Personalising cross-population respiratory motion models using anatomical features

Devis Peressutti
devis.peressutti@kcl.ac.uk
Graeme P. Penney
graeme.penney@kcl.ac.uk
Christoph Kolbitsch
christoph.kolbitsch@kcl.ac.uk
Andrew P. King
andrew.king@kcl.ac.uk

Division of Imaging Sciences and
Biomedical Engineering,
King's College London, U.K.

Abstract

Subject-specific motion models have been proposed to address the problem of respiratory motion in image acquisition and image-guided interventions, but the need for a dynamic calibration scan to form the model can interrupt the clinical workflow. Cross-population models require no such calibration scan but lack the accuracy of subject-specific models. To address these problems, we propose a novel personalisation method for cross-population respiratory motion models. Unlike previous approaches, our method selects a subset of the population sample that is more likely to have similar respiratory motion to that of a new subject. The selection is based on anatomical features and therefore exploits inter-subject variability in motion to improve the accuracy of the resulting model. We present results on cardiac respiratory motion using a sample of 23 MRI datasets from healthy volunteers. Results show improvements in the median/95^{th} quantile of the motion estimation error of 20/17.2% compared to a standard cross-population model and accuracy comparable to subject-specific models for some subjects.

1 Introduction

Respiratory motion currently limits the accuracy of image-guided interventions applied to organs in the chest and abdomen, causing misalignments between the static images used for guidance and the moving anatomy. A similar problem exists in image acquisition where respiratory motion can cause artefacts in acquired images. As described in [6], subject-specific respiratory motion models represent a promising solution. Motion models describe the relationship between the motion of the anatomy and some measurable surrogate data. When forming the model, the surrogate data are acquired contemporary to dynamic calibration images depicting the respiratory motion of the anatomy, and the motion is then modelled as a function of the surrogate data. During model application, only the surrogate data are acquired, and the model estimates the motion given the current surrogate data [6].

The dynamic calibration scan used to build the subject-specific model is typically acquired using Computed Tomography (CT) or Magnetic Resonance Imaging (MRI), depending on the application. However, the calibration scan is often impractical or even impossible to acquire, due to dose issues, high cost, and patient considerations, such as bariatric

patients or patients with MRI-incompatible implants. To overcome these limitations, cross-population models have been proposed for the lungs [2, 5] and liver [8, 9]. These models require no dynamic calibration scan and are formed from data acquired from different subjects, averaging out the inter-subject variation in motion [6]. Typically, the cross-population average motion model is personalised to an individual by registering a static population anatomy image to a corresponding image of the new subject, and transforming the motion model accordingly. However, since respiratory motion can differ dramatically between subjects, cross-population models are currently not as accurate as subject-specific models. In [9] a technique was proposed for more selective personalisation based on surrogate similarity for the purpose of making more accurate estimates of respiratory drift. However, to date no work has demonstrated a personalisation technique that results in more accurate motion models based on information from static images alone. Because of these limitations of subject-specific and cross-population models, there is still only one example (the Cyberknife Synchrony system) of clinical translation of a motion model-based technique [10].

We present a framework for the personalisation of cross-population models that addresses these limitations. Our framework eliminates the need for a dynamic calibration scan and provides motion estimates that are more accurate than those produced by a standard cross-population model. This is achieved by learning the relationship between a vector of anatomical features and the respiratory motion.

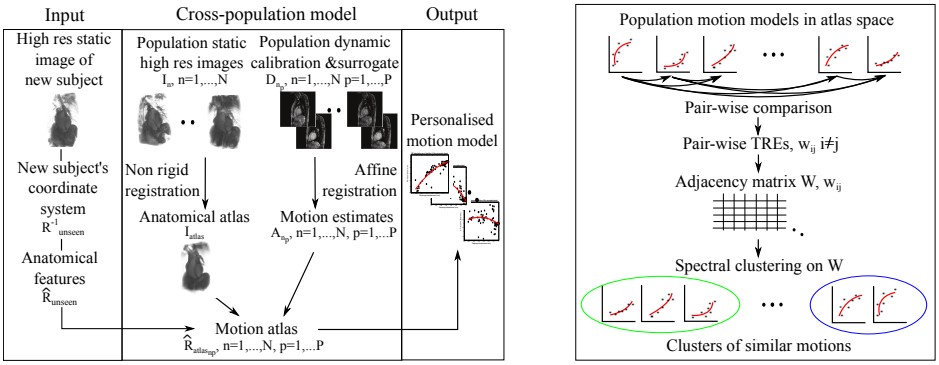
2 Methods and Materials

Our method is schematically represented in Figure 1(a). The input of the personalisation process is a high resolution image of the anatomy of a new unseen subject, while the output is a personalised respiratory motion model.

The cross-population model is formed as follows. Given a population sample of N datasets consisting of a high resolution image of the anatomy, a dynamic calibration scan depicting the respiratory motion and some surrogate data (see Section 2.1), an average atlas of the anatomy is built using the N high resolution images, as proposed by [2, 5]. The respiratory motion estimates derived from the N dynamic calibration scans and surrogate data are then transformed to the atlas coordinate system to produce a motion atlas. This motion atlas can subsequently be used to form a respiratory motion model. Unlike the cross-population models proposed thus far [2, 5, 6, 8] where an average motion model of the N datasets is used for any new unseen subject, we select a sub-set K of the population sample which is more likely to represent the respiratory motion of the new unseen subject. In order to determine K, we compare the respiratory motions of the N subjects in the motion atlas and cluster them according to their similarity. A classifier is then trained to learn and exploit the relationship between anatomical features derived from a static image and the respiratory motions. The underlying hypothesis is that anatomical features can be used as predictors of respiratory motion. To the authors' knowledge, this is the first work to investigate such a hypothesis. The idea is that, given a high resolution image of a new unseen subject, the classifier will return the sub-set K that best describes the new subject's respiratory motion. This way, the inter-subject motion variation will be exploited to obtain motion estimates that are more accurate than standard cross-population model estimates.

2.1 Materials

A sample of 23 cardiac MRI datasets acquired from healthy volunteers was used for this study. All images were acquired using a 1.5T Philips Achieva MRI scanner. The details of



(a) (b)

Figure 1: (a) Overview of the proposed personalisation framework. (b) Clustering of the N population respiratory motions based on their similarity.

the high resolution 3D MRI volume and dynamic 3-D MRI calibration scan used for forming the motion model are:

- **Dynamic 3-D calibration scan**: 3-D TFEPI, ECG-triggered and gated at late diastole, typically 20 slices, $TR = 10ms$, $TE = 4.9ms$, flip angle $= 20^\circ$, acquired voxel size 2.7 x 3.6 x 8.0mm^3, reconstructed voxel size 2.22 x 2.22 x 4.0mm^3, TFE factor 26, EPI factor 13, TFE acquisition time 267.9ms.
- **High resolution 3-D:** 3-D balanced TFE, cardiac gated at late diastole, respiratory gated at end-exhale, 5mm navigator window, typically 120 sagittal slices, $TR = 4.4ms$, $TE = 2.2ms$, flip angle$= 90^\circ$, acquired voxel size 2.19 x 2.19 x 2.74mm^3, reconstructed voxel size 1.37 x 1.37 x 1.37mm^3, the acquisition window was optimised for each volunteer and was on average 100ms, scan time approximately 5 minutes.

The dynamic calibration scan was ECG-triggered and gated, so one volume was acquired for each heart beat. The images therefore represented the motion of the heart due to respiration only. The scan acquired 40 images while the volunteer was breathing normally. The superior-inferior (SI) displacement of the left hemi-diaphragm was employed as the respiratory surrogate. The high resolution MRI image is a standard pre-procedure acquisition in many clinical protocols and provides high spatial resolution information about the anatomy and pathology of the heart.

2.2 Methods

Anatomical atlas. To eliminate anatomical variation from the comparison of the different motions, an average shape atlas in its natural coordinate system was formed, using the approach described in [3]. To remove positional differences from the registrations, the high resolution images were first translated so that the centres of mass of the heart of each subject were aligned. The centres of mass were computed over a manually positioned binary mask covering the main cardiac chambers and vessels. The same mask was then employed for motion estimation and evaluation purposes. Given the population high resolution images $I_n, n = 1,\ldots,N$, one image was randomly selected as a starting reference I_{ref} and all remaining images were non-rigidly registered to it [1]. An average intensity image I_{avg_0} was computed using all N warped images. I_{avg_0} was then employed as the new reference and $I_n, n = 1,\ldots,N$ were non-rigidly registered to it. By averaging the intensities of the new set of warped images, I_{avg_1} was obtained and used as the new reference image. The registration and averaging processes were repeated until the similarity measure between I_{avg_t} and $I_{avg_{t-1}}$

was higher than a predefined threshold. We used Normalised Cross-Correlation (NCC) as a similarity measure and 0.99 as the threshold. To remove any remaining bias towards I_{ref}, I_{avg_t} was non-rigidly registered to $I_n, n = 1,\ldots,N$ and then warped using the mean values of the N resulting deformation fields. In this way, I_{avg_t} is warped to its *natural coordinate system* [3], which requires the minimal non-rigid deformation to explain the anatomical inter-subject variability. I_{atlas} is the final average intensity image in its natural coordinate system.

Cross-population model formation. Denoting by D_{n_p} the dynamic calibration image p of subject n, the dynamic image of subject n having the highest surrogate value was selected as the reference end-exhale image $D_{n_{ref}}$. As described in [4], the images D_{n_p} were registered to $D_{n_{ref}}$ using an affine registration algorithm and a set of P affine transformations A_{n_p} was obtained for each subject n. To localise the registration to the heart only, the dynamic images were masked using the binary mask used for the atlas building. To compare the motions of the different subjects, the transformations A_{n_p} were all transformed to the coordinate system of I_{atlas} as follows. We denote by R_n the non-rigid transformation that maps each high resolution image $I_n, n = 1,\ldots,N$ to I_{atlas}. The transformation $R_{atlas_{n_p}} = R_n \circ A_{n_p} \circ R_n^{-1}$ describes the respiratory state p of subject n in the atlas natural coordinate system [2, 5]. Since $R_{atlas_{n_p}}$ results in a non-rigid transformation, but an affine transformation is considered sufficient to model cardiac respiratory motion [4], a point-based minimisation algorithm was employed to linearise $R_{atlas_{n_p}}$, resulting in $N \times P$ affine transformations $\hat{R}_{atlas_{n_p}}$. These, together with the corresponding surrogate data for each subject, form the motion atlas.

Motion clustering. This paragraph details the clustering of the N subjects' respiratory motions based on their similarity, as shown in Figure 1(b). In order to compare the N motions, the respiratory surrogates were normalised, so that their ranges were [-1,0] for any subject n. Subject-specific affine motion models of the transformations $\hat{R}_{atlas_{n_p}}$ were then built as described in [4]. To quantify motion similarities, 10 evenly distributed surrogate values in [-1,0] were used to compute 10 motion model estimates for each subject n. Target Registration Errors (TRE) between each pair of motion models were computed over the 10 motion estimates using all voxels in the binary mask covering the heart of the atlas as target points. These TRE values were used to cluster the subjects into groups with similar motions as follows. Using the 95^{th} quantile of the pair-wise TREs, a $N \times N$ adjacency matrix W was built, where the entries $w_{i,j}$ represent the TRE between subject i and j, indicating the degree of similarity in their average respiratory motions. By employing a spectral clustering technique [7], the N respiratory motions were grouped into clusters. The number of clusters was chosen to be the maximum number of clusters for which all clusters contained at least 2 subjects.

Personalisation. In principle, a wide range of image-based and non-image-based data could be used for the personalisation of the cross-population model. However, in this preliminary work we used only image-based features, namely the affine parameters that relate the new subject's high resolution image to the atlas average image I_{atlas}. To compute the anatomical feature vector for each subject, the non-rigid anatomical registrations R_n were linearised, again using a point-based minimisation algorithm, to obtain affine anatomical transformations $\hat{R}_n$. The non-translational components of $\hat{R}_n$ were considered only (3 rotations, 3 scalings and 3 shear angles), describing the different shapes and poses of the hearts. The feature vectors were formed from the coefficients of the affine matrix representing this transformation normalised by their standard deviation. A supervised random forest classifier was trained providing the clusters C as outputs and the anatomical feature vectors as predictors. Once the classifier was trained, the anatomical feature vector $\hat{R}_{unseen}$ for a new unseen subject was classified into one of the clusters C_k. The P respiratory affine transformations A_{n_p} of

the K datasets in the cluster C_k were warped to the coordinate system of the new unseen subject using the non-rigid transformation R^{-1}_{unseen}. The personalised cross-population motion model was then built as described in [4]. The final motion estimates were obtained using the personalised model and the surrogate data of the new unseen subject.

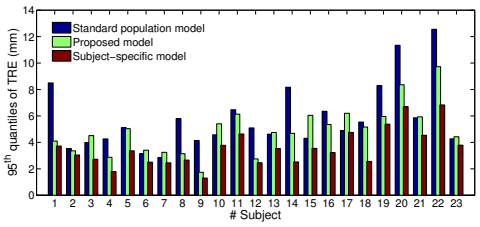
2.3 Evaluation

For evaluation, a leave-one-out cross-validation was employed: each of the 23 subjects was left out in turn and the remaining 22 datasets were used to construct the cross-population model. An anatomical feature vector derived from the left-out subject's high resolution image was used to personalise the cross-population model. For a thorough accuracy evaluation, we non-rigidly registered the dynamic images D_{LO_p} of the left-out dataset to the dynamic end-exhale reference image $D_{LO_{ref}}$ [1]. This process resulted in P gold-standard non-rigid motion fields which were employed to evaluate the accuracy of the personalised motion model. The TRE between the affine motion fields estimated by the motion models and the non-rigid motion fields was computed over the binary mask covering the heart (see Sec 2.2).

We compared our personalised model with a standard cross-population model [2, 5, 8], where all 22 datasets were used in the computation of the motion model for the left-out subject, and to a subject-specific motion model. To build all models, we warped $D_{LO_{ref}}$ using the gold-standard non-rigid motion fields, obtaining a set of artificial images with known, realistic motion fields. Affine registration and model building [4] was applied to the artificial images. For completeness sake, the TRE of no respiratory motion estimate is also computed. Results are provided in Section 3 .

3 Results

Results of the leave-one-out cross-validation are shown in Figure 2. The median and 95^{th} quantile of the TREs were computed for each left out subject for each technique compared. For compactness sake, the mean and standard deviation are computed over all 23 subjects, both for median and 95^{th} quantile. Our method is more accurate than a standard cross-population model proposed to date, with motion estimates closer to the subject-specific estimates. The last row of the table in Figure 2 shows an average improvement of 20% for medians and 17% for 95^{th} quantiles of TREs achieved by our method compared to an average cross-population model. The highest improvements of the 95^{th} quantile were achieved for subject 9 and 1 (58% and 52% respectively).



Method	Median Mean/std (mm)	95^{th} quantile Mean/std (mm)
No estimate	3.5/2.3	11.6/4.5
Average pop model	2.0/1.2	5.8/2.5
Proposed pop model	1.6/0.8	4.8/1.8
Subject-specific model	1.1/0.6	3.6/1.4
Improvement (%)	20.0/33.3	17.2/28 .0

Figure 2: Results of the leave-one-out cross-validation. On the left, 95^{th} quantiles of TRE for each subject are shown. On the right, the table reports mean/std deviation of median and 95^{th} quantile TRE values for the methods compared over all subjects.

4 Discussion and Conclusions

We have proposed a novel personalisation method for cross-population respiratory motion models. Our method exploits inter-subject motion variability by investigating the relationship between the anatomy and its respiratory motion. We have presented results for cardiac respiratory motion derived from MRI. Results showed the proposed model to be more accurate than a standard cross-population model, with accuracy for some subjects comparable to subject-specific motion models, but without the need for a dynamic calibration scan. The proposed personalisation is particularly effective for those subjects with a respiratory motion that differs significantly from the average cross-population motion.

Healthy volunteer datasets were considered in this work. Future work will investigate the application of the technique to clinical data. For patients, a richer source of predictors may be necessary to describe anatomical changes of the heart due to pathology, and we plan to investigate the use of non-imaging data from the patient record for this purpose.

In this work we modelled and compared normal respiratory motion. Future investigation might extend the method to different breathing patterns, as can often be the case during acquisitions/interventions, and different clustering and classification techniques. Moreover, different modalities such as CT or 3-D echocardiography could be employed to build and personalise the proposed cross-population model.

Acknowledgements

This work was funded by EPSRC programme grant EP/H046410/1 and supported by the National Institute for Health Research (NIHR) Biomedical Research Centre award to Guy's and St Thomas' NHS Foundation Trust in partnership with KCL and King's College Hospital NHS Foundation Trust.

References

[1] C. Buerger et al. Hierarchical adaptive local affine registration for fast and robust respiratory motion estimation. *Med. Image Anal.*, 15:551–564, 2011.

[2] J. Ehrhardt et al. Statistical modeling of 4D respiratory lung motion using diffeomorphic image registration. *IEEE Trans. Med. Imaging*, 30:251–265, 2011.

[3] A.F. Frangi et al. Automatic construction of multiple-object three-dimensional statistical shape models: application to cardiac modeling. *IEEE Trans. Med. Imaging*, 21:1151–1166, 2002.

[4] A.P. King et al. A subject-specific technique for respiratory motion correction in image-guided cardiac catheterisation procedures. *Med. Image Anal.*, 13:419–431, 2009.

[5] T. Klinder and C. Lorenz. Respiratory motion compensation for image-guided bronchoscopy using a general motion model. In *9th IEEE ISBI, 2012*, pages 960–963, 2012.

[6] J.R. McClelland et al. Respiratory motion models: A review. *Med. Image Anal.*, 17:19–42, 2013.

[7] A.Y. Ng et al. On spectral clustering: Analysis and an algorithm. In *Adv. Neural Inf. Process. Syst.*, pages 849–856. MIT Press, 2001.

[8] F. Preiswerk et al. Robust tumour tracking from 2D imaging using a population-based statistical motion model. In *IEEE Workshop on MMBIA*, pages 209–214, 2012.

[9] G. Samei et al. Predicting liver motion using exemplar models. In *Abdominal Imaging. Computational and Clinical Applications*, pages 147–157. Springer, 2012.

[10] A. Schweikard et al. Robotic motion compensation for respiratory movement during radiosurgery. *Comp. Aided Surg.*, 5:263–277, 2000.