Bulletin of the *Transilvania* University of Braşov • Vol. 8 (57) No. 1 - 2015
Series I: Engineering Sciences

ROBUST ANATOMY DETECTION USING TRACKING ALGORITHMS IN 4D ECHOCARDIOGRAPHY

Mihai SCUTARU[1]

Abstract: *Cardiac diseases are considered to be the most lethal classes of diseases, impacting approximately 1 in 3 individuals in western countries. Currently, the vast majority of surgical procedures targeting these afflictions involve percutaneous, non-invasive methods, which rely heavily on various imaging techniques, along with advanced image guidance through computer vision. This article aims to present an assessment of how various detection and tracking techniques such as Marginal Space Learning, Optical Flow and Belief Propagation can be combined to identify anatomical structure locations in 4D echocardiography, at high speed (34 ms) and with great accuracy (2.5 ± 1.2 mm mean error).*

Key words: *Computer Vision, Belief Propagation, Optical Flow, 4D TEE.*

1. Introduction

Cardiac disease diagnostic and treatment preparation relies heavily on modern imaging technologies. In order to have the correct course of action, the geometry of the heart during the heartbeat must be quantified. One of the most frequently used technologies is 4D (3D + time) trans-esophageal echocardiography (TEE). It relies on ultrasound to observe and record the heart motion. Physicians then either manually measure the obtained images or rely on automated tools to provide the necessary information.

Recently, several approaches have been proposed for anatomy detection from ultrasound images. In [2] the authors developed a method to segment the left-ventricle heart valves using Marginal Space Learning (MSL) [9]. An extended variant of this method is described in [7], which uses bio-mechanical constraints to further refine the obtained results. A successful detector plus Optical Flow tracker based approach is presented in [8] applied to the contour of the left ventricle, and highlights the drifting encountered when using trackers. Also Belief Propagation trackers have been successfully used to identify structures which are not visible in every stage of the cardiac cycle in [5].

This paper proposes a novel, fully automated approach for the identification of cardiac sub-anatomies by using 3D detection techniques and augmenting them with a composite tracking mechanism using Belief Propagation and Optical Flow. During the experimentation phase a comparison was done with respect to accuracy and speed of the utilized algorithms. Based on the obtained results,

[1] Automatics and Information Technology Dept., *Transilvania* University of Braşov.

the proposed composite approach has proven to be a viable method in the position identification of cardiac sub-anatomies.

2. Method

The basis for the proposed algorithm extensions is MSL. This method enables rigid detection of anatomical structures by identifying their position, orientation and scale. From this mechanism, the focus was the position detector which employs a Probabilistic Boosting Tree [6] classifier using 3D Haar-like features to identify a configurable number of candidates for the location of an anatomical structure. The correct selection of one of these candidates is essential in the accuracy of the detection process. Typically 60 position candidates are obtained, each one having a probability factor attached. This factor is used by the classifier to rank the candidates and therefore output only the top 60 starting from the one with the highest value. When the targeted sub-anatomy is clearly visible in the TEE data, such as a high contrast area, these candidates are generally clustered around a central position. This makes the decision process straightforward. A simple averaging operation is in most cases enough to produce an accurate result.

The important challenges are faced when the PBT classifier does not provide a singular cluster of candidates and there are no clear outliers. From such a sparse candidate cloud, in most situations it is extremely hard to make a clear, accurate choice. This is the situation in which refinement methods play a key role.

Unlike other imaging technologies used in cardiac intervention and therapy planning, 4D echocardiography has the distinct advantage of capturing the motion of the cardiac structures. When doing 3D detection, as is the case of MSL, all the information gathered for the decision process belongs to a single TEE frame. This comes as a downside of the fact that the method is also applicable on other imaging modalities that lack temporal information, such as Computerized Tomography. A possible extension of the capabilities of MSL would be to introduce a range of spatial constraints which limit the output to key regions. This has proved not to be robust enough, firstly due to high dynamics of the anatomic structures during the cardiac cycle, and also due to the high position variability between different patients.

Another approach would be to leverage information from neighbouring frames. The basic premise is that by running the PBT classifier on a range of frames and analysing the relationship between the obtained candidate clouds, a more accurate decision can be drawn for the final position result. In this context, two tracking algorithms were deployed and their individual performance measured.

The first algorithm introduced is Belief Propagation (BP) [3]. This works by arranging all the available candidates in a graphical model and associating a random variable $x(v)$ for each node. Nodes from candidates on adjacent frames are fully interconnected, and costs are attached to each node and edge. The nodes have the unary cost $\varphi(x_v)$ represented by the probability outputted by the PBT classifier. The edges have pair-wise cost of $\varphi(x_u,x_v)$ representing the Euclidian distance between candidates on neighboring frames. Having these defined, we can obtain the joint probability distribution of all the variables in the graph is expressed as a pair-wise Markov Random Field:

$$p(X) = \prod_{v \in V} \varphi_v(x_v) \prod_{u,v \in E} \varphi_{u,v}(x_u x_v) . \qquad (1)$$

In order to obtain the best set of candidates for the described graph, max product belief propagation is used, by defining m as the max-product message in (2).

$$m_{u\to v}^{(t)}(x_v) = \max_{x_u \in X_u}\left[\varphi_v(x_v)\varphi_{u,v}(x_u x_v)\left(\prod_{w\in\Gamma(u)\backslash v} m_{w\to u}^{(t-1)}(x_{uv})\right)\right]. \tag{2}$$

Let $\mu_v(x_v)$ be the estimated belief at node x_v and defined by the max-marginal of x_v:

$$\mu_v(x_v) = \max_{x'|x'_v = x_v} p(x'_1, x'_2, ..., x'_N), \tag{3}$$

which can be approximated as:

$$\mu_v(x_v) \ni \varphi_v(x_v)\left(\prod_{w\in\Gamma(u)\backslash v} m_{w\to u}^{(t-1)}(x_{uv})\right). \tag{4}$$

Given these max-marginals, the MAP (maximum a posteriori estimation) estimation is computed such that:

$$\hat{x} \in \arg\max_{x'} p(x'_1, x'_2, ..., x'_N). \tag{5}$$

Based on this result, we obtain the most efficient path through the graph.

Practically this method ensures that in adjacent frames, the candidates selected have the closest Euclidian distance between the two frames. Therefore if on one frame we have a clustered PBT result, this will also reflect on close frames, by propagating the more accurate location through distance differences.

The second tracking algorithm that has been used to refine the detection results is Optical Flow. The main functionality of this algorithm is to estimate the frame-to-frame motion of each pixel in an image, and is mostly used in cases where the motion has reduced amplitude and the pixel brightness is close to constant. In the first implementation proposed by Horn and Schunck [1] the equation to be minimized is:

$$E = \iint \Big[\left(I_x u + I_y v + I_t\right)^2 + \alpha^2\left(\|\nabla u\|^2 + \|\nabla v\|^2\right)\Big] \mathrm{d}x\,\mathrm{d}y. \tag{6}$$

In (6) the first term represents the brightness consistency constraint, and the second term is the smoothness constraint which uses the gradient of the optical flow vector. The equation also uses a regularization constant, α, which weights the contribution of the smoothness constraint.

Another popular variant of the Optical Flow algorithm is proposed by Lucas and Kanade [4] which extends the original concept by assuming that over a local neighbourhood, the flow is constant. Therefore the optical flow equations can be solved by using the least squares criterion.

The proposed method is based on the Lucas-Kanade method and consists of a multi-stage approach. On the first available frame, the Optical Flow tracker is initialized using the candidate range outputted by the position detector from MSL. On subsequent frames the tracker functions independently and tracks the position of the candidates frame by frame. Using just this approach is unfortunately subject to tracker drift issues so a feedback loop was implemented. This is where the second stage of the proposed method comes in. In order to permanently correct the tracker drift for each frame, the available candidates from the position detector can be leveraged. But using these directly as input for the tracker would negate any benefits of using optical flow in the first place. Therefore a result fusion stage was created. On every frame, this module takes the output of the position detector and the optical flow tracker and performs a weighted averaging operation. This can be configured with a specific weight which favours the results from the tracker. The candidates outputted from this result fusion stage are used as feedback input for the tracker on the next frame.

Using this multistage approach with weighted feedback loop ensures that during the tracking process we avoid drifting, and also retain the temporal information deduced from the Optical Flow algorithm.

3. Experiments and Results

As mentioned in the introduction the experiments done have focused on identifying the most suitable tracking algorithm to be used to augment the performance already obtained through Marginal Space Learning. First the accuracy of the mechanism was assessed on a testing set containing 26 expertly annotated TEE data sets. Another important factor that was analysed was the speed at which each mechanism variant was able to produce results. To that end all the tests were run on the same testing machine using an Intel Core i7-4800MQ 2.70 GHz processor and 16 GB of RAM.

The cardiac sub-anatomies used to benchmark all algorithm variants are the left (L), right (R) and non-coronary (N) aortic valve leaflet hinge points. These points mark the lower-most region of attachment of the aortic leaflets to the valve, and mark the beginning of the Left Ventricle Outflow Tract. These structures can be observed in Figure 1.

Since the position detector is an assisted learning algorithm, it was necessary to train it specifically for the target anatomies.

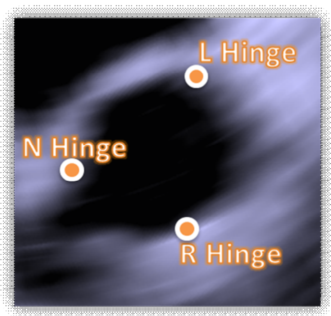

Fig. 1. *2D short-axis view of the anatomies used for algorithm evaluation*

In order to achieve this, a training set of 680 annotated data sets was assembled. This training set does not include the 26 datasets used for testing.

3.1. Quantitative Assessment

For a complete benchmarking of the algorithm performance, the first assessment was done using just the position detector. The second experiment attached the BP module after the position detector. The third experiment consisted of using the Optical Flow module in conjuncture with the position detector. Finally, the fourth configuration tested involved the position detector followed by the BP *n*+ode as well as the Optical Flow modules, named Composite Tracker.

The last configuration, aside from being able to assess the possible cumulative improvement brought by using both techniques, should also have a distinct performance advantage. As mentioned in the description of each algorithm, they rely on a set of input candidates which are processed over the available frames. By using the BP node first, we practically eliminate a large number of outliers and present only one candidate for the Optical Flow mechanism to track, instead of the initial 60. Therefore the further refinement of the results should not affect the overall speed of the system.

Alternatively, using the Optical Flow module first and then the BP would not produce any meaningful results. Due to the nature of the BP mechanism, it relies on a set of candidates for each individual frame, in order to optimize the result. Having a single candidate per frame as would be outputted by the Optical Flow nodes would prove futile since no optimization can be applied and the outputs would be identical to the input data.

The actual evaluation process takes the detection result for each of the targeted sub-anatomies and measures its Euclidian distance to the annotated position, which is considered the detection error. By collecting all these values over the entire testing set, the mean error, standard deviation and the maximum error of 90% of results were computed for all the four configurations. Also since the position detector always generates a set of candidates, in the evaluation process, the mean position was computed in order to have a single result. The evaluation results, represented by average results for the three targeted anatomies are presented in Table 1.

Algorithm Accuracy Table 1

Error[mm]	OP	BP	Comp	PosDet
Mean	2.10	2.45	2.55	2.25
Std Dev	1.33	1.13	1.13	1.33
90%	3.21	3.51	3.35	4.28

Since the primary purpose of utilizing tracking algorithms for detection performance augmentation is to reduce the number of outlier cases, the mean 90% error between the three target sub-anatomies has been computed and plotted in Figure 2.

From the plotted results it is obvious that the tracker based algorithms obtain significantly better results with respect to outlier removal. This, over time has the effect of reducing noise of the detected population, and keeping more in line with the dynamic of the anatomy.

By analysing the performance of the tracker augmented mechanism variants some important conclusions can be draw. As expected, the Optical Flow based mechanism is more efficient at removing the outlier cases than the BP method. This stems from the fact that the BP solution relies completely on the results of the position detector and only improves the selection process. In contrast, the Optical Flow solution generates its own set of candidates which are fused with the position detector results through the weighted averaging mechanism described above. Therefore we see less outlier when using the later solution.

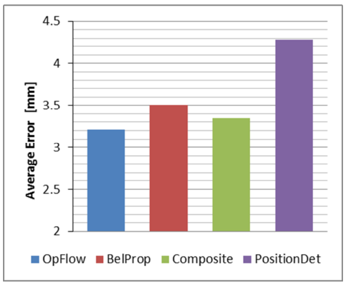

Fig. 2. *Average 90% outlier error rates [mm] for the tracker based solutions (blue, red, green) and position detector (purple)*

The Composite approach performance is situated between the other two solutions and it correlates with the other two results. Since the initial BP algorithm has proven to have slightly worse results, it was expected that the second stage tracker improves the outcome.

3.2. Detection Speed Assessment

Another very important factor in the decision between which mechanism is more suited for sub-anatomy location identification, is the speed at which it produces results. In order to accurate determine this, the algorithms were run on all available frames from the testing set and each individual run was measured. In total 1686 operations were performed with each algorithm and the average run time and its Standard Deviation were computed. The obtained results are shown in Figure 3.

As expected, the fastest method is the position detector as it only analyses one frame at a time.

The second fastest is BP as it takes into account all available candidates for the position detector. Also the variation in run time is visibly increased with the tracker, since they will run slower as more frames are analysed.

The important difference is observed between the Optical Flow tracker and the Composite mechanism. As detailed in the

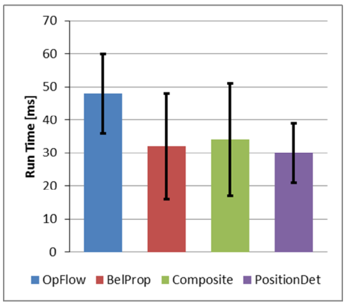

Fig. 3. *Average runtime [ms] with Standard Deviation (black) for the trackers (blue, red, green) and position detector (purple)*

Method section, a speed difference was expected since the number of processed candidates is dramatically reduced.

It is worth noticing that the highest encountered runtimes for the Composite tracker are near the mean value of the Optical Flow tracker, which should attest the significance of the performance improvement.

4. Conclusions

The goal of the presented work was to provide an overview of how tracking algorithms can improve the robustness of sub-anatomy detection from Ultrasound data. Also a novel approach was presented that aims to combine the strengths of two tracking algorithms. During the experimentation phase, the proposed composite approach has proven to efficiently eliminate outliers while also obtaining results without a significant speed decrease compared to the baseline non-tracking solution.

Acknowledgements

This paper is supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), ID134378 financed from the European Social Fund and by the Romanian Government.

References

1. Horn, B.K.P., Schunck, B.G.: *Determining Optical Flow*. In: Artificial Intelligence **17** (1981), p. 185-203.
2. Ionasec, R., Voigt, I., et al.: *Patient-Specific Modeling and Quantification of the Aortic and Mitral valves from 4D Cardiac CT and TEE*. In: IEEE Trans. Medical Imaging **29** (2010) No. 9, p. 1636-1651.
3. Kothapa, R., Pacheco, J., et al.: *Max-Product Particle Belief Propagation*. In: Master Thesis, Brown, 2011.
4. Lucas, B.D., Kanade, T.: *An Iterative Image Registration Technique with an Application to Stereo Vision*. In: Proceedings of Imaging Understanding Workshop, p. 121-130, p. 1981.
5. Scutaru, M., Voigt, I., et al.: *Robust Detection of Mitral Papillary Muscle from 4D Transesophageal Echocardiography*. In: STACOM 2015, LNCS **8896**, p. 249-256, 2015.
6. Tu, Z.: *Probabilistic Boosting-Tree: Learning Discriminative Models for Classification, Recognition, and Clustering*. In: IEEE ICCV 2005, p. 1589-1596, 2005.
7. Voigt, I., Mansi, T., et al.: *Robust Physically-Constrained Modeling of the Mitral Valve and Subvalvular Apparatus*. In: MICCAI 2011, Part III. LNCS **6893**, p. 504-511, 2011.
8. Yang, L., Georgescu, B., et al.: *Prediction-based Collaborative Trackers: A Robust and Accurate Approach toward 3D Medical Object Tracking*. In: IEEE Trans. Medical Imaging **30** (2011) No. 11, p. 1921-1932.
9. Zheng, Y., Georgescu, B., et al.: *Constrained Marginal Space Learning for Efficient 3D Anatomical Structure Detection in Medical Images*. In: IEEE Conf. Computer Vision and Pattern Recognition (CVPR'09), Miami, FL, 2009.