An Adaptive Method of Tracking Anatomical Curves in X-Ray Sequences

Yu Cao[1] and Peng Wang[2]

[1] Department of Computer Science and Engineering, University of South Carolina, Columbia, SC 29208, USA
[2] Corporate Research and Technology, Siemens Corporation, 755 College Road East, Princeton, NJ 08540, USA

Abstract. Tracking anatomical structures in X-Ray sequences has broad applications, such as motion compensation for dynamic 3D/2D model overlay during image guided interventions. Many anatomical structures are curve-like such as ribs and liver dome. To handle various types of anatomical curves, a generic and robust tracking framework is needed to track shapes of different anatomies in noisy X-ray images. In this paper, we present a novel tracking framework, which is based on adaptive measurements of structures' shape, motion, and image intensity patterns. The framework does not need offline training to achieve robust tracking results. The framework also incorporates an online learning method to robustly adapt to anatomical structures of different shape and appearances. Experimental results on real-world clinical sequences confirm that the presented anatomical curve tracking method improves the tracking performance compared to a baseline performance.

1 Introduction

Anatomical curve tracking in X-ray sequences have important applications in image guided surgery, such as motion compensations for 3D/2D dynamic model overlap, and needle insertion guidance. In such applications, it is desirable to compensate organ motions, in order that models can be properly visualized on X-ray images to guide the interventional procedures. Since many anatomical structures are curve-like, e.g., ribs and liver dome, as shown in Fig. 1, this paper presents a generic framework that tracks a wide variety of anatomical curve structures in a fluoroscopic sequence, to provide motion compensation information for image guided interventions.

The generic anatomical curve tracking in X-ray sequences is challenging. Shown in Fig. 1, the image quality is usually poor due to preferred low radiation dose. Anatomical structures have different shapes which undergo continuous changes due to breathing and cardiac motions during interventions. The motion could cause image blurs, and occlusions between different structures. Traditional tracking methods that are based on the constant image intensity assumptions, such as optical flow based method [1], will suffer from severe drifting when being used in such challenging situations. A variety of curve tracking methods have

N. Ayache et al. (Eds.): MICCAI 2012, Part I, LNCS 7510, pp. 173–180, 2012.

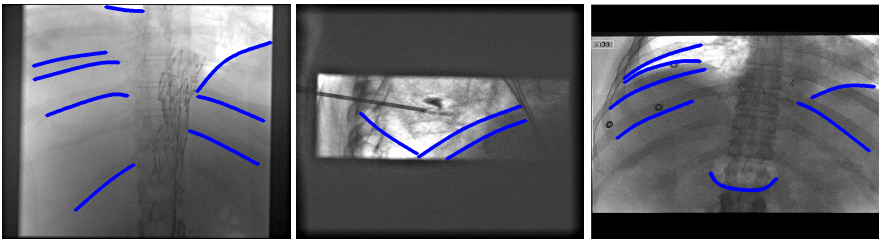

Fig. 1. Exemplar anatomical curves in X-ray image sequences, denoted by blue curves. The curves include ribs, liver dome, and diaphragm.

been developed for tracking anatomical structure in [2,5,10] or devices in [9,7,4]. In [2], an endocardium tracking method based on fusing optical flows with learning based shape subspace is used. In [5], a shape-based segmentation and tracking of deformable anatomical structures method is presented. In [9], a probabilistic tracking method is presented in order to track the guidewire in fluoroscopy images. This method requires offline learning to build classifiers specifically for guidewires. In [7], a deformable guidewire tracking method is proposed based on offline training. [4] proposed a graph-based guidewire tracking method which relies on B-spline curve model and strong geometric interests points. However, all the above methods focus on tracking specific anatomical structures or devices, rather than a generic form of anatomical structures.

In this paper, we present a probabilistic framework for anatomical curves tracking in X-ray image sequences. Through novel measurement models and probabilistic measurements fusion, the framework can capture the shape and image intensity variations of generic anatomical curves, and adapt to different tracking situations. Compared with existing methods [2,5,9,7,4], the presented framework makes the following contributions: 1) It is a generic approach and can be applied to track a variety of anatomical curve structures; 2) It introduces novel measurements in a Bayesian framework, including a novel formalization of combining optical flow, binary image patterns, and an online learning method as measurements for curve tracking; and 3) by the fusion of multiple measurements, the method is adaptive to motions, shapes and intensity pattern changes during tracking. The details of the tracking framework are explained in Section 2, and the experimental results in Section 3 demonstrates the effectiveness of presented tracking method.

2 Anatomical Curve Tracking Framework

2.1 Framework Overview

In this paper, the anatomical curve tracking is formalized with a Bayesian inference framework. A spline Γ is used to represent an anatomical curve structure to be tracked. To simplify the representation, a spline curve is sampled and noted as $\Gamma(\mathbf{x})$, where $\mathbf{x} = \{x_1, \ldots, x_N\}$ are the set of N uniformly sampled landmarks.

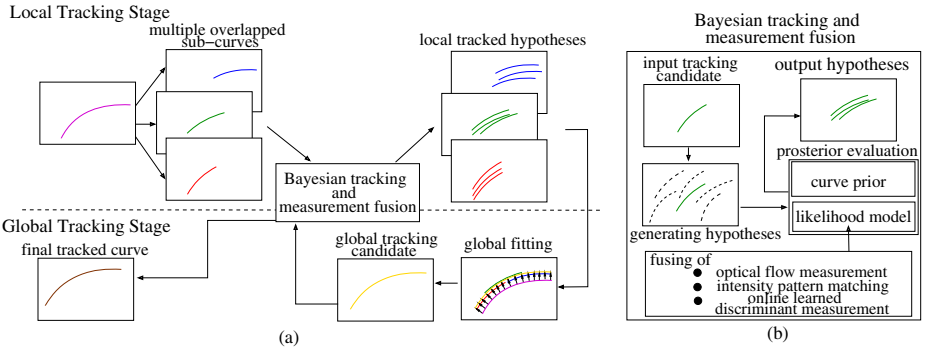

Fig. 2. (a) The hierarchical tracking framework; (b) The Bayesian framework for curve tracking

At the t-th frame, a curve candidate is noted as $\Gamma_t(\mathbf{x}_t^k)$ and the landmarks are $\mathbf{x}_t^k$. Base on the Bayesian rule and a commonly assumed Markov property for tracking, the posterior probability of a curve to be tracked has the form of

$$P(\Gamma_t(\mathbf{x}_t^k)|\mathbf{Z}_t) \propto P(\Gamma_t(\mathbf{x}_t^k))P(\mathbf{Z}_t|\Gamma_t(\mathbf{x}_t^k)) \tag{1}$$

where $\mathbf{Z}_t$ is the image observation on the t-th frame. The tracking result $\hat{\Gamma}_t(\mathbf{x}_t)$ is the curve that maximizes the posterior probability, i.e.,

$$\hat{\Gamma}_t(\mathbf{x}_t) = \arg\max_{\Gamma_t(\mathbf{x}_t^k)} P(\Gamma_t(\mathbf{x}_t^k)|\mathbf{Z}_t) \tag{2}$$

In Eqn. (1), $P(\Gamma_t(\mathbf{x}_t^k))$ is a prior probability of a curve candidate $\Gamma_t(\mathbf{x}_t^k)$. The prior probability imposes a constraint on 2D motions between a curve candidate $\Gamma_t(\mathbf{x}_t^k)$ and the previous tracked curve $\hat{\Gamma}_{t-1}(\mathbf{x}_{t-1})$. The likelihood model $P(\mathbf{Z_t}|\Gamma_t(\mathbf{x}_t^k))$ measures the likelihood of the tracking candidate $\Gamma_t(\mathbf{x}_t^k)$ based on the observation at the t-frame. To adaptively track anatomical curves in X-ray images, carefully designed prior models and likelihood measurement models are applied in our framework, with more details provided in Section 2.2 and Section 2.3, respectively.

Our tracking framework follows a hierarchical scheme, i.e, from a local scale to a global scale, as illustrated in Fig. 2.(a). At the local scale, a curve is divided into several sub-curves (each two neighbouring sub-curves have 50% overlaps with each other), and then each sub-curve is independently tracked. The local tracking provides multiple candidates for each local curve, and their combinations (using global fitting, Fig. 2.(a)) provide hypotheses for the subsequent global tracking, which is to maximize the posterior probability of the whole curve. Both local and global tracking follow the same Bayesian tracking framework, as shown in Fig. 2.(b). The hierarchical tracking scheme allows the adaptive and effective tracking of an anatomical curve: first, the local tracking allows flexible affine deformation for each sub-curve whose motion may be different from the other part of the curve; second, the global tracking combines all the possible

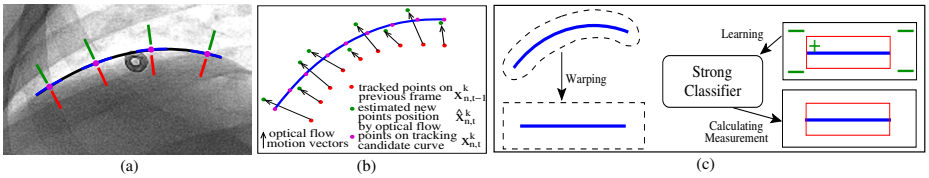

Fig. 3. Illustration of: (a) intensity pattern; (b) optical flow measurement; (c) online learned discriminant measurement

combinations from local tracking results and impose constraints from a global shape, so it can prevent possible drifting of local sub-curves.

2.2 Curve Prior

Most motions of anatomical structures such as ribs and livers, are mainly caused by breathing motions, and less impacted by cardiac motions. So we can conveniently assume that the motion can be approximated by an affine transformation. At the t-th frame, a curve candidate $\Gamma_t(\mathbf{x}_t^k)$ is transformed from the tracked curve at a previous frame $\hat{\Gamma}_{t-1}(\mathbf{x}_{t-1})$, via an affine transformation:

$$\Gamma_t(\mathbf{x}_t^k) = A_k \hat{\Gamma}_{t-1}(\mathbf{x_{t-1}}) \tag{3}$$

A_k is an affine transformation matrix. Through our experiments, we find such motions can well describe the anatomical curve movements in X-ray images.

By further decomposing A_k with the QR decomposition $A_k = Q_k R_k$ [8], we can retrieve the affine motion parameters set $\mathbf{M}_k$. Based on $\mathbf{M}_k$ of each curve candidate, we define the curve prior as:

$$P(\Gamma_t(\mathbf{x_t^k})) \propto G(\mathbf{M}_k|\mathbf{0}, \Sigma) \tag{4}$$

where $G(\mathbf{M}_k|\mathbf{0}, \Sigma)$ is a Gaussian distribution with zero mean. Σ is the diagonal covariance matrix. Without sacrificing the generalization of the algorithm, Σ is empirically set and the same parameters are applied to all the data.

2.3 Likelihood Measurements

It is challenging to robustly and adaptively model the shape and appearance of a generic curve, as it can be of a variety of continuously changing shapes and image intensities. In this work, we achieve the robust tracking in two ways: 1) designing novel measurements that can effectively model curves' shape and appearance during tracking; 2) fusing multiple measurements. The measurement models used in this framework include optical flow measurement noted as $P^O(\mathbf{Z}_t|\Gamma_t(\mathbf{x}))$, intensity pattern matching noted as $P^I(\mathbf{Z}_t|\Gamma_t(\mathbf{x}))$, and online learned discriminant measurement noted as $P^B(\mathbf{Z}_t|\Gamma_t(\mathbf{x}))$. Each of the measurement models is explained at subsequent sections. By fusing multiple measurements, the likelihood measurement model can be written as

$$P(\mathbf{Z}_t|\Gamma_t(\mathbf{x}_t)) = P_O P^O(\mathbf{Z}_t|\Gamma_t(\mathbf{x}_t)) + P_I P^I(\mathbf{Z_t}|\Gamma_t(\mathbf{x}_t)) + P_B P^B(\mathbf{Z}_t|\Gamma_t(\mathbf{x}_t)) \tag{5}$$

where P_O, P_I, P_B are the prior probability of each measurement model, and $P_O + P_I + P_B = 1$. In the experiments, P_O, P_I and P_B can be equally weighted.

Optical Flow Measurement. Optical flow assumes constant image intensity of corresponding pixels, and provides motion estimation of individual pixels. However, optical flow suffers "aperture" and "drifting" problems on a homogeneous region, thus we use the optical flow results as one of measurements, instead of fully depending on optical flow results. Assuming that the new position of a point $x^k_{n,t-1}$ in the curve $\hat{\Gamma}_{t-1}(\mathbf{x}^k_{t-1})$ estimated by the optical flow method is $\hat{x}^k_{n,t}$, illustrated in Fig. 3.(b), we can define the likelihood of $x^k_{n,t}$, the new position of $x^k_{n,t-1}$, as:

$$P(\mathbf{Z}_t|x^k_{n,t}) = G[\hat{x}^k_{n,t}; \sigma_o](x^k_{n,t}) \tag{6}$$

where $G[\hat{x}^k_{n,t}; \sigma_o]$ is a Gaussian distribution with $\hat{x}^k_{n,t}$ as the mean, and σ_o as the standard deviation. The measurement of the a curve is therefore the integration of the measurements of all the N points along the curve $\mathbf{Z}_t$, given that the landmarks are uniformly sampled along the curve:

$$P^O(\mathbf{Z}_t|\Gamma_t(\mathbf{x}^k_t)) = \frac{1}{N}\sum_{n=1}^{N} P(\mathbf{Z}_t|x^k_{n,t}) \tag{7}$$

Intensity Pattern Matching. In the anatomical curve tracking, it is observed that the intensity patterns of a curve needs to remain similar between two successive images. However, directly using image intensity for template matching leads to poor results due to low image quality of X-ray. In this method, three intensity patterns, similar to the LBP (Local Binary pattern) [6], are defined to describe the curve intensity. For simplicity, we name them SLBP, Spline and Bar pattern. As shown in Fig. 3.(a), given a tracking candidate $\Gamma_t(\mathbf{x}^k_t)$ shown as the black curve and its landmarks shown as magenta points, then for each landmark $x^k_{n,t}$ on the curve, we define three profiles: (1) along positive curve norm with average intensity $I^{\mathcal{P}}_n$ (green bar); (2) along $\Gamma_t(\mathbf{x}^k_t)$ and centered at $x^k_{n,t}$ with average intensity $I^{\mathcal{S}}_n$ (blue curve); and (3) along negative curve norm with average intensity $I^{\mathcal{N}}_n$ (red bar). Then a binary 3-tuple (a^1_n, a^2_n, a^3_n) can be defined as

$$a^1_n = \begin{cases} 1, & I^{\mathcal{P}}_n > I^{\mathcal{N}}_n \\ 0, & \text{else} \end{cases} \qquad a^2_n = \begin{cases} 1, & I^{\mathcal{N}}_n > I^{\mathcal{S}}_n \\ 0, & \text{else} \end{cases} \qquad a^3_n = \begin{cases} 1, & I^{\mathcal{P}}_n > I^{\mathcal{S}}_n \\ 0, & \text{else} \end{cases} \tag{8}$$

Then we can define SLPB, Spline and Bar intensity patterns (I_L, I_S and I_B) as follows respectively:

$$\begin{aligned} &\mathbf{I}_L(\Gamma_t(\mathbf{x}^k_t)) = (a^1_1, a^2_1, a^3_1, \cdots, a^1_N, a^2_N, a^3_N), \quad \mathbf{I}_S(\Gamma_t(\mathbf{x}^k_t)) = (I^{\mathcal{S}}_1, \cdots, I^{\mathcal{S}}_N), \\ &\mathbf{I}_B(\Gamma_t(\mathbf{x}^k_t)) = (I^{\mathcal{P}}_1, I^{\mathcal{S}}_1, I^{\mathcal{N}}_1, \cdots, I^{\mathcal{P}}_N, I^{\mathcal{S}}_N, I^{\mathcal{N}}_N). \end{aligned} \tag{9}$$

The measurement of intensity pattern $P^I(\mathbf{Z}_t|\Gamma_t(\mathbf{x}_t^k))$ is then defined based on the correlation between each curve and the template (tracked curves at a previous frame) as:

$$P^I(\mathbf{Z}_t|\Gamma_t(\mathbf{x}_t^k)) = 0.5 * (\frac{|\mathbf{I}(\Gamma_t(\mathbf{x}_t^k)) \cdot \mathbf{I}_{template}|}{|\mathbf{I}(\Gamma_t(\mathbf{x}_t^k))| \cdot |\mathbf{I}_{template}|} + 1) \tag{10}$$

Online Learned Discriminant Measurement. To further improve the tracking robustness, we introduce a discriminative measurement model. Different from previous two measurement models, the discriminant measurement further distinguish the curve from backgrounds. Since we aim at tracking generic anatomical curves, we use the online boosting method [3] (rather than offline methods) to build the online discriminant model. Given a tracked curve $\hat{\Gamma}_{t-1}(\mathbf{x})$ in $(t-1)$-th frame, its neighboring region in a X-ray image can be warped into a straightened image, as illustrated in Fig. 3.(c). The straightened image together with positive (sampled within upper red box in Fig. 3.(c)) and negative samples (randomly sampled surround the red box) are used to train an online-boost tracker during the tracking [3]. Through online updating, the discriminant measurement model enables itself to adapt to the appearance changes during tracking. During tracking, for each curve candidate, we input its straightened image patch (lower red box in Fig. 3.(c)) to online-boost tracker. The probabilistic score from the online-boost tracker is used as the online discriminant measurement:

$$P^B(\mathbf{Z}_t|\Gamma_t(\mathbf{x}_t)) = \sum_{i=1}^{N'} \alpha_i \cdot h_i^{sel} \tag{11}$$

where h_i^{sel} are the selectors of the online boost tracker, α_i are the linear combination weights, and N' is the number of selectors in the tracker [3].

3 Experiments

Our framework is validated on a set of X-ray image sequences acquired in clinical scenarios. The dataset includes 22 sequences, more than $2,000$ frames in total. The anatomical structures in the dataset include lung, ribs, diaphragm, and liver. The frame sizes range from 512×512 to 1024×1024 where the physical distance between neighbouring pixels is 0.1mm. The dataset is further divided into two sub-datasets: Dataset-1 and Dataset-2. Dataset-1 contains 22 sequences which are acquired under normal radiation dose and have reasonably distinguishable curves. Dataset-2 is only consist of challenging sequences (11 sequences) in Dataset-1, where the images are acquired under lower radiation dose which makes curves even hard for human to observe.

For each sequence, we annotate multiple (2 to 8) curves along anatomical structures throughout the image sequence as the ground-truth. The annotation on the first frame is used to initialize tracking (physicians we collaborated with agree to manually specify the curves of interests to track), and the annotations

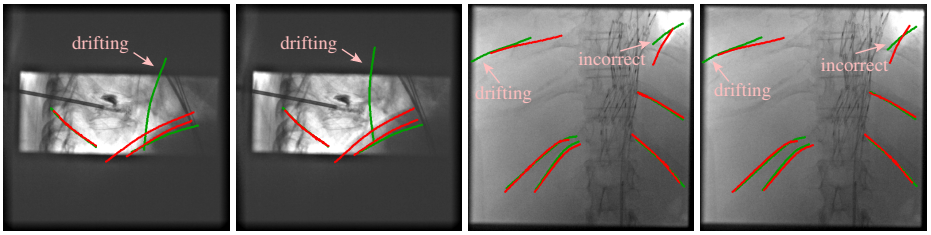

Fig. 4. Qualitative tracking results comparison between only using optical flow measurement (green curves) and using SLBP+optical flow measurements+discriminative online-learned measurement (red curves)

Table 1. Comparison of tracking error rates of different methods and combinations: (I)SLBP+Spline, (II)SLBP+Bar, (III)Spline+Bar, and (IV)SLBP+Spline+Bar

Evaluation on Dataset-1									
Threshold τ	Physical Distance	Optical flow	SLBP	Spline	Bar	(I)	(II)	(III)	(IV)
10	1mm	10.0%	7.7%	9.5%	8.1%	7.7%	8.6%	8.5%	8.9%
Evaluation on Dataset-2									
Threshold τ	Physical Distance	Optical flow	SLBP	Spline	Bar	(I)	(II)	(III)	(IV)
10	1mm	12.3%	8.6%	10.9%	9.2%	8.7%	9.5%	9.6%	10.4%

on the rest frames are used for quantitative evaluations. We define the following quantitative evaluation metrics. For a landmark on a curve, we calculate the shortest distance d from this landmark to the corresponding ground-truth curve. For a pre-defined threshold τ, if $d <= \tau$, we consider the landmark as being correctly tracked. We use the averaged tracking error rate (rate of incorrectly tracked landmarks number over total landmarks number) as tracking performance.

We tested the proposed framework and compared the tracking performance between different measurement combinations. Figure 4 shows some visual tracking results. We notice that the results from only using optical flow measurement suffer from drifting problem, especially when the image quality is low. The average tracking error rate curves of using different measurement combinations are shown in Fig. 5. From the figure, using SLBP+optical flow+online-learned discriminative measurements achieves the best performance on both Dataset-1 and Dataset-2 from thresholds 5 to 15, except only a few thresholds.

Further from Table 1, at $\tau = 10$ where the physical distance is 1mm, using SLBP+optical flow+online-learned discriminative measurements achieves the best performances on Dataset-1 and Dataset-2 with error rates 7.7% and 8.6%, respectively. This is much better than only using optical measurements whose tracking error rates are 10% and 12.3%, respectively. This demonstrates that the presented tracking framework improves the baseline performance (when only using optical flow) by a significant amount.

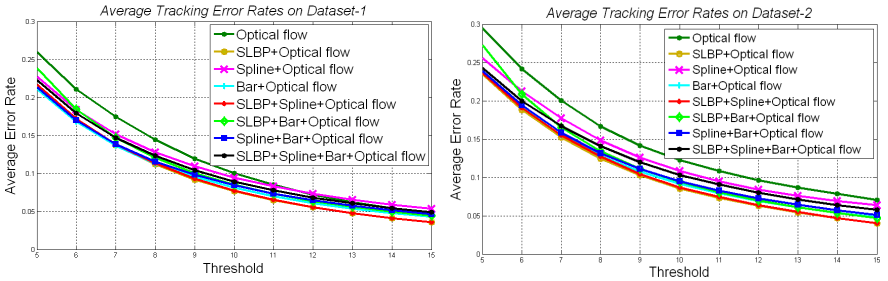

Fig. 5. The quantitative comparisons between the results of using different measurement combinations on: (a)Dataset-1 and (b)Dataset-2

4 Conclusion

In this paper, we present a probabilistic framework for adaptively tracking generic anatomical curve structures in X-ray images. We demonstrate the effectiveness of integrating appearance pattern, shape and motion information in tracking curve structures. We plan to extend the method to other image modalities such as ultrasound, and to further explore its usages in clinical applications.

References

1. Comaniciu, D.: Nonparametric information fusion for motion estimation. In: CVPR (2003)
2. Georgescu, B., Zhou, X.S., Comaniciu, D., Rao, B.: Real-Time Multi-model Tracking of Myocardium in Echocardiography Using Robust Information Fusion. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) MICCAI 2004, Part II. LNCS, vol. 3217, pp. 777–785. Springer, Heidelberg (2004)
3. Grabner, H., Grabner, M., Bischof, H.: Real-time tracking via on-line boosting. BMVC 1, 47–56 (2006)
4. Honnorat, N., Vaillant, R., Paragios, N.: Graph-Based Geometric-Iconic Guide-Wire Tracking. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part I. LNCS, vol. 6891, pp. 9–16. Springer, Heidelberg (2011)
5. Mignotte, M., Meunier, J., Tardif, J.C.: Endocardial Boundary Estimation and Tracking in Echocardiographic. Pattern Analysis & Applications 4, 256–271 (2001)
6. Ojala, T., Pietikainen, M., Harwood, D.: Performance evaluation of texture measures with classification based on kullback discrimination of distributions. In: ICPR, pp. 582–585 (1994)
7. Pauly, O., Heibel, H., Navab, N.: A Machine Learning Approach for Deformable Guide-Wire Tracking in Fluoroscopic Sequences. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part III. LNCS, vol. 6363, pp. 343–350. Springer, Heidelberg (2010)
8. Shoemake, K., Duff, T.: Matrix animation and polar decomposition. In: Proceedings of the Conference on Graphics Interface 1992, pp. 258–264. Morgan Kaufmann Publishers Inc. (1992)
9. Wang, P., Chen, T., Zhu, Y., Zhang, W., Zhou, S.K., Comaniciu, D.: Robust guidewire tracking in fluoroscopy. In: CVPR, pp. 691–698 (2011)
10. Wang, P., Zhou, S., Szucs, M.: Endocardium tracking by fusing optical flows in straightened images with learning based detections. In: ISBI (2011)