Seventh Hungarian Conference on Computer Graphics and Geometry, Budapest, 2014

Analysis of Image Descriptors for Zebrafish Toxicity Testing

László Szirmay-Kalos[1], Bence Parajdi[1], and Zsolt Csenki[2]

[1]: Dept. of Control Engineering and Information Technology, Budapest University of Technology and Economics
[2]: Dept. of Aquaculture, Institute of Environmental and Landscape Management, Szent István University

Abstract

This paper evaluates image descriptors for the determination whether zebrafish embryos are dead or alive. The image processing pipeline includes image stitching, illumination compensation, transformation to polar coordinates to emphasize the circular geometry of embryos, extraction of darker chorion and yolk boundary and yolk darkness features, and finally classification with a support vector machine. Individual descriptors should be strongly discriminative and preferably independent to allow the combination of their strengths by the support vector machine.

1. Introduction

The zebrafish (Danio rerio in Latin), also called the *Zebra Danio* is a tropical freshwater fish belonging to the minnow family (Cyprinidae) of order Cypriniformes. It is an important vertebrate model organism in scientific research because of its regenerative abilities and its transparent body which allows in vivo optical inspection. Another advantages for research include the facts that the generation of identical Zebrafish twins is an easy and inexpensive process, and animal protection laws do not apply to fish embryos before the fish first starts eating. Zebrafish embryo toxicity testing is used for routine sewage surveillance in many countries[2].

This paper proposes an image processing system that automatically determines whether 0–16 hour old zebrafish embryos are dead or alive. In this phase, embryos are still unhatched and have a circular like shape (Fig. 1). The system does not require manual intervention and is prepared for arbitrary embryo position and orientation. High resolution images of living and dead embryos are shown by Fig. 2.

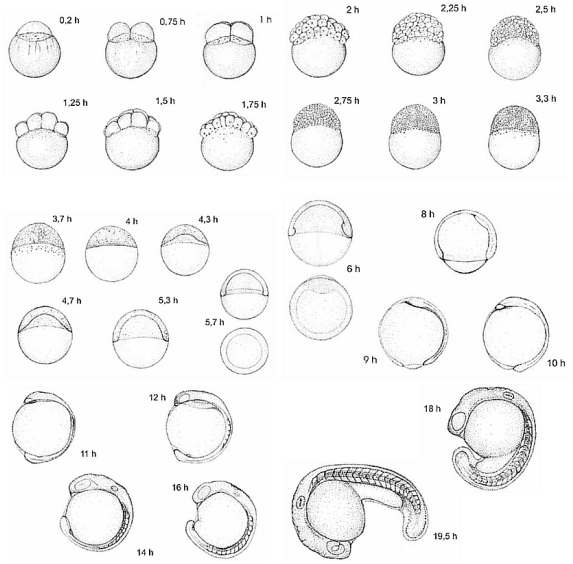

Figure 1: *Early development of healthy zebrafish embryos*[2].

2. Problem analysis

Examining a larger collection of zebrafish images, we can make the following observations:

1. Zebrafish is photographed by trans-illuminating the body and the plate by area light sources of not uniform but slowly changing intensity. The observed illumination intensity is not white and is different in every image.
2. Both healthy and dead embryos have well defined outer dark boundary surrounding the chorion, which has roughly circular shape.
3. The size of the embryos is well defined by biological constraints and has small variation. The diameters of the yolk

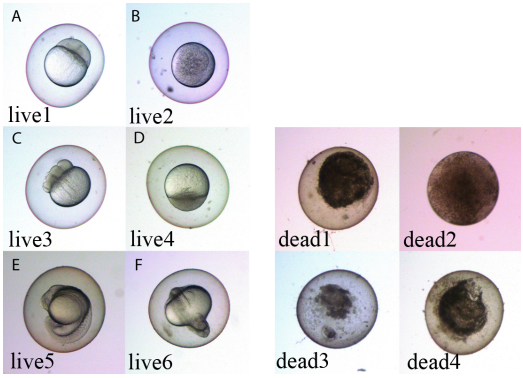

Figure 2: *High resolution images or living (left) and dead (right) zebrafish embryos.*

and the chorion are typically 0.6 mm and 1.2 mm, respectively.

4. The chorion includes the yolk that is also roughly circular in early stages, but gets less circular later. The yolk has also a well defined darker boundary curve in healthy embryos, but the boundary is not darker than the yolk itself in case of mortality.
5. The chorion can be of low noise both in alive and dead embryos.
6. The yolk is a little darker and noisier in dead embryos. The shape of the alive and dead embryos may be similar. The yolk of dead embryos has high variation, including fungi and darker spot like features.

An image processing system is usually built of three main phases, *image enhancement* that eliminates noise and additional effects corrupting the information of our interest and prepares the image for information extraction, *descriptor generation* that reduces the dimension of the information, i.e. extracts the most distinctive features, and finally *classification* that makes the decision based on the descriptors. In our case, image enhancement should eliminate the effects of non-constant and unknown background lighting and of the plate and convert the image into a form that is appropriate for recognizing circular features. For classification we use *Support Vector Machines* (SVM). The key part of the method is the definition of distinctive descriptors.

Our descriptors should be *translation and rotation invariant* since zebrafish can be anywhere and with any orientation in its well of the plate. The descriptor should also be *resolution independent* but *not scale invariant*, since images are made with different magnification and resolution parameters but these are known and the size of the embryo is also well defined (the diameter of the chorion and the yolk are about 1.2 mm and 0.6 mm, respectively). The size can be exploited to increase the robustness of recognition. To make the algorithms physical size dependent but resolution independent, all descriptors are computed in physical space where the unit is [mm]. Whenever a pixel space integral, e.g. convolution is computed, the size of the pixel in physical space is taken into account.

To find good descriptors, we can note that the chorion and the background may be similar for both healthy and dead embryos, so the descriptors should concentrate on the yolk region only. The shape of the chorion is not distinctive since again both healthy and dead embryos can have similar shape. We can check whether it is darker or of higher variation than expected in healthy embryos and also whether its boundary is darker than the yolk itself. In order to improve the discrimination power of image descriptors, the background should be identified and descriptors should be computed only for the foreground characterizing the yolk, which requires the identification of the embryo and the yolk. The circular like shape is emphasized by converting the image from Cartesian to polar coordinates.

A robust and accurate classification requires the identification of features that are significantly different in alive and dead embryos. We have decided to use the features that the yolk boundary of living embryos is a dark curve that separates the lighter chorion and yolk while the boundary is not significantly darker than the yolk in dead embryos. These boundaries are found by *ridge/valley detection* algorithms. Additionally, we examine descriptors that focus on the texture of the yolk, focusing on the darkness and the noise.

3. Previous work

A comprehensive review of Zebrafish image processing is [3], addressing applications like tracking of cells during embryogenesis, heartbeat detection, identification of dead embryos, recognition of tissues and anatomical landmarks, and quantification of behavioral patterns. The identification of dead embryos falls into the category of *phenotype recognition*.

In the method of [10], the image is binarized using an adaptive thresholding, in which a local threshold is set to the mean value of its local neighbors. The binary image is eroded to remove small spurious features and then dilated to connect broken segments. Of the connected objects, the one with the maximum area is recognized as the chorion. The second largest object in the image is the cytoplasm, the boundary of which is represented by a chain code contour. The boundary of the cytoplasm is often not fully connected, thus, a convex hull of the contour is constructed and used as initial positions for subsequent snake tracking. The centroid of the contour is recognized as the cytoplasm center. In order to distinguish the yolk from the cell portion, the cytoplasm contour after snake tracking is fitted into an ellipse using a least squares method, and intercepted into two parts by the minor axis of the fitted ellipse. Based on the fact that the cell portion always has greater convex deficiency, the cell and yolk portions are distinguished.

Liu et al.[6] classified three zebrafish phenotypes: *hatched*, *unhatched* and *dead* by evaluating six intensity and texture descriptors, including Local Edge Histogram Descriptor (LEHD), Color Layout Descriptor (CLD), Scalable Color Descriptor (SCD), Global and Semi-global Edge Histogram Descriptor (GSEHD), Representative Color Descriptor (RCD), and Color Histogram Descriptor (CHD). The LEHD descriptor provides texture information in terms of the spatial distribution of five types of edges, i.e., vertical, horizontal, forward diagonal, backward diagonal, and undirectional edge. LEHD comprises $16 \times 5 = 80$ histogram bins corresponding to the distribution of the five different edge types over 4×4 non-overlapping image blocks of equal size. Global and Semi-global Edge Histogram Descriptors (GSEHD) are constructed by aggregating the block histograms of the entire image and five sub-image groups comprised by 4 blocks. The CLD descriptor captures the local spatial distribution of intensity by using the coefficients of the 8×8 Discrete Cosine Transformation (DCT). The Representative Color Descriptor (RCD) comprised of 64 representative colors. SCD is a Haar transform encoded intensity histogram, which characterizes an image by the global color distribution. The GSEHD, SCD, and CHD are global descriptors capturing overall information about the images.

Jeanray et al.[4] computed pixel-based image descriptors and extremely randomized trees to recognize basic phenotypes and more subtle ones such as pericardial edema and curved tails.

Alshut[1] transformed the images into a 2D feature space where the y-axis gives values for the center of mass of the gray value histogram, and the x-axis indicates the mean intensity value in the chorion center.

4. The proposed system

We should identify first the embryo, then the yolk. The process contains a sequence of operations forming a pipeline. In the following, we discuss the elemental steps one by one.

4.1. Initial stitching

A zebrafish image should be first obtained from a collection of images of the plate that are neither stitched nor positioned on individual wells (Fig. 3). Fig. 4 shows how significant the error is when images are just placed one after the other. To reduce this error, we implemented a stitching algorithm that translates individual images along the border to minimize the difference, i.e. the L_2 error between the two edges.

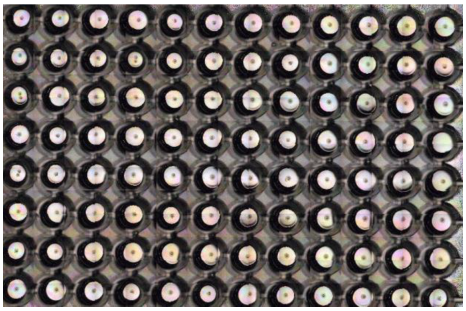

Figure 3: *Collection of images made by stepping the camera over a plate.*

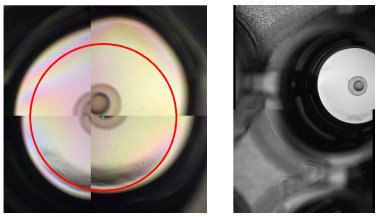

Figure 4: *Neighboring images before stitching and having executed the stitching operation.*

4.2. Image enhancement

4.2.1. Illumination compensation

The objective of this step is to make the image independent of the actual lighting, camera setup and plate/water effects in order to make the image characterize solely the examined object. The compensation is done according to an illumination model that assumes that the embryo is trans-illuminated by an area light source having bi-linearly varying but unknown wavelength dependent intensity $L_0(x,y,\lambda)$ where x and y are pixel coordinates and λ is the wavelength corresponding to the additive RGB primaries. Intensity $L_0(x,y,\lambda)$ depends on light source and also on the glass container as well as the water which may modify the light intensity.

The trans-illuminated embryo is assumed to be absorbing participating medium of wavelength independent absorbing cross section $\sigma_a(\vec{s})$. We assume that emission and scattering are negligible. Thus, the radiance perceived by the camera is

$$L(x,y,\lambda) = L_0(x,y,\lambda) \cdot a(x,y) = L_0(x,y,\lambda) \cdot \exp\left(-\int \sigma_a(\vec{s}) \mathrm{d}s\right).$$

where $a(x,y)$ is the total attenuation, which is the negative exponential of the optical depth, which in turn, is the line integral of the absorption cross section[8]. The embryo is characterized by the total attenuation

$$a(x,y) = \frac{L(x,y,\lambda)}{L_0(x,y,\lambda)}$$

which is independent of the illumination conditions and results in a gray-scale image.

The key part of this step is the determination of unknown

illumination intensity $L_0(x,y,\lambda)$, which is seen directly by the camera where no embryo part is visible. The edges of the individual wells cause strongly non-linear variation, so we consider those points that are outside of the embryo but inside its own well. Such regions are identified by fitting circles on the well and the embryo by Hough transform and considering only those pixels that are in between these circles. For the selected pixels, we fit a bi-linear radiance function $L_0(x,y,\lambda)$ for each representative wavelength λ, i.e. for red, green and blue. Let us consider just the red component, in the implementation the same method should be repeated three times. The samples taken outside the embryo are $\bar{x}_i, \bar{y}_i, R_i$ $(i = 1,\ldots,n)$ where $\bar{x}_i, \bar{y}_i$ are the normalized coordinates of the sample i that are zero at the left bottom corner of the image and are 1 at the right upper corner, and R_i is the red intensity of the sample pixel, and n is the number of samples.

The distribution of the red intensity $R(\bar{x},\bar{y})$ is searched in a bi-linear form:

$$R(\bar{x},\bar{y}) = R_{00}(1-\bar{x})(1-\bar{y}) + R_{01}(1-\bar{x})\bar{y} + R_{10}\bar{x}(1-\bar{y}) + R_{11}\bar{x}\bar{y},$$

where $R_{00}, R_{01}, R_{10}, R_{11}$ are the unknown intensities at the four corners of the image. These unknown values are determined with least-square fitting of the samples, i.e. we solve the following overdetermined system with minimizing the quadratic error:

$$R_i = R_{00}(1-\bar{x}_i)(1-\bar{y}_i) + R_{01}(1-\bar{x}_i)\bar{y}_i + R_{10}\bar{x}_i(1-\bar{y}_i) + R_{11}\bar{x}_i\bar{y}_i,$$

or in matrix form

$$\mathbf{S}_n = \mathbf{P}_{n\times 4} \cdot \mathbf{R}_4$$

where $\mathbf{S}_n$ is the n element vector of intensity samples R_i, $\mathbf{R}_4$ is the four element vector of unknown parameters $R_{00}, R_{01}, R_{10}, R_{11}$, and $\mathbf{P}_{n\times 4}$ is the $n \times 4$ element matrix of sample coordinates

$$\mathbf{P}_{n\times 4} = \begin{bmatrix} (1-\bar{x}_1)(1-\bar{y}_1) & (1-\bar{x}_1)\bar{y}_1 & \bar{x}_1(1-\bar{y}_1) & \bar{x}_1\bar{y}_1 \\ \vdots & & & \\ (1-\bar{x}_n)(1-\bar{y}_n) & (1-\bar{x}_n)\bar{y}_n & \bar{x}_n(1-\bar{y}_n) & \bar{x}_n\bar{y}_n \end{bmatrix}.$$

The least-square solution can be obtained with the Moore-Penrose pseudo inverse:

$$\mathbf{R}_4 = \left(\mathbf{P}_{4\times n}^T \cdot \mathbf{P}_{n\times 4}\right)^{-1} \cdot \mathbf{P}_{4\times n}^T \cdot \mathbf{S}_n.$$

4.2.2. Localization of the embryo

Having executed the illumination compensation, the resulting gray scale image can be imagined as a height field where intensity defines the height of a terrain. The boundary of the chorion is roughly a circle of approximately known radius and is darker than its neighborhood. Circles can be extracted by Hough transform. Making it more robust, the search space is restricted to biologically valid sizes of the embryo. Additionally, we enforce Hough transform to consider only valleys and not arbitrary edges as potential circle points, thus first run a ridge/valley detection algorithm on the gray scale image and apply Hough transform on its result. Ridges and valleys are locally ellipse like features where one axis is significantly larger than the other. A valley point is also local minimum in the maximum curvature direction.

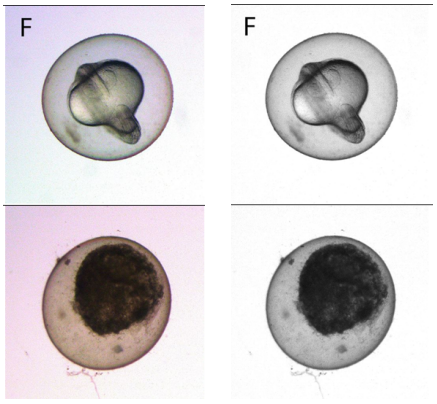

Figure 5: *Results of illumination compensation.*

Ellipses are obtained from the second order approximation of the image function $a(x,y)$:

$$a(x,y) \approx a(x_0,y_0) + \mathbf{g} \cdot [x-x_0, y-y_0]^T + \frac{1}{2}[x-x_0, y-y_0] \cdot \mathbf{H} \cdot [x-x_0, y-y_0]^T,$$

where

$$\mathbf{g} = [\mathbf{g}_x, \mathbf{g}_y], \quad \mathbf{g}_{c_1} = \frac{\partial a(x,y)}{\partial c_1}$$

is the gradient, and

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{xx} & \mathbf{H}_{xy} \\ \mathbf{H}_{xy} & \mathbf{H}_{yy} \end{bmatrix}, \quad \mathbf{H}_{c_1c_2}(x,y) = \frac{\partial^2 a(x,y)}{\partial c_1 \partial c_2}$$

is the Hessian matrix comprising of the second derivatives. In these equations c_1 and c_2 can stand for any of coordinates x and y.

Before computing the derivatives, the noise should be reduced either by regression[7] by convolving with e.g. a Gaussian filter

$$G(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right),$$

thus we compute the derivative of filtered image

$$\tilde{a}(x,y) = a(x,y) * G(x,y).$$

Variance σ of the Gaussian controls the scale of the fea-

tures that should be selected to correspond to the width of the boundary line[5].

The filtered derivative can be obtained by convolving the unfiltered function with the derivatives of the Gaussian.

$$\mathbf{g}_{c_1} = \frac{\partial \tilde{a}(x,y)}{\partial c_1} = a(x,y) * \frac{\partial G(x,y)}{\partial c_1},$$

$$\mathbf{H}_{c_1 c_2}(x,y) = \frac{\partial^2 \tilde{a}(x,y)}{\partial c_1 \partial c_2} = a(x,y) * \frac{\partial^2 G(x,y)}{\partial c_1 \partial c_2}.$$

The axis directions of the ellipse approximation are determined by the eigenvectors of the Hessian and the lengths of the axes by the eigenvalues. A pixel is classified as a valley point if one of the eigenvalues (defining the curvature) is a large positive value while the other eigenvalue is close to zero. Additionally, a valley is a local minimum with respect to the direction of the steepest hillside, thus the derivative in the direction of the eigenvector corresponding to the high curvature must be close to zero, i.e we should find a zero crossing between the current pixel and its neighboring pixels.

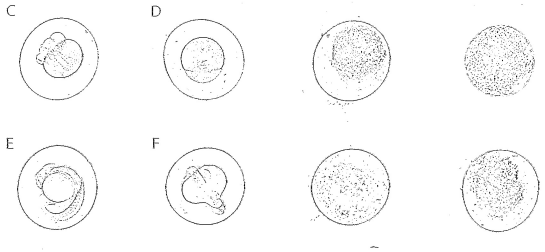

Figure 6: *Results of valley detection.*

The valley image is a binary image showing valley points (Fig. 6). The circle resulting from circle identification in the valley image is drawn over the attenuation image in Fig. 7.

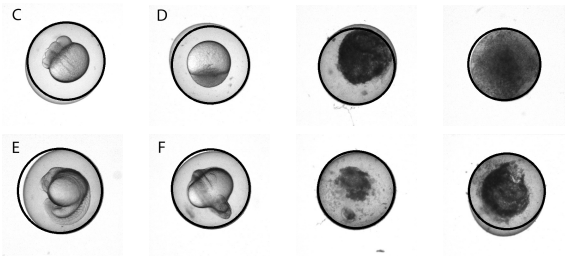

Figure 7: *Circles computed by the Hough transformation of the valley image drawn on the illumination compensated original images. We keep only the strongest circle from candidates where the radius is approximately similar to the expected embryo radius.*

4.2.3. Transformation to polar coordinates

The boundaries of the chorion and the yolk are roughly concentric circles, thus the algorithm extracting these curves should prefer such geometric arrangements. The arrangement of two circles is complicated, thus we transform the images into a space where ideal concentric circles become two horizontal lines. In the transformed space, curve extraction prefers horizontal lines. To set up a polar coordinate system where circular geometry is easier to handle, we have to find the center of the embryo.

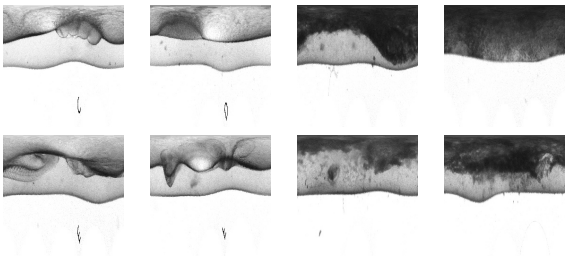

Figure 8: *Embryos of Fig. 2 in polar coordinates.*

Having identified the circle of center o_x, o_y in physical space approximating the boundary of the embryo, the polar map $p(\theta, r)$ is obtained as

$$p(\theta, r) = a(o_x + (r_{\max} - r)\cos(\theta), o_y + (r_{\max} - r)\sin(\theta))$$

where $\theta \in [0, 2\pi]$ and $r \in [0, r_{\max}]$, $r_{\max} = 2$ [mm]. The Jacobi determinant of this mapping is $(r_{\max} - r)^2$, thus it expands regions in the center, i.e. the yolk region, which is particularly useful when the yolk is our region of interest. The polar map is generated at $N = 256$ resolution.

4.3. Image descriptors based on the intensity map

Descriptors can be computed on the original attenuation map as well as its on the polar map. We expect polar map descriptors to be more distinctive since here the yolk represents a larger part of the image.

In polar coordinates the top of the figure corresponds to the center of the embryo where the yolk is located (Fig. 8). Here, the texture of dead embryos is darker and noisier.

To compute mechanical parameters like the mass and moment of inertia, a mass density parameter is needed, which is defined as the intensity of the inverted image:

$$\rho(x,y) = 1 - I(x,y).$$

In this way darker regions are taken into account in the mechanical oriented descriptors.

4.3.1. Mass

The mass is the integral of the mass density over the image, which can be computed for the attenuation image and also for the polar map. As the polar map is a distorted image, the two values will be different.

4.3.2. Moment of inertia

For the original image the moment of inertia is computed for an axis that is perpendicular to the image plane and crosses it in the center of the embryo:

$$\Theta = \int_x \int_y \rho(x,y)((x-o_x)^2 + (y-o_y)^2)\mathrm{d}x\mathrm{d}y.$$

Based on the recognition that dead embryos have darker yolk, we compute the inertia of the image emphasizing the contribution more toward the top. An appropriate measure is the moment of inertia of the inverted image with respect to the axis of the lower edge of the image:

$$\Theta_{\text{polar}} = \int_\theta \int_r \rho(x,y) r^2 \mathrm{d}r\mathrm{d}\theta.$$

4.3.3. Intensity histogram based descriptors

The intensity histogram is computed for pixels inside the circle corresponding to the chorion boundary. The histograms of living and dead embryos and also their average histograms are shown by Fig. 9.

Note that the histograms of dead embryos peak around darker colors while the histograms of living embryos around lighter colors. Based on the Alshut's approach[1], we calculate the *center of mass of the intensity histogram* and use it as an image descriptor.

Another descriptor can be obtained by calculating the "distance" between the current histogram and the histograms of average-living and average-dead. The average-living and average-dead histograms are obtained during the training of the system. There are different possibilities to characterize the difference of two distributions or functions. We could use, for example, the sum of the absolute differences (L_1 distance) or the square root of the sum of the squared differences (L_2 distance), but these metrics do not exploit the fact that we compare two distributions.

For two distributions, the *Kullback–Leibler (K-L) divergence* can be used, which is a measure of the information lost when distribution q_i is used to approximate distribution h_i. By definition, the K-L divergence of distribution q_i from h_i is

$$D_{KL}(h||q) = \sum_i \log\left(\frac{h_i}{q_i}\right) h_i.$$

K-L divergence is not a real metric since it is not symmetric and does not satisfy the triangle inequality. We select the average histogram to be q_i and require all bins to have a non-zero value. The current embryo is characterized by h_i. If h_i is zero, then $0 \log 0$ is replaced by zero.

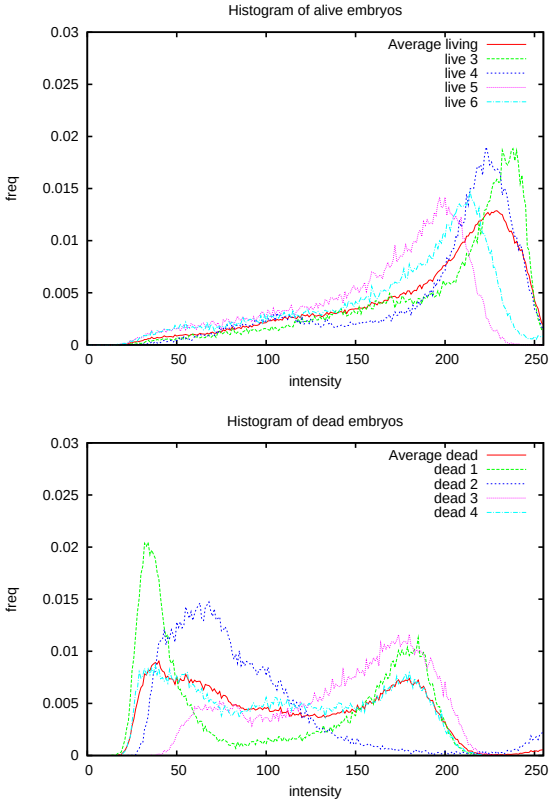

Figure 9: *Intensity histograms and average histograms of living and dead embryos, respectively.*

4.3.4. Intensity in the center of the chorion

Having identified the circle approximating the chorion boundary, the intensity value averaged in a small neighborhood of the center is computed.

4.3.5. Entropy of the gradient histogram

The yolk texture seems to be noisier in dead embryos, thus the noise level needs to be characterized. There are several options for this. We can use the Fourier descriptors or examine histograms. We examined both gray level histograms and gradient direction histograms. Histograms can be used directly as image descriptors in classification, or we can further compress information by calculating a single value from each histogram.

We used the entropy to characterize a gradient direction histogram with a single value:

$$S = -\sum_i p_i \log p_i,$$

where p_i is the probability of bin i that is calculated as the

relative frequency. Entropy is large when the distribution is more uniform, i.e. the image is noisier[9].

4.4. Image descriptors based on valley detection

Having transformed the image into polar coordinates, the boundary of the chorion and the healthy yolk can be imagined as horizontal valleys of the depth field where intensity $p(x,y)$ corresponds to height. Based on the first and second derivatives, a binary image is generated where a pixel is 1 if the first derivatives in this and its upper neighbor have different signs (there is a zero crossing) and the second derivative is larger than an appropriate threshold (Fig. 10).

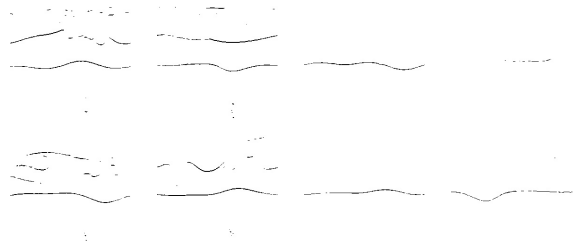

Figure 10: *Binary image of valley detection in the polar map.*

The classification of dead and healthy embryos is based on whether the valley around the yolk can be clearly identified. Note that the boundary of the chorion is always quite clear independently of the state of the embryo, and in addition to the boundaries there are shorter curve segments due to noise. Having identified the pixels that correspond to horizontal valleys, we check whether these pixels form two long mainly horizontal curves.

4.4.1. Lengths of yolk boundary segments

The binary image of valley detection is processed by a modified *flood filling* algorithm that prefers horizontally expanding curves.

This algorithm moves horizontally on white pixels until it can. When the next horizontal neighbor is black, it checks the upper right and lower right pixels and continues if one of them is white. If they are also black, the same check is repeated by only limited number of times, which defines the maximum gap this algorithm can fill. The algorithm stops when it reaches the right side or when the current curve cannot be continued. During flood filling, we report the lengths of the identified curves and keep the five largest values (Tab. 1).

We note that the largest value is close to 256, the resolution of the valley map, when the boundary of the chorion can be clearly identified. The remaining segments may belong to the boundary of the yolk. If they are longer or their sum is about 256, the yolk boundary is strong, which indicates a healthy embryo. To further compress information, we can find the segment length sufficit or deficit after removing the length of the chorion boundary which is expected to be 256 in both healthy and dead embryos. For living embryos, we expect a positive result, which is equal to 256 in the ideal case, and a negative or zero result for dead embryos.

Embryo	Segment lengths
Live1	199, 181, 79, 26, 13
Live2	255, 120, 111, 24, 12
Live3	241, 114, 45, 18, 17
Live4	238, 153, 39, 20, 17
Live5	232, 81, 34, 29, 25
Live6	249, 69, 26, 24, 12
Dead1	249, 0, 0, 0, 0
Dead2	20, 15, 5, 5, 2
Dead3	189, 34, 4, 3, 0
Dead4	53, 50, 48, 28, 21

Table 1: *Valley segment lengths for the embryos of Fig. 2. The polar map has 256 pixel resolution.*

5. Results

Tab. 2 summarizes the discussed descriptor values for the embryos of Fig. 2, including the *mass* of the inverted intensity image, *inertia* of both the inverted intensity image and of the polar map, the *intensity in the center of the chorion* (Int in Cent), the *center of mass of the intensity histogram* (Center Hist), the *Kullback-Leibler divergence* from the average alive and average dead, respectively, the *entropy of the gradient histogram of the polar map* and the *sufficit/deficit of the valley lengths in the polar map*. We can conclude that all these descriptors are able to deterministically separate the given 10 test cases. However, their performance is different on larger test sets.

The mass and inertia of the inverted intensity map is robust to the detection of the chorion but is less discriminative than the inertia of the polar map. The transformation to the polar map distorts the image emphasizing pixels in the center where the yolk is, which automatically improves the discriminative power. The intensity of the center and the center of mass of the intensity histogram together form a robust discriminator. Similarly, the Kullback-Leibler distances from the average dead and average living histograms should be considered together and classify an embryo based on their

Embryo	Mass	Inertia	Polar Inertia	Int in Cent	Center Hist	KL Live	KL Dead	Entropy	Valley length
Live1	1838	28	48	239	210	0.263	2.805	5.07	242
Live2	3103	46	116	115	186	0.134	2.36	5.11	266
Live3	2878	47	85	209	196	0.051	2.276	5.05	179
Live4	2727	41	92	159	194	0.05	2.35	5.12	211
Live5	4387	75	109	212	160	0.371	0.609	5.11	145
Live6	3776	61	103	200	172	0.138	1.114	5.12	124
Dead1	6571	103	220	50	103	1.507	0.228	5.16	-7
Dead2	6643	94	225	40	88	1.649	0.417	5.17	-209
Dead3	4806	78	169	82	142	0.677	0.282	5.13	-26
Dead4	6336	95	212	71	109	1.037	0.023	5.13	-56

Table 2: *Image descriptors for the embryos of Fig. 2.*

relation. The entropy of the gradient histogram is not too strong, so we plan to use other measures for the characterization of the gradient histogram. Finally, the valley length sufficit/deficit is also good, but it is sensitive to the threshold setting determining the minimum curvature for valleys.

6. Conclusions

In this paper we discussed and analyzed several image descriptors to classify alive and dead zebrafish embryos. The power of individual descriptors is improved further by applying all of them in a support vector machine if descriptors are "orthogonal" in the sense that they focus on different image features. This is the main reason, we proposed descriptors falling into three categories, characterizing the darkness, nosiness and the boundary of the yolk.

Acknowledgement

This work was supported by OTKA K–104476 and the Hungarian Government, managed by the National Development Agency, and financed by the Research and Technology Innovation Fund (grant no. KMR 12-1-2012-0221).

References

1. R. Alshut, J. Legradi, R. Mikut, U. Strahle, and M. Reischl. Robust identification of coagulated zebrafish eggs using image processing and classification techniques. In *19th Workshop on Computational Intelligence*, pages 9–21, 2009.
2. Thomas Braunbeck and Eva Lammer. Fish embryo toxicity assays. Technical Report UBA Contract Number 203 85 422, Aquatic Ecology and Toxicology, Department of Zoology, University of Heidelberg, 2006.
3. Ralf Mikut et al. Automated processing of zebrafish imaging data: a survey. *Zebrafish*, 10(3):401–421, 2013.
4. N. Jeanray, R. Maree, B. Pruvot, O. Geurts P. Stern, L Wehenkel, and et al. Phenotype classification of zebrafish embryos by supervised learning. *Toxicol Lett*, pages 211–215, 2012.
5. József Koloszár, László Szirmay-Kalos, Zsolt Tarján, and Dávid Jocha. Shape based computer aided diagnosis and automated navigation in virtual colonoscopy. In *Spring Conference on Computer Graphics*, pages 113–119, Budmerice, 2006.
6. R. Liu, S. Lin, R. Rallo, Y. Zhao, R. Damoiseaux, and et al. Automated phenotype recognition for zebrafish embryo based in vivo high throughput toxicity screening of engineered nanomaterials. *PLoS ONE*, 7(4):e35014, 2012.
7. L. Neumann, B. Csébfalvi, A. König, and E. Gröller. Gradient estimation in volume data using 4D linear regression. In *Computer Graphics Forum*, volume 19(3), pages 351–358, 2000.
8. L. Szirmay-Kalos, L. Szécsi, and M. Sbert. *GPU-Based Techniques for Global Illumination Effects*. Morgan and Claypool Publishers, San Rafael, USA, 2008.
9. Balázs Teréki, Pirkko Oittinen, and László Szirmay-Kalos. Informational aesthetic measure for 3d stereoscopic imaging. In *KÉPAF*, pages 57–66, 2013.
10. W. Wang, X. Liu, D. Gelinas, B. Ciruna, and Y. Sun. A fully automated robotic system for microinjection of zebrafish embryos. *PLoS ONE*, 2(9):e862, 2007.