

# Deformable 2D-3D Registration of the Pelvis with a Limited Field of View, Using Shape Statistics

Ofri Sadowsky<sup>1</sup>, Gouthami Chintalapani<sup>1</sup>, and Russell H. Taylor<sup>1</sup>

Department of Computer Science, Johns Hopkins University, USA  
ofri@cs.jhu.edu

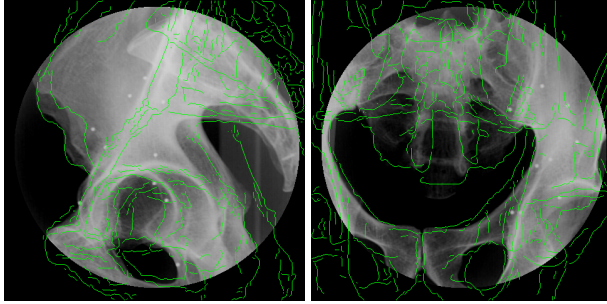
**Abstract.** Our paper summarizes experiments for measuring the accuracy of deformable 2D-3D registration between sets of simulated x-ray images (DRR's) and a statistical shape model of the pelvis bones, which includes x-ray attenuation information ("density"). In many surgical scenarios, the images contain a truncated view of the pelvis anatomy. Our work specifically addresses this problem by examining different selections of truncated views as target images. Our atlas is derived by applying principal component analysis to a population of up to 110 instance shapes. The experiments measure the registration error with a large and truncated FOV. A typical accuracy of about 2 mm is achieved in the 2D-3D registration, compared with about 1.4 mm of an "optimal" 3D-3D registration.

## 1 Introduction

Deformable statistical models of anatomy are useful tools for recovering the shape of patient anatomy when only partial information is available. In a previous publication [1], we showed that a deformably registered anatomical model can be combined with a limited-trajectory set of x-ray images to create a CT-like "hybrid reconstruction". The current paper focuses on assessing the accuracy of our deformable registration.

A frequent challenge in registering data to intra-operative fluoroscopic images is truncation, or a limited field of view (FOV) in the x-ray. Fig. 1 shows examples of fluoroscopic images of a dry cadaveric pelvic girdle (our target anatomy), taken with a common 9" C-arm. Only parts of the large and complex bone appear in the image, and this may affect the registration results. A main goal of this paper is to characterize a trade-off between different combinations of target images: a large FOV and a small number of projections, or a small FOV and varying numbers of projections. In comparison, previous works on deformable 2D-3D registration, such as [2,3,4], usually applied to smaller bones, such as the proximal femur, and the view included the full anatomy.

We show in this paper that our shape atlas and the deformable 2D-3D registration method we created are relatively robust to the FOV truncation problem, if the imaging poses are carefully selected. Our registration is based on normalized mutual information (NMI) similarity, and does not require edge detection



**Fig. 1.** Truncated fluoroscopic images of the pelvic bones. The images were taken with a 9" OEC 9600 C-arm. Contours of a deformably registered atlas are overlaid in green. The contours are part of an ongoing work discussed in Section 4.

(compare, for example, with [3]) or segmentation of the target images. The use of NMI simplifies the assumptions on the image content, and appears to contribute to the robustness to imaging constraints, such as noise, occlusions, and truncated view.

## 2 Method

### 2.1 Atlas Creation

Our anatomical atlas [12] is modeled after the work of Yao [5,4] with some modifications (the main differences are in the methods used to generate the shape mesh and to co-register the template to the individual subjects), and essentially following the *Active Shape Models* paradigm [6]. To create the atlas, we start by labeling a selected CT study – the *template* CT, denoted  $CT_0$  – for the anatomy of interest, i.e. the pelvis bones. From the labeled voxels, we create a template tetrahedral mesh [7], denoted  $S_0$ . Given a collection of *subject* CT studies,  $\{CT_j\}_{j=1}^N$ , we compute for each subject instance  $j$  a deformation that maps the voxels of  $CT_0$  to corresponding voxels of  $CT_j$  [8]. This deformation is applied to the vertices of  $S_0$  to obtain a subject shape  $S_j$ . Finally, all the shape instances are aligned using a similarity transformation, and modes of variation are extracted from this population using principal component analysis. A new shape instance can be generated from the atlas as

$$\hat{S} = \bar{S} + \sum_{i=1}^L w_i V_i \quad (1)$$

where  $\bar{S}$  is the “mean shape” obtained from aligning all the instance shapes;  $\{w_i\}$  are scalar weights; and  $\{V_i\}_{i=1}^L$  are the first  $L$  modes in the distribution of shapes.

Every tetrahedral cell in the atlas mesh is assigned with a CT density attribute, in the form of a 3<sup>rd</sup> order barycentric Bernstein polynomial. The polynomial coefficients are obtained by fitting to the CT intensity numbers in the

template study voxels covered by that cell. At this stage, we use the densities of a single subject in the atlas, though Yao performed a statistical analysis of the densities as well.

## 2.2 2D-3D Registration

For the registration experiments, we use Take software [9] to create DRR's of segmented CT datasets. For each target image, denoted  $I_k$ , the intrinsic and extrinsic camera parameters are given. In this paper's DRR's, the parameters are user-defined. In a clinical scenario, a calibrated C-arm and a form of tracking, e.g. optical markers or encoders, would be used to image the patient.

The registration algorithm creates DRR's of our atlas using a fast rendering algorithm [10]. For each camera pose  $k$ , the *normalized mutual information* similarity measure

$$NMI_k = (H(I_k) + H(DRR_k))/H(I_k, DRR_k) \quad (2)$$

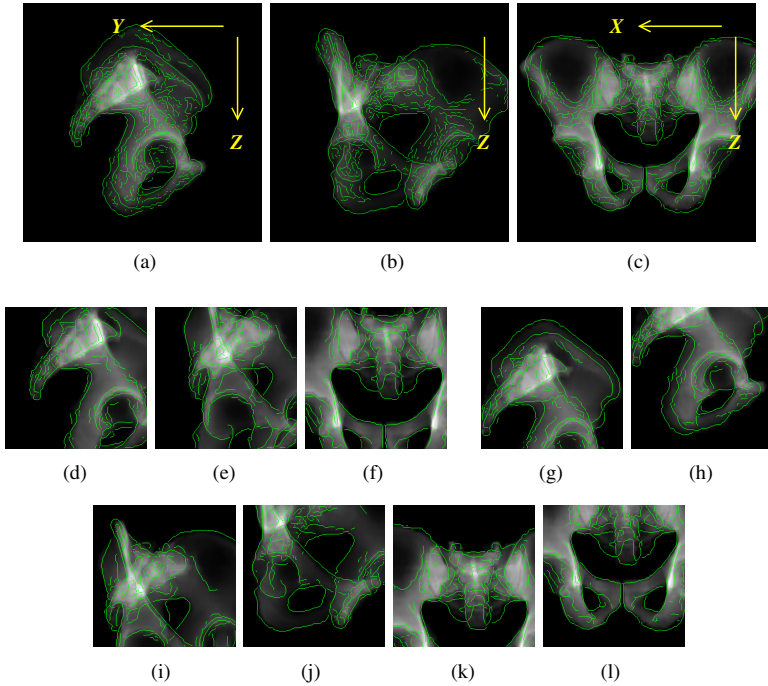
is computed (selected based on [11]), where  $H$  is the entropy of pixel intensity distribution (or joint distribution). The final similarity score for the entire set of projections is  $Sim = \sum_{k=1}^K NMI_k$ . Our method searches for the maximum of this score using the Nelder-Mead downhill-simplex algorithm, over the space spanned by the following parameters: translation ( $\mathbf{d}$ ), rotation ( $\mathbf{R}$ ), isotropic scale ( $s$ ), and shape mode weights  $\{w_i\}$ . The estimated shape has the form in Equation 1. The final, registered model is the outcome of applying a similarity transformation to  $\hat{S}$ :  $S^{reg} = s(\mathbf{R} \cdot \hat{S} + \mathbf{d})$ . We alternate subsets of parameters, such as a translation vector or rotation "Rodrigues" vector, search for the optimal value on each subset, and fix the result when searching the next subset.

## 2.3 Imaging Parameters

This paper examines two sizes of the field of view: 270 mm and 160 mm (see Fig. 2). These sizes relate to a "virtual" detector plane located at the isocenter of camera motion (typically inside the object), and reflect approximately the diameter of the visible portion of anatomy. The 270 mm FOV covers either the full or nearly-full anatomy of the pelvic bones. The 160 mm FOV is roughly comparable to a 9" C-arm image intensifier (about 216 mm) lying outside of the object (Fig. 1). The source-to-detector distance was set to 800 mm.

By our experience (and Yao's [5]), when a large field of view is available, two or three viewing directions, such as antero-posterior (AP), lateral, and oblique, are sufficient for a good registration. But when only a truncated view is available, more views may be required. To examine this, we run the registration using three, six, or eight truncated views, which cover increasing portions of the pelvis's anatomy.

Fig. 2 demonstrates the target images and the registration results for one subject. For the 270 mm images, (a)-(c), we rotated the camera in a circular arc about the the object's  $Z$  axis, and imaged at angles  $0^\circ$  (lateral view),  $45^\circ$



**Fig. 2.** Reference images and deformable registration results. The contours of a registered atlas are overlaid in green on the reference images. (a)-(c) FOV=270 mm. (d)-(l) FOV=160 mm. Images (d)-(f) are used in the three-view registration. Images (g)-(l) are used in the six-view registration. All of (d)-(l) except (f) are used in the eight-view registration. The axis directions of the model are shown in yellow.

(oblique view), and  $90^\circ$  (AP view). The model axis directions are illustrated in the Figure. For the 160 mm images, (d)-(l), we included rotation and translation in the camera trajectory. The translations included moving the camera 40 mm up and down the  $Z$  axis, and moving 30 mm up the  $Y$  axis. The view angles included  $0^\circ$ ,  $30^\circ$ ,  $45^\circ$ , and  $90^\circ$ . The images were selected to include features such as the proximal and distal ends of bones, the ilium, and the sacrum.

## 2.4 Accuracy Measure

Because of the deformable component, it is hard to define the “exactly correct” registration parameters, unless we use artificially created instances. Hence, in this paper, we measure the effective error by surface-to-surface distance between the registered shape of the atlas and a “ground truth” shape.

We performed leave-one-out tests by randomly selecting datasets from the study population. For each selected subject,  $j$ , we used the remaining datasets to create a shape atlas,  $A_j$ . The subject shape  $S_j$  (left out of the atlas) was regarded as the “ground truth”, and used as a segmentation mask over the CT

study  $CT_j$ . The masked subject dataset,  $CT_j^{Sgmt}$ , was passed as input to the DRR generator to create the target images  $\{I_k^{(j)}\}$ .

Next, the atlas  $A_j$  was deformably registered with  $\{I_k^{(j)}\}$  to create a result shape  $S_j^{reg}$ . Finally, the outer surface of  $S_j^{reg}$  was compared with the  $S_j$  by way of projection on the nearest neighbor: For each vertex  $\mathbf{v}_l$  on the surface of  $S_j^{reg}$ , we define  $\mathbf{s}_l$  as the nearest point on  $S_j$  ( $\mathbf{s}_l$  is not necessarily a vertex of  $S_j$ ). We defined  $\mathbf{e}_l = \mathbf{s}_l - \mathbf{v}_l$  as the error vector at the vertex  $\mathbf{v}_l$ , and computed statistics on  $\|\mathbf{e}_l\|$ .

For comparison, we estimated the “optimal” 3D-3D registration that our atlas may achieve for the subject  $j$  using *mode matching* [5]. Based on the correspondence of vertices between the atlas  $A_j$  and the shape  $S_j$ , we computed the deformable registration parameters that minimize the sum of squared distances between the corresponding vertices. This produced an “optimal” shape  $U_j$ , which we compared with  $S_j$  using the same surface-to-surface error metric as above.<sup>1</sup>

### 3 Results and Discussion

Our study population consisted of 110 CT scans of the pelvis, taken from prostate cancer patients. Eleven datasets were selected at random as experiment targets. They were resampled to a uniform voxel size of 1.875 mm cube before DRRs were generated. The mesh population was the output of the first pass of a bootstrapping experiment, presented separately [12]. All the registration experiments used the first  $L = 15$  shape variation modes.

In all the experiments, we started from an arbitrary initial guess of zero translation, rotation, and deformation magnitude. The final translation magnitudes were between 13 mm and 59 mm, in different directions; the final rotation magnitudes were between  $10^\circ$  and  $17^\circ$ , about different axes. Table 1 summarizes the surface distances of each registration experiment from the “ground truth” shape  $S_j$ .

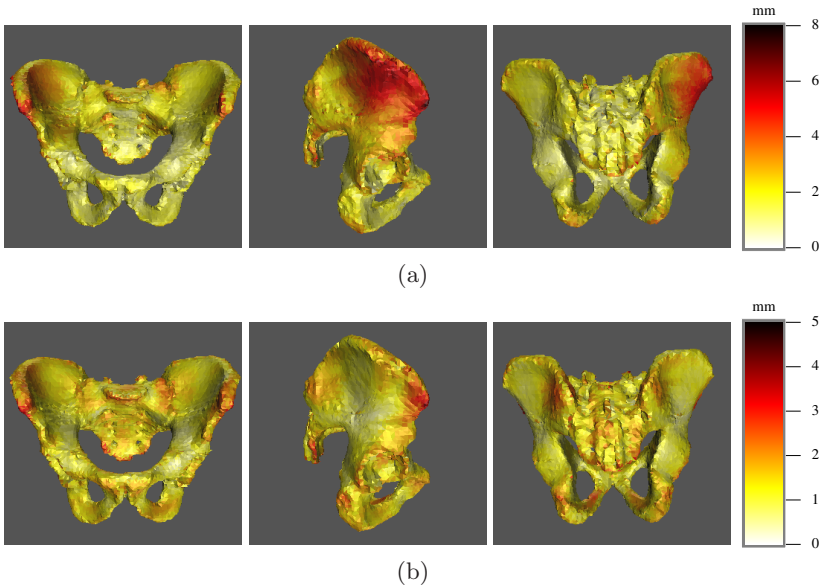
According to Table 1, there is a slight decline in the mean registration error of the truncated images when more views are used. This suggests diminishing returns on the number of truncated views. For some subjects (e.g. 10, 26, 44) the registration error with three truncated views than with six or eight. This may be due to the assumptions underlying our choice of objective function in relation to the evaluation criterion. The best 2D-3D registration results, in almost all cases, were achieved with the larger FOV of 270 mm. For comparison, the mean error of the 3D-3D registration was about 1.4 mm.

Fig. 3 shows the distribution of registration errors on the surface of a representative bone, computed as a mean per vertex over the population of Table 1. We compare the results of (a) 2D-3D registration using eight views, and (b) 3D-3D

<sup>1</sup> In a previous series of experiments, we compared  $U_j$  with a hand-segmented shape. Due to shape uncertainties in both  $S_j$  and the hand segmentation, we decided to use  $S_j$  as the ground truth.

**Table 1.** Summary of surface distances (in mm) between the registration results and subject-specific (“ground truth”) shapes

Sub- ject	FOV=270 mm 3 views		FOV=160 mm 3 views		FOV=160 mm 6 views		FOV=160 mm 8 views		3D Mode Matching	
	Mean	Max	Mean	Max	Mean	Max	Mean	Max	Mean	Max
10	1.94	12.06	2.05	10.77	2.20	11.00	2.09	11.48	1.59	6.97
26	2.00	8.31	2.02	13.78	2.10	12.45	2.10	12.83	1.59	8.13
31	1.77	7.93	2.37	17.48	1.86	10.53	1.72	8.87	1.63	10.37
41	1.71	9.35	1.85	7.91	1.75	8.04	1.82	8.25	1.02	5.68
43	3.44	17.83	3.42	20.59	3.26	19.51	2.98	17.68	1.08	6.00
44	1.92	10.26	2.34	13.62	2.37	14.02	2.40	13.45	1.50	7.73
58	2.23	9.08	2.43	13.15	2.31	12.81	2.48	11.59	1.70	7.67
60	1.70	9.01	1.58	10.13	1.65	7.40	1.67	7.71	1.49	6.09
66	1.59	8.80	1.50	7.73	1.74	9.75	1.81	9.21	1.18	6.14
68	1.48	7.05	1.85	7.17	1.66	7.49	1.71	8.61	1.27	8.48
76	1.92	16.34	2.48	13.74	2.02	13.66	2.13	14.68	1.52	11.82
Mean	1.97	10.55	2.17	12.37	2.08	11.51	2.08	11.31	1.42	7.73

**Fig. 3.** Distribution of surface registration errors on the pelvis. (a) 2D-3D registration, FOV=160 mm, 8 views. (b) Mode-matching 3D-3D registration. The error is the mean surface distance per vertex over the studied population. The color scales are adjusted to highlight each registration’s individual distribution.

mode matching. The color scales in the two parts are adjusted to highlight the distribution, not to compare the errors. For the 2D-3D registration, the error distribution is asymmetric, with a significant tilt to the right hemipelvis. This may be a side-effect of the selection of imaging trajectory (Fig. 2), which preferred one side of the subject over another. In the mode matching registration, some areas of the bone have higher errors than others, for example: the anterior spine of the ilium, the vicinity of the obturator foramen, and the tip of the sacrum. Some of these “common” errors also appear in the 2D-3D registration.

They may be regions of higher variation among the studied datasets, where the statistical modes cannot approximate the shape as well as in other locations.

Overall, the spatial distribution of errors for both the 2D-3D and 3D-3D appears, visually, quite similar, with the exception of the asymmetry. In future research, we are planning to see if the asymmetry can be reduced by changing the imaging trajectory.

## 4 Conclusion and Future Work

We have shown that a 2D-3D deformable registration of a shape and density atlas of the pelvis to x-ray images, based on normalized mutual information image similarity, can be robust to image truncation. A typical expected surface registration error is about 2 mm, compared with an “ideal” registration error of about 1.4 mm. A larger FOV can usually yield a better registration accuracy, and the use of more truncated views can improve the accuracy to some degree. This result has important applications in recovering a patient’s anatomical shape in clinical scenarios, when small FOV C-arms are used to image the patient.

While the results we present are encouraging, applying the method on real fluoroscopic images is still a challenge. In Fig. 1, we show the result of a deformable 2D-3D registration as green contours of the registered model overlaid on x-ray target images of a cadaveric bone specimen. Qualitatively, we can observe that our model aligns with the bone as a whole, but does not align very well with specific features, such as the obturator foramen and the acetabular rim. It may be that this particular specimen is not a good representative of the studied population, which would decrease the overall quality of the registration. In addition, the MI similarity measure did not produce a good registration with these images, possibly because of different imaging characteristics of dry bone compared with live patient DRR. The results shown were obtained by maximizing the *structural similarity index* (SSIM) [13].

Human subject images also contain other anatomical detail than the pelvis bones: femurs, spine, and other organ tissues. These affect the image quality significantly, adding clutter and reducing contrast. In initial experiments we conducted, we included a crude “soft tissue” model and used the SSIM index as to achieve a coarse-level registration. Again, the MI measure did not perform as well.

It has been suggested (e.g. in [11]) that NMI as an image similarity measure may be more robust to FOV truncation than mutual information. In a previous series of experiments, we used MI as the objective function. While the rate of successful registrations was higher with NMI, the final accuracy seems roughly the same for both functions. In some of the previous experiments, the asymmetry of error was not as pronounced as in the results here.

We would like to thank the anonymous reviewers of this paper for their suggestions, which led to the new experiments in this paper. Further study of different similarity measures for the registration with simulated and real images, and their applicability, is continuing. We are also continuing a search for representative cadaver bones which match well with the training population.

## Acknowledgments

We would like to thank Lotta Ellingsen and Pauline Pelletier for assisting in creating the atlas. The pelvis datasets were given to us by Dr. Ted DeWeese and Dr. Lee Myers from Johns Hopkins Radiation Oncology. This work was supported in part by NSF ERC Grant EEC9731748, and by NIH/NIBIB research grant R21-EB003616.

## References

1. Sadowsky, O., Ramamurthi, K., Ellingsen, L.M., Chintalapani, G., Prince, J.L., Taylor, R.H.: Atlas-assisted tomography: Registration of a deformable atlas to compensate for limited-angle cone-beam trajectory. In: IEEE International Symposium on Biomedical Imaging (ISBI), IEEE Computer Society Press, Los Alamitos (2006)
2. Fleute, M., Lavallée, S.: Nonrigid 3-d/2-d registration of images using statistical models. In: Taylor, C., Colchester, A. (eds.) MICCAI 1999. LNCS, vol. 1679, pp. 138–147. Springer, Heidelberg (1999)
3. Zheng, G., Ballaster, M.A.G., Styner, M., Nolte, L.P.: Reconstruction of patient-specific 3d bone surface from 2d calibrated fluoroscopic images and point distribution model. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 25–32. Springer, Heidelberg (2006)
4. Yao, J., Taylor, R.H.: Assessing accuracy factors in deformable 2d/3d medical image registration using a statistical pelvis model. In: Ninth Int. Conference on Computer Vision, Nice, France, pp. 1329–1334 (2003)
5. Yao, J.: A statistical bone density atlas and deformable medical image registration. PhD thesis, Johns Hopkins University (2002)
6. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Activeshape models—their training and application. *Computer Vision and Image Understanding* 6(1), 38–59 (1995)
7. Mohamed, A., Davatzikos, C.: An approach to 3d finite element mesh generation from segmented medical images. In: IEEE International Symposium on Biomedical Imaging (ISBI), IEEE Computer Society Press, Los Alamitos (2004)
8. Ellingsen, L.M., Prince, J.L.: Mjolnir: Deformable image registration using feature diffusion. In: Proceedings of the SPIE Medical Imaging Conference. vol. 6144, pp. 329–337 (2006)
9. Muller-Merbach, J.: Simulation of x-ray projections for experimental 3d tomography. Technical report, Image Processing Laboratory Department of Electrical Engineering Linköping University, SE-581 83, Sweden (1996)
10. Sadowsky, O., Cohen, J.D., Taylor, R.H.: Projected tetrahedra revisited: A barycentric formulation applied to digital radiograph reconstruction using higher-order attenuation functions. *IEEE Transactions on Visualization and Computer Graphics (TVCG)* 12(4), 461–473 (2006)
11. Studholme, C., Hill, D., Hawkes, D.: An overlap invariant entropy measure of 3d medical image alignment. *Pattern Recognition*, 71–86 (1999)
12. Chintalapani, G., Ellingsen, L.M., Sadowsky, O., Prince, J.L., Taylor, R.H.: Statistical atlases of bone anatomy: Construction, iterative improvement and validation. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007. LNCS, vol. 4791, pp. 499–506. Springer, Heidelberg (2007)
13. Zhou, W., Bovik, A.C., Seikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing* 13(4), 600–612 (2004)