Comparison of Different Metrics for Appearance-model-based 2D/3D-registration with X-ray Images

Philipp Steininger[1], Karl D. Fritscher[1], Gregor Kofler[1], Benedikt Schuler[1], Markus Hänni[2], Karsten Schwieger[3], Rainer Schubert[1]

[1]Institute for Biomedical Image Analysis, UMIT Hall in Tirol, Austria
[2]AO Development Institute, AO Foundation Davos, Switzerland
[3]AO Research Institute, AO Foundation Davos, Switzerland
philipp.steininger@umit.at

Abstract. The general idea of the presented work is to overcome known problems with segmentation and analysis of 2D radiographs by registering a 3D appearance-model. Therefore this paper introduces a novel method to register 2D x-rays with 3D appearance-models by optimizing the appearance and pose of the model until a virtual radiograph of the generated model-instance optimally fits the investigated x-ray. The approach was tested on a sample set of 15 human femur specimen using different metrics and optimization techniques to investigate the impact on the resulting implicit 2D-segmentation. The first promising results are presented and discussed in detail.

1 Introduction

2D radiographs suffer from considerable information loss compared to the associated real patient anatomy. This is mainly due to overlaps and summation of organs as a consequence of an x-ray's projective nature. In addition, the missing depth information results in ambiguity concerning the origin of 2D shape variation which is a mixture of pose and 3D shape variation leading to insufficient segmentation results. To overcome these problems registration of radiographs with CT-datasets of the patient can help and methods for this approach have been proposed frequently. For instance, a pre-operative CT-scan of the investigated anatomical region can be aligned with an intra-operative x-ray of the same patient to verify the 3D patient position during interventions [1, 2]. Nevertheless, in many clinical situations a CT-image is not indicated and therefore not available. In these cases, it would be desirable to reconstruct the 3D anatomy based on x-ray images without relying on pre-acquired CT-scans of the same patient by substituting the CT-image with a statistical 3D model. In this context Lamecker et al. [3] presented a semi-automatic approach for extracting the 3D geometry of the pelvis from coronal and sagittal x-rays based on a 3D statistical shape model and a known camera calibration. As a consequence of the usage of a pure shape model in this and similar approaches the 3D intensity-distribution of the modeled region is not reconstructed. Hence it is often desired to derive not only 3D shape information on certain structures of a patient from radiographs, but also 3D intensity reconstruction. This is useful in situations were the intensity values are used, e.g. estimating the local anchorage quality of osteoporotic bone in order to enable efficient surgical planning of femoral fracture fixations.

To overcome these limitations of using shape models an approach to use a 3D appearance-model (AM) [4] of the investigated anatomical region for intensity-based 2D/3D-registration will be proposed in this work. By continuously varying the appearance of the model, the 3D anatomy is successively approximated. This approach is expected to allow analyses directly based on the shape, pose and intensity-distribution of the resulting 3D model-instance. In the following we will focus on the human femur. More precisely, we zoom in on the femur's implicit 2D-segmentation in x-rays which can be achieved by projecting the resulting model-instance onto the x-ray-plane.

In the past, different metrics for intensity-based registration [1, 5] were suggested which significantly influence the registration accuracy. Therefore our goal was to study these metrics regarding their suitability for the described approach.

2 Materials and Methods

The underlying AM is a so-called 'InShape'-model [4] of the human femur. It is a combined shape-intensity model S generated from 15 manually segmented CT-scans S_i of femur specimen. By parameterizing the model $S(\alpha) = \bar{S} \cdot U_k \alpha$ with a k-dimensional coefficient vector α the shape-intensity instances S_i of the training set can be approximated. $\bar{S}$ denotes the mean shape-intensity pair and U_k defines a matrix whose columns represent the first k orthogonal modes of shape-intensity variation within the training set.

In our approach the registration process is divided into two stages: In a first step the spatially initialized mean shape-intensity instance of the 'InShape'-model is projected onto a virtual image-plane using ray-casting. This digitally reconstructed radiograph (DRR) is then rigidly aligned with the investigated x-ray during an intensity-based 2D/2D-registration sub-procedure (e.g. [1, 2]). The DRR's image-plane is pre-configured and does not need any preparative calibration procedures. The resulting 2D transformation of the 2D/2D-registration is used for initializing the model's pose in the next step. In this second step, a 2D/3D-registration process, the model's pose and appearance is optimized until the DRR of the generated model-instance optimally matches the examined x-ray as illustrated in Fig. 1(a). The six degrees of freedom of the involved rigid 3D transformation together with k model parameters used for determining the shape and intensity-distribution of the actual model-instance are optimized. This is done with respect to an intensity-based measure calculated from the overlay of the actual DRR and the investigated radiograph.

In both registration stages different metrics and optimization methods have been applied including powell optimization (P) [6], 1+1 evolutionary optimization (OE) [7], mean reciprocal square difference metric (MD) [8], mattes mutual information metric (MM) [9] and normalized mutual information metric (NM) [5]. The result of the complete registration process is a geometrical 3D transformation (3 rotational and 3 translational components) on the one hand and a specific appearance of the model (k-dimensional vector) on the other hand.

3 Results

The approach described in section 2 was tested on a training set of 15 human femur specimen. Both CT-scans and x-ray images of all specimen have been available at original resolutions of 0.254 mm and 0.1 mm respectively. An 'InShape'-model [4] of the femur with a final resolution of 1.0 mm was generated based on these manually segmented CT-scans. The resulting AM comprised the upper regions of the femur and a short part of the shaft (Fig. 1(a)).

Two study settings were used for evaluation. In the first setting, the spatial position of the mean-shape instance for 2D/2D-registration was manually initialized as accurate as possible using the femoral head midpoint as reference. In the second setting the initial spatial position of the mean-shape instance for the first stage was interactively altered up to 6.0 mm in x- and y-direction in order to simulate manual initialization by an expert. For both settings the introduced two-stage registration process was applied utilizing different combinations of optimization techniques and metrics.

After a total number of 600 optimization-iterations (termination condition) the final output of the presented two-stage strategy is a specific model-instance whereof a DRR is calculated. The x-ray of the investigated femur specimen as

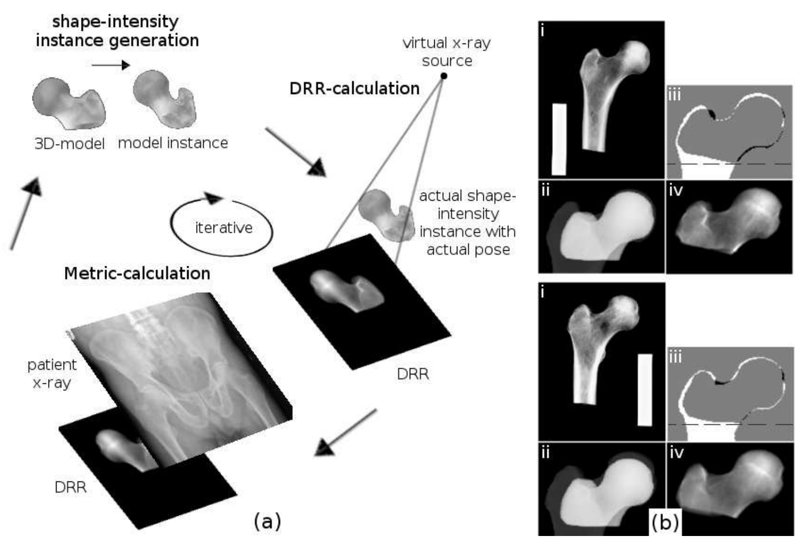

Fig. 1. (a) 2D/3D-registration scheme and (b) two exemplary qualitative results

Table 1. Results of the *first setting* on a set of femur specimen radiographs ($n = 15$)

Registration		SI [1]			HD [mm]			CD [mm]		
2D/2D	2D/3D	mean	SD	max	mean	SD	max	mean	SD	max
MD,OE	MM,OE	0.902	0.038	0.953	7.114	2.655	12.011	3.741	1.351	6.094
MD,OE	MD,OE	0.911	0.036	0.959	5.498	2.197	11.517	3.371	1.405	5.610
MD,OE	NM,OE	0.933	0.024	0.968	4.239	1.502	6.223	2.778	1.000	4.748
NM,OE	MM,OE	0.891	0.041	0.972	6.188	2.548	10.644	4.197	1.519	6.092
NM,OE	MD,OE	0.899	0.042	0.962	4.913	1.465	6.958	3.778	1.450	7.341
NM,OE	NM,OE	0.895	0.046	0.969	5.597	1.891	10.035	4.127	1.758	6.375
MM,P	MM,OE	0.913	0.038	0.961	4.970	1.770	7.685	3.618	1.129	6.105
MM,P	MD,OE	0.839	0.246	0.962	10.057	16.546	67.006	6.888	12.646	50.417
MM,P	NM,OE	0.926	0.031	0.965	4.018	1.203	6.223	3.008	1.365	5.507

well as the generated output DRR can be binarized using a threshold. Furthermore the investigated x-ray must be cropped to the size of the femur's bounding box in the DRR (see dashed lines in images iii in Fig. 1(b)). Three common similarity measures can be calculated from the superimposed cropped binarized x-ray A and the binarized DRR B: the κ-statistic-based similarity index $SI = 2|A \cap B|/(|A| + |B|)$ where $|\cdot|$ denotes the size of a set, the directed hausdorff distance $HD = \max_{a \in A} \min_{b \in B} ||a - b||$ and the contour mean distance $CD = \max(h(A, B), h(B, A))$ where $h(A, B) = \max_{a \in A} \min_{b \in B} ||a - b||$. These measures rate the 2D-segmentation quality of our approach.

Table 1 lists the results of the first study setting. The first two columns show the applied combinations of metrics and optimizers in the two registration stages. Next, the values of arithmetic mean, standard deviation and maximum for each considered similarity measure over the complete set of investigated femur specimen radiographs ($n = 15$) are depicted. The best results were achieved by applying the combinations MD/OE (2D/2D) and NM/OE (2D/3D) resulting in $SI = 0.933 \pm 0.024$, $HD = 4.239 \pm 1.502$ mm, and $CD = 2.778 \pm 1.000$ mm.

In the second setting with less accurate initialization the combinations MD/OE (2D/2D) and NM/OE (2D/3D) also performed best. Over the specimen training set the investigated measures amounted to $SI = 0.928 \pm 0.026$, $HD = 4.736 \pm 1.959$ mm and $CD = 3.406 \pm 0.693$ mm.

Fig. 1(b) shows two exemplary qualitative results (upper images: first setting, lower images: second setting). The investigated femur specimen images (i), logarithmic overlays showing the initial spatial position of the mean shape-intensity instance (ii), the resulting segmentations as binary difference images (iii) and the DRRs of the final model-instances (iv) are depicted.

4 Discussion

The combinations of MD/OE for 2D/2D-registration and NM/OE for 2D/3D-registration perform best in both studies. Furthermore it can be stated that the loss of precision in the second study setting, due to a less accurate initialization, is only marginal. These results show that the described two-stage registration in general is suited for the segmentation of femur structures in radiographs. As seen on Fig. 1(b), the 2D-segmentation quality in the femoral head region tends to be better than in the region of the greater trochanter which may be a consequence of stronger appearance variations in this anatomical area. Behiels et al. [10] compared three different 2D-segmentation approaches based on active shape models (ASM) over a set of proximal femur x-rays. Our segmentation method's performance is comparable to the results presented in [10]. As a consequence of the low number of outliers, the maximum segmentation errors (HD) of our method are relatively small. Therefore our approach appears to be more robust compared to [10]. Another advantage of our approach is that variations in shape (actually appearance) and changes in pose of the investigated anatomical structure are modeled separately. Hence the result also reflects these two aspects explicitly in contrast to ASMs. Currently, the DRR quality is still not optimized. Thus, the next step in this project will be the generation of more realistic DRRs by considering the implicitly modeled electron densities of the investigated anatomical structure. This enhancement of the virtual x-ray is expected to further increase the total reconstruction quality. Since the sample size is too low for validated statements, we currently started to evaluate the presented two-stage registration in the course of a clinical study ($n = 340$). Besides the 2D-segmentation quality we will also evaluate the accuracy and robustness of the reconstructed 3D model appearances compared to the acquired CT-scans of the patients. First results of this study will be available in the near future.

References

1. Maintz JBA, Viergever MA. A survey of medical image registration. Med Image Anal. 1998;2(1):1–36.
2. Birkfellner W, Seemann R, Figl M, et al. Wobbled splatting: A fast perspective volume rendering method for simulation of x-ray images from CT. Comp Prog Med Biol. 2003;50:N73–N84.
3. Lamecker H, Wenckebach T, Hege H. Atlas-based 3D-shape reconstruction from x-ray images. Proc ICPR. 2006; p. 371–4.
4. Fritscher KD, Gruenerbl A, Schubert R. 3D image segmentation using combined shape-intensity prior models. J Comp Assist Rad Surg. 2007;1:341–50.
5. Pluim JPW, Maintz JBA, Viergever MA. Mutual information based registration of medical images: a survey. IEEE Trans Med Imag. 2003;22:986–1004.
6. Press WH, Flannery BP, Teukolsky SA, et al. Numerical Recipes in C. 2nd ed. Cambridge University Press; 2002.
7. Rechenberg I. Evolutionsstrategie. Optimierung technischer Systeme nach Prinzipien der biologischen Evolution. Frommann Holzboog; 1973.
8. Holden M, Hill DLG, Denton ERE, et al. Voxel similarity measures for 3D serial MR brain image registration. Proc IPMI. 1999; p. 472–7.
9. Mattes D, Haynor DR, Vesselle H, et al. PET-CT image registration in the chest using free-form deformations. IEEE Trans Med Imag. 2003;22(1):120–8.
10. Behiels G, Maes F, Vandermeulen D, et al. Evaluation of image features and search strategies for segmentation of bone structures in radiographs using active shape models. Med Imag Anal. 2002;6(1):47–62.