Comparison of Segmentation Algorithms for the Zebrafish Heart in Fluorescent Microscopy Images

P. Krämer[1], F. Boto[1], D. Wald[1], F. Bessy[1], C. Paloc[1], C. Callol[2], A. Letamendia[2], I. Ibarbia[2], O. Holgado[2], and J.M. Virto[2]

[1] VICOMTech, Paseo de Mikeletegi 57, 20009 Donostia – San Sebastián, Spain
[2] Biobide, Paseo de Mikeletegi 58, 20009 Donostia – San Sebastián, Spain

Abstract. The zebrafish embryo is a common model organism for cardiac development and genetics. However, the current method of analyzing the embryo heart images is still mainly the manual and visual inspection through the microscope by scoring embryos visually a very laborious and expensive task for the biologist. We propose to automatically segment the embryo cardiac chambers from fluorescent microscopic video sequences, allowing morphological and functional quantitative features of cardiac activity to be extracted. Several methods are presented and compared within a large range of images, varying in quality, acquisition parameters, and embryos position. Despite such variability in the images, the best method reaches a 70% of accuracy, allowing reducing biologists workload by automating some of the tedious manual segmentation tasks.

1 Introduction

The zebrafish (danio rerio) is a widely used model for the study of vertebrate development. Due to its prolific reproduction, small size and transparency, the zebrafish is a prime model for genetic and developmental studies, as well as research in toxicology and genomics. While genetically more distant from humans, the vertebrate zebrafish nevertheless has comparable organs and tissues, such as heart, kidney, pancreas, bones, and cartilage.

During the last years tremendous advances in imaging system have been made allowing the acquisition of high-resolution images of the zebrafish. Anyhow, the processing of such images is still a challenge [1]. To date, only few work has been presented addressing the analysis of zebrafish images [2,3,4]. For instance [4] presents a method to acquire, reconstruct and analyze 3D images of the zebrafish heart. The reconstruction of the volume is based on a semi-automatic segmentation procedure and requires the help of the user. The approach of [3] avoids segmentation and instead derives a signal of the heart of the images itself to quantify heartbeat parameters.

Despite such isolated research initiatives, the current method of analyzing the embryo heart images in laboratories is still mainly the manual and visual processing using commercial software, offering limited and traditional analysis methods.

G. Bebis et al. (Eds.): ISVC 2009, Part II, LNCS 5876, pp. 1041–1050, 2009.

Some morphological and functional quantitative features of cardiac activity can be extracted from the images. However, the heart itself and the cardiac chambers have to be segmented. Such segmentation is usually done manually, on each image composing a sequence - a very tedious and laborious task for the biologist. In order to provide the biologist with a tool reducing its workload, we propose to attempt segmenting automatically the cardiac chambers of the zebrafish embryo heart from microscopic video sequences. Several methods are described and compared. In our experiment, transgenic embryos expressing fluorescent protein in the myocardium were placed under light microscopy allowing to capture fluorescent images of the heart at video rate. In particular, we are interested in segmenting the heart of zebrafish embryos after two days of post-fertilization (2 dpf). In early stages of the zebrafish development the primitive heart begins a simple linear tube. This structure gradually forms into two chambers, a ventricle and an atrium. At 2 dpf the heart tube is already partitioned into atrium and ventricle as depicted in Figure 1. They are separated by a constriction which will later form the valve. At this stage the heart is already beating. More information on zebrafish heart anatomy can be found in [5].

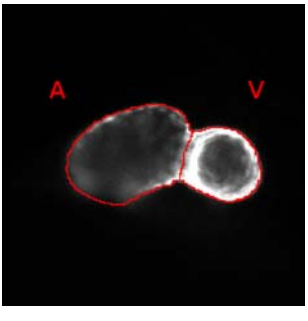

Fig. 1. The 2 dpf zebrafish heart already consists of two chambers: the atrium (A) and ventricle (V) (fluorescent image courtesy of Biobide, Spain)

The remainder is organized as follows: In section 2 we present several approaches to segment the zebrafish heart. Then, we present in section 3 two methods to identify the chambers. In section 4 we show some results and compare the segmentation methods respectively for the heart and its chambers. We give a conclusion of our work and outline future research in section 5.

2 Segmentation of the Zebrafish Heart

In this section we outline different approaches to segment the shape of the zebrafish heart. For the methods of subsection 2.3, 2.4, 2.5, we cast the images to 8-bit gray level images and stretch the gray level range into $[0, 255]$.

2.1 Adaptive Binarization

The Adaptive Binarization method is based on the hypothesis that the image of the heart consists of three brightness levels such as illustrated in Fig. 2: one corresponding to the background and two corresponding to the fluorescent heart where strong contracted regions appear brighter due to a higher concentration of fluorescent cells.

For preprocessing, we smooth the image using a Gaussian kernel of aperture size 7×7 to remove noise. Then, the region of the heart with highest brightness is segmented by first applying a Contrast-Limited Adaptive Histogram Equalization (CLAHE) [6] using a uniform transfer function and then the automatic threshold method from Otsu [7]. In order to segment the second, less brighter region of the heart, we exclude the previous segmented region and apply CLAHE and Otsu again. The final segmentation is obtained by combining both segmentation results. Postprocessing includes the filling of holes which can appear inside in the heart.

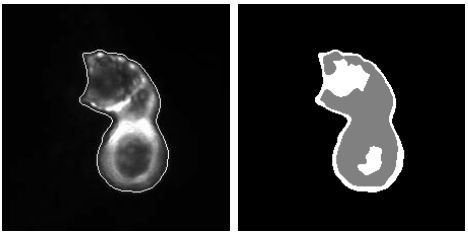

Fig. 2. The image of fluorescent heart (courtesy of Biobide, Spain) consists of three brightness levels: one corresponding to the background and two to the heart

2.2 Clustering

This methods relies like the previous one on the assumption that there are three different brightness levels, although it is an statistical approach. It is based on unsupervised classification in order to distinguish between object and background pixels. First, each pixel is characterized by the mean luminance value of the 3×3 mask centered at the pixel. A unidimensional feature space results. Then, we use a k-means (k = 3) classifier in order to separate the pixels into three clusters. The cluster to which belongs the pixel at position $(0, 0)$ is then defined as the background and others as the region of the heart. Similarly than above, we apply hole filling as postprocessing. For more information on the k-means clustering can be found in [8].

2.3 Voronoi-Based Segmentation

In this case the classification is based on the mean and standar deviation value to classify the pixels in two regions. The Voronoi diagram-based segmentation [9] divides an image into regions using Voronoi diagram and classifies the regions

as either inside or outside the target based on classification statistics (mean and standard deviation), and then break up the regions on the boundary between the two classifications into smaller regions and repeat the classification and subdivision on the new set of regions. The classification statistics can be obtained from an image prior which is a binary image of preliminary segmentation.

In order to compute the image prior, we apply first a bilateral filter to smooth the image while preserving edges. Afterwards, the gradient magnitude is computed using a recursive Gaussian filter and Sigmoid filter to map the intensity range into $[0, 255]$. Then a threshold is applied to the gradient magnitude to obtain a binary mask. As the binary mask may contain holes, we apply a morphological closing operation and fill the holes to complete the object's shape. Then the main region of the heart is isolated from noise in the binary image by a region growing algorithm to the binary with the brightest pixel in the image as seed point. Typically, the brightest pixel in the gray-level image belongs to the region of the heart.

After the Voronoi segmentation we apply again morphological closing, hole filling, isolation of the main region, and morphological erosion to smooth the contours.

2.4 Level Set

The main idea of this method is different to the previous methods. First a presegmentation accomplished which is then refined, but here we use the level set approach [10] for refinement. We choose this method because of its fast performance and the availability of source code which can be found at [11].

The method starts with a morphological reconstruction to suppress structures that are lighter than their surroundings and that are connected to the image border. Then, edges are detected using the Canny edge detector. Dilation, hole filling, and erosion are applied to the contour image. The biggest region is considered as the region of the heart while the others are considered as noise. We complete the process by applying again dilation and hole filling.

A Gaussian filter is applied to smooth the original gray level image for noise removal. Then we apply the level set method [10] with contours of the binary mask as initialization. We chose the edge indicator function $1/(1+g)$ as suggested by the author where g is the gradient magnitude of the Gaussian filtered gray level image.

2.5 Watershed

This approach is different to the previous one as it does not rely on a presegmentation by binarization and it is based on the topology of the image [12].

First the border structures are suppressed by morphological reconstruction. This is followed by a strong low-pass filtering (Gaussian filter) in a morphological reconstruction by erosion using the inverse of morphological gradient. This attenuates unwanted portions of the signal while maintaining the signal intensity as the Watershed method is known to oversegment the image. Afterwards, a

small threshold is applied to set the background to zero and the image intensity is adjusted so that such that 1% of data is saturated at low and high intensities. We apply to this gradient magnitude the watershed segmentation. An oversegmented image may result with typically one region belonging to the background and several regions belonging to the heart. The latter ones are joined to form the region of the heart.

3 Identification of the Chambers

The objective is now to divide the heart into the chambers based on the results of the methods presented in the previous section.

3.1 Convexity Defects

The method assumes that there is a constriction between the two chambers (see Figure 1) causing two convex points in the contour of the heart's shape. Therefore, we compute the convexity defects of the contour using its convex hull. Generally more than two convexity defects are found due to irregularities in the contour caused by the segmentation as depicted in Figure 3. Moreover, we assume that the convexity defects denoting the constriction between the chambers are parallel. Thus, we choose the four most important convexity defects, i.e. the four points with the highest distance from the convex hull, and compute the angle for each pair as:

$$\theta = \arccos\left(\frac{\mathbf{v_1} \cdot \mathbf{v_2}}{|\mathbf{v_1}||\mathbf{v_2}|}\right) \tag{1}$$

where $\mathbf{v_1}, \mathbf{v_2}$ are respectively the vectors between the start and end points of the first and second convexity defects. If the angle is lower than a small threshold, then the pair of convexity defects is considered as a possible candidate for the constriction, otherwise it is rejected. Finally, we choose the pair with the highest mean distance as the points of the constriction from the remaining. We compute the straight line interpolating the points which separates both chambers. As there is often a high variation of the straight line along the image sequence, we correct it by Double Exponential Smoothing-Based Prediction (DESP) [13] using the results of the previous images.

3.2 Watershed

This method is based on the results of the Watershed segmentation of subsection 2.5. The general idea is to divide the segmented shape into the two chambers by applying a second watershed segmentation. Therefore, the background is masked out and a watershed segmentation is applied in this area after a strong low-pass filtering. If two regions result, then they correspond to a rough identification of two chambers. Otherwise the regions have to be joined until only two regions remain. Therefore, we use the chamber identification of the previous image. We compute the intersection of a region in the current image with the

Fig. 3. The segmented heart (red line) and the convex hull (yellow line) with convexity defects of the shape (yellow points) (fluorescent image courtesy of Biobide, Spain)

identified chambers of the previous image. Then, the region is identified to belong to the chamber where the intersection is maximal. It can happen that only one region is obtained by the Watershed segmentation. Then, the segmentation of the previous image is used for further processing of the current image.

This chamber identification is very rough whereas the outline is not coincident with that one of subsection 2.5 as can be seen in Figure 4. Thus unassigned pixels remain. In order to assign them to one of the chambers, an Euclidean distance transform is computed for each chamber. Then, the non-assigned pixels of the segmentation are joined with the chamber for which the distance transform is smaller.

4 Results and Discussion

In this section, we show and discuss the results obtained with the methods presented above. First, we compare the algorithms for segmenting the shape of the heart from section 2 using an accuracy measure. Then, we evaluate visually

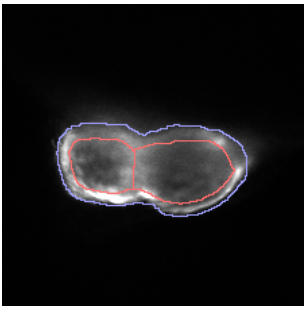

Fig. 4. The Watershed segmentation (blue line) and the first rough identification of the chambers (red line) (fluorescent image courtesy of Biobide, Spain)

Table 1. Mean, Std, Max and Min accuracy for the segmentation algorithms

Method	Mean accuracy	Standard deviation	Max accuracy	Min accuracy
Adaptive binarization	0.870	0.052	0.946	0.725
Clustering	0.876	0.057	0.954	0.759
Voronoi segmentation	0.890	0.046	0.937	0.769
Level set	0.850	0.062	0.935	0.724
Watershed	0.856	0.044	0.896	0.717

the results of chamber segmentation algorithms from section 3. Moreover, we give computational times for all methods.

4.1 Comparison of Segmentation Algorithms

There are several methods to measure the performance of segmentation algorithms. In this work, we compare the segmented images with ground truth images which were obtained by manual segmentation. Although manual segmentations may be very subjective, we consider unsupervised evaluation methods such as [14,15] unsuitable for this comparison. Furthermore, the authors of [14] state that existing unsupervised approaches are most effective at comparing different parameterizations of the same algorithm and that they are less effective at comparing segmentations from different algorithms.

We used the Jaccard coefficient [16,17] as performance measure for each segmentation method. This coefficient measures the coincidence between the segmentation result R and the ground truth A. Then, the segmentation accuracy is measured as:

$$P(R, A) = \frac{|R \cap A|}{|R \cup A|} = \frac{|R \cap A|}{|R| + |A| - |R \cap A|} \tag{2}$$

with $|\cdot|$ as the number of pixels of the given region. The nominator $|R \cap A|$ means how much of the object has been detected while the denominator $|R \cup A|$ is a normalization factor to scale the accuracy measure into the range of $[0, 1]$. Likewise pixels falsely detected as belonging to the object (false positives) are penalized by the normalization factor. Thus, this accuracy measure is insensitive to small variations in the ground truth construction and incorporates both, false positives and negatives, in one unified function [17].

In our experiments we used 26 image sequences with a resolution of 124×124 pixels and with a sufficient image quality for a objective manual evaluation. For each image sequence we segmented manualy the first 20 images with the above presented methods and compared them with a ground truth segmentation. The results of the mean Jaccard coefficient for each sequence are presented in Table 1. The Voronoi segmentation and both thresholding methods outperform the watershed and level set methods. A visual inspection of the segmentation results reveals similar results.

The level set method gives good results on high contrast edges, but in regions where edges are blurred, the level set does not approach well the shape of the heart resulting in holes in the object shape or a too large shape. Moreover, we found it difficult to determine a common set of parameters suitable for all sequences. Finally we used the parameter set used in the example code [11] and fixed the number of iterations to 20. A specific choice of parameter set and number of iterations for each sequence/image might improve the results.

The contours of the watershed method appear very rough and are often too tight. This might be due to the strong low-pass filtering in the post-processing which causes an edge mismatch. Equally a false classification of the regions into background and foreground may cause an inaccurate segmentation.

The Voronoi segmentation method reveals the best results in term of accuracy measure. The contours are typically a little bit irregular, some postprocessing could be applied to smooth them. In case of low-contrast contours it may behave similar to the level set method. The overall results is quite satisfying.

The adaptive binarization tends to have a slightly larger contours, but approaches well the object shape. This might cause the lower accuracy results, but the overall segmentation results are good. Sometimes in case of low-contrast edges the object shape may be incomplete.

The clustering method tends also to larger contours, but slightly tighter than the adaptive binarization method. Therefore, a higher accuracy is achieved. However, in case of low-contrast edges it reveals more often incomplete shapes than the adaptive binarization. Note that the accuracy can vary as the randomized choice of initial cluster may result in slightly different segmentation results.

The computational cost can not be directly compared as the implementations use different programming languages and libraries (the adaptive binarization, clustering, and Voronoi methods are implemented in C++ using respectively OpenCV, OpenCV and Torch, and ITK; the level set and watershed methods are implemented in Matlab), but the average execution time for each image is about one second.

4.2 Chamber Identification

In this section, we present some results for the convexity defects and watershed methods used to divide the heart into its chamber. We only evaluated the convexity defects method in combination with the adaptive binarization and clustering methods as they have the shortest execution times and we obtained good results in terms of accuracy and visual inspection.

Evaluation of the results was realized by determining manually how many images were correctly divided into two chambers with respect to all segmented images. During evaluation we became aware of the fact that in some cases it is very difficult to decide if the chambers have been separated accurately enough or not, because the line segmenting both chamber might be displaced with respect to the restriction. Here, we adopted a hard line and classified such images as not correctly segmented.

Table 2. The ratio of correct chamber identification for the algorithms presented

Method	Ratio
Adaptive binarization + convexity defects	0.704
Clustering + convexity defects	0.577
Watershed	0.456

For evaluation we used only 24 out of the 26 sequences from above, because in two other ones the chambers are superimposed. Hence, the evaluation in those cases is very difficult and we chose to discard those sequences, we visually inspected 480 images. Our results are shown in Table 2, where the best result is obtained for the adaptive binarization method which also has the shortest execution time.

5 Conclusion and Future Work

We presented different methods to automatically segment the cardiac chambers of the zebrafish embryo from fluorescent microscopy video sequences. First, we implemented and compared various methods to extract the heart as a whole. The Voronoi-based segmentation gives the best results in terms of accuracy, followed by thresholding methods, such as the adaptive binarization and clustering method. Other methods, such as level set and watershed were also implemented but showed worse results they were also found more difficult to configure because of the variability in the images.

We then compared various methods to identify the two chambers from the whole heart segmentation. This processing task can be very useful for cardiac study, allowing to analyze morphological and functional activity of each chamber separately. Cardiac pathology, such as fibrillation for example, can affect the atrium (atrial fibrillation) or the ventricle (ventricular fibrillation). It is therefore important to be able to extract not only the whole heart, but also each chamber from the microscopic video sequence. Our comparative study showed that the adaptive binarization method in combination with the detection of convexity defects outperforms clearly the other methods.

While current methods of analyzing the embryo heart images are still mainly based on manual and visual inspection through the microscope, we have proposed image processing methods so that to substitute manual segmentation with automatic process. Although manual control and visual assessment are still necessary, our methods have the potential to drastically reduce biologist workload.

References

1. Vermot, J., Fraser, S., Liebling, M.: Fast fluorescence microscopy for imaging the dynamics of embryonic development. HFSP Journal 2, 143–155 (2008)
2. Luengo-Oroz, M.A., Faure, E., Lombardot, B., Sance, R., Bourgine, P., Peyriéras, N., Santos, A.: Twister segment morphological filtering. A new method for live zebrafish embryos confocal images processing. In: ICIP, pp. 253–256. IEEE, Los Alamitos (2007)

3. Fink, M., Callol-Massot, C., Chu, A., Ruiz-Lozano, P., Belmonte, J.C., Giles, W., Bodmer, R., Ocorr, K.: A new method for detection and quantification of heartbeat parameters in drosophila, zebrafish, and embryonic mouse hearts. BioTechniques 46(2), 101–113 (2009)
4. Liebling, M., Forouhar, A., Wolleschensky, R., Zimmermann, B., Ankerhold, R., Fraser, S., Gharib, M., Dickinson, M.E.: Rapid three-dimensional imaging and analysis of the beating embryonic heart reveals functional changes during development. Developmental Dynamics 235, 2940–2948 (2006)
5. Hu, N., Sedmera, D., Yost, H., Clark, E.: Structure and function of the developing zebrafish heart. The Anatomical Record 260, 148–157 (2000)
6. Zuiderveld, K.: Contrast Limited Adaptive Histogram Equalization. In: Graphics Gems IV, pp. 474–485. Academic Press, London (1994)
7. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. on Systems, Man and Cybernetics 1, 62–69 (1979)
8. Bishop, C.M.: Pattern Recognition and Machine Learning (Information Science and Statistics). Springer, Heidelberg (2007)
9. Imelinska, C., Downes, M., Yuan, W.: Semi-automated color segmentation of anatomical tissue. Computerized Medical Imaging and Graphics 24, 173–180 (2002)
10. Li, C., Xu, C., Gui, C., Fox, M.: Level set evolution without re-initialization: A new variational formulation. In: CVPR, vol. 1, pp. 430–436. IEEE, Los Alamitos (2005)
11. Li, C.: Home page, `http://www.engr.uconn.edu/~cmli/`
12. Vincent, L., Soille, P.: Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE Transactions on Pattern Analysis and Machine Intelligence 13, 583–598 (1991)
13. LaViola Jr., J.: Double exponential smoothing: An alternative to kalman filter-based predictive tracking. In: Immersive Projection Technology and Virtual Environments, pp. 199–206 (2003)
14. Zhang, H., Fritts, J.E., Goldman, S.A.: Image segmentation evaluation: A survey of unsupervised methods. Computer Vision and Image Understanding 110, 260–280 (2008)
15. Sezgin, M., Sankur, B.: Survey over image thresholding techniques and quantitative performance evaluation. Journal of Electronic Imaging 13, 146–168 (2004)
16. Cox, T., Cox, M.: Multidimensional Scaling, 2nd edn. Chapman & Hall/CRC, Boca Raton (2001)
17. Ge, F., Wang, S., Liu, T.: New benchmark for image segmentation evaluation. Journal of Electronic Imaging 16, 33011 (2007)