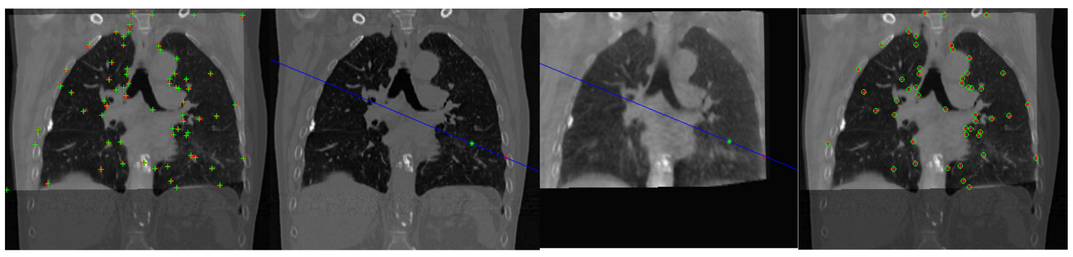

Fig. 3: Representative slices containing features used by visual inspection like bifurcation of bronchus, base of the ribs, heart and diaphragm. In first image, a.) the presumtive matches of manual landmarks are represented on overlayed CT and CBCT images. Images b.) c.) shows the RANSAC line fitting for eliminate the outliers. The last image, d.) shows the corresponding landmarks after eliminating the outliers by RANSAC.

3 - Results

In Table 1. the fiducial registration error is compared by the accuracy measure based on features descriptors for CT, MR-T1, MR-T2 weighted and CBCT images.

Each image has isotropic 1mm^3 voxel size, therefore the metric of accuracy is mm.

Keeping the matching part of the algorithm, we compared our accuracy results with manual landmark based accuracy values for ten brain patiens. (Table 2.)

Accuracy measure	**Fiducial based**									**SURF based**								
Modalities	MR- T1 – CT			MR-T2 – CT			CBCT – CT			MR-T1 – CT			MR-T2 – CT			CBCT – CT		
Views	S	C	A	S	C	A	S	C	A	S	C	A	S	C	A	S	C	A
Pig head	1.1 ± 0.6	0.9 ± 0.4	1.6 ± 0.7	1.2 ± 0.3	1.9 ± 0.7	1.2 ± 0.7	1.1 ± 0.5	1.8 ± 0.7	1.9 ± 0.7	2.6 ± 1.3	2.6 ± 1.2	2.0 ± 1.0	2.0 ± 1.0	2.6 ± 1.6	2.0 ± 1.0	2.1 ± 0.9	2.9 ± 0.9	

Table 1: Comparison of accuracy result by the automatic method with fiducial registration error (FRE).

Accuracy measure	**Manual landmark based**									**SURF based**								
Modalities	***MR-T1 – CT***			***MR-T1Gd – CT***			***MR-T2 – CT***			***MR-T1 – CT***			***MR-T1Gd – CT***			***MR-T2 – CT***		
Views	S	C	A	S	C	A	S	C	A	S	C	A	S	C	A	S	C	A
Patient1	1.9± 0.4	1.6± 0.6	1.4± 0.8	-	2.0± 0.8	2.4± 0.7	–	–	–	1.4± 0.8	1.6 ± 0.6		2.4 ± 0.7	2.0± 0.8	2.4 ± 0.7	–	–-	–
Patient2	1.7± 0.7	2.0± 0.9	1.4± 0.8	2.0± 0.7	1.4± 0.7	2.0± 0.9	–	–	–	1.4± 0.8	2.0± 0.9	1.7± 0.7	2.0± 0.9	1.4 ± 0.7	2.0± 0.7	–	–	–
Patient3	1.9± 0.7	1.6± 0.7	1.9± 0.8	1.5± 0.9	1.8± 0.9	1.4± 0.8	2.2± 1.9	2.0± 0.8	1.4± 0.8	1.9± 0.8	1.6± 0.7	1.9± 0.7	1.4± 0.8	1.8± 0.9	1.5 ± 0.9	1.4± 0.8	2.0± 0.8	2.2± 1.9
Patient4	1.7± 0.6	1.7± 0.8	1.2± 0.8	2.0± 0.8	1.7± 0.7	1.7± 1.1	1.7± 0.5	1.7± 0.6	1.7± 0.5	1.2± 0.8	1.7± 0.8	1.7± 0.6	1.7± 1.1	1.7± 0.7	2.0± 0.8	1.4± 0.8	2.0± 0.6	1.9± 0.5
Patient5	2.1± 0.7	1.4± 0.8	1.2± 0.8	1.9± 0.8	2.0± 0.8	1.9± 0.7	2.0± 0.7	2.0±	1.5± 0.6	1.4± 0.8	1.4 ± 0.8	2.1 ± 0.7	1.9± 0.7	2.0± 0.8	1.9± 0.8	1.5± 0.6	2.0±	2.0± 0.7
Patient6	1.8± 0.7	1.8± 0.7	1.8± 0.7	1.8± 0.8	1.8± 0.6	1.8± 0.7	–	–	–	2.0± 1.1	1.8± 1.8	1.9± 1.1	1.8± 1.1	1.7± 0.8	1.8± 1.0			
Patient7	1.4± 0.7	1.7± 0.6	1.7± 0.6	1.7± 0.7	1.3± 0.6	1.7± 0.6	2.1± 0.5	2.0±	2.0±	1.0± 0.4	1.0± 0.4	1.8± 1.0	1.0± 0.5	1.0± 0.6	2.0± 0.8	1.0± 0.3	1.4± 0.2	1.0± 0.2
Patient8	1.8± 0.5	0.9± 0.6	1.7± 0.5	1.2± 0.3	0.6± 0.3	1.2± 0.5	–	–	–	1.7 ± 0.5	0.9± 0.6	1.8± 0.5	1.2 ± 0.3	0.6± 0.3	1.2± 0.3	–	–	–
Patient9	1.4± 1.0	1.4± 1.0	2.0± 0.4		2.0± 0.4	1.0± 0.7	1.8± 0.3	1.0± 0.9		2.0± 0.4	1.4± 1.0	1.4± 1.0	1.0± 0.7	2.0± 0.4		1.0± 0.9	1.8± 0.3	
Patient 10	2.0± 1.0	2.2± 0.5	1.0± 0.2	1.9± 0.5	1.9± 0.4	2.4± 0.6	–	–	--	1.0± 0.2	2.2± 0.5	2.0± 0.8	2.4 ± 0.6	1.9± 0.4	1.9± 0.5	–	–	--
Total	1.8 ± 0.7	1.7 ± 0.7	1.6 ± 0.8	1.9 ± 0.8	1.8 ± 0.7	1.8 ± 0.7	2.1 ± 0.6	2.0 ± 0.8	1.6 ± 0.6	1.7 ± 0.6	1.8 ± 0.9	1.7 ± 0.5	1.8 ± 0.7	1.8 ± 0.5	1.8 ± 0.4	1.3 ± 0.7	1.2 ± 0.7	1.6 ± 0.7

Table 2. Comparison of accuracy results obtained based on automatic method based on feature descriptors and manual landmark based accuracy for brain clinical dataset. The "--" cell values adduce missing data.

4 - Conclusion

Based on the results obtained, defining an automatic accuracy measure using feature descriptors can be considered a promising method.

In the pig dataset case, the difference between the fiducial registration error (FRE) and our accuracy measurement is in range of 1±0.5 mm on sagittal view, 1.1 ± 0.5 mm on coronal and 0.4 ± 0.2 mm on axial view. This difference can be considered adequate for a non-invasive, automatic measure.

We need to prove our method robustness and accuracy also for misregistered cases.

As future perspective the validation method based on 2D features detection could also be elaborated to 3D futures.

References

[1] Birkfellner, W. (2011). Applied Medical Image Processing: A Basic Course, Taylor & Francis Group, LLC: 231-247

[2] Pawiro, S.; Markelj, P.; Pernuš, F.; Gendrin, C.; Figl, M.; Weber, C.; Kainberger, F.; Nöbauer-Huhmann, I.; Bergmeister, H.; Stock, M.; Georg, D.; Bergmann, H. & Birkfellner, W. (2011). Validation for 2D/3D registration I: A new gold standard data set, Medical Physics 38 : 1481-1490

[3] Bay, H.; Tuytelaars, T. & Van Gool, L. (2006). SURF: Speeded up robust features, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 3951 LNCS : 404-417.

[4] Evans, C. (0000). , OpenSURF - Open Source SURF Feature Extraction Library

[4] Zuliani M.; Kenney C. S.; Manjunath, B. S. (2005). The Multiransac Algorithm and its Application to Detect Planar Homographies, IEEE International Conference on Image Processing

Acknowledgment

This work is part of the SUMMER Marie Curie Research Training Network (PITN-GA-2011-290148), which is funded by the 7th Framework Programme of the European Commission (FP7-PEOPLE-2011-ITN). The information and views set out in this publication are those of the authors and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

Furtado H. is supported by Christian Doppler Laboratory for Medical Radiation Research, Medical University Vienna. We gratefully thank for visual inspection and choosing the features for manual annotation for brain patients to Umberto Sabatini from Fondazione Santa Lucia, Rome, for provided brain clinical datasets to Anne Laprie from Institute Claudius Regaud, for five lung patients and visual inspection to Urusula Nestle from University Klinikum Freiburg.