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

Automated Analysis of Zebrafish Images for Screening Toxicants

Charu Hans, Catherine W. McCollum, Maria B. Bondesson, Jan-Ake Gustafsson, Shishir K. Shah, *Senior Member, IEEE*, and Fatima A. Merchant, *Senior Member, IEEE*

***Abstract*— An important factor facilitating the application of zebrafish in biomedical research is high throughput screening of vertebrate animal models. For example, being able to model the growth of blood vessel in the vasculature system is interesting for understanding both the circulatory system in humans, and for facilitating large scale screening of the influence of various chemicals on vascular development. Compared to other models, the zebrafish embryo is an attractive alternative for environmental risk assessment of chemicals since it offers the possibility to perform high-throughput analyses *in vivo*. However the lack of an automated image analysis framework restricts high throughput screening. In this paper, we provide a method for quantitative measurements of zebrafish blood vessel morphology since it is difficult to assess changes in vessel structure by visual inspection. The method presented is generalized, i.e. it is not restricted to any specific chemically treated zebrafish, and can be used with wide variety of chemicals.**

I. Introduction

The vascular system is a vital component of all vertebrate animals, supplying oxygen and essential nutrients to every tissue and organ. Concurrent with the increasing levels of environmental pollution, accumulating evidence shows that today's exposure levels of industrial chemicals have hazardous effects on human health. One of the main concerns of pollutants is that they may act as developmental toxicants, perturbing normal development of the fetus. It has been suggested that several environmental pollutants could act as vascular disrupting compounds by targeting blood vessel development [3]. A wide range of congenital diseases are associated with blood vessel formation [4], development of the cardiovascular system and associated vascular defects [5], and the adverse effects of exposure to toxic elements during development. Thus, there is a critical need to identify the vascular disruptors out of thousands of industrial chemicals.

Manuscript received February 4, 2013.

C. Hans is with the Department of Computer Science, University of Houston, Houston, TX 77204 USA (e-mail: chans@cs.uh.edu).

C. W. McCollum is with the Department of Biology and Biochemistry, University of Houston, Houston, TX 77204 USA (e-mail: cwmccoll@central.uh.edu).

M. Bondesson is with the Department of Biology and Biochemistry, University of Houston, Houston, TX 77204 (e-mail: ebondess@central.uh.edu).

J. Gustafsson is with the Department of Biology and Biochemistry, University of Houston, Houston, TX 77204 (e-mail: jgustafs@Central.UH.EDU).

S. K. Shah is with the Department of Computer Science, University of Houston, Houston, TX 77204 USA (e-mail: sshah@central.uh.edu).

F. A. Merchant is with the Department of Engineering Technology, University of Houston, Houston, TX 77204 USA (phone: 713-743-8292; fax: 713-743-0172; e-mail: fmerchant@uh.edu).

Zebrafish has recently emerged as an invaluable vertebrate system for toxicity assessment [10]. Particular strengths of zebrafish are its small size, transparency, and the existence of a large number of transgenic fish expressing fluorescent proteins in specific cells or tissues, making it a good model for direct observation of the development of organs and tissues during normal and perturbing conditions. Despite the potential of the zebrafish as a model, engineering quantification is still a fledgling field in zebrafish research, primarily due to the lack of tools that can yield objective and quantitative measurements from imaging. The actual number of screens for toxic assessment is very limited [8]. A major limitation is an automated image analysis approach that can facilitate large scale screening. While many studies have been dedicated to the analysis of blood cells, and morphological analysis of neurons [9], few studies have been done for automated quantification of vascular growth in zebrafish [2].

Most of the previous reports on zebrafish image analysis are based on visual inspection, which is a tedious and time consuming process. Moreover, with the increasing capacity to acquire image data, there is now a pressing need for image processing algorithms to facilitate such analysis in an automated manner [9].

There has been limited research to facilitate image analysis of zebrafish images. Liu et al. [9] have performed automatic detection of neuron loss and defective somites, as well as quantitative measurement of gene expression levels in zebrafish embryos. The gene expression level measurement was performed by image segmentation using seed detection and region growing. Vogt et al. [8] implemented a user guided image interpretation tool to generate rule-based hierarchical image segmentation. These pre-processed images were then used for blood vessel quantification. Numerical measurements of blood vessels development were captured. Feng et al. [7] developed a 3D attributed vessel represent graph (AVRG) approach to reconstruct caudal vasculature of zebrafish embryo. Chen et al. proposed a robust automatic segmentation scheme based on a modified edge-following technique to identify ROIs for gene expression quantification [2]. Researchers have also used commercial software packages such as Cellomics ArrayScan II [8] to perform aided image interpretation. In summary, although methods have been developed to process zebrafish images, as the applications of the zebrafish model expands, there is a concurrent demand for a variety of image processing methods and new image processing algorithms are constantly being requested [2]. The methodology presented in this paper is quantitative, can be utilized with wide varieties of toxin treated zebrafish, and capable of

quantifying changes in fine structure not quantifiable by the human eye. In the past, automatic identification of the applicable region of interest (ROI), and accurate quantification of ROI features has been very challenging.

In this paper, we focus on developing an image processing algorithm for automated analysis of compound-treated zebrafish embryos. We have used the transgenic zebrafish (Tg(kdrl:GFP)), that expresses green fluorescent protein (GFP) in vascular endothelial cells as a model to identify vasculature disrupting compounds. The algorithm extracts valid zebrafish embryos from images, while excluding other objects, then aligns zebrafish embryos of different orientations so that the embryo is positioned with its head region to the upper right corner and its dorsal area facing up (Fig. 4(b)). Next, a global threshold is applied to the ROI to segment the back, head, dorsal aorta (DA) and tail region. Next, the algorithm searches for the boundaries corresponding to the dorsal aorta of each embryo to determine the valid ROI that encloses the intersegmental vessels (ISV) (Fig. 5). After the ROI is found, we calculate the ISV area, the mean ISV length and the normalized ISV area. This paper presents a pipeline of image processing and analysis algorithm that enables quantitative analysis of the vascular system in the zebrafish embryo. To evaluate the developed approach, we used embryos treated with two different drugs to quantify the changes in the ISV region.

The remainder of the paper is organized as follows. In Section II, we introduce the data acquisition and the proposed automated analysis method. Section III presents the experimental results and conclusions are drawn in Section IV.

II. Materials and Methods

A. Data Acquisition

Transgenic zebrafish Tg(kdrl:GFP) embryos were exposed to two different chemicals, from 3 hours post fertilization (hpf) to 3 days post fertilization (dpf) and assessed for vascular abnormalities. Images were acquired with an Olympus IX-51 fluorescent microscope with a 4X objective and GFP filter (Fig. 1).

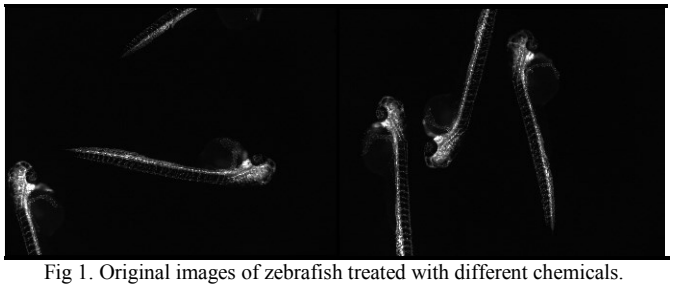

Fig 1. Original images of zebrafish treated with different chemicals.

B. Methods

The procedure for image analysis is outlined in Fig. 2. It consists of three major steps as follows: alignment of the zebrafish embryo, modeling of the ROI, and quantitative measurement of the vasculature.

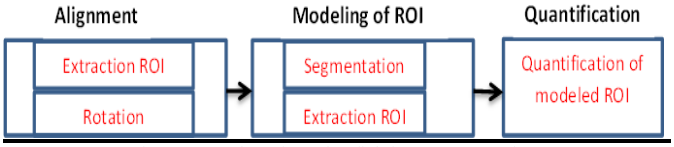

Fig 2. Procedure for zebrafish image processing

The first step consists of two parts. Stage one involves extracting the relevant zebrafish embryo. An image may have multiple zebrafish embryos as shown in Fig. 1. Our aim is to capture the complete anatomic structure of each zebrafish. For, this we smooth the image with Gaussian filter and perform global thresholding, followed by blob labeling for extracting the each of the zebrafish embryos. We discard blobs under 3 conditions (Fig. 3): (i) if blobs are touching the boundaries of image (ii) if area of blobs is above certain threshold and (iii) if area of blobs is below certain threshold. Area threshold was empirically determined between 65000-12500 pixels.

Fig 3. Perform Gaussian filtering with radius 6.0. Use global thresholding of 6 to isolate zebrafish embryo from background. Embryos in red indicate the ROI, whereas regions in white were discarded.

The second stage consists of the rotation of the detected blob. Each blob is bounded by a rectangle [6]. Segmented blobs are rotated around its center, with the angle that the bounding box makes with the horizontal axis (Fig. 4).

Consider $(x_i, y_i), i = 1, 2, \dots, n$ be the n boundary points of zebrafish embryo. Then the centroid is given by

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

Let θ be the angle of the major axis with the horizontal axis. The equation of the line passing through $(\bar{x}, \bar{y})$ at an angle θ is

$$x \tan\theta - y + \bar{y} - (\bar{x}) \tan\theta \qquad (1)$$

The perpendicular distance of an edge point $(x_i, y_i), i = 1, 2, \dots, n$ to the line in Eq. (1) is

$$p_i = (x_i - \bar{x}) \sin\theta - (y_i - \bar{y}) \cos\theta$$

In order to compute the angle θ, we minimize sum of the square of the perpendicular distances P with respect to θ.

$$P = \sum_{i=1}^{n} [(x_i - \bar{x}) \sin\theta - (y_i - \bar{y}) \cos\theta]^2$$

Therefore, $\frac{\partial P}{\partial \theta} = 0$ gives,

$$tan2\theta = \frac{2\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}[(x_i - \bar{x})^2(y_i - \bar{y})^2]}$$

In order to set the head region towards upper right corner, the bounding box is divided half way along vertical axis. The area enclosing the head region is determined based on number of pixels in each half (more pixels on head side, as shown in Fig. 4(c), based on the prior information). The image is flipped accordingly. In order to make sure yolk region is placed in the lowermost right quadrant region, the region with the head is divided into three regions from top to bottom, parallel to horizontal axis, as depicted in Fig. 4(d). Then on the basis of prior knowledge, the region with less number of pixels is selected as the yolk region and image is flipped accordingly. The area under the blob is masked out from the original image and subject to same steps of rotation and flipping as the blob (Fig. 4(e)).

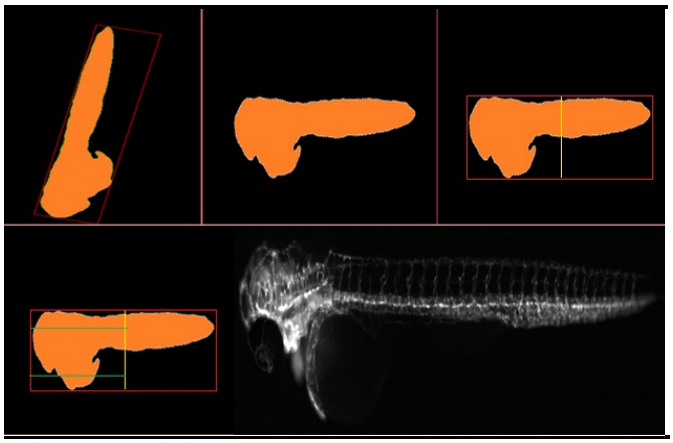

Fig 4. (a) Bounded rectangle on zebrafish embryo. (b) Zebrafish embryo rotated in upright position. (c) Decision for position of head region. (d) Decision for direction of yolk region. (e) Masked out area under the blob from original zebrafish image.

After the position of the zebrafish embryo in the image is normalized, the zebrafish embryo is preprocessed and segmented to separate ISV area. The large vessels (dorsal aorta, tail) and head structure are separated from ISV based on intensity values. All the pixels from the top boundary of segmented region (green region in Fig. 5(b)) are stored in an array.

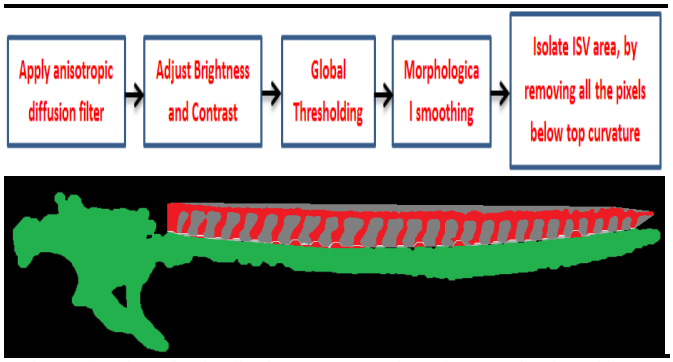

Fig 5. (a) Gives the details about segmentation of ISV region in red from rest of region (head, dorsal aorta, and posterior cardinal vein in green with the average intensity value of 75) (b) Separated regions of ISV, and remaining zebrafish anatomy. Gray colored region is the envelope for ISV.

In the original zebrafish image we retain only those pixels that lie above the pixels stored in the array, to obtain the ISV (pixels above threshold obtained by Otsu's method of segmentation [1], red region in Fig. 5(b)). The algorithm then removes small and isolated objects by morphological operations.

The final step involves quantification of intersegmental vessels. We reported average values for ISV length, area and relative area for all of the ISVs. ISV length was computed via network analysis. The network analysis algorithm operates in two major steps: (1) Thinning of the binary input for centerline extraction; (2) Identification of junction points, end-points (vertices) and edges (branches). For first step, binary thinning is used for finding the centerlines of objects in the input image. Algorithmic details are explained in [11]. The general idea is to erode the ISV's surface iteratively until only the skeleton remains. The structure obtained from the first step is a set of connected pixels as shown in Fig. 6(b). We decompose the complex vascular network into individual vessel segments suitable for quantitative analysis.

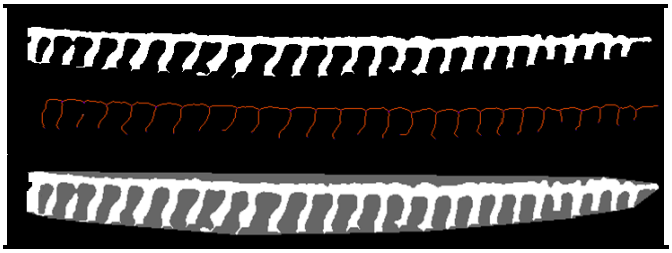

Fig 6. (a) Separated ISV region of zebrafish embryo. (b) Skeleton of ISV, used for calculating length of ISV. (c) ISV region shown bounded by a convex hull, used to calculate normalized ISV area.

Second step involves finding junction and end-points. Pixels with less than 2 neighbors are marked as end-point and pixels with more than 2 neighbors are marked as junction. We store the edges between junctions and end-points, where end-points lie vertically below junctions. Each edge represents an ISV, and the number of pixels in each ISV gives length of ISV. Average ISV length is computed by averaging over all edges (Fig. 6(a)).

ISV area (Fig. 6(b)) is the number of pixels spanned by ISV and normalized ISV area is ISV area divided by area given by the convex hull of ISV region (Fig. 6(c)).

The above method provides a quantitative measurement of blood vessel structure in zebrafish that is difficult to observe by visual inspection. The method is not restricted to any particular kind of chemically treated, or normal zebrafish embryos, hence can be generalized for use with any images of the zebrafish vasculature.

III. Experiments and Results

Automated image analysis of the original images was performed using the method in Section II. 27 images were used to evaluate our proposed algorithm. 13 of them belonged to group 1, and other 14 belonged to group 2. Out of the 13 images in group 1, 6 were controls and 7 were treated. In group 2, 6 were controls, and 8 were treated. From group 1 we could not use two treated images as the thresholded ISV region had overlapping structures and the skeletonization algorithm failed to get disjointed ISV (Fig. 7).

Figure 8 and 9 show the box plot of quantified measures for group 1 and 2, respectively. Data sets were analyzed by

two-tailed Student's t-test assuming unequal variance ($p < 0.1$) to understand the difference between the two groups. Group 1 showed substantial difference between ISV area, normalized ISV area, and mean ISV length (Fig. 8). For group 2, ISV area and mean ISV length showed significant difference, whereas normalized ISV area failed to show significant amount of difference between controls and treated (Fig. 9).

Fig 7. Image analysis algorithm failed to give ISV skeleton

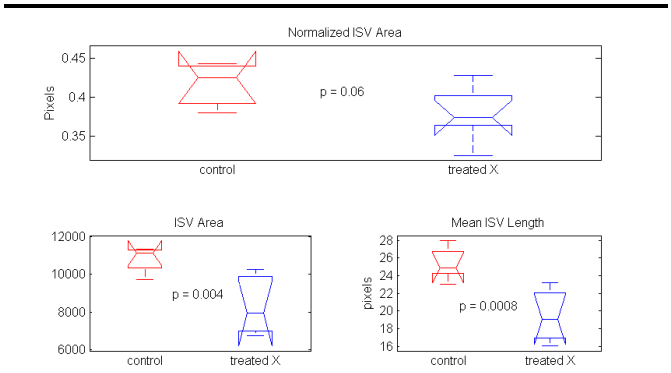

Fig 8. Box plot for group 1.

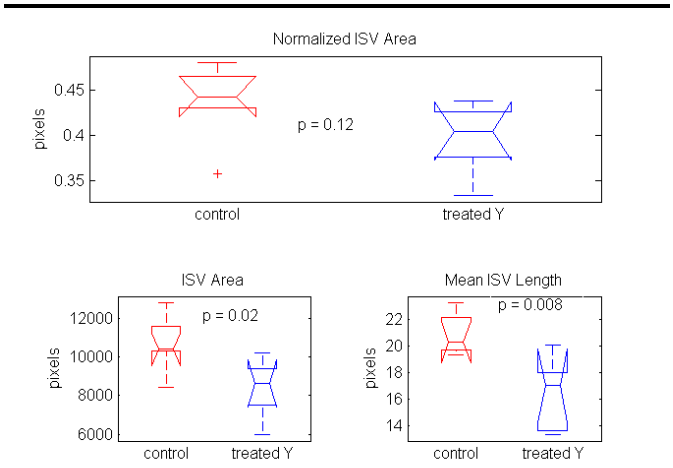

Fig 9. Box plot for group 2.

These differences would be a challenge to notice by visual inspection. Further, visual inspection varies with observer. Compared with manual analysis, an automated method also has the advantage of speed, uniformity and sensitivity. Finally, image-processing algorithms can be easily combined with statistical tools for further analysis.

IV. Conclusion

The large amount of image data obtained from the experiments makes manual analysis a time-consuming and error prone process. In this paper we have shown a fully automated image analysis pipeline that segments individual zebrafish embryos and analyzes their vascular structures. This method is applicable for facilitating large scale screening of the influence of various chemicals on vascular development in zebrafish.

Table I. Group 1 statistics

Group 1 Average Statistics					
ISV Area		***Normalized ISV Area***		***Mean ISV Length***	
Controls	*Treated*	*Controls*	*Treated*	*Controls*	*Treated*
10793	8328.3	0.41	0.38	25.37	19.42

Table II. Group 2 statistics

Group 2 Average Statistics					
ISV Area		***Normalized ISV Area***		***Mean ISV Length***	
Controls	*Treated*	*Controls*	*Treated*	*Controls*	*Treated*
10644.8	8397.8	0.43	0.39	20.84	16.51

Acknowledgment

The biological experimentation was supported by a grant from EPA (grant # R834289).

References

[1] Otsu N, "A thresholding selection method from gray-level histograms", *IEEE Trans System Man Cybernetic*, pp. 62-66, 1979.

[2] Chen S, Zhu Y, Xia W, Xia S, Xu X. "Automated analysis of zebrafish images for phenotypic changes in drug discovery". *J Neurosci Method* , vol. 200, no. 2, pp. 229-236., 2011.

[3] Kleinstreuer N.C., Judson R.S., Reif DM, Sipes N.S., Singh A.V., Chandler K.J., Dewoskin R., Dix D.J., Kavlock R.J., Knudsen T.B., "Environmental impact on vascular development predicted by high-throughput screening", *Environ Health Perspect*, vol. 119, no. 11, pp. 1596-603, Nov. 2011.

[4] Folkman J., "Angiogenesis in cancer, vascular, rheumatoid and other disease", Nature *Med*., vol. 1, pp.27–31, 1995.

[5] Shalaby F., Rossant J., Yamaguchi T.P., Gertsenstein M., Wu X.F., Breitman M.L., Schuh A.C., "Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice", *Nature 376*, 1995, pp. 62–66.

[6] Chaudhuri D., Samal A.. 2007. "A simple method for fitting of bounding rectangle to closed regions", *Pattern Recogn.,* vol. 40, no. 7, pp. 1981-1989, 2007.

[7] Feng J., Cheng S.H., Chan P.K., and H.H.S. IP, "Reconstruction and representation of caudal vasculature of zebrafish embryo from confocal scanning laser fluorescence microscopic images*," Comput. Biol. Med*., vol. 35, pp. 915–931, 2005.

[8] Vogt A., Cholewinski A., Shen X.Q., Nelson S.G., J.S. Lazo, M. Tsang, and N.A. Hukriede, "Automated image-based phenotypic analysis in zebrafish embryos.," *Developmental Dynamics*, vol. 238, pp. 656–663, 2009.

[9] Liu T., Lu J., Wang Y., Campbell W.A., Huang L., Zhu J., Xia W., Wong S.T., "Computerized image analysis for quantitative neuronal phenotyping in zebrafish," *J Neurosci Methods*, vol. 153, pp. 190–202, 2006.

[10] Cheng S.H., Chan P.K., Wu R.S., "The use of microangiography in detecting aberrant vasculature in zebrafish embryos exposed to cadmium", *Aquatic Toxicology*, vol. 52, no. 1, pp. 61-7, 2001.

[11] Lee T.C., Kashyap R.L., and Chu C.N., "Building skeleton models via 3-d medial surface axis thinning algorithms," *CVGIP: Graphical Models and Image Processing*, vol. 56, pp. 1049–9652, 1994