

An Application of Shape-Based Level Sets to Fish Detection in Underwater Images

Mehdi Ravanbakhsh (mehdi.r@rmit.edu.au)
Mark R. Shortis (mark.shortis@rmit.edu.au)
RMIT University, GPO Box 2476, Melbourne, VIC 3001 Australia

Faisal Shaifat (faisal.shafait@uwa.edu.au)
Ajmal Mian (ajmal.mian@uwa.edu.au)
The University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009 Australia

Euan S. Harvey (euan.harvey@curtin.edu.au)
Curtin University, GPO Box U1987, Perth, WA 6845 Australia

James W. Seager (jseager@seagis.com.au)
SeaGIS P/L, PO Box 1085, Bacchus Marsh, VIC 3340 Australia

Abstract

Underwater stereo-video technology systems are used widely for measurement of fish. However the effectiveness of the stereo-video measurement has been limited because most operational systems still rely on a human operator. In this paper, an automated approach for fish detection using a shape-based level sets framework is presented. Shape knowledge of fish is modelled by Principal Component Analysis (PCA). The Haar classifier is used for precise position of the fish head and snout in the image, which is vital information for close proximity initialisation of the shape model. The approach has been tested on under-water images representing a variety of challenging situations typical of the underwater environment, such as background interference and poor contrast boundaries. The results obtained demonstrate that the approach is capable of overcoming these limitations and capturing the fish outline at sub-pixel accuracy.

Keywords: image segmentation, fish detection, under-water image, level sets, prior shape knowledge, registration

Introduction

The monitoring of fish for stock assessment in aquaculture, commercial fisheries and in the assessment of the effectiveness of biodiversity management strategies such as Marine Protected Areas and closed area management is essential for the economic and environmental management of fish populations. Video based techniques for fishery independent and non-destructive sampling are now widely accepted. The advantages of using stereo-video for counting the numbers of fish, measuring their lengths and defining the sample area have been well demonstrated (Shortis et al., 2009). However the effectiveness of the stereo-video measurement has been limited because most operational systems still rely on a human operator to identify and measure the snout and tail of the fish in order to determine the length by intersection. Whilst automation of identification of objects and image measurement processes have been demonstrated in many other contexts, due to the uncontrolled underwater environment combined with the loss of contrast because of attenuation through the water, an automated solution for fish sizing has been elusive. Whilst automation of some aspects of the process has been established for at least 15 years (Lines et al., 2001), only recently have fully operation systems that identify, delineate, track and measure fish in an uncontrolled environment been reported (Shortis et al., 2013).

The ultimate aim of this research is to develop a general approach to the automatic measurement of fish in underwater environments. The focus of this work will be on identification and delineation of Southern Bluefin Tuna (SBT). In context of this research, automated detection methodologies comprise two steps: identification and subsequent delineation of the fish outline. The existing literature on fish detection has mainly focused on the identification step where the presence of fish is recognised in the scene followed by the estimation of the fish location (Palazzo et al., 2013; Spampinato et al., 2008; Walther et al., 2004; Zhou and Clark, 2006; Morais et al., 2005; Evans et al., 2003). In contrast, relatively few approaches have been reported that deal with both identification and the following delineation of the fish silhouette (Khanfar et al., 2010; Lines et al., 2001; Hariharakrishnan & Schonfeld, 2005). Most of these approaches use low-level image features such as colour, texture, intensity and motion to detect fish. However, in a real life, the uncontrolled underwater environment

produces images that are characterised by low contrast, background clutter and interference, partial occlusion caused by adjacent or foreground objects, varied illumination conditions and shadows. The aforementioned research works fail to produce high quality results mainly due to misleading low-level features resulting from image noise and occlusion, or lack of sufficient low-level features necessary for object modelling. High-level knowledge of the shape of the fish can significantly aid in providing an efficient solution to these problems.

In this paper, an automated approach for fish detection using a shape-based level sets framework is presented. An example of under-water stereo images used is shown in Figure 1. The prior knowledge of the shape of the fish is modelled using Principal Component Analysis (PCA) (Leventon et al., 2000) and this knowledge is used to guide the level set curves. PCA enables the representation of global shape variation of the object of interest through a training set of shape templates. The global shape information is incorporated into the Mumford-Shah functional, as reported by Chan and Vese (2001), which can detect objects in strongly cluttered scenes. A Haar-like detector method (Lienhart and Maydt, 2002) is used to identify the existence of fish and determine their locations in the image. This information is vital to place the initial shape in close proximity to the object to be segmented, which increases the success rate and requires less iteration for convergence. Once the fish are independently identified on the left and right images, stereo intersections for the snout and tail is computed based on the well-established approach of a geometrically constrained epipolar search and template match between the two images.



Figure 1: Typical stereo image pair captured during a transfer from the purse-seine net to the grow-out cage. The water surface is to the right of the images and the apparent vertical orientation of the fish is caused by the mounting of the stereo-video system on the side of the transfer gate.

The outline of the paper is as follows. In the following section, a short review of level sets is given followed by the description of the individual steps of the proposed detection strategy along with mathematical equations in the subsequent section. Then, experimental results using underwater sample image sequences recorded in cages are presented and evaluated. The paper concludes with a discussion of the progress and results achieved, and an outlook for future work.

Level Set Representation

The core idea of level sets is to implicitly represent a contour C as the zero level curve of a function ϕ of higher dimension (Figs. 2-a & 2-b). An initialisation of ϕ can be constructed in the following way: Let C be a closed curve representing the boundary between two regions, one region inside the curve and another region outside the curve. ϕ is then defined as the signed distance $\pm d(x)$ to the curve, negative inside and positive outside. The definition is illustrated:

$$\phi(x) = \begin{cases} -d(x) & \text{if } x \text{ is inside } C \\ +d(x) & \text{if } x \text{ is outside } C \end{cases} \quad (1)$$

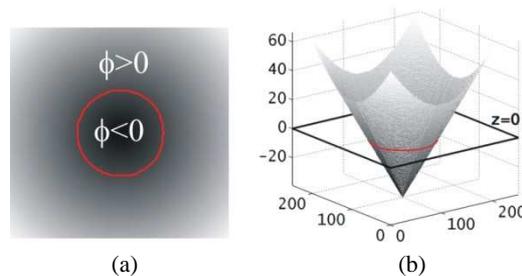


Figure 2: Illustrating level sets. (a) The curve C (red) is used to construct the level set function ϕ such that ϕ is negative inside and positive outside the curve. Distance values d are grey value coded. (b) A plane at zero level ($Z=0$) intersects the level set function ϕ , and thus the zero level curve C is obtained.

While the use of the distance $d(x)$ is not mandatory when using level sets, it assures that ϕ does not become too flat or too steep near C and subsequently can be differentiated across the zero level curve without running into numerical problems.

In order to combine the characteristics of the level set function, image information and shape knowledge of the known object, an energy functional can be set up and consequently minimised using the calculus of variations. Minimising the energy functional is performed in an iterative process moving the initial curve towards the object boundaries.

Detection Strategy

The fish detection strategy comprises three primary steps (Figure 3). First, the presence of fish is recognised and the initial locations are determined using segmentation of a frame difference from an averaged background image. A Haar like detector is then employed to estimate the snout and tail locations, from which the initial position and orientation of each fish in the image can be derived. Subsequently, a shape prior model is constructed by PCA using a set of training samples. The level sets curve is then initialised and evolved to locate the fish boundary. The result consists of the detected fish.

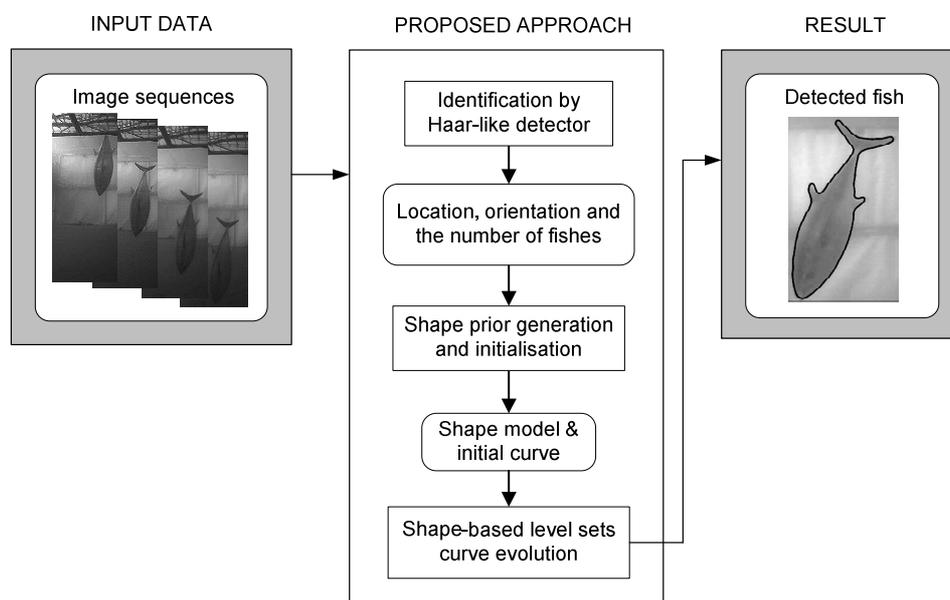


Figure 3: Workflow of fish detection

Identification

In this stage, the location of fish snout and tail in the image are determined. Precise localisation of the snout and tail leads to the estimation of pose parameters in 2D space, these being two rotations, two translations and one scale parameter.

In this research, the Haar classifier is used to locate the fish snout and tail. To train the classifier, 200 manually cropped images of the target object (snout or tail) are used so that the classifier can learn which features (among a set of possibly thousands of features) can locate the target with high accuracy. These features, once learned, are then used to construct the object classifier that can locate the presence of the object in cluttered scenes. Due to their high detection speed and ability to perform a scale-space search, Haar classifiers are employed in this research for locating snout and tail of fish in underwater image sequences. The results of independent detection of the snout and tail using Haar detectors are further improved by using the expected distance and angle relationships between the detected snouts and tails. The search space for tail detection is based on the results of the snout detection and vice versa. Figure 4 shows an example of a the result from the Haar classifier used to identify the snouts and tails and of Southern Bluefin Tuna (SBT) during a transfer.

Precise localisation of the tip of the snout and the valley point of the tail, used as reference points, are used to estimate the rigid transformation parameters. These transformation parameters are then used to first generate the reference shape and subsequently initialise the shape model, two crucial steps in accurate and correct delineation of fish.

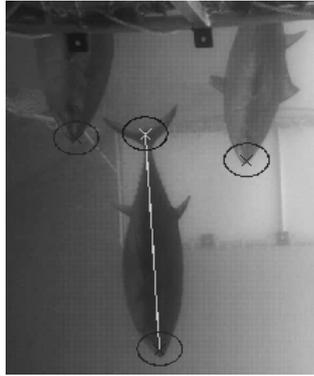


Figure 4: Shows the identification of SBT snouts and tails marked by circles using Haar classifier.

Shape Prior Generation

The generation of initial shape, also called *shape prior*, comprises two steps: first, the training samples need to be geometrically aligned, and subsequently, the shape model is constructed from the aligned shapes. The alignment involves matching shapes of training samples that differ in size, orientation and translation. In the literature, a large number of shape matching methods have been reported. A complete review of those methods is given in Veltkamp and Hagedoorn (1999).

In this paper, the alignment of training samples is realised using the method introduced in Chen et al. (2002). Suppose that the training set contains n given curves C_1, \dots, C_n with their corresponding interior regions A_1, \dots, A_n . The shape similarity measure of the shapes C_1 and C_2 is defined as:

$$a(C_1, C_2) = \text{area of } (A_1 \cup A_2 - A_1 \cap A_2) \quad (2)$$

In the alignment process, the pose parameter of C_1 is considered to be fixed, and the rest of samples (C_2, \dots, C_n) are jointly aligned to C_1 through the solution of the rigid transformation $C_j^{\text{new}} = s_j R_j C_j + T_j$ ($j=2, \dots, n$) such that the area $a(C_1, C_j^{\text{new}})$ is minimised. These values are obtained by a global optimisation algorithm called the genetic algorithm (Davis, 1991), which makes it less likely for the underlying function to be trapped in suboptimal local minimum compared with purely local methods such as gradient descent.

The shapes are encoded in binary images to simplify the alignment task. Figure 5 shows a set of 20 training samples manually digitised and the result of their alignment. The first sample (bottom-row, left-most), which is the scaled, shifted and rotated version of the corresponding sample manually digitised (top-row, left-most), is adopted as the reference. It has fixed pose parameters estimated in the identification process and to which the rest of samples are registered. Figure 6-a & 6-b show the amount of shape variability depicted in the overlap images before and after the alignment. It can be seen that even large shape discrepancies can often exist in real fish images. These shape differences can be removed successfully which demonstrates the effectiveness of the alignment method. Furthermore, model variability is represented in Figure 6-c showing that the areas around the boundaries of the fish fin and tail experience the largest deformations in the fish body outline. It is interesting to note that key regions that could be used for species identification, such as the dorsal and anal fins and the tail, are the profile sections which show the greatest variability.

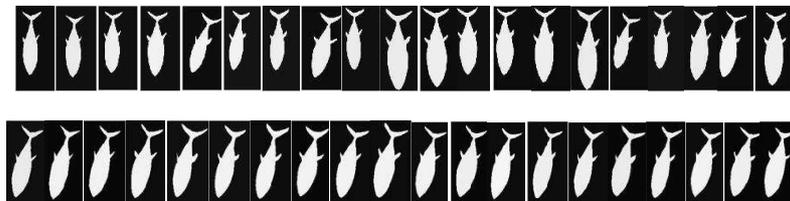


Figure 5: Top-row shows binary representation of training samples of fish shapes. Bottom-row presents the training samples after geometric alignment.

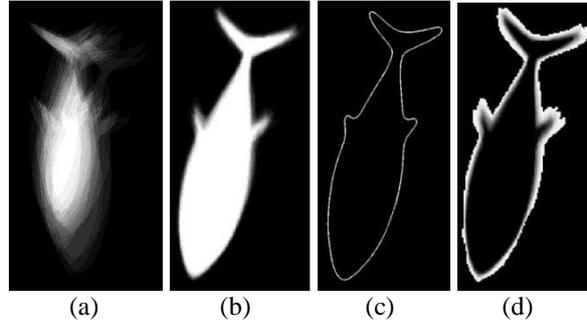


Figure 6: (a) Overlaid training samples with varying degrees of overlap before alignment; (b) Aligned samples; (c) Average of aligned shapes; (d) Showing model variability which are gray-value coded with white and black representing highest and lowest variability respectively.

In the next step, a shape model is constructed using the aligned shapes. The PCA method is selected to construct the shape model due to its efficiency at capturing the main variations of a training set while removing redundant information. Similar to Leventon et al. (2000), the boundaries of each of the training shapes are represented in the training dataset as the zero level set of n Signed Distance Functions (SDFs) $\{\phi_1 \dots \phi_n\}$ with negative distances assigned to the inside and positive distances assigned to the outside of the shape boundary.

Suppose M is a matrix whose column vectors are the n aligned training SDFs $\{\phi_i\}$, PCA is then applied to these SDFs to compute eigenvalues and eigenvectors of the covariance matrix:

$$\Sigma = \frac{1}{n} M M^T \quad (3)$$

and the mean level set function of the training set

$$\bar{\phi} = \frac{1}{n} \sum_{i=1}^n \phi_i \quad (4)$$

The eigenvectors are called principal components or eigenshapes. In practice, the first k principal components ($k \leq n$) are sufficient to model the major shape variations in the training samples. In Milka et al. (1999), a method is proposed for determining the value of k by examining the eigenvalues of the corresponding eigenvectors. This approach however cannot be adopted here as the value of k varies in different applications (Tsai et al., 2003). In this work, the value of k was set empirically. Then, shape can be represented as zero level set of the following function

$$\phi(\mathbf{w}) = \bar{\phi} + \sum_{i=1}^k w_i \phi_i \quad (5)$$

where $\mathbf{w} = \{w_1 \dots w_k\}$ denote the weights for the k eigenshapes with the variances of these weights $\{\sigma_1^2 \dots \sigma_k^2\}$ given by the eigenvalues. In the equation (5), the shape variability is restricted to the variability given by the eigenshapes. To accommodate wider range of shape variability, pose parameters \mathbf{p} , these being translation, scale, orientation, are incorporated to the level set function of (5). With the addition of \mathbf{p} , the implicit description of shape is given by the zero level set of the following function

$$\phi(\mathbf{w}, \mathbf{p}) = \bar{\phi} + \sum_{i=1}^k w_i \phi_i \quad (6)$$

where $\bar{\phi}$ and each ϕ_i are now a function of \mathbf{p} .

Once the shape model is generated, an initial level set function is constructed using a rectangle curve around the detected fish. Then, the zero level set of the level set function is evolved towards the fish boundary according to the energy functional. The energy functional is described in the following section.

Shape-Based Level Sets Energy Functional

The energy functional is based on the segmentation model proposed by Chan and Vese (2001) in an effort to overcome limitations found with the previous edge-based strategies. Unlike edge-based methods where the provision of close initialization to the object of interest and good contrast boundaries are necessary to locate those boundaries, region-based methods used in this work are independent of image gradients and less likely to converge to local minima if an undesirable feature or image noise is present.

Let I be a given image and C the evolving curve defined as $C = \{(x,y) \in \mathbb{R}^2: \phi(x,y) = 0\}$, with u and v denoting two constants representing the averages of I inside and outside the curve C . Assume that the image I is formed by two regions of approximately piecewise-constant intensities with distinct values of I_0^i and I_0^o , and that the object to be detected is represented by the region with value I_0^i and boundary C . Then, $I_0 \approx I_0^i$ inside the object (inside C) and $I_0 \approx I_0^o$ outside the object (outside C). By minimizing the following energy equation, the boundary of the object of interest C is obtained (Chan and Vese, 2001)

$$E_{cv} = \int_{\text{inside } C} |I - u|^2 dx dy + \int_{\text{outside } C} |I - v|^2 dx dy \quad (7)$$

which is equivalent to the energy functional below (Tsai et al., 2001)

$$E_{cv} = -(u^2 A_u + v^2 A_v) = -\left(\frac{S_u^2}{A_u} + \frac{S_v^2}{A_v}\right) \quad (8)$$

where A_u and A_v denote areas, and S_u and S_v represent the sum intensity of areas inside and outside C . Then, the gradient descent is employed to search for the parameters \mathbf{w} and \mathbf{p} that minimise E_{cv} to implicitly determine the segmenting curve C . The parameters A_u, A_v, S_u and S_v can be expressed in terms of $\phi(\mathbf{w}, \mathbf{p})$

$$S_u = \int_{\Omega} I H(\phi(\mathbf{w}, \mathbf{p})) dA; A_u = \int_{\Omega} H(\phi(\mathbf{w}, \mathbf{p})) dA \quad (9)$$

and

$$S_v = \int_{\Omega} I H(-\phi(\mathbf{w}, \mathbf{p})) dA; A_v = \int_{\Omega} H(-\phi(\mathbf{w}, \mathbf{p})) dA \quad (10)$$

where Ω defines a bounded and open subset of \mathbb{R}^2 and H denotes the Heaviside function

$$H(\phi(\mathbf{w}, \mathbf{p})) = \begin{cases} 1 & \text{if } \phi(\mathbf{w}, \mathbf{p}) < 0 \\ 0 & \text{if } \phi(\mathbf{w}, \mathbf{p}) \geq 0 \end{cases} \quad (11)$$

The energy function (8) is minimised with respect to \mathbf{w} and \mathbf{p} using gradient descent optimisation

$$\nabla_{\mathbf{w}} E_{cv} = -2(u \nabla_{\mathbf{w}} S_u + v \nabla_{\mathbf{w}} S_v) + (u^2 \nabla_{\mathbf{w}} A_u + v^2 \nabla_{\mathbf{w}} A_v) \quad (12)$$

$$\nabla_{\mathbf{p}} E_{cv} = -2(u \nabla_{\mathbf{p}} S_u + v \nabla_{\mathbf{p}} S_v) + (u^2 \nabla_{\mathbf{p}} A_u + v^2 \nabla_{\mathbf{p}} A_v) \quad (13)$$

where the gradient parameters are given as

$$\nabla_{w_i} A_u = -\nabla_{w_i} A_v = -\int_C \phi_i(\mathbf{p}) ds \quad (14)$$

$$\nabla_{w_i} S_u = -\nabla_{w_i} S_v = -\int_C I \phi_i(\mathbf{p}) ds \quad (15)$$

$$\nabla_{p_i} A_u = -\nabla_{p_i} A_v = -\int_C \nabla_{p_i} \phi(\mathbf{w}, \mathbf{p}) ds \quad (16)$$

$$\nabla_{p_i} S_u = -\nabla_{p_i} S_v = -\int_C I \nabla_{p_i} \phi(\mathbf{w}, \mathbf{p}) ds \quad (17)$$

where the segmenting curve C is given by the zero level set of $\phi(\mathbf{w}, \mathbf{p})$, and $\nabla_{p_i} \phi(\mathbf{w}, \mathbf{p})$ is the gradient of $\phi(\mathbf{w}, \mathbf{p})$ taken with respect to the i th component of the transformation matrix \mathbf{p} that includes translation, rotation and scale. The gradient descent optimisation of the equations (12&13) leads to the parameters \mathbf{w} and \mathbf{p} . The updated \mathbf{w} and \mathbf{p} parameters, which are iteratively computed during the optimisation, are then used to implicitly determine the location of the segmenting curve C .

The curve evolution is terminated when the overall change in the evolving curve positions per iteration is less than 0.1 pixels. A smaller threshold considerably increases the computation cost, although the quality of the final result is the same.

Experimental Evaluation

Underwater image sequences recorded at the transfer gate between two cages have been used to test the fish detection algorithm. From the large number of video samples recorded for 8 transfers, 35 sample images have been chosen to represent the variable and uncontrolled nature of the marine environment. These images include a

varying number of SBT with a range of illumination changes, background interference and occlusions caused by adjacent fishes. Moreover, SBT appear in the image sequences with missing or poor contrast boundaries which further exacerbates the challenging conditions.

In Fig 7, an example of results is shown where the initial curve is placed as a rectangle around the fish of interest and subsequently converged to the fish boundary by minimising the energy functional presented in the previous section. Further example results are shown in Figure 7 where, in the four right-most samples, SBT are partially occluded by other neighbouring fishes in foreground and background. Almost in all samples, fish boundaries are of low contrast especially in areas around the tail and fin. The detection results shown in Fig.7 demonstrate that the approach is capable of overcoming those limitations typical of the underwater environment and capturing the fish outline accurately.

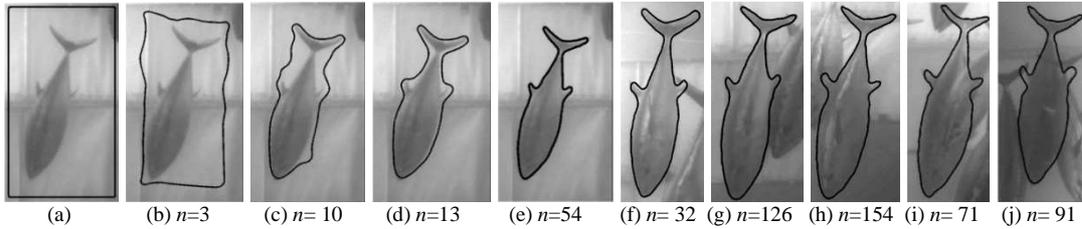


Figure 7: Fish detection result. (a) Initial curve; (b), (c) and (d) show the intermediate curves and (e) represents the final detection result. (f), (g), (h), (i) and (j) show the detection results of different fish in the presence of a range of background interference and foreground occlusions by other fish (two rightmost samples). n denotes the number of iterations in the intermediate and the final results.

In order to quantitatively evaluate the performance of the approach, the detection results were compared to manually plotted fish used as reference data. The comparison was carried out by matching the detection results to the reference data using the so-called *buffer method* (Heipke et al., 1998). A detected object is assumed to be correct if the maximum distance between the detected object and its corresponding reference does not exceed the buffer width. Furthermore, a reference object is assumed to be matched if the maximum deviation from the detected object is within the buffer width. Based on these assumptions the following quality measures were used in our work:

- *Completeness*: is the ratio of the number of matched reference objects to the whole number of objects.
- *Correctness*: is the ratio of the number of correctly detected objects to the number of detected objects.
- *Geometric accuracy*: is the average distance between the correctly detected objects and its corresponding reference expressed as root mean square (RMS) value.

Table 1 shows the evaluation result of the fish detection. The buffer width can be defined according to the required detection accuracy for a specific application. In our tests, the buffer was set to 3, 5 and 8 pixels according to the range of accuracy achievable at the identification step. Furthermore, this selection allows assessment of the relevance of the approach for applications that demand varying degrees of accuracy. From the buffer width value 3 pixels to 8 pixels, both the completeness and correctness have increased implying that the results are more complete and correct for higher buffer width values. The geometrical accuracy increases in inverse proportion to the buffer width value, so that results obtained with a value of 3 pixels are more accurate than those obtained with a larger buffer width value.

Buffer width (pixel)	Correctness (%)	Completeness (%)	Geometric accuracy (pixel)
3	89.6	91.4	0.7
5	94.3	94.3	0.8
8	100	100	0.9

Table I: Evaluation results for fish detection applied on 35 samples

As expected, the results are encouraging, but whilst sub-pixel geometric accuracy has been achieved in all experiments with high rates of completeness and correctness, severe deformation taking place around the fins and the tail of the fish cannot be absorbed with the current approach. The table nevertheless shows that the developed approach is in principle capable of extracting fish accurately under occlusion and within variable underwater environments.

Accurate extraction of the shape is important for fish biomass estimation, length measurement and species recognition (Shortis et al., 2013). In each case an accuracy of one pixel would be sufficient to establish the

initial conditions, so even the least favourable accuracy result in the table above would still be acceptable and simultaneously provide a high level of correctness and completeness.

Conclusion and Outlook

In this paper, an automated approach for the detection of fish from under-water images has been proposed, developed and tested. It comprises a region-based level set method that enables the delineation of the fish outline. The shape information of fish is incorporated into the level sets formulation through the PCA method to overcome such limitations as poor contrast boundaries, background clutter and occlusions caused by neighbouring fish. To provide a close initialisation for the shape model, the pose of fish in the image is determined using the Haar classifier. The results of the developed approach have been applied to 35 samples of varying quality and occlusion level and presented a quantitative evaluation of the results using three buffer width values.

The presented results show that level sets can be used to delineate fish outlines from under-water images if the shape information of the fish species is incorporated into the level sets energy functional. Furthermore, it was found that an energy function that is independent of image gradients and includes the shape model is able to overcome various kinds of disturbances and the problems related to low quality images recorded in the underwater environment, such as poor contrast and uneven illumination.

The current approach has been developed to detect SBT in an aquaculture environment. The techniques developed here have clear potential to be extended to wild habitats provided that the perspective deformation of the fish body and movement information derived from image sequences are taken into account. In wild habitats, fish can move in any direction with large deformations occurring in the image of the body, causing this fish detection approach to break down.

For the technique to be successful in wild habitats, varying rates of deformation and fish orientation need to be modelled. The detection of different fish species in addition to SBT is another goal that will be pursued in future research, as in reef and other underwater habitats many fish species are present. Furthermore, investigation into the possibility of using colour information in the level sets formulation will be carried out.

References

- Chan, T.F. and Vese, L.A., 2001. Active contours without edges. *IEEE Trans. on Image Processing*, 10(2): 266–277.
- Chen, Y., Tagare, H., Thiruvankadam, S., Huang, F., Wilson, D., Gopinath, K., Briggsand, R. and Geiser, E., 2002. Using prior shapes in geometric active contours in a variational framework. *International Journal of Computer Vision*, 50(3): 315–328.
- Davis, L., 1991. Handbook of Genetic Algorithms. Van Nostrand: 100 pages.
- Evans, F., 2003. Detecting fish in underwater video using the EM algorithm. Proceedings of the 2003 *IEEE International Conference on Image Processing*, 3: III – 1029–1032.
- Hariharakrishnan, K. and Schonfeld, D., 2005. Fast object tracking using adaptive block matching. *IEEE Transactions on Multimedia* 7(5): 853–859.
- Heipke, C., Mayer, H., Wiedemann, C. and Jamet, O., 1998. External evaluation of automatically extracted road axes. *Photogrammetrie, Fernerkundung, Geoinformation*, 2: 81–94.
- Khanfar, H., Charalampidis, D., Ioup, G., Ioup, J. and Thompson, C. H., 2010. Automated recognition and tracking of fish in underwater video. Final Report, LA Board of Regents Contract NASA(2008)-STENNIS-08: 40 pages.
- Leventon, M., Grimson, W. and Faugeras, O., 2000. Statistical shape influence in geodesic active contours. *IEEE International Conference of Computer Vision and Pattern Recognition*, 1: 316–323.
- Lienhart, R. and Maydt, J., 2002. An extended set of Haar-like features for rapid object detection. Proceedings, *IEEE International Conference on Image Processing*, 1:900-903. doi: 10.1109/ICIP.2002.1038171
- Lines, J.A., Tillett, R.D., Ross, L.G., Chan, D., Hockaday, S. and McFarlane, N.J.B., 2001. An automated image-based system for estimating the mass of free-swimming fish. *Journal of Computers and Electronics in Agriculture*, 31(2): 151–168.
- McInerney, T. and Terzopoulos, D., 1995. Topologically adaptable snakes. *Proceedings of the Fifth IEEE International Conference on Computer Vision*: 840–845.
- Mika, S., Schölkopf, B., Smola, A., Müller, K.R., Scholz, M. and Rätsch, G., 1999. Kernel PCA and de-noising in feature spaces. *Advances in Neural Information Processing Systems*, MIT Press, 11: 536–542.
- Morais, E.F., Campos, M.F.M., Padua, F.L.C. and Carceroni, R.L., 2005. Particle filter-based predictive tracking for robust fish counting. *18th IEEE Brazilian Symposium on Computer Graphics and Image Processing*: 367–374.
- Palazzo, S., Kavassidis, I. and Spampinato, C., 2013. Covariance based modeling of underwater scenes for fish detection. *Proceedings of IEEE International Conference on Image Processing*, Melbourne, Australia. Paper 3591, 5 pages. Available at <http://groups.inf.ed.ac.uk/f4k/PAPERS/ICIPs13.pdf>

- Shortis, M. R., Harvey, E. S. and Abdo, D. A., 2009. A review of underwater stereo-image measurement for marine biology and ecology applications. In *Oceanography and Marine Biology: An Annual Review*, Volume 47, Gibson, R. N., Atkinson, R. J. A. and Gordon, J. D. M. (Editors). CRC Press, Boca Raton FL, USA. ISBN 978-1-4200-9421-3. 342 pages.
- Shortis, M.R., Ravanbakhsh, M., Shafait, F., Harvey, E.S., Mian, A., Seager, J.W., Edgington, D, Cline, D. and Culverhouse P., 2013. A review of techniques for the identification and measurement of fish in underwater stereo-video image sequences. *Videometrics, Range Imaging, and Applications XII*, SPIE Vol. 8791, paper 0G. The International Society for Optical Engineering, Bellingham WA, USA.
- Spampinato, C., Chen-Burger, Y.-H. , Nadarajan, G. and Fisher, R., 2008. Detecting, Tracking and Counting Fish in Low Quality Unconstrained Underwater Videos, 2: 514–519.
- Tsai, A., Yezzi, A., Wells, W., Tempny, C., Tucker, D., Fan, A., Grimson, W.E. and Willsky, A., 2003. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Transactions on Medical Imaging*, 22(2): 137–154.
- Tsai, A., Yezzi, A., Wells, W., Tempny, C., Tucker, D., Fan, A., Grimson, W. and Willsky, A., 2001. Model-based curve evolution techniques for image segmentation. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1: 463–468.
- Veltkamp, R. and Hagedoorn, M., 1999. State-of-the-art in shape matching. Technical Report UU-CS-1999-27, Utrecht University, Sept. 1999.
- Walther, D., Edgington, D. and Koch, C., 2004. Automated video analysis for oceanographic research. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*: 544–549.
- Zhou, J. and Clark, C., 2006. Autonomous fish tracking by ROV using monocular camera. *The 3rd Canadian Conference on Computer and Robot Vision*: 8 pages.