

Database guided detection of anatomical landmark points in 3D images of the heart

Thomas Karavides^{a,b}, K. Y. Esther Leung^a, Pavel Paclik^c, Emile A. Hendriks^b, Johan G. Bosch^{*a}

^a Thoraxcenter Biomedical Engineering, Erasmus MC Rotterdam, the Netherlands;

^b Information and Communication Theory Group, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, the Netherlands;

^c PR Sys Design, Delft, the Netherlands

ABSTRACT

Automated landmark detection may prove invaluable in the analysis of real-time three-dimensional (3D) echocardiograms. By detecting 3D anatomical landmark points, the standard anatomical views can be extracted automatically in apically acquired 3D ultrasound images of the left ventricle, for better standardization of visualization and objective diagnosis. Furthermore, the landmarks can serve as an initialization for other analysis methods, such as segmentation. The described algorithm applies landmark detection in perpendicular planes of the 3D dataset. The landmark detection exploits a large database of expert annotated images, using an extensive set of Haar features for fast classification. The detection is performed using two cascades of Adaboost classifiers in a coarse to fine scheme. The method is evaluated by measuring the distance of detected and manually indicated landmark points in 25 patients. The method can detect landmarks accurately in the four-chamber (apex: 7.9 ± 7.1 mm, septal mitral valve point: 5.6 ± 2.7 mm; lateral mitral valve point: 4.0 ± 2.6 mm) and two-chamber view (apex: 7.1 ± 6.7 mm, anterior mitral valve point: 5.8 ± 3.5 mm, inferior mitral valve point: 4.5 ± 3.1 mm). The results compare well to those reported by others.

Keywords: Classification, pattern recognition, landmark detection, echocardiography

1. INTRODUCTION

Real-time 3D echocardiography offers new possibilities in the analysis of cardiac function. The interpretation of the echocardiographic images, however, is critically dependent on the experience of the observer as well as the standardization of views. Variability in choice of anatomical views results, unfortunately, in high interobserver and interinstitution variability in the diagnosis. Additionally, quantitative analysis methods, like segmentation, often need a good initialization. As a result, a method that detects left ventricular landmark points (apex, mitral valve hinge points) in an automated way would be of great importance.

Several methods for automated landmark detection have been proposed for echocardiographic images. For example, Van Stralen et al.¹ describe a technique for detection of the left ventricular (LV) long axis and mitral valve plane in 3D ultrasound images. This is achieved through several consecutive steps. After taking 2D slices perpendicular to the acquisition axis over time, a circular Hough transform is applied to detect the centre of the endocardium border. In that way possible centres are detected for each slice, and through dynamic programming, a path through all slices approximates the left ventricular long axis. Further steps provide the final computation both of the axis and the mitral valve plane.

The Fourier Mellin Transform (FMT) has also been proposed for localization of the left ventricle². The main idea behind this approach is to build an appearance template of the object of interest, and to express both the template and the search image in a rotation and scale invariant representation using the FMT. This way, the estimation of rotation, scale and position converts into a cascade of linear shift detections, enabling a fast estimate of the object pose.

*j.bosch@erasmusmc.nl; phone +31-10-7038088; fax +31-10-7044720; www.erasmusmc.nl/thoraxbme

Recently, classification approaches have been proposed for automated object detection. Classification makes use of prior knowledge embedded in expert annotated databases. The objects are represented using many image features and detected by a classifier (e.g. support vector machine, linear discriminant analysis etc) trained from labeled examples. Georgescu et al.³ introduce a classification method, which they named database-guided paradigm, to detect structures of interest in medical images. A robust classifier is constructed using a boosted cascade of weak classifiers. Each weak classifier uses a feature from a large set of Haar features. A similar approach is used by Carneiro et al.⁴ for the detection and measurement of fetal structures in ultrasound images.

Lu et al.⁵ presented a classification approach for the detection of standard view planes, similar to the one proposed in the current study. They use a cascade of three classifiers, which are trained based on boosting techniques. The first classifier uses Haar wavelet-like feature types and the next two use steerable features⁶.

In this study, a classification based method was developed to detect the standard anatomical landmark points in two-chamber (2C) and four-chamber (4C) view planes in a fully automated way. The method uses smaller databases than those that have been used to similar approaches⁵, since large databases of training data were not available, which is often the case for medical images. To make the classifier more robust, artificial variations are used. The small database available was also the reason to use 2D instead of 3D features, since the latter would require an enormous database to model the 3D LV adequately. Boosting techniques are used to train the classifier. In initial experiments, we saw that small tree classifiers performed better than weak classifiers consisting of a single feature, so we decided to use such classifiers in the boosting scheme. Similar to the procedures described by Nemes et al.⁷ and Leung et al.⁸ the method will be applied to single 3D images at end-diastole (ED).

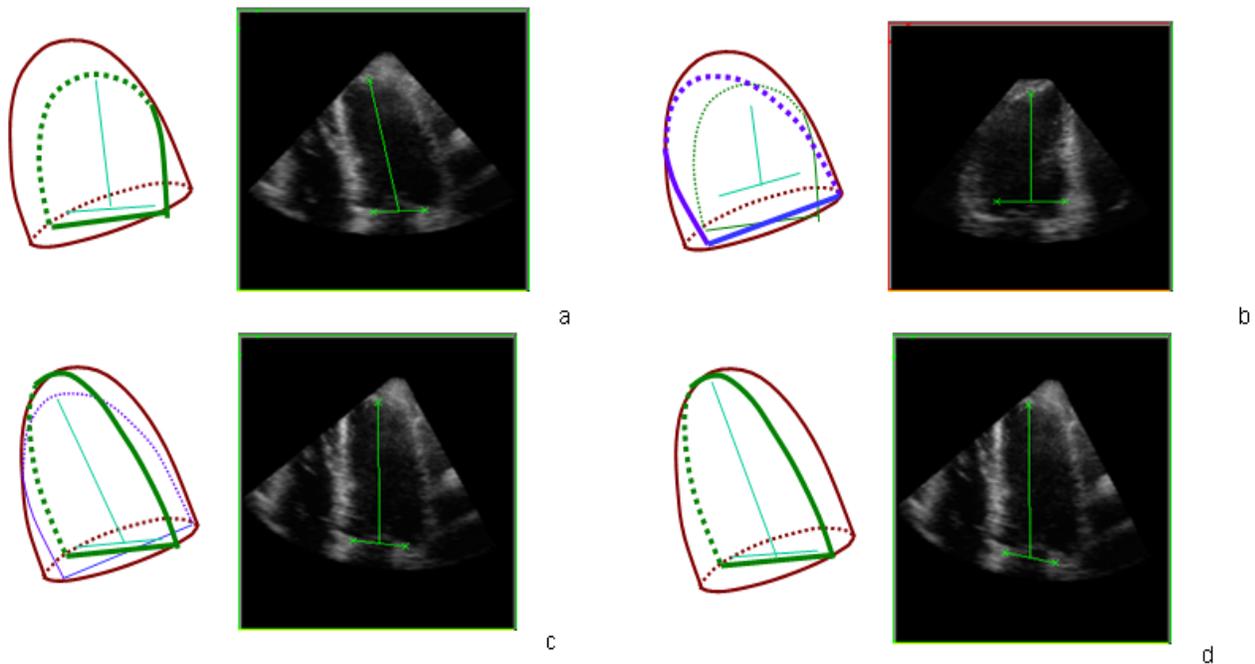


Figure 1. Manual iterative 3D landmark detection approach.
a) After the apex and mitral valve markers are placed in an approximate 4C plane, the LAX is calculated,
b) the perpendicular 2C approximation plane is taken and the LAX is adjusted,
c) the perpendicular 4C approximation is used to adjust the LAX again,
d) step b) and c) are repeated until the true LAX is found.

2. MATERIALS AND METHODS

2.1 Manual scheme

Previously, we proposed a manual method to find the anatomical views 2C and 4C, in which perpendicular 2D planes from a 3D data set were annotated iteratively^{7,8}. By annotating the apex and mitral valve centers in perpendicular long-axis (LAX) planes, the 3D LAX can be found easily (Figure 1).

The LAX can be re-estimated in the updated views to converge to the LAX, which usually takes only 4-6 iterations. In this study we propose to replace the manual annotation in 2D cross-sections using the fully automated landmark detection.

2.2 Landmark detection

Instead of detecting the landmark points individually, we try to locate the region containing those landmarks. The region is defined by a box, in which the positions of these landmarks are fixed (Figure 2). To do this, we place boxes of different sizes and orientations at different positions of the image. For each box, we determine whether it contains these landmarks at fixed positions within the box using classifiers.

We model the left ventricle using simple and easy to compute Haar wavelet-like features, that have been used in similar applications⁴ (Figure 3). The output value of each feature is the difference between the sum of the image pixels in the white section and the sum of image pixels in the black section. To compute these features efficiently, we used the algorithm of Viola and Jones⁹. The pixels of the resulted filtered image (Figure 4) are the features used to train the classifier.

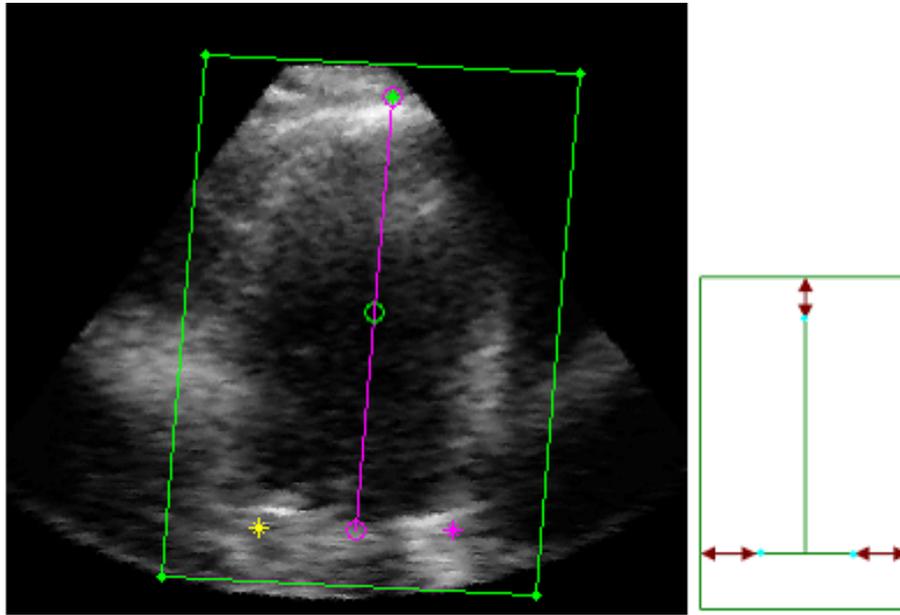


Figure 2. Annotated image of a 2-chamber view (left) with a rotated and skewed rectangle box, where the landmarks have fixed relative positions, within the box. The template box is shown on the right.



Figure 3. Types of Haar features used by the classifiers.

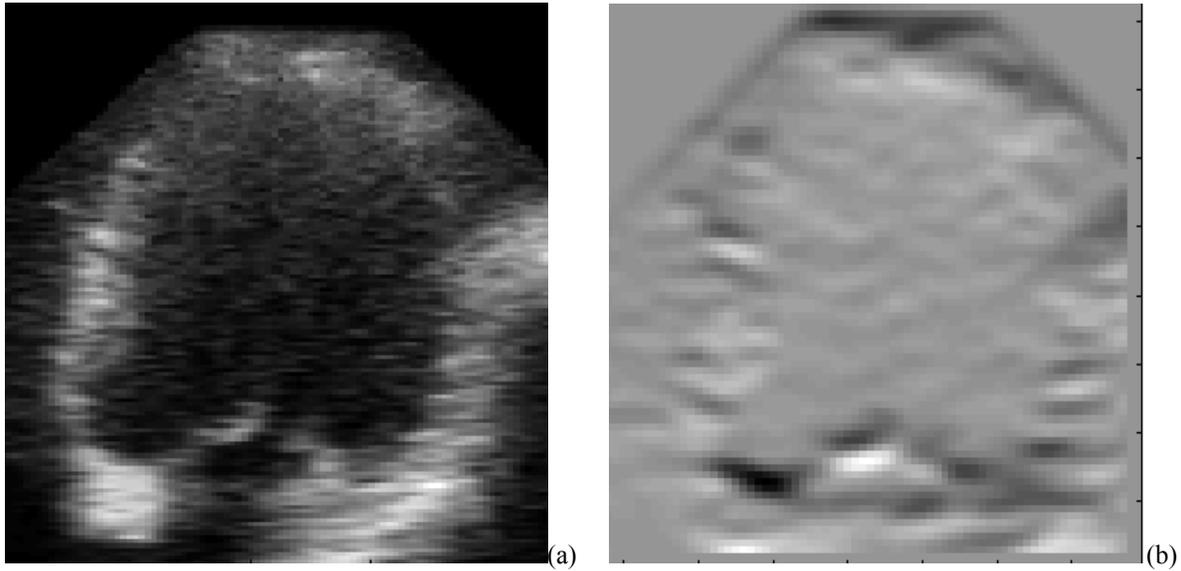


Figure 4. Example of (a) an image box and (b) its filtered result using the first Haar feature from fig. 3.

For the calculation of these features in our ultrasound images, we had to deal with missing image information, in regions where the template exceeds the actual image boundaries. Georgescu et al.³ proposed a mask to exclude invalid pixel values. Apart from using that method, we explicitly exclude these pixels. By summing all training images from all patients, we use this summed image to know which regions were invalid for all our patients. We used this information to exclude the feature values of these regions from both our training and testing set (Figure 5).

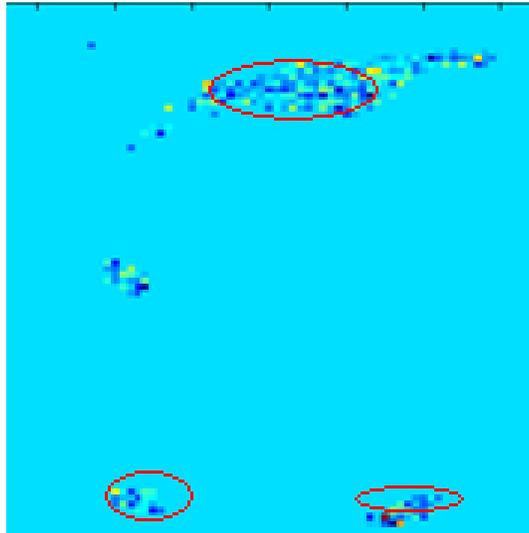


Figure 5. Map of features used by the third classifier after explicitly excluding invalid region values. Positions correspond to an image box as in fig. 4. The classifier focuses near the anatomical landmark locations (red circles). Colors denote the number of features used at a certain position.

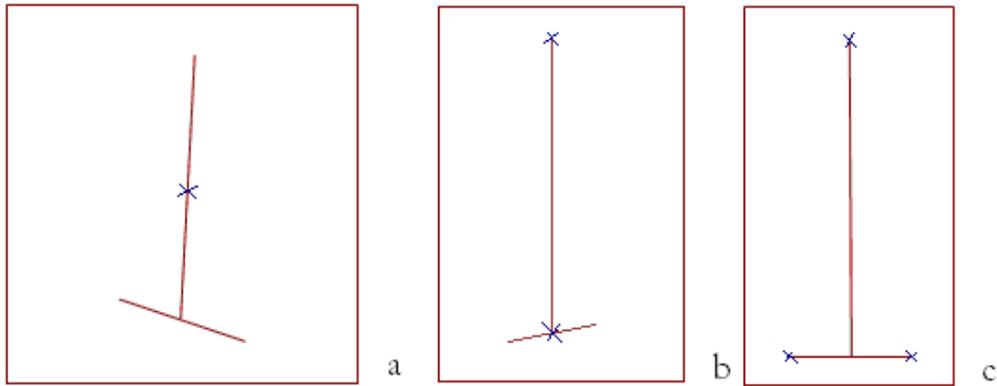


Figure 6. Template boxes. a) Template of first classifier, only LV center has fixed position. b) Template of second classifier, fixed positions for apex and mitral valve center. c) Template of third classifier, fixed positions for all anatomical landmark points.

The method works in a coarse to fine approach with more pose parameters and finer steps, using a cascade of three classifiers. The first classifier detects the approximate center of the left ventricle (irrespective of size and rotation). The second one detects the apex and mitral valve center (so, approximate LV position, size and rotation, but not shear). The third classifier detects all three apex and mitral valve point positions (anterior/inferior in 2C and septal/lateral in 4C), i.e. the complete pose, with more precision. To train the three classifiers of the cascade, different positive and negative examples are generated from the annotated database, to match the three templates in Figure 6. From each patient, we can generate several positive examples, by applying controlled amounts of translation, rotation and shear to the ideal box. That way we model the natural variation of LV size and pose and we have the opportunity to use larger number of examples for the training phase, which is important for obtaining a robust classifier.

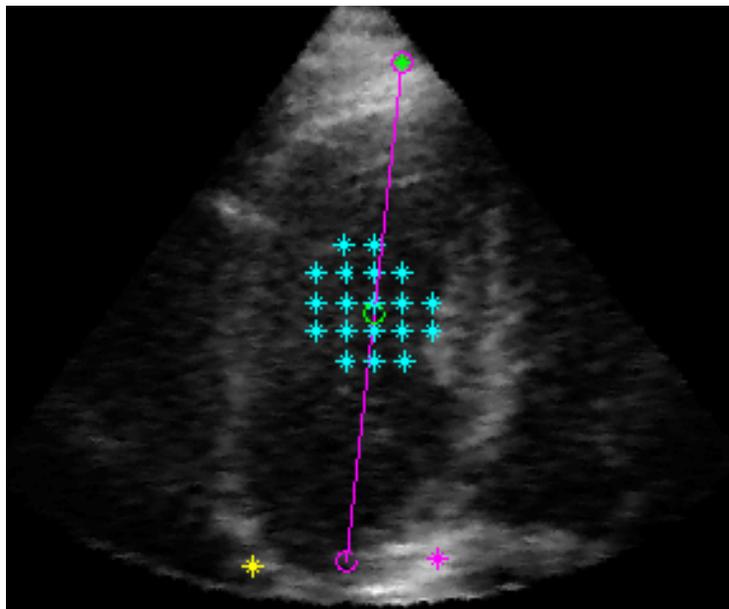


Figure 7. Example of detected LV centers for the 2 chamber view by the first classifier. The true center is marked with a green circle and the hits with cyan asterisks. The inferior (magenta), anterior (yellow) and apex (cyan) markers are also indicated.

We train classifiers to distinguish between boxes with the landmarks at the correct positions (positive examples) and those with landmarks at the wrong places (negative examples). The classifiers are trained using Adaboost, a machine learning algorithm which combines smaller, weak classifiers into a robust, strong one. This is accomplished by iteratively selecting the best weak classifiers. After each iteration, training examples which are misclassified by the current strong classifier are upweighted, to allow the classifier to adapt more to these misclassified examples during the next iteration. Here, we use small tree classifiers, each using a small number of Haar features, as weak classifiers.

The cascade is then used for detection. First a rectangular box is scanned over the 2D image. Each box is a first candidate for containing the center of the left ventricle. The first classifier will evaluate each box and determine the most probable “hits” (detected positives), as shown in Fig. 7. These will be used to generate new candidates to match the second template, by translating and rotating the region around the detected LV centers. After classifying them with the second classifier, the new “hits” will be used to give us new candidates for the third classifier. This time, the boxes will undergo finer translation, rotation, scale and shear, to create sub images to detect all landmark points (Figure 6). The third classifier is used to give us the final estimate of the landmark position, which is the average of all given “hits” for each patient. To limit our “hits” we pruned the classified candidates from stage to stage, as we progress through the cascade to exclude spatial outliers, as a post-processing step.

2.3 Experimental setup

The evaluation is done in three stages: training the classifiers, tuning them to a desired operating point by ROC analysis (tradeoff between correctly classified sub images and misclassified ones) and finally, testing the algorithm. 85 manually annotated 3D patient datasets were used in the project. 60 patients are used to train the classifiers. The remaining 25 are split into batches of five for cross-validation: each time, 20 are used for the ROC analysis and five for evaluation, by measuring the detection error in the original 3D set.

The basic idea of ROC analysis is very simple: for a given trained classifier and a labeled test set, define a set of possible operating points and estimate different type of classifier errors at these points. For training our classifier we used prTools (<http://www.prtools.org/>), a toolbox developed in Matlab, while the ROC analysis was done using PRSD Studio (<http://prsdstudio.com>) from PR Sys Design, which runs in Matlab (version R2007b, MathWorks Inc).

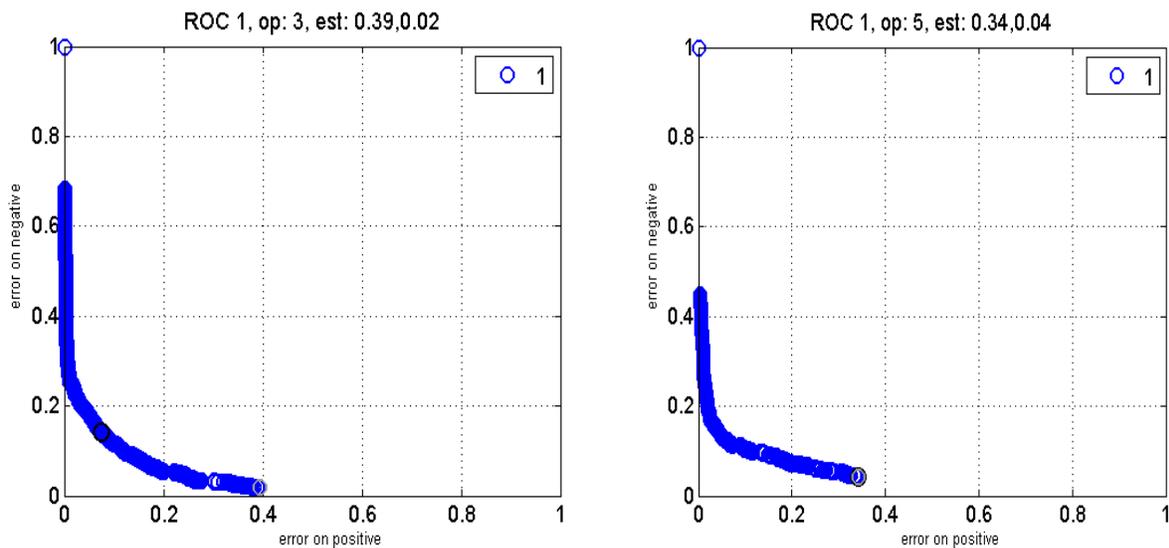


Figure 8. ROC curves of the second classifier in the 2C (left) and 4C view (right).

3. RESULTS

We present the detection results in the four and two chamber views. A first impression of the classifiers' performance can be given from the ROC curves (Figure 8). The curves indicate that our classifiers perform quite well. The detection accuracy was evaluated by measuring point-to-point errors, meaning the distance between the manually indicated position of the sought point and the one detected by the classifier. The detection errors are summarized in Table 1.

Table 1. Results on 2C and 4C view. * denotes statistically significant (paired t-test $p < 0.05$) with respect to the previous classifier. † denotes statistically insignificant difference with respect to the previous classifier.

		Average point errors 2C (mm)	Average point errors 4C (mm)
Step 1	Image center	6.5±2.5	7.9±4.1
Step 2	Image center	5.8±2.4*	7.6±4.8*
	Apex	7.2±7.0	8.0±5.0
	Mitral valve center	5.6±5.2	8.2±4.7
Step 3	Apex	7.1±6.7†	7.9±7.1†
	Mitral valve center	5.2±2.8†	4.8±2.3*
	Anterior/ Septal	5.8±3.5	5.6±2.7
	Inferior/ Lateral	4.5±3.1	4.0±2.6

4. DISCUSSION

We decided to follow a multilevel approach, in which the global location of the left ventricle is found first, and then rotation and scaling are found and further refined, to locate the landmark positions. The method produces a reasonably accurate estimation for the locations of the points. We can detect the points in both the 2C and 4C views (Table 1). The multilevel approach is justified, since the results are improving from stage to stage. The results improved significantly from Step 1 to Step 2, while from Step 2 to Step 3, although not always significant, there is still some improvement.

Some approaches which also identify the sought landmark points in addition to the long axis or the apical views' orientation angles have been published (Table 2). As already mentioned, Van Stralen's method¹ uses a combination of the circular Hough transform and dynamic programming. We used this approach to detect the apex and mitral valve center in exactly the same database. This approach does not detect rotation, resulting in larger errors than our method (Table 2).

Table 2. Comparisons with existing methods.

Point errors (av±SD)	Apex	Mitral valve center
Proposed	7.5±3.3mm	5.0±2.5mm
Orderud ¹⁰	8.4±3.5mm	3.6±1.8mm
Lu ⁵	4.5±3.5mm	3.6±3.1mm
Leung ⁸	7.6±4.8mm	4.5±2.9mm
Van Stralen ¹	14.7±8.6mm	8.4±5.7mm
Interobserver ⁸	7.1±2.9mm	3.8±1.3mm
Intraobserver ⁸	5.2±2.0mm	3.3±1.5mm

Lu's method⁵ results in smaller errors than our method does, probably due to better performance of the steerable features or the structure of the cascade. We have to note though, that their database was much larger (326 patients from whom 244 were used for training). Compared to the results given by Orderud et al.¹⁰ we find lower errors for the apex and mitral valve center. Leung et al.⁸ outperform our method, but we have to take into consideration that they aimed at registering markers in stress images, given the markers in the rest image, instead of detecting the standard views, which might have influenced the results. Finally, Table 2 shows comparisons with the interobserver and intraobserver variation. The fact that our errors are close to those provided by the interobserver and intraobserver variability is very promising, since these are the ground truth of manual annotation.

5. CONCLUSIONS

We showed that using our proposed approach, the anatomical landmark points can be detected accurately. Our errors were comparable to the interobserver and intraobserver variability of the manually indicated landmarks (following the same protocol as in this study), meaning that our approach is really promising. Additionally, the errors rates show that the original limitation of using a small database was surpassed by using artificial variations. The algorithm proved to be accurate and robust enough in both planes, so it could be combined with the already existing manual technique to extract the views and result in a fully automated method of detecting anatomical landmarks of the left ventricle.

REFERENCES

- [1] Van Stralen, M., Leung, K. Y. E., Voormolen, M. M., De Jong, N., Van der Steen, A. F. W., Reiber, J. H. C. and Bosch, J. G., "Time Continuous Detection of the Left Ventricular Long Axis and the Mitral Valve Plane In 3-D Echocardiography," *Ultrasound in Medicine & Biology* 34(2), 196-207 (2008).
- [2] Ma, M., Bosch, J. G., Reiber, J. H. C. and Lelieveldt, B. P. F., "Fully Automatic Estimation of Object Pose for Segmentation Initialization: Application to Cardiac MR and Echocardiography Images," *Proc. SPIE Medical Imaging*, 6512, 65123J, 1-8 (2007).
- [3] Georgescu, B., Zhou, X. S., Comaniciu, D. and Gupta, A., "Database-Guided Segmentation of Anatomical Structures with Complex Appearance," *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2, 429-436 (2005).
- [4] Carneiro, G., Georgescu, B., Good, S. and Comaniciu, D., "Detection and Measurement of Fetal Anatomies from Ultrasound Images using a Constrained Probabilistic Boosting Tree," *IEEE Transactions on Medical Imaging* 27(9), 1342-1355 (2008).
- [5] Lu, X., Georgescu, B., Zheng, Y., Otsuki, J. and Comaniciu, D., "AutoMPR: Automatic detection of standard planes in 3D echocardiography," *Proc. 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, ISBI 2008, 1279-1282 (2008).
- [6] Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M. and Comaniciu, D., "Fast Automatic Heart Chamber Segmentation from 3D CT Data Using Marginal Space Learning and Steerable Features," *Proc IEEE 11th International Conference on Computer Vision, ICCV 2007*, 1-8 (2007).
- [7] Nemes, A., Leung, K. Y. E., Van Burken, G., Van Stralen, M., Bosch, J. G., Soliman, O. I. I., Krenning, B. J., Vletter, W. B., Ten Cate, F. J. and Geleijnse, M. L., "Side-by-Side Viewing of Anatomically Aligned Left Ventricular Segments in Three-Dimensional Stress Echocardiography," *Echocardiography* 26(2), 189-195 (2009).
- [8] Leung, K. Y. E., Van Stralen, M., Nemes, A., Voormolen, M. M., Van Burken, G., Geleijnse, M. L., Ten Cate, F. J., Reiber, J. H. C., De Jong, N., Van der Steen, A. F. W. and Bosch, J. G., "Sparse Registration for Three-Dimensional Stress Echocardiography," *IEEE Transactions on Medical Imaging* 27(11), 1568-1579 (2008).
- [9] Viola, P. and Jones, M., "Rapid Object Detection using a Boosted Cascade of Simple Features," *Proc. IEEE Conf Computer Vision and Pattern Recognition* 2001, 511-518 (2001).
- [10] Orderud, F., Torp, H. and Rabben, S. I., "Automatic alignment of standard views in 3D echocardiograms using real-time tracking," *Proc. SPIE Ultrasonic Imaging and Signal Processing*, 7265, 72650D, 1-7 (2009).