

HHS Public Access

Author manuscript

Comput Methods Biomech Biomed Eng Imaging Vis. Author manuscript; available in PMC 2020 January 01.

Published in final edited form as:

Comput Methods Biomech Biomed Eng Imaging Vis. 2019 ; 7(3): 297–301. doi: 10.1080/21681163.2018.1501765.

Automatic segmentation of the thumb trapeziometacarpal joint using parametric statistical shape modelling and random forest regression voting

Marco T. Y. Schneider^a, Ju Zhang^a, Joseph J. Crisco^b, Arnold-Peter C. Weiss^b, Amy L. Ladd^c, Poul M. F. Nielsen^{a,d}, and Thor Besier^{a,d}

^aAuckland Bioengineering Institute, The University of Auckland, Auckland, New Zealand ^bDepartment of Orthopedics, Warren Alpert Medical School of Brown University, Rhode Island Hospital, RI, USA ^cDepartment of Orthopedic Surgery, Stanford, Stanford University, CA, USA ^dDepartment of Engineering Science, The University of Auckland, Auckland, New Zealand

Abstract

We propose an automatic pipeline for creating shape modelling suitable parametric meshes of the trapeziometacarpal (TMC) joint from clinical CT images for the purpose of batch processing and analysis. The method uses 3D random forest regression voting (RFRV) with statistical shape model (SSM) segmentation. The method was demonstrated in a validation experiment involving 65 CT images, 15 of which were randomly selected to be excluded from the training set for testing. With mean root mean squared (RMS) errors of 1.066 mm and 0.632 mm for the first metacarpal and trapezial bones respectively, and a segmentation time of ~2 minutes per CT image, the preliminary results showed promise for providing accurate 3D meshes of TMC joint bones for batch processing.

Keywords

automatic segmentation; statistical shape model; random forest; regression voting; trapeziometacarpal; model generation

1. Introduction

The trapeziometacarpal (TMC) joint is highly susceptible to osteoarthritis (OA), and can impair the upper extremity by up to 50 % (Pellegrini Jr 2005). Several factors have been implicated in its pathogenesis, with biomechanical factors paramount (Hunter et al. 2005). Morphology of the first metacarpal and trapezial bones are important biomechanical factors that must be considered as they affect the moment arms of muscles and ligaments, joint kinematics, posture during tasks, and the cartilage stresses and strains during contact(Arnold and Delp 2001, Halilaj et al. 2014, Nanno et al. 2006).

Corresponding Author: Marco Tien-Yueh Schneider, Auckland Bioengineering Institute, University of Auckland, Level 6, 70 Symonds Street, Auckland 1010, New Zealand, Tel: +64 9 3737599, msch153@aucklanduni.ac.nz.

Schneider et al.

Anatomically accurate 3D models of the TMC joint can be obtained through segmentation of clinical CT volumes. However, traditional manual and semi-automatic segmentation methods are time-consuming and require manual intervention to complete the segmentation of an image volume. The segmented data, usually in the form of point clouds, require meshing or triangularisation prior to musculoskeletal modelling, finite element analysis, or shape modelling. Therefore, these methods are unsuitable in workflows that involve batch processing, where the speed of segmentation is important, minimal user input is desired, and data may be required in certain formats for further processing.

Automatic segmentation methods address the main disadvantages of both manual and semiautomatic methods by removing the requirement of manual intervention; they are time efficient, and can be used in batch processes. This would allow for a potential clinical tool such as an automated pipeline that processes CT volumes for useful information such as the corresponding stress distributions. However, automatic segmentation methods such as region growing, active shape models (ASM), point distribution model based statistical shape model (SSM) segmentation methods, may suffer from decreased robustness because of their reliance on correct initialisation, linear search spaces perpendicular to the model surface, high contrast edges, and high resolution relative to object feature size. In the wrist, these automatic segmentation methods struggle with the crowding of carpal bones and small joint spaces relative to voxel dimensions that create low contrast edges, especially in lower resolution clinical CT.

Three-dimensional random forest regression voting (RFRV) automatic shape model segmentation uses randomly sampled image features (e.g. 3D Haar-like features) to train a forest of decision trees to predict the most likely image location of the desired model. This has been demonstrated in 2D with facial recognition (Cootes et al. 2012) and 2D segmentation of the proximal femur (Lindner et al. 2013), and more recently demonstrated in 3D in the Liver (Norajitra et al. 2015). This method can be combined with parametric statistical shape modelling to create an automatic segmentation pipeline that has increased robustness to initialisation, increased speed of segmentation, and automatic meshing for downstream analysis.

The purpose of this paper is to present a pipeline for automatically creating parametric meshes of the TMC joint from clinical CT images for the purpose of batch processing, shape modelling, and analysis. We detail the pipeline below then present quantitative segmentation results on a set of 65 clinical CT images of the wrist.

2. Methods

The method uses a SSM combined with RFRV to perform automatic segmentation of the TMC joint. For each bone, a SSM is first created and trained with segmented point clouds. Random forest (RF) regressors are trained on CT image data for each node. During segmentation, RF regressors predict the location of nodes in new CT images and the SSM is fit to the predicted nodes.

2.1 Statistical shape model

A parametric shape model is required to train the RF regressors, constrain the possible segmentation geometries, and to yield a parametric mesh of the desired geometry.

The parametric SSM is generated using a published method (Schneider et al. 2015) that is based on techniques described by Zhang et al. (2014). In summary, this method uses a custom template mesh consisting of cubic Lagrange elements (Nielsen 1987) to represent the morphology. The mesh is designed to capture the morphological variation in anatomical landmark regions across the population. The template mesh is then fit to all point clouds in the training set through an iterative fitting process involving a series of coarse and fine fits. The fitted meshes of the training set are then rigidly aligned with a partial Procrustes alignment that conserves size.. Principal component analysis (PCA) (section 2.2) is performed on the mesh nodal coordinates to produce an initial shape model. This process is repeated to propagate correspondence by using the shape model instead of the template mesh, until the RMS error is less than the in-plane pixel resolution (Zhang, Malcolm, Hislop-Jambrich, Thomas and Nielsen 2014). The final shape model is produced by performing a final PCA on the nodal coordinates of the correspondent training set meshes.

2.2 Principal component analysis

Principal component analysis is performed to remove linear dependencies and for dimensionality reduction and allows any shape, *x*, in the training set to be reconstructed by the weighted sum of *n* principal components, ϕ , and the mean shape, \bar{x} (Heimann and Meinzer 2009, Schneider, Zhang, Crisco, Weiss, Ladd, Nielsen and Besier 2015, Zhang, Malcolm, Hislop-Jambrich, Thomas and Nielsen 2014):

$$x = \overline{x} + \sum_{i=0}^{n} \omega_i \phi_i \quad (1)$$

2.3 Random forest regression voting

Segmentation can be achieved by automatically locating the position of each mesh node in the CT image. This detection is performed by a RF regressor trained for each node. In this work, we used the scikit-learn¹ implementation of the RF regressor (Breiman 2001, Pedregosa et al. 2011).

For each regressor, 3D Haar-like features (section 2.4) are randomly sampled about the corresponding mesh node in an omni-directional search space. Each regressor learns the spatial distribution of the feature response about each node by associating the feature response with the corresponding displacement vector between the feature centroid and the node coordinates (Figure 1). When features are sampled from an unseen image, the regressors can predict its displacements based on the feature response sampled, allowing for prediction of nodal locations.

¹http://scikit-learn.org/stable/

Comput Methods Biomech Biomed Eng Imaging Vis. Author manuscript; available in PMC 2020 January 01.

During segmentation, the mean mesh of the SSM is initialised, ideally in a neighbourhood of the desired object. Each tree in the RF regressor of each node votes by estimating the displacement vector that points to the location of the mesh node from the sampled 3D Haarlike feature. This creates a distribution of votes that predict the location of the node in the CT volume. The mean location of the votes is taken as the final prediction of the bestmatched image location of the mesh node. This process is performed for each node in each mesh to obtain the predicted nodal locations of the desired image geometry in a single iteration. The mean mesh is then fit to the predicted nodal locations using deformations permitted by the SSM (Equation 1), resulting in a segmented mesh. Meshes segmented with this shape model are also parametric, allowing subsequent analysis of morphology and stress analysis to be directly compared in the same frame of reference.

2.4 3D Haar-like features

Haar-like features are used in image object recognition and have been used in image segmentation in 2D (Cootes, Ionita, Lindner and Sauer 2012) and 3D (Norajitra and Maier-Hein 2017, Norajitra, Meinzer and Maier-Hein 2015). Haar-like features are calculated by comparing the difference in summed pixel intensity between regions of pixels (in 2D) or voxels (in 3D) within a bounding box These regions may be labelled as 'dark' and 'light', where the bounds of the 'light' region can be randomised to create an infinite set of 3D Haar-like features (Lindner, Thiagarajah, Wilkinson, Consortium, Wallis and Cootes 2013, Norajitra and Maier-Hein 2017). Due to this formulation, 3D Haar-like features do not support more complex features that have more than one 'light' region, such as chequered 3D Haar-like features, which can provide important textural information. In this study, we used a fixed set of eight 3D Haar-like features (Figure 2), including three features with one 'light' region, three features with two 'light' regions, three features with axis aligned chequered regions (two 'light' and two 'dark' regions), and one feature with completely chequered regions (four 'dark' and four 'light' regions). These features can be calculated efficiently by precomputing the integral image using the equation 2. The difference in summed pixel intensity can then be calculated using equation 3 (Norajitra and Maier-Hein 2017).

$$I_{\sum(x, y, z)} = \sum_{\substack{x' \le x \\ y' \le y \\ z' \le z}} i(x', y', z') \quad (2)$$

$$\begin{split} \Sigma_{c} &= \Sigma_{x_{max}, y_{max}, z_{max}} - \Sigma_{x_{min} - 1, y_{max}, z_{max}} \\ &- \Sigma_{x_{max}, y_{min} - 1, z_{max}} + \Sigma_{x_{min} - 1, y_{min} - 1, z_{max}} \\ &- \Sigma_{x_{max}, y_{max}, z_{min} - 1} + \Sigma_{x_{min} - 1, y_{max}, z_{min} - 1} \\ &+ \Sigma_{x_{max}, y_{min} - 1, z_{min} - 1} - \Sigma_{x_{min} - 1, y_{min} - 1, z_{min} - 1} \end{split}$$
(3)

3. Validation experiment

The method described above was applied to a set of 65 CT images from healthy adult males and females. The wrist was imaged in clinical neutral position with a 16-slice CT scanner (GE LightSpeed 16, General Electric, Milwaukee, WI). The scanner was set with the following settings: tube voltage at 80 kVp, tube current at 80 mA, slice thickness of 0.625 mm, and in-plane resolution of 0.4 mm \times 0.4 mm. The trapezia and first metacarpals were segmented semi-automatically using Mimics v12.11, and exported as triangulated surfaces. The vertices of the triangulated surfaces were extracted to obtain a training set of point clouds for the SSM.

Fifty CT images were randomly selected to be the training data set for the SSM and the RF regressors, and the remaining 15 CT images were used to evaluate the accuracy and performance of the method. Two parametric template meshes were created, one for the first metacarpal, consisting of 398 nodes and 97 cubic Lagrange elements, and one for the trapezium, consisting of 344 nodes and 52 cubic Lagrange elements. These template meshes were used to create two SSMs, one for each bone in the joint, which were consolidated into a complete model of the TMC joint, consisting of 742 nodes. 50 CT sampling windows, with a size of $10 \times 10 \times 10$ voxels, were randomly sampled within an omnidirectional search space of 30 voxels about each node in the CT image. Integral images were computed, and 3D Haar-like features were calculated and used to train the RF regressors that consisted of 20 random decision trees. The time required to train 742 RF regressors took approximately 3 hours on an Intel Xeon quad core computer.

To evaluate the accuracy and performance of the method, we applied the trained RF regressors to the remaining 15 CT images not used in training. The TMC joint model was initialised in the centre of the CT images, near the in-image TMC joint bones. The RF regressors then predicted the locations of all 742 mesh nodes. The meshes of each bone were fit to the corresponding predicted node locations with the weighted sum of principal components; by minimising the least squared error between the predicted nodal coordinates and shape model node coordinates, producing a segmented parametric mesh of the TMC joint bones.

4. Results

The segmentation time for all 15 datasets was approximately 30 minutes on an Intel Xeon quad core computer. The underlying image size was approximately $512 \times 512 \times 346$ voxels. The integral image was computed in less than 5 seconds. The average surface-to-surface RMS error between automatically segmented mesh and manually segmented ground truth for all 15 datasets was 1.066 mm in the first metacarpal and 0.632 mm in the trapezium (Table 1). Figure 3 shows the mean pointwise error distribution of all 15 datasets. The largest errors appear to occur in regions of high curvature. Figure 4 shows the segmentation of a randomly selected dataset overlapped with the ground truth.

5. Discussion

These results show promise for a fully automatic segmentation pipeline that creates anatomically accurate parametric meshes from CT image data. Since the deformation of the mesh geometry was restricted by the SSM, the segmented shapes of the bones were constrained to be anatomically accurate. Pointwise error distribution plots of the automatically segmented surface compared with the manually segmented ground truth indicated that the error was highest about regions of high curvature (Figure 3). The mean error in the first metacarpal was 0.997 mm \pm 0.372 mm, and 0.564 mm \pm 0.282 mm in the trapezium, which may be considered reasonable depending on the purpose of the model. The mean % volume overlap was 84.12 % in the first metacarpal and 86.03 % in the trapezium, is comparable to the intermediate rigid model fitting results on the liver reported by Norajitra, Meinzer and Maier-Hein (2015), except in a much smaller joint. However, for the purpose of contact biomechanics at the articular surfaces of the joint, these errors may require further reduction, as the maximum errors (2.110 mm in the first metacarpal and 1.776 mm in the trapezium) are comparable to the size of the joint space in the TMC joint (~ 2 mm).

This pipeline appears to be robust to variation in the location of initialisation of the shape model. The success of ASMs and region growing algorithms are highly dependent on correct initialisation. In our test scenarios, as the TMC joint bones were the subject of interest, we assumed that the in-image position of the TMC joint bones would be close to the centre, and thus, it seemed reasonable to initialise the shape model in the centre of the CT images. However, the TMC joint was not perfectly located in the centre in any of the datasets, and was located up to 30 mm away. Despite this, the pipeline managed to segment the TMC joint bones to RMS errors of 1.066 mm in the first metacarpal and 0.632 mm in the trapezium. We expect that these errors can be improved by using a multi-scale approach (Norajitra and Maier-Hein 2017). First, RF regressors trained with large windows and distances can be used to initialise the mesh closer to the joint. Afterwards, incremental improvements in segmentation can be attempted by using RF regressors trained on smaller windows and distances. As the pipeline performs quickly (~2 minutes per segmentation) this step can be repeated a number of times to increase the accuracy of segmentation without sacrificing too much time. Furthermore, Norajitra and Maier-Hein (2017) showed in three organs (liver, spleen, and kidney) that a multi-scale approach could remove the need for careful model initialisation. Their success may also be owing to the use of randomised features which can give a richer description of encountered anatomical structures. This could be implemented in our pipeline and theoretically would allow the random forest to make better classifications and further reduce the segmentation error. Further improvements in segmentation accuracy could be achieved by relaxing the SSM constraints in the final iteration of fitting. In another study, Norajitra, Meinzer and Maier-Hein (2015) reported an improvement in segmentation accuracy with the application of a deformable surface algorithm during the final step in fitting. This was to allow the surface to deviate from the rigid shape constraint of the SSM, and to compensate for the liver's high variability in shape. Although the variability in the TMC joint is low (Schneider et al., 2015), we expect that the accuracy of TMC segmentation would improve with this approach.

6. Conclusion

We have presented a pipeline for automatically creating parametric meshes of the TMC joint from CT images of the wrist, using RFRV combined with parametric statistical shape modelling. This method has demonstrated increased flexibility in terms of segmentation error with regards to the location of initialisation compared to methods such as active shape modelling and region growing. The use of 3D Haar-like features were used to teach the spatial characteristics of the images to the RF regressors, and to predict the location of mesh nodes in new images. Our results were promising, with mean RMS errors of 1.066 mm and 0.632 mm for the first metacarpal and trapezial bones respectively. We propose a multi-scale approach involving a series of coarse to fine segmentations in concert with feature randomization to improve the segmentation accuracy. This pipeline has the potential to be used for rapid segmentation of clinical CT images, for the purpose of implant design, biomechanics analysis, and in surgical planning.

Acknowledgements

This work was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health (Award number AR059185) as well as the Auckland Bioengineering Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Arnold AS, Delp SL. 2001 Rotational moment arms of the medial hamstrings and adductors vary with femoral geometry and limb position: implications for the treatment of internally rotated gait. Journal of biomechanics.34:437–447. [PubMed: 11266666]
- Breiman L 2001 Random Forests. Machine Learning. 10 01;45:5-32.
- Cootes TF, Ionita MC, Lindner C, Sauer P. Robust and accurate shape model fitting using random forest regression voting Proceedings of the European Conference on Computer Vision; 2012: Springer.
- Halilaj E, Moore DC, Laidlaw DH, Got CJ, Weiss A-PC, Ladd AL, Crisco JJ. 2014 The morphology of the thumb carpometacarpal joint does not differ between men and women, but changes with aging and early osteoarthritis. Journal of biomechanics.47:2709–2714. [PubMed: 24909332]
- Heimann T, Meinzer H-P. 2009 Statistical shape models for 3D medical image segmentation: a review. Medical image analysis.13:543–563. [PubMed: 19525140]
- Hunter D, Zhang Y, Sokolove J, Niu J, Aliabadi P, Felson D. 2005 Trapeziometacarpal subluxation predisposes to incident trapeziometacarpal osteoarthritis (OA): the Framingham Study. Osteoarthritis and cartilage.13:953–957. [PubMed: 16139531]
- Lindner C, Thiagarajah S, Wilkinson J, Consortium T, Wallis G, Cootes T. 2013 Fully automatic segmentation of the proximal femur using random forest regression voting. IEEE transactions on medical imaging.32:1462–1472. [PubMed: 23591481]
- Nanno M, Buford WL, Patterson RM, Andersen CR, Viegas SF. 2006 Three-dimensional analysis of the ligamentous attachments of the first carpometacarpal joint. The Journal of hand surgery. 31:1160–1170. [PubMed: 16945721]
- Nielsen PMF. 1987 The anatomy of the heart: a finite element model ResearchSpace@ Auckland.
- Norajitra T, Maier-Hein KH. 2017 3D Statistical Shape Models Incorporating Landmark-Wise Random Regression Forests for Omni-Directional Landmark Detection. IEEE transactions on medical imaging.36:155–168. [PubMed: 27541630]
- Norajitra T, Meinzer H-P, Maier-Hein KH. 3D statistical shape models incorporating 3D random forest regression voting for robust CT liver segmentation. Proceedings of the SPIE Medical Imaging; 2015: International Society for Optics and Photonics.

- Pedregosa F, Varoquaux G, Gramfort A, Michel VT, hirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al. 2011 Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research.12:2825–2830.
- Pellegrini VD Jr. 2005 The ABJS 2005 Nicolas Andry Award: osteoarthritis and injury at the base of the human thumb: survival of the fittest? Clinical orthopaedics and related research.438:266–276. [PubMed: 16131901]
- Schneider M, Zhang J, Crisco J, Weiss A, Ladd A, Nielsen P, Besier T. 2015 Men and women have similarly shaped carpometacarpal joint bones. Journal of biomechanics.48:3420–3426. [PubMed: 26116042]
- Zhang J, Malcolm D, Hislop-Jambrich J, Thomas CDL, Nielsen PM. 2014 An anatomical region-based statistical shape model of the human femur. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization.2:176–185.



Figure 1.

3D Haar-like features (red boxes) and corresponding displacements (arrows) are used to train a RF regressor on the spatial distribution of features around each node of the shape model mesh. During segmentation, the shape model mesh (white dotted line) is initialised and 3D Haar-like features (red boxes) are sampled (dashed lines) from the mesh node (red point on white dotted line). The regressor for each node uses the features to predict the correct location (arrow) of the node (red point) in the CT image.

Schneider et al.



Figure 2.

The ten types of 3D Haar-like features that were used to train the RF regressors. Feature values were calculated by comparing the difference in summed pixel intensities in the dark and light bounding boxes.

Schneider et al.



Figure 3.

Mean pointwise error distribution. The volar-radial view (A), and dorsal view (C) are shown for the first metacarpal. The volar view (B) and dorsal view (D) are shown for the trapezium.



Figure 4.

Segmented TMC joint bones. Red shows ground truth. Yellow shows automatically segmented mesh.

Table 1.

Calculated error distributions (in mm) for test population (n = 15). Percentage volume overlap was calculated using the tanimoto metric $T(A, B) = (A \cap B)/(A \cup B)$.

	Mean Error	RMS Error	Max Error	Min Error	Mean % Volume Overlap
1 st Metacarpal	0.997 ± 0.372	1.066	2.110	0.0457	84.12%
Trapezium	0.564 ± 0.282	0.632	1.776	0.013	86.03%