Towards Automated Spine Mobility Quantification: A Locally Rigid CT to X-ray Registration Framework

David Drobny[1,2](✉) , Marta Ranzini[2] , Amanda Isaac[2] ,
Tom Vercauteren[2] , Sébastien Ourselin[2] , David Choi[3] ,
and Marc Modat[2]

[1] Wellcome/EPSRC Centre for Interventional and Surgical Sciences,
University College London, Gower Street, London WC1E 6BT, UK
`d.drobny.17@ucl.ac.uk`
[2] School of Biomedical Engineering & Imaging Sciences, King's College London,
King's Health Partners, St Thomas' Hospital, London SE1 7EH, UK
`{marta.ranzini,a.isaac,tom.vercauteren,`
`sebastien.ourselin,marc.modat}@kcl.ac.uk`
[3] University College Hospital, London NW1 2BU, UK
`d.choi@ucl.ac.uk`

Abstract. Different pathologies of the vertebral column, such as scoliosis, require quantification of the mobility of individual vertebrae or of curves of the spine for treatment planning. Without the necessary mobility, vertebrae can not be safely re-positioned and fused. The current clinical workflow consists of radiologists or surgeons estimating angular differences of neighbouring vertebrae from different x-ray images. This procedure is time consuming and prone to inaccuracy. The proposed method automates this quantification by deforming a CT image in a physiologically reasonable way and matching it to the x-ray images of interest. We propose a proof of concept evaluation on synthetic data. The automatic and quantitative analysis enables reproducible results independent of the investigator.

Keywords: Spine · Vertebra · Mobility quantification · 3D-2D registration · Volume-projection registration

1 Introduction

Spine mobility quantification (SMQ) describes the measurement of angles between vertebrae and their change between different positions, in order to evaluate the mobility of individual vertebrae. It is, for example, performed for scoliosis patients to determine whether vertebrae have the required mobility to be

realigned and fused to improve the curvature of the spine. SMQ is typically evaluated manually based on x-ray images. The clinician measures the angles between neighbouring vertebrae in x-ray images acquired in two opposite and extreme positions, for example flexion and extension. The change of the angular difference from one position to the other informs about the mobility of those vertebrae. Manual measurements are subjective and prone to errors, especially when a 3D movement is assessed by rotations within one plane only.

In many clinical cases a CT image is available which could be used for 3D-2D registration. This could help automating SMQ by aligning the CT image to the X-ray target pose and deriving the SMQ values from the applied deformation.

Registration of CT volumetric images to x-ray projection images is a relevant task also for several other clinical applications, such as the automated localisation of an interventional image with respect to the pre-operative CT data. Specifically, this approach can be used in image-guided spine surgery where the target vertebra has to be identified reliably [3]. According to Mody *et al.* more than 0.3% of spine surgery procedures have been affected by wrong-level errors which can have a severe negative impact on the patient [5]. This demonstrates the difficulty of vertebra identification even for experienced surgeons. After robust identification, a precise registration further enables navigated surgery which typically makes use of intra-operative CT (iCT) imaging. A reliable and accurate 3D-2D registration for interventional x-ray images could thus decrease errors in surgical procedures, improve accuracy of image-guided interventions or otherwise maintain accuracy while decreasing radiation exposure to patients and clinicians, compared to using iCT imaging.

Previous work in volume-projection registration (of the spine) predominantly considers rigid registration frameworks [3,4,8,9,11]. This reduces the complexity of this ill-posed problem but limits the potential accuracy and applications, as it cannot capture local deformations of soft tissue related to different positions of the subject. Non-rigid advances have been mostly focused on other body parts, for example, on lung movement [12], and might be unrealistic in rigid anatomical structures such as bones.

To overcome the limitations of existing methods used for spine applications, we suggest a locally rigid registration framework based on the works of Arsigny *et al.* [1]. The proposed framework extends the poly-rigid transformation model to ensure the preservation of local rigidity during articulated movement and relies on a novel regulariser to enforce physiologically reasonable transformations. It enables better matching of images acquired with different patient position, e.g. flexion and extension, or pre-operative supine and intra-operative prone. Furthermore, the parametrisation of our model can be directly used to quantify relative positions of the vertebrae, for example for SMQ.

We demonstrate a proof of concept for spine mobility quantification based on simulated data using our registration framework. Rotation angles can be determined with high precision which motivates further development of this approach.

2   Methods

The proposed registration framework enables the recovery of spine movement via 3D-2D registration and the quantification of differential movement between neighbouring vertebrae. The underlying transformation model uses the poly-rigid approach which combines n local sets of rigid parameters $\boldsymbol{l} = \{\boldsymbol{l}_1, .., \boldsymbol{l}_n\}$ into a single smooth deformation field [1]. It requires a set of non-overlapping regions R_i which, in our case, are given by a binary mask for each vertebra i. Based on the regions, individual weight maps w_i are computed that define the local influence of each rigid transformation T_i.

2.1   Registration Framework

The registration algorithm finds an optimal choice of parameters $\boldsymbol{p} = (\boldsymbol{g}, \boldsymbol{l})$, with global affine transformation parameters $\boldsymbol{g}$ and local rigid transformation parameters $\boldsymbol{l}$, that describes the relationship between the reference x-ray image $\mathcal{X}$ and the CT volume $\mathcal{V}$ via the transformation $\mathcal{T}$. The transformation $\mathcal{T}$ combines the poly-rigid transformation $\mathcal{T}_{pr}(x, \boldsymbol{l})$ with a global affine transformation $\mathcal{A}(x, \boldsymbol{g})$:

$$\mathcal{T}(x, \boldsymbol{p}) = \mathcal{A}(x, \boldsymbol{g}) \circ \mathcal{T}_{pr}(x, \boldsymbol{l}). \tag{1}$$

To enable the comparison of 3D and 2D images, a projection $\mathcal{P}$ is used:

$$\mathcal{P} : \mathbb{R}^3 \mapsto \mathbb{R}^2. \tag{2}$$

The distance measure $\mathcal{D}$ describes the dissimilarity of the images and is combined with a regulariser $\mathcal{R}$ to form the objective function $\mathcal{F}$:

$$\mathcal{F}[\mathcal{X}, \mathcal{V}, \boldsymbol{p}] = \mathcal{D}\Big[\mathcal{X}, \mathcal{P}\big[\mathcal{T}(\mathcal{V}, \boldsymbol{p})\big]\Big] + \alpha \, \mathcal{R}(\boldsymbol{p}). \tag{3}$$

The parameter α controls the ratio between both terms and thus adjusts the influence of the regularisation. Minimisation of $\mathcal{D}$ makes the appearance of the images more similar while $\mathcal{R}$ promotes a anatomically realistic transformation.

The numerical optimisation of $\mathcal{F}$ is performed iteratively via a gradient descent solver. In each step a subset of transformation parameters is evaluated, global rigid followed by local rigid parameters for each region, one at a time. Optimising only a subset of parameters reduces the complexity of the partial derivatives and empirically leads to a more stable behaviour of the model.

For increased robustness to local minima and to increase the capture range, a multi-resolution approach is applied that performs the optimisation on multiple levels – from coarse to fine. The framework also enables 3D-3D registration of a volume $\mathcal{V}$ to a reference volume $\mathcal{V}_{ref}$. In this case the projection step is omitted and the objective function is therefore:

$$\mathcal{F}_{3D-3D}[\mathcal{V}_{ref}, \mathcal{V}, \boldsymbol{p}] = \mathcal{D}\big[\mathcal{V}_{ref}, \mathcal{T}(\mathcal{V}, \boldsymbol{p})\big] + \alpha \, \mathcal{R}(\boldsymbol{p}). \tag{4}$$

2.2   Poly-rigid Transformation Model

The poly-rigid transformation model is based on a set of local rigid transformation parameters l_i with associated weight maps w_i defining its local influence for each region i. To combine the different local rigid transformations T_i into a single one, an ad hoc solution is the weighted average of the transformations' displacement fields. This does not preserve local behaviours and can lead to folding, i.e. breaking of the topology, which is to be avoided in most medical image processing.

Arsigny *et al.* proposed a method that provides a combined transformation that is invertible and thus does not lead to folding [1]. The local rigid transformations T_i are by definition homomorphisms, i.e. invertible. The average of infinitesimal small downscaled homomorphisms is also invertible as the displacement gets close to zero. The composition of homomorphisms is also a homomorphism so that the full scale transformation recovered by composition of the infinitesimal transformations is a homomorphism as well. The transformation $T(t, x)$ is parametrised as a velocity field and integrated over time $t \in [0, 1]$. At time $t = 0$ the transformation is the identity and at $t = 1$ it is $\mathcal{T}_{pr}$.

The weighted average of the downscaled transformations is computed as

$$T\left(\frac{1}{m}, x\right) = \sum_{i=1}^{n} w_i(x) T_i(x)^{\frac{1}{m}}, \tag{5}$$

where $T^{\frac{1}{m}}$ describes the m-th matrix square root of the rigid transformation matrix T. m has to be chosen such that the downscaled velocity field is close enough to zero [1]. This average function needs to be upscaled to get the wanted poly-rigid transformation $\mathcal{T}_{pr}$:

$$\mathcal{T}_{pr}(x) = T(1, x) = T\left(\frac{1}{m}, x\right)^{m}. \tag{6}$$

Arsigny *et al.* use an efficient scaling-and-squaring scheme to integrate the final transformation [1]. However, to maintain the desired rigid properties for each region, the weights need to be updated after each integration step as suggested by Porras *et al.* [6]. Therefore we use the following Euler integration scheme:
For $j = 2, .., m$

$$w_i^{(j)}(x) = w_i\left(T^{(j-1)}\left(\frac{1}{m}, x\right)\right)$$
$$T^{(j)}\left(\frac{1}{m}, x\right) = \sum_{i=1}^{n} w_i^{(j)}(x) T_i^{\frac{1}{m}}(x) \tag{7}$$
$$T\left(\frac{j}{m}, x\right) = T\left(\frac{j-1}{m}, x\right) \circ T^{(j)}\left(\frac{1}{m}, x\right)$$

The superscript $\cdot^{(j)}$ indicates that the corresponding term is updated in step j.
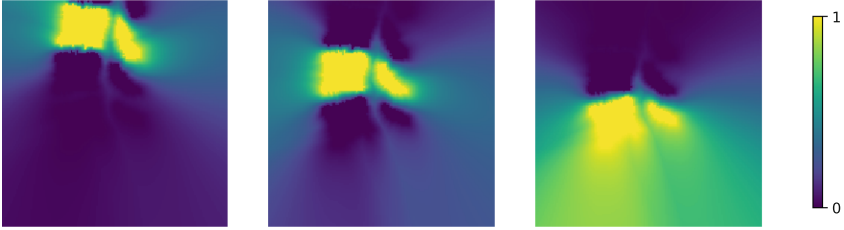
Fig. 1. Weight maps of the three lowest lumbar vertebrae. A vertebra in between other vertebrae has mostly influence perpendicular to the spine. The bottom vertebra (in this experiment the sacrum is not modelled as a rigid region) dominates the weights in the lower part of the image as the other vertebrae are further away and affect it less.

2.3   Weight Map Computation

Poly-rigid transformation models require weight maps w_i which define the spatial influence of multiple transformations T_i. Our model's weight computation is based on non-overlapping regions R_i, that represent a spine segmentation.

In the region based weight computation of Porras *et al.* [6], each weight w approximates the following values: $w(x) = 1$ if x is within the region, $0 < w(x) < 1$ otherwise. This implies that each region has a global influence and the rigid properties of a region are not preserved anywhere. To guarantee the rigid transformation of a region, weights with following properties are required.

$$w_i(x) = \begin{cases} 1, & \text{if } x \in R_i \\ 0, & \text{if } x \in R_k, \ k \in \{1,..,n\} \setminus i \\ v, v \in \mathbb{R}, \ 0 < v < 1, & \text{elsewhere} \end{cases} \tag{8}$$

The computation of the weight map can be split into three steps. First, creating a smooth weight map that is 1 within a region and decreases outside:

$$w_i^a(x) = \frac{1}{\exp\left(s * EDT_i(x)\right)}, \tag{9}$$

where EDT denotes the Euclidean Distance Transform of the binary mask of region R_i in mm (i.e. the distance to the nearest voxel of this region). The slope constant s controls how quickly the weight decays and is fixed for all regions throughout the computations.

To set the influence of all other regions R_k to 0 within region R_i, we multiply each region's weight by the complementary weight of all other regions:

$$w_i^b(x) = w_i^a(x) \prod_{\substack{k=1 \\ k \neq i}}^{n} \left(1 - w_k^a(x)\right). \tag{10}$$

Normalising the weights w_i^b to recover the maximal value of 1 within each region yields the final weight maps w_i that fulfil Eq. (8):

$$w_i(x) = \frac{w_i^b(x)}{\sum\limits_{k=1}^{n} w_k^b(x)}. \tag{11}$$

Figure 1 shows example weight maps computed for a CT slice containing the lowest lumbar vertebrae. This example visualises that within a region (i.e. a vertebra) the weight is one, smoothly decaying when moving away from it, and zero in each other region.

2.4   Regularisation

In order to guide the behaviour of the transformation model in a way that is closer to anatomical spine movement, we introduce a set of regularisation terms. These soft constraints penalise parameters that are in unexpected value ranges. The regulariser must have little influence while the parameters are within the expected value range and steeply increase when outside. Therefore the general form of each soft constraint C used is:

$$C(q) = \gamma \, |q - b|^c \,, \tag{12}$$

where b is the absolute upper bound for the value q and the exponent c controls how steep the penalty term increases.

Firstly, the parameters of the global affine registration $\boldsymbol{g}$ are limited as we assume the general direction of x-ray acquisition is known and thus, the volume should not translate or rotate too much away from the expected view.

$$\mathcal{R}^g(\boldsymbol{g}) = C(|\boldsymbol{g}^{trans}|) \; + \; C(|\boldsymbol{g}^{rot}|) \tag{13}$$

Secondly, the local rigid parameters $\boldsymbol{l}$ are constrained via the parameter difference of pairs of neighbouring vertebrae, separate for rotation and translation:

$$\mathcal{R}^l(\boldsymbol{l}) = \sum_{i=2}^{n} C(|\boldsymbol{l}_i^{trans} - \boldsymbol{l}_{i-1}^{trans}|) \; + \; C(|\boldsymbol{l}_i^{rot} - \boldsymbol{l}_{i-1}^{rot}|) \tag{14}$$

This way neighbouring vertebrae cannot translate or rotate too far from their initial relative position. This constraint guarantees that the spine moves consistently and retains its integrity.

An unregularised optimisation step might lead to local rigid parameters that would move regions into another. The overall poly-rigid transformation avoids this kind of folding but compromises the rigidity of the underlying transformations to achieve this. The third regulariser $\mathcal{R}^o$ is used to discourage such optimisation steps by penalising transformations that cause voxel overlap in order to achieve diffeomorphic transformations while preserving the local rigidity.

$$\mathcal{R}^o(\boldsymbol{p}) = C\left(\left|\{x | x \in R_u \cap R_v, u, v \in \{1, .., n\}, u \neq v\}\right|\right) \tag{15}$$

The sum of all individual terms leads to the overall regulariser $\mathcal{R}$ as in Eq. (3):

$$\mathcal{R}(\boldsymbol{p}) = \mathcal{R}^g(\boldsymbol{g}) + \mathcal{R}^l(\boldsymbol{l}) + \mathcal{R}^o(\boldsymbol{p}) \tag{16}$$

3   Experiments

Data from the Spineweb database which consist of CT images and corresponding vertebra segmentation have been used for the experiments described in this work [13]. A CT image is cropped to the lumbar area including six vertebrae above the sacrum and resampled to a 1 mm isotropic image of $162 \times 162 \times 312$ voxel. The rigid parameters of the vertebrae are user-defined to simulate two movements: flexion and extension. The original CT image is considered as the starting position and angles are interpreted as differential angles to this position. The image axes x, y, and z correspond to the lateral, anterior-posterior, and superior-inferior patient axis, respectively. X-ray projections are simulated from the deformed CT images in lateral direction (along the x-axis) and are used as the reference images $\mathcal{X}$ while the initial CT volume is then used as moving image $\mathcal{V}$. The projection method $\mathcal{P}$ used for the synthetic experiments is a basic averaging of intensity values along a given direction, which can be easily replaced by other approaches to account for more clinically realistic scenarios, like those presented by Unberath *et al.* [7].

Table 1 shows the rotation angles chosen for each vertebra in the flexion and extension simulation. For this experiment only rotations around the x- and y-axis are considered, as those are the angles clinically measured for SMQ and rotations around z-axis are minor in the lumbar spine. The values are in the typical range observed in clinical practice, for example as reported by Wilke *et al.* [10]. Figure 2 shows the flexion, extension and initial position of the synthetic data as 3D rendered visualisations. The vertebrae are deformed rigidly while the surrounding soft tissue, like the arteries, are non-linearly deformable.

In our experiment, images in two different poses are target of the registration: flexion and extension. This resembles the application of SMQ where usually two opposite extremal positions are compared. For each pose, two experiments are evaluated: (1) 3D-3D registration—matching the original CT $\mathcal{V}$ to the synthetic CT image $\mathcal{V}_{ref}$ and (2) 3D-2D registration—matching the CT image $\mathcal{V}$ to the synthetic x-ray image $\mathcal{X}$. Two metrics are used to evaluate the registration results: the error of the estimated angle difference and the average error of the recovered displacement field at the location of the vertebrae. The displacement field error gives information about whether the vertebrae are in the right position while the angle errors describe how well the orientation of the vertebrae was recovered, which is the focus of SMQ.

For performance reasons, the poly-rigid transformation computed by Eq. (7) is used only as a post processing step and we use the direct weighted average during the optimisation, as was suggested by Commowick *et al.* [2]. We used the mean of squared difference as similarity measure $\mathcal{D}$ during our experiments. The parameters are either motivated by clinical movement ranges or chosen empirically: four levels of multi-resolution approach, with the finest resolution of 1 mm, the weight constant $s = 0.05$, the constraint scaling parameter $\alpha = 100$. Referring to Eqs. 12 and 16, the first two constraint functions have the exponent $c = 6$, scaling $\gamma = 1$ and the boundaries b of 20 mm, 30° for the global rigid
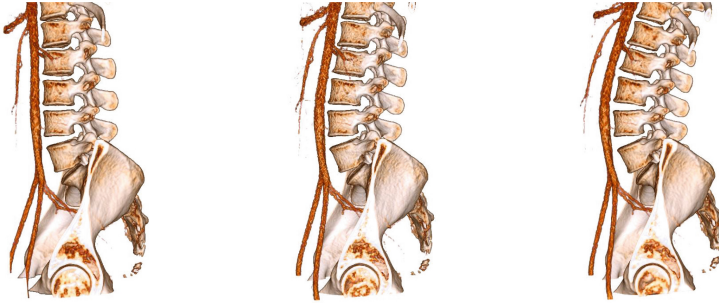
Fig. 2. 3D render of the original CT scan (centre), the synthetic CT scan in flexion (left) and extension (right) (see Supplementary material for animated figure).

Table 1. Rotation angles in degree per vertebra used to generate the synthetic images.

vertebra	flexion			extension		
	x	y	z	x	y	z
L1	− 11.00	−5.00	0.00	+10.00	+5.00	0.00
L2	− 8.50	−4.00	0.00	+7.50	+4.00	0.00
L3	− 5.00	−2.00	0.00	+4.00	+2.00	0.00
L4	0.00	0.00	0.00	0.00	0.00	0.00
L5	+ 4.00	+1.00	0.00	−3.50	−1.00	0.00

parameters and 8 mm, 7° for the local differences, the overlap constraint has $b = c = 1$ and $\gamma = 0.001$.

4   Results

Our experiments show that with both 3D-3D as well as 3D-2D registration, the parameters of the synthetic images can be recovered. In Table 2 the quantitative results of all experiments are summarised. Using the 3D-3D registration higher accuracy was achieved compared to the 3D-2D cases which is due to the higher information retained in a volume compared to its projection. In the 3D-3D experiments, the average of the mean absolute angle error was reduced from the initial 2.47° to 0.12° and from 2.70° to 0.16° for extension and flexion respectively. In the 3D-2D case, the respective average errors are 0.59° and 0.42°. The deformation field error (DFE) is also reported as a measure of registration accuracy in Table 2. The DFE confirms the higher accuracy of the 3D-3D approach compared to the 3D-2D registration.

For the spine mobility quantification the angular difference between the two positions for each vertebra is of interest. This can be directly derived from the individual registration results as the difference of the respective angles.

Table 2. For each experiment, this table shows the mean absolute angle error [in degrees] per rotation axis, averaged over all vertebrae, as well as combined into a single average error, the mean and maximum displacement field errors (DFE) [in mm] before and after registration. The DFEs are evaluated within the spine mask only.

	flexion			extension		
	initial	3D-3D	3D-2D	initial	3D-3D	3D-2D
error x	5.70	0.11	0.45	5.00	0.06	0.37
error y	2.40	0.06	0.55	2.40	0.07	0.99
error z	0.00	0.31	0.30	0.00	0.23	0.41
average error	2.70	0.16	0.43	2.47	0.12	0.59
average DFE	4.37	0.05	0.52	3.11	0.10	1.71
max DFE	24.32	0.50	2.06	16.13	2.49	11.31

5 Discussion

The non-rigid alignment of a volumetric to a projection image is challenging because several degrees of freedom affecting the 3D volume have to be recovered by only comparing 2D images. We presented a framework based on a locally rigid transformation model that enables 3D-3D and 3D-2D registration of spine images. These synthetic experiments show that both 3D-3D and 3D-2D registration can be recovered, demonstrating that the registration with our transformation model can lead to clinically useful results for spine mobility quantification.

This framework offers the full processing pipeline needed for different applications and is flexible to adaptations required by the use of real clinical data. We also introduced simple constraints to guide a consistent movement of the spine while optimising individual vertebrae parameters iteratively. Our experiments are limited by only considering synthetic x-ray images generated from deformed CT scans. Clinical x-ray images are more challenging to register and thus need specialised image similarity measures. Future work will thus focus on the identification of a measure suitable for clinical data, and on the benchmarking of the proposed approach against the current clinical practice on real patient data. Furthermore we will use a higher level parametrisation of the spine shape, e.g. using principal components of a spine shape model, to optimise the position of multiple vertebrae simultaneously in a consistent way. This will be helpful to capture larger deformations especially in an extended field of view. Such a model could also be used to extract population statistics like the principal modes of spine variability and movement patterns. As the proposed framework proved effective also for 3D-3D registration, it will be tested on further clinical applications, such as spine CT-MRI registration. Finally, as this pipeline requires vertebra segmentations, automated segmentation techniques (*e.g.* using deep learning approaches) will be explored to complete the registration framework.

References

1. Arsigny, V., Commowick, O., Ayache, N., Pennec, X.: A fast and log-euclidean polyaffine framework for locally linear registration. J. Math. Imag. Vis. **33**(2), 222–238 (2009)
2. Commowick, O., et al.: An efficient locally affine framework for the smooth registration of anatomical structures. Med. Image Anal. **12**(4), 427–441 (2008)
3. De Silva, T., et al.: 3D–2D image registration for target localization in spine surgery: investigation of similarity metrics providing robustness to content mismatch. Phys. Med. Biol. **61**(8), 3009–3025 (2016)
4. Duménil, A., Kaladji, A., Castro, M., Göksu, C., Lucas, A., Haigron, P.: A versatile intensity-based 3D/2D rigid registration compatible with mobile C-arm for endovascular treatment of abdominal aortic aneurysm. Int. J. Comput. Assisted Radiol. Surg. **11**(9), 1713–1729 (2016)
5. Mody, M.G., Nourbakhsh, A., Stahl, D.L., Gibbs, M., Alfawareh, M., Garges, K.J.: The prevalence of wrong level surgery among spine surgeons. Spine **33**(2), 194–198 (2008)
6. Porras, A.R., et al.: Locally affine diffeomorphic surface registration and its application to surgical planning of fronto-orbital advancement. IEEE Trans. Med. Imag. **37**(7), 1690–1700 (2018)
7. Unberath, M., et al.: DeepDRR – a catalyst for machine learning in fluoroscopy-guided procedures. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11073, pp. 98–106. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00937-3_12
8. Varnavas, A., Carrell, T., Penney, G.: Fully automated 2D–3D registration and verification. Med. Image Anal. **26**(1), 108–119 (2015)
9. Wang, L., Gao, X., Zhang, R., Xia, W.: A comparison of two novel similarity measures based on mutual information in 2D/3D image registration. In: ICMIPE 2013 - Proceedings of 2013 IEEE International Conference on Medical Imaging Physics and Engineering, pp. 215–218. IEEE, October 2013
10. Wilke, H.J., Schmidt, R., Richter, M., Schmoelz, W., Reichel, H., Cakir, B.: The role of prosthesis design on segmental biomechanics semi-constrained versus unconstrained prostheses and anterior versus posterior centre of rotation. Eur. Spine J. **21**(SUPPL. 5), 577–584 (2012)
11. Wu, J., Kim, M., Peters, J., Chung, H., Samant, S.S.: Evaluation of similarity measures for use in the intensity-based rigid 2D–3D registration for patient positioning in radiotherapy. Med. Phys. **36**(12), 5391–5403 (2009)

12. Xia, W., Jin, Q., Ni, C., Wang, Y., Gao, X.: Thorax x-ray and CT interventional dataset for nonrigid 2D/3D image registration evaluation. Med. Phys. **45**(11), 5343–5351 (2018)
13. Yao, J., Burns, J.E., Munoz, H., Summers, R.M.: Detection of vertebral body fractures based on cortical shell unwrapping. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012. LNCS, vol. 7512, pp. 509–516. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33454-2_63