Robust Live Tracking of Mitral Valve Annulus for Minimally-Invasive Intervention Guidance

Ingmar Voigt[1], Mihai Scutaru[2], Tommaso Mansi[3], Bogdan Georgescu[3], Noha El-Zehiry[3], Helene Houle[4], and Dorin Comaniciu[3]

[1] Imaging and Computer Vision, Siemens Corporate Technology, Erlangen, Germany
[2] Imaging and Computer Vision, Siemens Corporate Technology, Brasov, Romania
[3] Imaging and Computer Vision, Siemens Corporate Technology, Princeton, NJ
[4] Siemens Corporation, Healthcare Clinical Products, Ultrasound

Abstract. Mitral valve (MV) regurgitation is an important cardiac disorder that affects 2-3% of the Western population. While valve repair is commonly performed under open-heart surgery, an increasing number of transcatheter MV repair (TMVR) strategies are being developed. To be successful, TMVR requires extensive image guidance due to the complexity of MV physiology and of the therapies, in particular during device deployment. New trans-esophageal echocardiography (TEE) enable real-time, full-volume imaging of the valve including 3D anatomy and 3D color-Doppler flow. Such new transducers open a large range of applications for TMVR guidance, like the 3D assessment of the impact of a therapy on the MV function. In this manuscript we propose an algorithm towards the goal of live quantification of the MV anatomy. Leveraging the recent advances in ultrasound hardware, and combining machine learning approaches, predictive search strategies and efficient image-based tracking algorithms, we propose a novel method to automatically detect and track the MV annulus over very long image sequences. The method was tested on 12 4D TEE annotated sequences acquired in patients suffering from a large variety of disease. These sequences have been rigidly transformed to simulate probe motion. Obtained results showed a tracking accuracy of 4.04mm mean error, while demonstrating robustness when compared to purely image based methods. Our approach therefore paves the way towards quantitative guidance of TMVR through live 3D valve modeling.

1 Introduction

The mitral valve (MV), which ensures the unidirectional flow from the left atrium (LA) to the left ventricle (LV), is often affected by heart failure or degenerative diseases [5]. One particular MV pathology is MV regurgitation, where the valve does not close properly and blood can flow back to the LA. Following the success of transcatheter aortic valve repair, transcatheter mitral valve repair (TMVR) strategies are being explored by the medical industry. MitraClip™ is today an established treatment, but approaches for minimally-invasive annuloplasty or

complete valve repair are being developed [1]. The MV physiology brings important challenges to solve: MV anatomy is more complex than the aortic valve, including papillaries, chordae, complex leaflets geometry and non-symmetrical annulus. The pathologies are also more heterogeneous. For these reasons, advanced, quantitative 3D imaging is required during TMVR.

So far TMVR imaging guidance is mostly performed under 2D trans-esophageal echocardiography (TEE), since high-temporal resolution and color flow quantification are required for a complete monitoring of device deployment and its impact on MV function. Recent breakthroughs in ultrasound hardware have made possible to acquire non-stitched, full-volume 3D images at high frame-rates while combining both anatomical images (B-mode) and color-flow Doppler [5]. Such new probes pave the way to 3D TMVR guidance through MV physiology imaging, which would make the current therapies easier to perform, but also opening new therapeutic possibilities. One requirement is to have live and continuous 3D MV modeling, quantification and tracking within the interventional setup.

Several approaches for MV modeling from 3D TEE have been proposed in the past [4, 8, 10]. Yet, all of them still require from seconds to minutes to process a single frame and are therefore not adapted for continuous, live 3D valve modeling. At the same time, very efficient object tracking methods have been developed in other fields. In [3] for instance, a graph-based approach was proposed to track devices in 2D X-ray images. In [6], the authors propose a real-time tracking of four MV annulus landmarks from 2D TEE. To the best of our knowledge, no solution is able to track and model the complete MV annulus continuously in live 3D TEE images.

This paper proposes an approach towards live, continuous 3D MV annulus modeling from 3D TEE to support TMVR. Tracking a structure in live images requires a fast but accurate algorithm (10-15 or more frame per seconds) that does not drift over time. To cope with potentially large deformations due to probe motion, a combination of machine-learning based detection algorithm and fast optical flow tracking method is employed, which both leverage non-stitch, full-volume 3D TEE imaging (Sec. 2). The method was tested on 12 synthetic sequences (up to 46s-long) obtained by concatenating fully annotated 3D TEE data acquired in patients, which are continuously deformed according to random rigid deformations that mimicked probe motion (Sec. 3). Our approach achieved a point-to-mesh accuracy of 4.04mm (3978 frames in total) at a frame-rate of 12.5fps. Sec. 4 concludes the manuscript.

2   Methods

The proposed approach combines two components that complement each other in robustness and speed (Fig. 1): 1) a **learning-based detector** of MV location, pose and size as well as landmarks and annulus, which is robust to image alterations from transducer motion (image translation and rotation from probe flexing) and artifacts; and 2) an **optical-flow tracker**, which is capable of run-
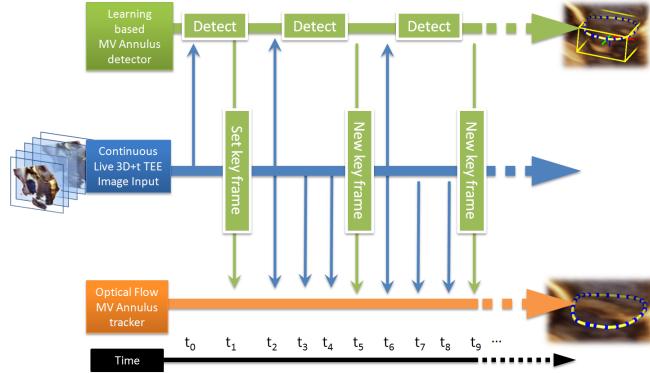
Fig. 1. System Overview. A continuous stream of images (**blue**) is being processed by two system components in parallel: a high frame rate optical flow tracker (**orange**), which is periodically re-initialized by a robust learning-based annulus detector (**green**).

ning at high frame rates, implements a key-frame approach for drift control while obtaining smooth and temporally-consistent motion estimates.

The two components, described in details in the following sections, are run in parallel. On system initialization, the detector component starts and determines valve presence in the ultrasound image. If the valve is found, the detector estimates the MV annulus curve, which is transfered to the flow tracker component for high frame-rate anatomy tracking. Subsequent images are being processed in parallel, i.e. the optical flow tracker processes each new image, while the detector runs in a separate thread to periodically re-detect valve presence and annulus, which is then fed back into the tracker to achieve robustness to large motion and ensure continuous control of tracking drift.

2.1 Mitral Valve Annulus Detector

The MV annulus detector is composed of a series of learning-based elements, as illustrated in Fig. 2. Learning-based detectors D_{rigid} of MV presence, location, orientation and size, as well as detectors of annulus landmarks and curve $D_{annulus}$, estimate model parameters $\phi_{rigid}(t)$ and $\phi_{annulus}(t)$ for an image $I(t)$ by maximizing the posterior probability p modeled by probabilistic boosting tree (PBT) [9] classifiers, $\phi(t) = argmax_{\phi}p(\phi|I(t))$. PBT classifiers are trained with Haar-like and steerable features from a manually generated database of ground truth locations of MV annulus and landmarks.

On system initialization, D_{rigid} is evaluated on the entire volume $I(t_0)$ using efficient search along increasing dimensionality of the parameter space employing the framework of marginal space learning (MSL) [11]. The search on the complete volume is repeated for subsequent images $I(t_1...t_n)$ until the MV is detected with high confidence ($p(\phi_{rigid}(t)) \geq 0.85$) on at least three consecutive images. Then the MV is assumed to be present within the volume and a region of interest (ROI)
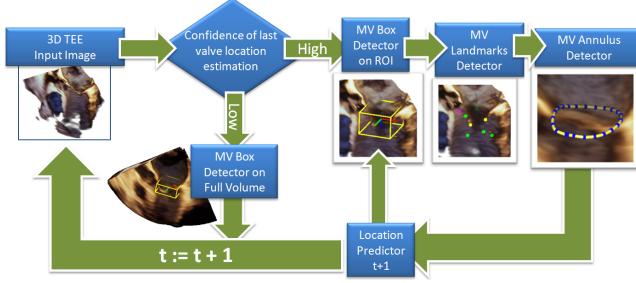
Fig. 2. Learning based hierarchical detector pipeline: on initialization the detector of MV presence, location, orientation and scale (box detector) runs on the full volume until at least three consecutive iterations were detected with high confidence estimates. Assuming MV presence, an ROI is computed and used for reducing the computational load. The ROI is updated at each iteration to account for probe motion.

$\Phi_{rigid}(t)$ is computed from the three last estimates to reduce the computational demand for estimating valve location. For subsequent detector invocations $t > t_n$, D_{rigid} is estimated by searching only within that ROI until the estimator confidence drops, i.e. $p(\phi_{rigid}(t)) < 0.85$, where the process is automatically reinitialized, i.e. runs again on the full volume.

To be robust to potential transducer motion, at each detector invocation a predictor P_{rigid} estimates the valve location for the next time the detector is invoked and updates the ROI center accordingly. In this work, P_{rigid} is empirically defined as the average trajectory over the last six iterations:

$$\Phi_{rigid}(t+1) = P_{rigid}(\phi_{rigid}) = \sum_{t-6}^{t} (\phi_{rigid}(t) - \phi_{rigid}(t-1)) \qquad (1)$$

Following the estimation of the rigid paramaters ϕ_{rigid}, $D_{annulus}$ detects anatomically defined landmarks – namely the left and right trigones as well as the postero-annular midpoint – by scanning respective classifiers over search ranges within ϕ_{rigid}. Finally the annulus is initialized as a closed curve by fitting a mean annulus shape composed of 58 points to the previously detected landmarks using thin plate splines (TPS). Specially trained PBT classifiers are evaluated by sampling the volume along planes that are transversal to the annulus curve at each curve point. The resulting curve $\phi_{annulus}(t)$ is spatially constrained using a point distribution shape model [2].

2.2   Optical Flow Key Frame Tracker

The optical flow key-frame tracker is a composite of two ordinary Lukas Kanade optical flow trackers [7]: a sequential tracker T_{seq}, which tracks landmarks from $I(t-1)$ to $I(t)$ and a second non-sequential key-frame tracker T_{key}, which registers the landmark defined on a past key frame $I(t_k < t)$ to the current frame

$I(t)$. The estimation results of both trackers are averaged to obtain the final estimate. In this way the tracker obtains smooth motion (via the frame by frame component T_{seq}) while reducing drift across cardiac cycles (via the key frame component T_{key}). The tracker estimates higher order terms iteratively by creating a warped image $I^1(t-1)$ out of the template image $I^0(t-1) = I(t-1)$ by applying the previously estimated motion vector field u^0 at locations x

$$I^1(x, t-1) = I^0(x + u^0, t-1)$$
$$\mathbf{M^1} u^1 = b^1$$
$$u^0 := u^0 + u^1$$

with $\mathbf{M^1}$ and b^1 computed from derivatives over space and time [7]. The scheme is repeated over six iterations, which was experimentally determined as point of convergence. In order to achieve high frame rates, the tracker runs directly on the spherical coordinate representation of the ultrasound image (acoustic space). Although the error is expected to increase with the distance to the transducer array due to the anisotropic image sampling, that limitation does not hinder our application as the mitral valve is typically located 50-70mm away from the transducer array, where the voxel distance is typically 1.2mm in the spherical coordinates representation of the image (assuming a typical angular resolution of about 1.3 degrees).

As the runtime of the detectors D (0.18 sec) exceeds the processing time of the optical flow trackers T (0.08 sec), both are run in parallel. The trackers are reinitialized each time the detector completes processing, by setting the respective annulus detection result $\phi_{annulus}(t_{key})$ and corresponding image $I(t_{key})$ as key frame for T_{key}, while T_{seq} restarts sequential tracking from the frame that D has finished at. For instance, following Fig. 1, let D start processing at time t_2 and complete at time t_5. The new key frame is set $t_{key} = t_2$ and T_{seq} restarts tracking by computing optical flow using $I(t_2)$, $I(t_6)$ and $\phi_{annulus}(t_2)$. While this means cardiac motion occurs in between t_2 and t_5, we observed that the annulus motion within 0.18s is typically small.

3   Experiments and Results

For our experiments, the detectors were implemented using CUDA version 3.2 and executed on a test machine using an nVidia Quadro K2100M graphics card, an Intel Core i7-4800MQ 2.70GHz processor and 16GB of RAM.

3.1   Dataset

The detector components were trained on 800 3D+t TEE volume sequences. To allow quantitative and thorough evaluation, the algorithm was fed with ever-looping recorded data from a separate set of 12 3D+t TEE volume sequences. The sets were manually annotated by an expert by manually fitting MV annulus ground truth models into the image data.

Fig. 3. Detection results for Mitral Valve annulus with 1.25mm mean error; yellow - ground truth, red: detector output

To test the method in operating room (OR) like conditions, the data were manipulated with rigid transformations that simulated probe motion based on the degrees of freedom that are typically observed during a clinical exam, i.e. transducer rotation along roll and yaw angles by 15 degrees (rotating and flexing) as well as shifts of 60mm collinear with the probe shaft (displacement along esophagus). The resulting sequences, obtained by looping, altering and concatenating an original sequence 26 times, ranged between 182 and 728 frames with frame rate between 5 to 32 fps, covering a total duration of 33 to 46 seconds. The volumes covered fields of view between $83° × 81° × 77$mm to $91° × 90° × 141$mm, typically covering both valves or mitral valve as well left and right ventricles. In total the resulting testing set comprised 3978 annotated 3D frames.

3.2 Quantitative Evaluation

For a quantitative analysis of the method, we evaluated the overall accuracy as well as tracking drift over time. Fig. 3 shows an example of ground truth annulus curve and obtained estimation result. As an accuracy metric the distance was computed from each point of the estimated MV annulus curve to the respective ground truth curve and vice versa, and finally averaged over the curve. Table 1 reports the overall accuracy of the proposed approach as well as detector and tracker components independently. The accuracy of the proposed approach ranges within the accuracy of the detector, the tracking components are subject to higher errors, due to drift. While the detector components ran with constant error and followed the probe motion, the trackers were subject to error accumulation over time, particularly in the presence of probe motion. This fact is particularly highlighted in Fig. 4, where the error distribution for the same categories showed significant amounts of outliers for the tracker only based approaches as a consequence of drift and changes in image appearance over time. On the other hand the detector ran at an average frame rate of 5.5fps, hence was not able to keep up with the imaging capabilities of state-of-the-art 4D imaging hardware. In contrast the tracker components operated at an average frame rate of 12.5fps, which is near interactive. Combining the two techniques, our approach was hence able to operate at high frame rates (12.5fps), while obtaining the same level of robustness as the detector, being robust to noise and probe motion.
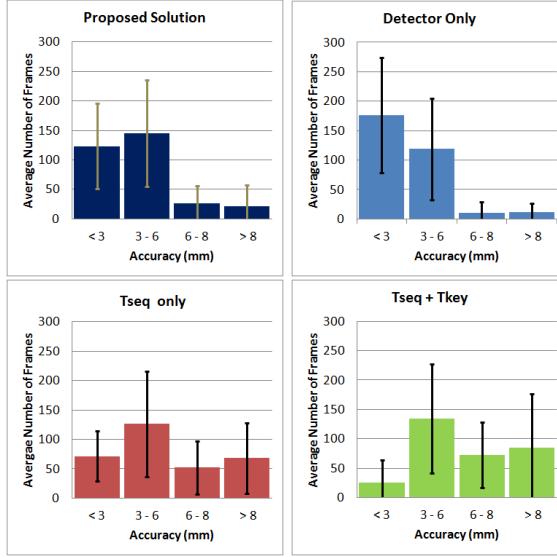
Fig. 4. Average error distributions over the testing set illustrated by error histograms for the proposed approach as well as isolated components. While the detector component operates with bounded error, the tracker components T_{seq} and T_{key} are subject to drift as can be seen by the higher number of error samples in histogram bins above 6mm error. Through the combination of all techniques, the proposed approach obtains the similar robustness as the detector while achieving the same frame rates an optical flow tracker

Table 1. Overall MV Annulus estimation accuracy reported in terms of mean $\pm$ std dev over the complete testing set.

Proposed approach	Detector only	$T_{seq} + T_{key}$	T_{seq} only
4.04±1.06 mm	3.37±0.69 mm	6.57±2.04 mm	5.28±1.19 mm

Finally we evaluated the average accuracy of the acoustic space tracking vs. Cartesian space tracking (same algorithm, but on Cartesian grids), particularly knowing that the tracking error could increase with the distance to the transducer array due to the non-linear acoustic sampling space. Both techniques exhibited similar performances with average errors of 4.36±2.2mm (Cartesian space tracking) vs. 4.13±1.43 (acoustic space tracking).

4   Conclusion

This paper presented an approach for robust tracking of MV annulus from 3D+t TEE volumetric data at high frame rates. It combines robust machine learning methods with image based tracking, hence enabling for robust live tracking within an interventional setting, where ultrasound imaging is subject to constant

changes of the field of view and motion. The approach was tested on a set of 3978 3D image frames generated out of 12 volume sequences of patient data, which included simulated probe motion to test the method in an operating room like setting. Reaching high frame rates and robustness, our approach enables real-time quantitative assessment of therapies and their impact on MV function, and could thus benefit emerging therapies such as TMVR.

References

1. Anyanwu, A.C., Adams, D.H.: Transcatheter mitral valve replacement: The next revolution?. Journal of the American College of Cardiology 64(17), 1820–1824 (2014)
2. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models—their training and application. Comput. Vis. Image Underst. 61(1), 38–59 (1995)
3. Heibel, H., Glocker, B., Groher, M., Pfister, M., Navab, N.: Interventional tool tracking using discrete optimization. IEEE Trans. Med. Imag.
4. Ionasec, R., Voigt, I., Georgescu, B., Wang, Y., Houle, H., Vega-Higuera, F., Navab, N., Comaniciu, D.: Patient-Specific Modeling and Quantification of the Aortic and Mitral Valves From 4-D Cardiac CT and TEE. Medical Imaging, IEEE Transactions on 29(9), 1636–1651 (2010)
5. Lancellotti, P., Zamorano, J.L., Vannan, M.A.: Imaging challenges in secondary mitral regurgitation unsolved issues and perspectives. Circulation: Cardiovascular Imaging 7(4), 735–746 (2014)
6. Li, F.P., Rajchl, M., Moore, J., Peters, T.M.: A mitral annulus tracking approach for navigation of off-pump beating heart mitral valve repair. Medical Physics 42 (2015)
7. Paragios, N., Chen, Y., Faugeras, O. (eds.): Mathematical Models in Computer Vision: The Handbook, chap. 15, pp. 239–258. Springer (2005)
8. Schneider, R.J., Perrin, D.P., Vasilyev, N.V., Marx, G.R., Pedro, J., Howe, R.D.: Mitral annulus segmentation from four-dimensional ultrasound using a valve state predictor and constrained optical flow. Medical image analysis 16(2), 497–504 (2012)
9. Tu, Z.: Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering. In: 10th IEEE International Conference on Computer Vision (ICCV 2005), 17-20 October 2005, Beijing, China. pp. 1589–1596 (2005), `http://doi.ieeecomputersociety.org/10.1109/ICCV.2005.194`
10. Weber, F.M., Stehle, T., Waechter-Stehle, I., Götz, M., Peters, J., Mollus, S., Balzer, J., Kelm, M., Weese, J.: Analysis of mitral valve motion in 4d trans-esophageal echocardiography for transcatheter aortic valve implantation. In: Statistical Atlases and Computational Models of the Heart-Imaging and Modelling Challenges, pp. 168–176. Springer (2015)
11. Zheng, Y., Georgescu, B., Ling, H., Zhou, S.K., Scheuering, M., Comaniciu, D.: Constrained marginal space learning for efficient 3d anatomical structure detection in medical images. In: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009), 20-25 June 2009, Miami, Florida, USA. pp. 194–201 (2009), `http://dx.doi.org/10.1109/CVPRW.2009.5206807`