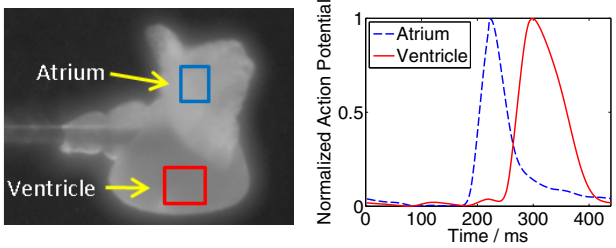
35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

On Identification of Sinoatrial Node in Zebrafish Heart Based on Functional Time Series from Optical Mapping

Weiguang Ding, Eric Lin, Amanda Ribeiro, Marinko V. Sarunic, Glen F. Tibbits, and Mirza Faisal Beg

***Abstract*— As a vertebrate cardiovascular model, the zebrafish heart has been used extensively in physiology research to study cardiac development and human cardiac disease. Optical mapping techniques provide an effective approach to record action potential propagation in the zebrafish heart. However, manual analysis of functional time series recorded from optical mapping can be laborious and time consuming. In this paper, a novel pipeline is proposed to assist physiologists in identifying the sinoatrial node (SAN) in zebrafish heart based on functional time series. First, the original optical mapping data are enhanced and averaged. Next, the heart is divided into small regions, and representative average time series are calculated. A 'discretization of derivative' (DoD) process is performed to model physiological similarity between signals. Finally, grouping is done on the DoD transformed representation, which is found to produce physiologically meaningful classification for SAN identification.**



(a) Optical mapping image (b) Typical signals

Fig. 1. Optical mapping zebrafish heart image and typical signals: (a) labelled optical mapping image zebrafish image, (b) plots of typical atria and ventricle signals from bounding boxes in (a).

I. INTRODUCTION

As a vertebrate cardiovascular model, the zebrafish heart has been used extensively in physiology research to study cardiac development and human cardiac disease [1]. The zebrafish heart is comprised of cells of different types, including nodal, atrial and ventricular cells. Each of these cell types has distinct mechanical and electrical characteristics that determine the heart's ability to provide adequate blood flow to the rest of the body. Accordingly, each cell type has a distinct action potential waveform, a specific time series of membrane depolarization that closely relates to its physiological function. The action potential waveform can be recorded by optical mapping techniques [2], [3], which makes the electrical activity of the heart visible and enables cell type determination by analyzing the time series of cardiac activation and membrane depolarization among different individuals and under various experiment conditions. Figure 1 shows the optical mapping zebrafish heart image and typical signals. Identification of sinoatrial node (SAN) cells is of particular interest since they are responsible for the heart's intrinsic rate and respond to extrinsic factors that modulate heart rate [4]. Recently, functional pacemaker of adult zebrafish heart has been identified as a ring around the venous pole [5] based on a complex combination of microscopic examination, gene expression pattern reconstruction, reporter transgenics and electrophysiology. However, this method is complicated, expensive, time consuming and still needs to be validated by other equally involved independent methods. On the other hand, considering that anatomical structure is hard to identify in optical mapping data due to the tissue transparency, a simple, cheap and fast method is desired for locating SAN node in every set of recorded optical mapping data. However, direct observation of time series from optical mapping data is difficult due to: 1. SAN cells are small and intermixed among larger atrial cells and 2. low signal-to-noise ratio and low intensity range of original optical mapping data. This makes preprocessing techniques essential even for manual identification. Furthermore, to enable fast identification of SAN based on functional time series, automatic analysis techniques are required. Previous analyses of cardiac optical mapping, including spatial-temporal filtering, activation map, conduction velocity and APD maps were reviewed in [6]. SAN was mainly identified as region with early depolarization from activation map, which is physiologically valid and applicable with fast cameras. However, when slower cameras are used as a trade-off for higher spatial resolution, activation map is less accurate, hence the resulting SAN identification. The shape of the entire action potential waveform may provide richer information for SAN identification, especially in the case of using low frame rate camera. For example, slow diastolic depolarization slope has been used in manual identification isolated coronary-perfused human sinus node from optical mapping data [7].

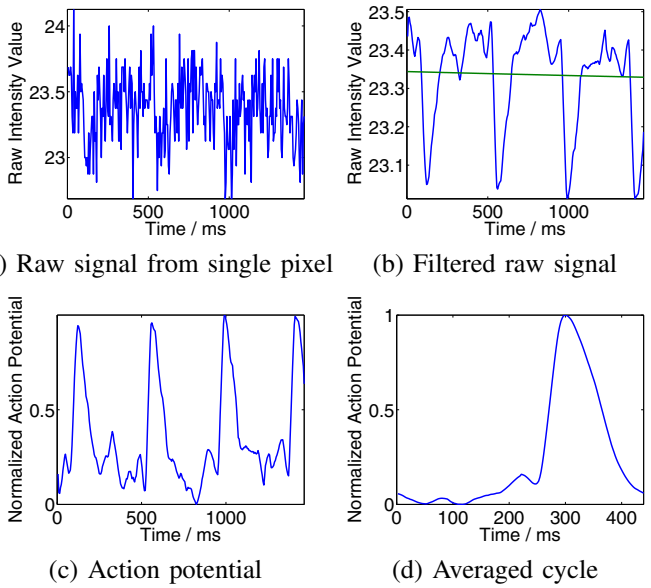
This paper describes a novel pipeline of automatic data analysis tools for analyzing functional optical mapping recordings to locate the SAN and other areas of interest in the

We acknowledge funding for this project from Canadian Institute of Health Research, Michael Smith Foundation for Health Research and the National Science and Engineering Research Council.

Weiguang Ding, Marinko V. Sarunic and Mirza Faisal Beg are with School of Engineering Science, Faculty of Applied Science, Simon Fraser University, 8888 University Dr, Burnaby, BC, Canada wding@sfu.ca, msarunic@sfu.ca, mfbeg@sfu.ca

Eric Lin, Amanda Ribeiro and Glen F. Tibbits are with Department of Biomedical Physiology and Kinesiology, Faculty of Science, Simon Fraser University, 8888 University Dr, Burnaby, BC, Canada elin@sfu.ca, alr3@sfu.ca, tibbits@sfu.ca

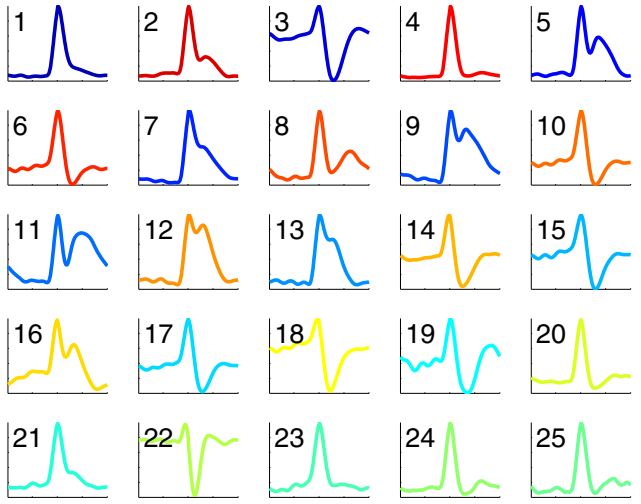
978-1-4577-0216-7/13/$26.00 ©2013 Crown



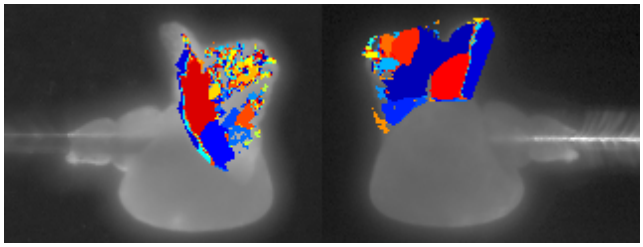
(a) Raw signal from single pixel (b) Filtered raw signal

(c) Action potential (d) Averaged cycle

Fig. 2. Preprocessing illustration: (a) raw signal from single pixel, (b) the spatial and temporal filtered signal; the green line is the fitted photobleaching trend, (c) photobleaching is removed, and signal is reversed and scaled, (d) averaged action potential from multiple cycles.



(a) Averaged signals from different clusters



(b) Area classfication on heart

Fig. 3. Result of the first stage of clustering: (a) averaged signals from each clusters, (b) zebrafish heart with color coded patches labelled. Each signal in (a) represent the the average in the cluster of the same color in (b). (b) contains images from anterior and posterior sides.

zebrafish heart. Representative signals from each identified region can then be used by scientists to investigate aspects of heart activity under different experimental conditions. This process will free physiologists from intensive and repetitive work of manually analyzing and comparing large amount of functional time series to locate the SAN.

II. METHODS

Zebrafish hearts were isolated and then labelled with the fluorescent voltage sensitive dye RH-237 [8] by immersion in 8 μM solution. 10 μM blebbistatin excitation-contraction uncoupler [9] was used to inhibit contraction and motion artifact. Excitation illumination was provided by a 200 mW 532 nm DPSS laser (Laserglow Technologies). Spectral filtering was provided by a 560 nm long pass dichroic and a 710 nm long pass emission filter (Omega Optical). Fluorescent images of the heart were acquired using two GE680 (Allied Vision Techologies) cameras at a frame rate of 205 fps and resolution of 640×480, followed by a 4 by 4 binning to reduce the resolution down to 160×120.

A. Preprocessing

Due to the low signal-to-noise ratio and low intensity resolution of the original optical mapping data, both spatial and temporal averaging were done on the original data. Then the drifting caused by photobleaching of the fluorophore [10] and other effects was compensated for by first fitting a quadratic curve [6] to each pixel time series and subtracting it from original signal followed by scaling to range [0, 1].

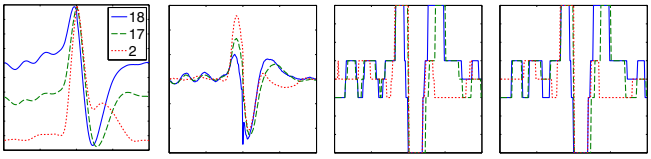
After enhancing, the time series waveform for each pixel in the atrium was taken and divided into constituent 'cycles' that make up the waveform, using the peaks of average atrium waveform. These cycles are then linearly interpolated to have the same length and averaged to get a single averaged cycle. This averaged cycle for each pixel is henceforth referred to as 'signal' or 'time series'. This preprocessing is illustrated in Figure 2.

In order to find the SAN, which is known to initiate the depolarization wave, the 'late' signals, defined as those that reached their peak value after the average atrial peak, were removed thereby restricting the processing only to those pixels in the atrium that depolarize early. However, the method for clustering described below is general, and can be used for clustering over any functionally interested region and is not specific to the atrium.

B. Clustering of Functional Time Series

The shape of the cardiac action potential waveform of different individual zebrafish hearts and under different experimental conditions shows considerable variability. This variability makes logical algorithm design and supervised learning difficult to perform, because 'new' waveforms will be out of view of predefined logic or encountered data. To make the algorithm robust to these sources of variability, we propose here a novel two-step unsupervised labeling framework.

For the first stage, each signal was treated as a point in high dimensional space, in which clustering will be done.



(a) (b) (c) (d)

Fig. 4. Procedure of performing discretization of derivative: (a) signals 2, 17, 18 from Fig.3. We anticipated that after DoD transformation, 17 and 18 will become similar and be different from 2, (b) derivative of these 3 signals, (c) discretized derivative, (d) After mode filter is performed. In (d), 17 and 18 are more similar compared to in the original vector space.

Common partional clustering algorithms such as k-means, fuzzy k-means and EM algorithm [11] can be used for this task. In this paper, we assume that cells of similar function will likely be spatially adjacent. Hence, the fuzzy clustering with spatial constraints algorithm [12] was adopted. This method penalizes spatially scattered clusters.

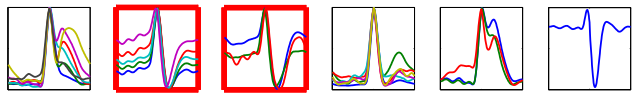
For the second stage of grouping, the clusters found in the first stage of clustering will be further grouped into larger clusters of similar shaped signals. Clustering the signals in their original vector space was found to be insufficient for this purpose. This is because the distance in original vector space does not represent the physiological similarity of signals from different cardiac cells. Instead, having similar rates of change and reaching 'hills' and 'valleys' in similar time, in other words, having similar 'trends', is likely correlated to similarities in physiology [13].

To represent the descriptive 'trends' by quantitative values, a novel 'discretization of derivative' (DoD) procedure is proposed. In DoD, the derivative of the original signal is calculated. Derivative is a good representation of 'trends'. However, sometimes signals sharing the same trend can be quite different in the magnitude of derivative values. This makes clustering in the derivative vector space also difficult. To cope with this problem, a discretization step is performed. The numerical derivative values were then thresholded to different integer values, which represent 'fast increasing', 'slow increasing', 'flat', 'slow decreasing' and 'fast decreasing'. Mode filter will be used to smooth the trends acquired from DoD procedure. The DoD process is illustrated in Fig.4.

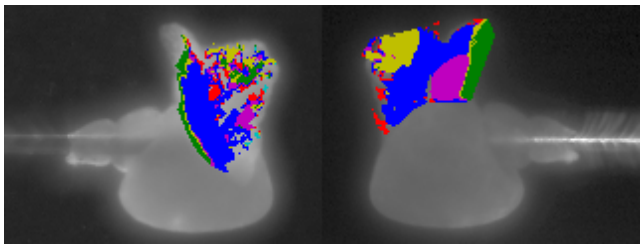
Next, simple k-means clustering with a smaller number of classes will be done on the DoD transformed representation to give the final grouping of functional time series, as shown in Figure 5. Time series in group 2 and 3 in Figure 5 (a) show a diastolic depolarization profile, therefore are chosen as the potential SAN signals, and the cells exhibiting these signals are labelled as the SAN as shown in Figure 5 (c).

III. RESULTS

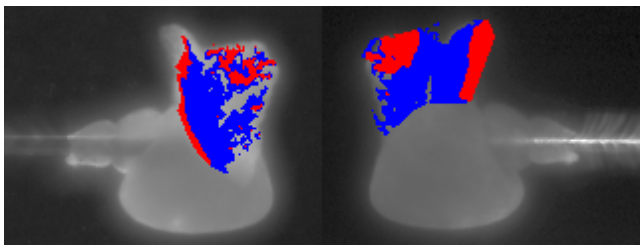
The result of two-stage grouping from four different hearts is shown in Figure 6. The number of clusters was chosen to be 25 in the first stage, and these were further merged in the second stage to give 6 clusters. The numbers 25 and 6 are chosen empirically.



(a) Further grouped average time series of small clusters



(b) Second step group labels



(c) Raw image with potential SAN labelled

Fig. 5. Illustration of second stage grouping: (a) further grouped time series from small clusters, thicker red frames represent manually identified SAN groups, (b) raw image labelled with new groups, (c) potential SAN labelled as red.

Within the same group, waveforms have similar trends, despite differences in original values. Also, variabilities of cardiac action potential morphology observed between different hearts is found not to affect the grouping performance. Among different individual hearts, the identified potential SAN area remains in similar position. According to visual examination validation done by trained cardiac physiology experts, the clustering and identification results are satisfactory and further physiology experiments will be done to verify the identified potential SAN regions.

IV. DISCUSSION AND CONCLUSION

In this paper, we present a novel pipeline of finding the potential SAN of the zebrafish heart based on functional time series data. Preprocessing including spatial and temporal averaging, drifting removal, scaling and cycle averaging is performed to get an averaged signal representing one cycle for each pixel. First stage clustering is then done on the original signal vector space to produce many small clusters, which are then averaged to get a representative cycle for each cluster. The proposed DoD method tranforms these average signals to a representation that attempts to model physiological similarity. Finally, k-means clustering is used to further combine smaller clusters into larger groups of similar pixels.

Our experiments and results show that the proposed method is able to divide the zebrafish heart into physiologically meaningful regions and identify potential SAN

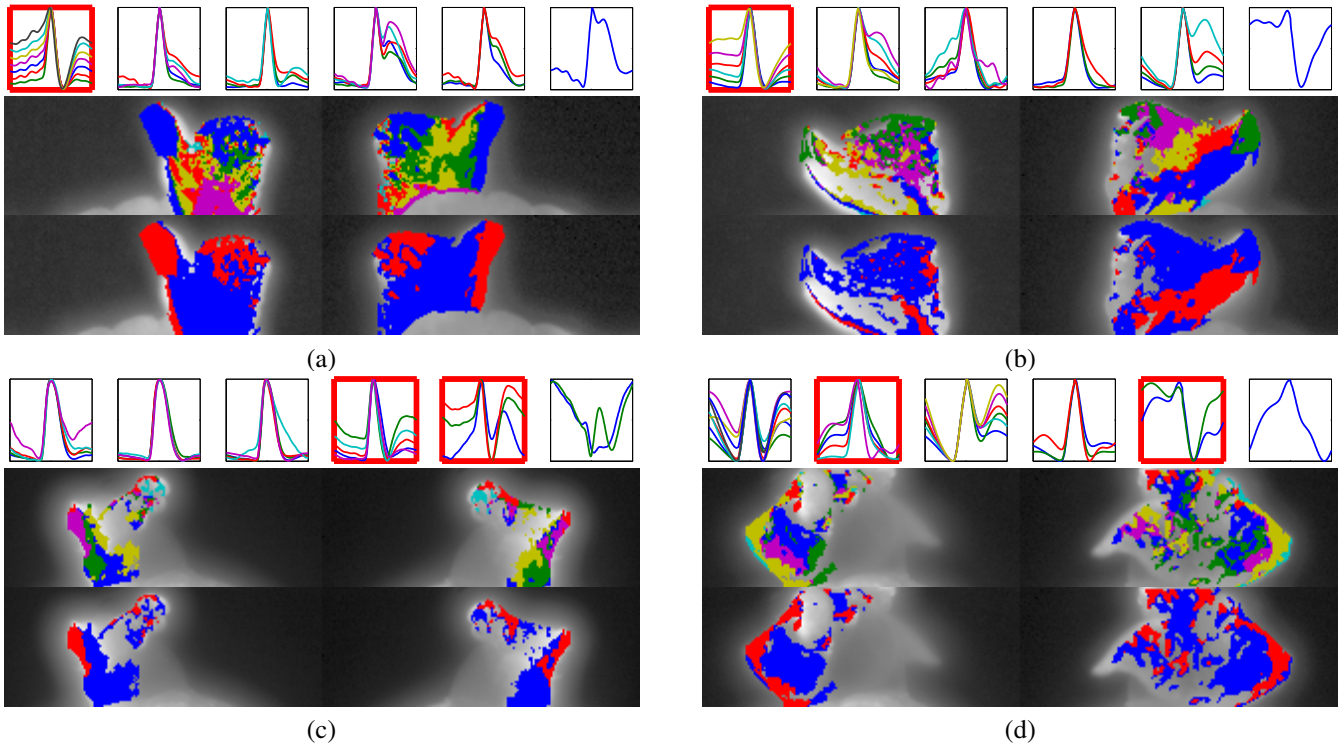


Fig. 6. Results of grouping on 4 different hearts following the novel two-stage clustering approach presented in this paper. In the first stage, the atrial region of interest was clustered into 25 regions (as shown in the middle row). The signals for these 25 clusters are shown in top row, where they have also been further grouped in the second step into a total 6 final clusters based on shape similiarity of the time-series in each cluster. Manually identified SAN signals and regions are labelled by thicker red frames (top row) and red patches (bottom row), respectively in each subfigure. Here, only the upper half of the image is shown, which contains atrial tissue.

regions. In terms of intra-group similarity, method is robust to waveform variations seen in different hearts. Due to lack of exact knowledge of the location of the zebrafish heart SAN in optical mapping data, quantitative validation of accuracy of SAN identification can not be provided at this stage and remains part of future work. Initial confidence in the accuracy of the algorithm in identifying SAN comes from visual identification of signal waveforms based on known physiological properties of SAN cells by trained cardiac physiology experts. Future improvements of the algorithm will include automatic parameter selection and incorporating more detailed physiological knowledge into the current processing pipeline.

REFERENCES

[1] Jeroen Bakkers, "Zebrafish as a model to study cardiac development and human cardiac disease.," *Cardiovascular research*, vol. 91, no. 2, pp. 279–88, July 2011.

[2] B.R. Choi and G. Salama, "Optical mapping of atrioventricular node reveals a conduction barrier between atrial and nodal cells," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 274, no. 3, pp. H829–H845, 1998.

[3] I.R. Efimov and T.N. Mazgalev, "High-resolution, three-dimensional fluorescent imaging reveals multilayer conduction pattern in the atrioventricular node," *Circulation*, vol. 98, no. 1, pp. 54–57, 1998.

[4] A.O. Verkerk, G.S.C. Geuzebroek, M.W. Veldkamp, and R. Wilders, "Effects of acetylcholine and noradrenalin on action potentials of isolated rabbit sinoatrial and atrial myocytes," *Frontiers in Physiology*, vol. 3, 2012.

[5] Federico Tessadori, Jan Hendrik van Weerd, Silja B Burkhard, Arie O Verkerk, Emma de Pater, Bastiaan J Boukens, Aryan Vink, Vincent M Christoffels, and Jeroen Bakkers, "Identification and functional characterization of cardiac pacemaker cells in zebrafish.," *PloS one*, vol. 7, no. 10, pp. e47644, 10 2012.

[6] Jacob I Laughner, Fu Siong Ng, Matthew S Sulkin, R Martin Arthur, and Igor R Efimov, "Processing and analysis of cardiac optical mapping data obtained with potentiometric dyes.," *American journal of physiology. Heart and circulatory physiology*, vol. 303, no. 7, pp. H753–65, Oct. 2012.

[7] Vadim V Fedorov, Alexey V Glukhov, Roger Chang, Geran Kostecki, Hyuliya Aferol, William J Hucker, Joseph P Wuskell, Leslie M Loew, Richard B Schuessler, Nader Moazami, et al., "Optical mapping of the isolated coronary-perfused human sinus node," *Journal of the American College of Cardiology*, vol. 56, no. 17, pp. 1386–1394, 2010.

[8] V.G. Fast and R.E. Ideker, "Simultaneous optical mapping of transmembrane potential and intracellular calcium in myocyte cultures," *Journal of cardiovascular electrophysiology*, vol. 11, no. 5, pp. 547–556, 2007.

[9] C.J. Jou, K.W. Spitzer, and M. Tristani-Firouzi, "Blebbistatin effectively uncouples the excitation-contraction process in zebrafish embryonic heart," *Cellular Physiology and Biochemistry*, vol. 25, no. 4-5, pp. 419–424, 2010.

[10] E.M. Goldys, *Fluorescence Applications in Biotechnology and Life Sciences*, Wiley, 2009.

[11] R. Xu and D. Wunsch, *Clustering*, IEEE Press Series on Computational Intelligence. Wiley, 2008.

[12] DL Pham, "Spatial Models for Fuzzy Clustering," *Computer Vision and Image Understanding*, vol. 84, no. 2, pp. 285–297, Nov. 2001.

[13] P. Nemtsas, E. Wettwer, T. Christ, G. Weidinger, and U. Ravens, "Adult zebrafish heart as a model for human heart? an electrophysiological study," *Journal of molecular and cellular cardiology*, vol. 48, no. 1, pp. 161–171, 2010.