Computers and Electronics in Agriculture 150 (2018) 52–61

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Enhanced fish bending model for automatic tuna sizing using computer vision

P. Muñoz-Benavent[a,*], G. Andreu-García[a], José M. Valiente-González[a], V. Atienza-Vanacloig[a], V. Puig-Pons[b], V. Espinosa[b]

[a] Institute of Control Systems and Industrial Computing (AI2), Universitat Politècnica de València (UPV), Camí de Vera (s/n), 46022 València, Spain
[b] Institut d'Investigació per a la Gestió Integrada de Zones Costaneres (IGIC), Universitat Politècnica de València (UPV), Camí de Vera (s/n), 46022 València, Spain

ARTICLE INFO

Keywords:
Underwater stereo-vision
Computer vision
Fisheries management
Automatic fish sizing
Biomass estimation

ABSTRACT

This paper presents a non-invasive fully automatic procedure to obtain highly accurate fish length estimation in adult Bluefin Tuna, based on a stereoscopic vision system and a deformable model of the fish ventral silhouette. The present work takes a geometric tuna model, which was previously developed by the same authors to discriminate fish in 2D images, and proposes new models to enhance the capabilities of the automatic procedure, from fish discrimination to accurate 3D length estimation. Fish length information is an important indicator of the health of wild fish stocks and for predicting biomass using length-weight relations. The proposal pays special attention to parts of the fish silhouette that have special relevance for accurate length estimation. The models have been designed to best fit the rear part of the fish, in particular the caudal peduncle, and a width parameter has been added to better fit the silhouette. Moreover, algorithms have been developed to extract snout tip and caudal peduncle features, allowing better initialization of model parameters. Snout Fork Length (SFL) measurements using the different models are extracted from images recorded with a stereoscopic vision system in a sea cage containing 312 adult Atlantic Bluefin Tuna. The automatic measurements are compared with two ground truths: one configured with semiautomatic measurements of favourable selected samples and one with real SFL measurements of the tuna stock collected at harvesting. Comparison with the semiautomatic measurements demonstrates that the combination of improved geometric models and feature extraction algorithms delivers good results in terms of fish length estimation error (up to 90% of the samples bounded in a 3% error margin) and number of automatic measurements (up to 950 samples out of 1000). When compared with real SFL measurements of the tuna stock, the system provides a high number of automatic detections (up to 6706 in a video of 135 min duration, i.e., 50 automatic measurements per minute of recording) and highly accurate length measurements, obtaining no statistically significant difference between automatic and real SFL frequency distributions. This procedure could be extended to other species to assess the size distribution of stocks, as discussed in the paper.

1. Introduction

Monitoring of wild fish stocks and inspection in aquaculture require extremely gentle handling of the target to avoid damage, but traditional sampling methods are usually invasive, expensive, time-consuming and laborious. Optical sensors and machine vision systems have proven to be very appropriate for developing faster, cheaper and non-invasive methods to work with live fish (in situ), as reported in recent years (Zion, 2012), (Shortis et al., 2016), (Mallet and Pelletier, 2014), (Boutros et al., 2015), (Hao et al., 2015), (Saberioon et al., 2017).

Automatic identification of a single fish is an essential step in achieving a fully automatic sizing process. Body bending while swimming means that the same individual is observed with very different shapes, sizes and orientations, depending on the video frame. So, robust fish detection methods dealing with these variations are required (Lines et al., 2001), (Rosen et al., 2013), (Rahim et al., 2012), (Atienza-Vanacloig et al., 2016). In (Atienza-Vanacloig et al., 2016) a deformable adaptive model based on computer vision methods that automatically fit the body ventral silhouette of adult Bluefin Tuna (*Thunnus Thynnus*) while swimming was proposed. This model achieved very high success rates (up to 90%) discriminating individuals in complex images acquired in real conditions, but it was not strong enough to estimate sizes.

* Corresponding author.
E-mail addresses: pmunyoz@upv.es (P. Muñoz-Benavent), gandreu@upv.es (G. Andreu-García), jvalient@upv.es (J.M. Valiente-González), vatienza@upv.es (V. Atienza-Vanacloig), vipuipon@upv.es (V. Puig-Pons), vespinos@upv.es (V. Espinosa).

https://doi.org/10.1016/j.compag.2018.04.005
Received 31 January 2018; Received in revised form 29 March 2018; Accepted 3 April 2018
0168-1699/ © 2018 Elsevier B.V. All rights reserved.

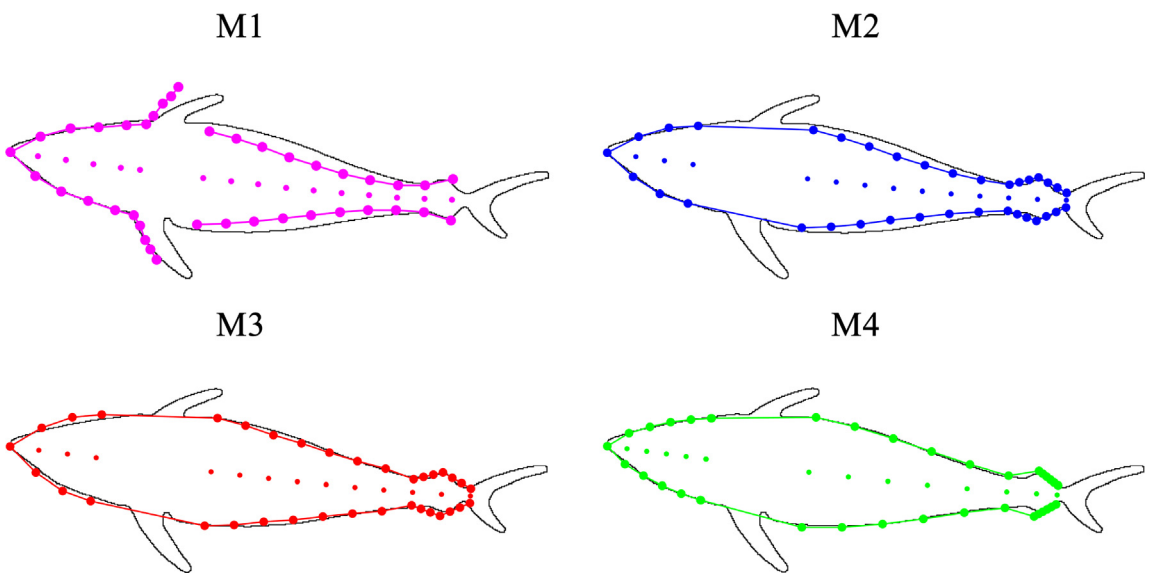

Fig. 1. Graphical representation of the four different geometric tuna models (**M1**, **M2**, **M3** and **M4**). Small dots representing model mid-body points and large dots representing model landmarks.

The purpose of this study is to define an effective geometric tuna model to automatically process the stereo videos and obtain accurate fish measurements. The present work takes the geometric model (**M1**) defined in (Atienza-Vanacloig et al., 2016) as a starting point and studies three new models (**M2**, **M3** and **M4**) to reach high similitude between models and real tuna silhouettes. Geometric models can provide a set of parameters and landmarks to capture the essential features of a tuna silhouette considering its variability. When the target has been identified and characterized in the images, 3D biometric measurements can be obtained from a calibrated stereo vision system. The models are compared from three points of view: quantity of successful fittings, computing time and accuracy of the length measurements.

To evaluate our proposals, fish length measurements using the different models are extracted from images recorded with a stereoscopic vision system under real conditions in Grup Balfegó growing farms, on the west Mediterranean coast. These measurements are compared with semiautomatic measurements of selected samples and with true data from Snout Fork Length (SFL) measurements of the tuna stock collected by Grup Balfegó at harvesting. The results confirm the potential of our fully automatic sizing method, which could be applied to monitor fish in aquaculture for growth management purposes and biomass estimation in fish transfers between cages and to monitor wild fish stocks.

1.1. State of the art

A variety of applications with optical sensors and machine vision systems have been developed to work in underwater conditions: fish sizing (Ruff et al., 1995), (Tillett et al., 2000), (Lines et al., 2001), (Harvey et al., 2003), (Costa et al., 2006), (Dunbrack, 2006), (Torisawa et al., 2011), (Letessier et al., 2015), (Williams and Lauffenburger, 2016); fish counting and sizing (Costa et al., 2009), (Rosen et al., 2013); fish sizing in combination with acoustic techniques (Sawada et al., 2009), (Espinosa et al., 2011), (Kloser et al., 2011); fish farm automation (Martinez-de Dios et al., 2003); wild fish stock assessment (Willis and Babcock, 2000), (Watson et al., 2009), (Harvey et al., 2012), (Langlois et al., 2012), (Seiler et al., 2012), (Zintzen et al., 2012), (Wakefield et al., 2013), (Santana-Garcon et al., 2014), (McLaren et al., 2015) and fish species classification (Hu et al., 2012), (Zion, 2012), (Huang et al., 2013), (Spampinato et al., 2010),(White et al., 2006).

Fish measurements, such as length, height and width, are commonly used for different purposes: as indicators of health in wild fish stocks (Dunbrack, 2006), (Shortis et al., 2016), (Rosen et al., 2013), (Shafait et al., 2017); for biomass estimation to control fishing quotas (ICCAT, 2015), to monitor growth rates in fish farms (Puig et al., 2012); and for fish sorting and grading (Hong et al., 2014), (Zion et al., 2007), (Hao et al., 2016), (Shafait et al., 2017). Measurements of live fish can be achieved with underwater stereoscopic vision systems, two cameras in a side-by-side arrangement, as recommended by the International Commission for the Conservation of Atlantic Tunas (ICCAT) in (ICCAT, 2015), to control catches for tuna farming. Nevertheless, vision sensors and image processing methods have to overcome difficulties such as limited visibility, temporal and spatial variations in lighting, varying distances and aspects between cameras and objects, motion and density of the monitored targets, and even lack of physical stability. Moreover, for the case of stereoscopic vision systems, the cameras must be synchronised to ensure temporal correspondence between both videos. In addition, underwater calibration of the system is required for accurate and reliable measurements. All these conditions represent a very demanding challenge, which have limited the development of fully automatic solutions. In fact, most of the aforementioned applications and the most widely used commercial systems for fish sizing, AQ1 AM100 (Phillips et al., 2009) and AKVAsmart, formerly VICASS (Shieh and Petrell, 1998), require human intervention in some of their stages, making the process slow and laborious, and introducing the variability of manual measuring. Some authors, such as (Lines et al., 2001), (Zion, 2012), (Shortis et al., 2016), (Atienza-Vanacloig et al., 2016), (Shafait et al., 2017), highlight the need for fully automatic methods for these tasks.

2. Geometric tuna models

Fish length and other features, such as bending angle, can be properly characterized using geometric models, since they are able to fit the tuna silhouette considering its variability due to different shapes, sizes and orientations. A geometric model for adult tunas (**M1**), formerly presented in (Atienza-Vanacloig et al., 2016) to discriminate individuals, is used in this paper to estimate fish length using a stereoscopic vision system. Based on that model (**M1**), three new models (**M2**, **M3** and **M4**) have been developed to improve the fit to the fish silhouette and the fish length estimation. Fig. 1 shows a graphical representation of these models.

The first main modification to model **M1** involves considering that areas with high variability, or with greater significance for providing accurate biometric measurements, should be represented by a greater

number or higher density of silhouette landmarks, which are estimated projecting the silhouette mid-body points. Both landmarks and mid-body points can be seen in Fig. 1 for the different models. Note that the area around the caudal peduncle has been modelled with a higher number of mid-body points in **M2**, **M3** and **M4**. The second main modification addresses the difficulties with modelling pectoral fins. In (Atienza-Vanacloig et al., 2016) they were included in the model because the proportionality between pectoral fin position and full body length helped to deliver better fitting in the model length parameter. However, they are seen in the images with highly variable shapes, which hinders silhouette fitting and results in fewer automatic measurements. Instead, in the new models, accuracy in the model length parameter is addressed by modelling the caudal peduncle, making the modelling of the pectoral fins unnecessary. The third main modification refers to the number of parameters used in each model, which implies different capabilities, from discrimination of individuals to highly accurate measurements. Moreover, feature extraction algorithms for caudal peduncle and snout tip recognition have also been developed to provide a better initialization of the model parameters, which is critical to avoid local minima and increase the number of successful fittings.

The parameters and features of the tuna models are summarized in Table 1, and the fish silhouettes for each model are shown in Fig. 1. In **M2**, the pectoral fins are removed and the caudal peduncle is modelled with a greater number of mid-body points for more accurate fish length estimation. In **M3**, the width feature is added to the fish model, providing a better fit to the fish silhouette. In **M4**, a different modelling of the caudal peduncle and an improved width parameter is applied to obtain highly accurate measurements. Note that the caudal fins are not included in any model due to their great variability, as mentioned in (Atienza-Vanacloig et al., 2016).

The different geometric models and the feature extraction algorithms are described in detail in the subsections below.

2.1. Base model with pectoral fins: M1 model

Fig. 2a shows the deformable model **M1** of tuna fish defined in (Atienza-Vanacloig et al., 2016) as a vector of five parameters $\mathbf{M1} = [s_x,s_y,l,\alpha,\theta]$, where: s_x and s_y give the image location of the snout tip; l is the length of the vertebral column; α denotes the angle of the fish head in relation to the horizontal axis, and θ is the global bending angle of the vertebral column.

The **M1** model is characterized by 20 mid-body points $v_i = (x_i^v, y_i^v)$, 16 of them distributed along the fish length l using a grid of size $S_l = l/16$, and 4 of them to model the pectoral fins (see Fig. 2b). The position of the mid-body points is computed according to the model parameters using the following equation:

$$\begin{pmatrix} x_i^v \\ y_i^v \end{pmatrix} = \begin{pmatrix} s_x \\ s_y \end{pmatrix} + \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} l_i\cos(\theta_i) \\ l_i\sin(\theta_i) \end{pmatrix} \quad (1)$$

where l_i is the length from the snout to the ith-vertebra and θ_i the bending angle of the ith-vertebra, which are calculated proportionally to model parameters l and θ using the following expressions:

$$l_i = \mathbf{c_l}(i)l; \quad \theta_i = \mathbf{c}_\theta(i)\theta; \quad i = 1...n \quad (2)$$

where n is the number of mid-body points and $\mathbf{c_l}$ and $\mathbf{c}_\theta$ are constant coefficient vectors.

The silhouette is modelled with 39 landmarks, one landmark for the snout tip and 19 landmarks for each side of the tuna body profile. The landmarks $k_i = (x_i^k, y_i^k)$ that configure the **M1** model silhouette are obtained from mid-body points v_i with the following expressions:

$$x_i^k = x_i^v \pm l_i w_i \sin\theta_i; \quad y_i^k = y_i^v \pm l_i w_i \cos\theta_i; \quad i = 1...n \quad (3)$$

where the positive or negative sign depends on the side of the tuna body profile, while w_i is the distance from vertebra to landmark, that is, the width of the ith-vertebra. As width is not a model parameter, w_i is calculated proportionally to model length l according to the following expression:

$$w_i = \mathbf{c_w}(i)l \quad ; \quad i = 1...n \quad (4)$$

where $\mathbf{c_w}$ is a constant coefficient vector.

Coefficient vectors $\mathbf{c_l}$, $\mathbf{c}_\theta$ and $\mathbf{c_w}$ are built empirically from a dataset of tuna silhouettes.

2.2. Model with caudal peduncle keel: M2 model

The capabilities of **M2** model have been increased compared to **M1** by introducing the following modifications: (i) more mid-body points are concentrated in the caudal peduncle, a crucial zone for length measurements; (ii) the area around the pectoral fin is not considered, as its many shapes can hinder model fitting because of the ambiguity between points on the silhouette of the body and these fins.

Deformable model **M2** is defined as a vector of five parameters $\mathbf{M2} = [s_x,s_y,l,\alpha,\theta]$, the same as **M1**, and is characterized by 18 mid-body points $v_i = (x_i^v, y_i^v)$, 14 of them distributed along the fish length using a grid of size $S_l = l/16$, and 4 of them to model the caudal peduncle (see Fig. 1). The position of the mid-body points and landmarks are calculated using Eqs. (1–4).

2.3. Model with caudal peduncle keel and width: M3 model

M3 model capabilities have been increased compared to **M2**, to include also a width parameter that allows a better fit to the fish silhouette. Consequently, **M3** is defined as a vector of six parameters $\mathbf{M3} = [s_x,s_y,l,\alpha,\theta,w]$, where the new parameter w is a coefficient that widens

Table 1
Parameters and main features of the proposed geometric tuna models (**M1**, **M2**, **M3** and **M4**).

Tuna model	M1	M2	M3	M4
Characteristic feature	pectoral fins	caudal peduncle	caudal peduncle and width	caudal peduncle back part and vector of widths
Parameters	s_x,s_y,l,α,θ	s_x,s_y,l,α,θ	$s_x,s_y,l,\alpha,\theta,w$	$s_x,s_y,l,\alpha,\theta,\mathbf{w},l_p,s_p$
	s_xs_y: Snout tip location	s_xs_y: Snout tip location	s_xs_y: Snout tip location	s_xs_y: Snout tip location
	l: Length	l: Length	l: Length	l: Length
	α: Angle to horizontal	α: Angle to horizontal	α: Angle to horizontal	α: Angle to horizontal
	θ: Bending angle	θ: Bending angle	θ: Bending angle	θ: Bending angle
			w: Width	$\mathbf{w}$: Vector of widths
				l_p: Caudal peduncle length
				s_p: Caudal peduncle slope
Silhouette landmarks	39	35	35	39
Mid-body points:	20	18	18	20
	4 to model the pectoral fins	4 to model the caudal peduncle	4 to model the caudal peduncle	4 to model the back part of the caudal peduncle
Shape of pectoral fins	Yes	None	None	None
Shape of caudal peduncle	None	Front and back part	Front and back part	Only the back part

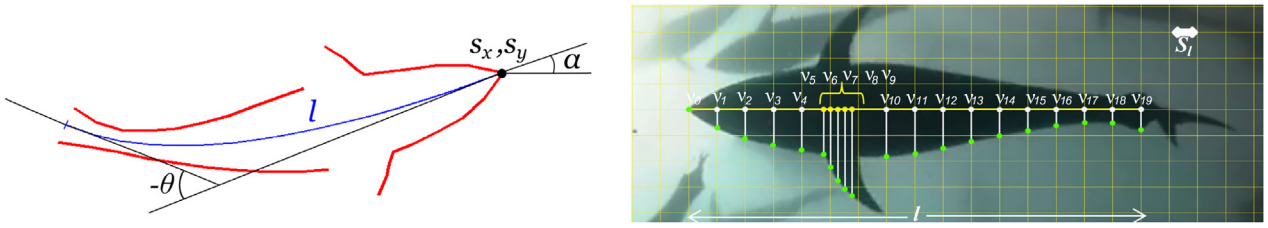

Fig. 2. Deformable tuna model **M1** presented in (Atienza-Vanacloig et al., 2016). On the left, s_x and s_y give the image location of the snout tip, l is the length of the vertebral column, α denotes the angle of the fish head in relation to the horizontal axis and θ is the global bending angle of the vertebral column. On the right, white and yellow dots represent the mid-body points, whereas green dots correspond to landmarks. Grid size $S_l = l/16$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

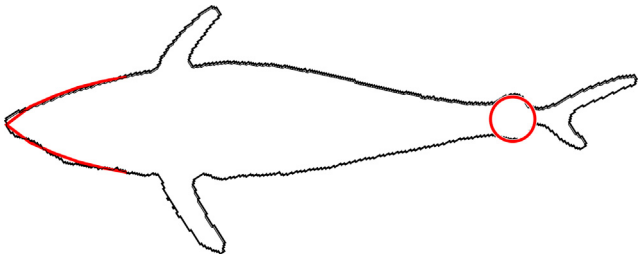

Fig. 3. Graphical representation of the feature extraction algorithms: caudal peduncle extraction using Hough transform and snout tip extraction using the geometric model of the fish head.

the model proportionally to a fattening factor. **M3** is characterized by the same 18 mid-body points $v_i = (x_i^v, y_i^v)$ as **M2** (see Fig. 1) and the position of mid-body points and landmarks is calculated using Eqs. (1–3). However, w_i is calculated in this case proportionally to the tuna width parameter, rather than tuna length, according to the following expression:

$$w_i = \mathbf{c_w}(i)w \quad ; \quad i = 1...n \tag{5}$$

2.4. Model with back peduncle part and width vector: M4 model

This **M4** model attempts to improve the fit to the fish silhouette by redesigning the caudal peduncle and the model width. In this case, the front part of the caudal peduncle is not modelled, and the back part is modelled as a segment (see Fig. 1). Moreover, changes in peduncle shape are considered by adding length and slope of the segment as parameters to the model. This modification improves the fit to the fish silhouette in the caudal peduncle, allowing more accurate fish length estimations. Regarding the model width, parameter w in **M3** is modified to become a vector of widths $\mathbf{w}$. While in the previous models, w_i is considered a function of length (in **M1** and **M2**) or width (in **M3**) with constant coefficients, **M4** assigns a variable-bounded width for each vertebral point, improving the fit to fish width.

Therefore, deformable model **M4** is defined as a vector of eight parameters $\mathbf{M4} = [s_x, s_y, l, \alpha, \theta, \mathbf{w}, l_p, s_p]$, where the new parameters are: $[l_p, s_p]$, length and slope of the segment representing the back part of the caudal peduncle, and $\mathbf{w}$, the widths vector. **M4** is characterized by 18 mid-body points $v_i = (x_i^v, y_i^v)$, 14 of them distributed along the fish length and 4 of them to model the back part of the caudal peduncle.

Position of the mid-body points and landmarks, excluding the ones that model the caudal peduncle, is calculated again using Eqs. (1–3), but in this case, the width coefficients $w_i = \mathbf{w}(i)$ are used. For the caudal peduncle, the mid-body points and landmarks are modified with $[l_p, s_p]$, length and slope of the segment, according to the equation of a straight line.

2.5. Feature extraction to initialize model parameters

Prior to applying the geometric model fitting procedure, a fast template matching procedure is applied to the fish candidates to deduce swimming direction and roughly initialize the model parameters according to the matching, as detailed in (Atienza-Vanacloig et al., 2016).

However, since a good initialization of the parameters is critical to avoid local minima and increase the number of successful fittings, further developments have been carried out. In particular, two feature extraction algorithms have been developed: caudal peduncle extractor and snout tip extractor. For the caudal peduncle, the algorithm searches for circles at the ends of the fish candidates using the Hough transform (Ballard, 1981). If successful, this detection is used to initialize geometric model length l. For the snout tip, the algorithm applies a fitting procedure for the geometric model of the fish head. If successful, snout tip parameters (s_x, s_y) are excluded from the complete model fitting procedure. Fig. 3 shows a successful case of features extraction, and its contribution to model initialization is analysed in Section 4.2.

3. Materials and methods

In this section, Fig. 4 describes and summarizes the computer vision algorithms involved in the process of fish sizing, as well as the offline manual and semiautomatic operations performed to generate ground truth data.

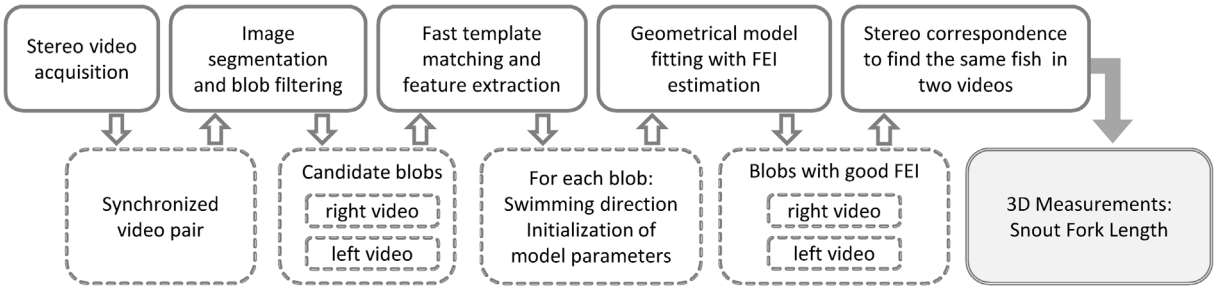

Fig. 4. In the first row, sequence of processes performed automatically in our proposal. In the second row, intermediate results of each step in dashed line and final result in continuous line. FEI refers to Fitting Error Index.

3.1. Stereo video acquisition

The recording was taken in the Grup Balfegó growing farms, on the west Mediterranean coast, using the AM100 stereovision system (www.aq1systems.com), consisting of two Gigabit Ethernet cameras, with image resolution of 1360 × 1024 pixels and frame rate of 12 fps, mounted in an underwater housing, with a baseline of 80 cm and an inward convergence of 6°.

The cameras were positioned 15 m deep in the grow-out cages and looking towards the surface to obtain a ventral silhouette of the fish. This camera arrangement has three advantages: first, with this orientation, sunlight acts like a backlight system so objects are always darker than water; second, in this set up, body bending can be clearly appreciated and dealt with; third, the most reliable measurements are obtained when fish are swimming in a plane orthogonal to the visual axis (Dunbrack, 2006). The acquired videos are processed automatically using the computer vision algorithms outlined in Fig. 4 and described below.

3.2. Image segmentation, blobs filtering and tuna model fitting

Image segmentation was implemented using local thresholding (Petrou and Petrou, 2011), a region-based technique for extracting compact regions (blobs) on each video frame, and morphological operations. The segmented blobs are geometrically characterized and sifted using shape (aspect ratio), pixel density and dimensional filters. Afterwards, the parameters of the geometric models are initialized as described in Section 2.5. An edge detection algorithm is then applied and a minimization algorithm is used to fit the deformable tuna models presented in Section 2 to the silhouette. A Fitting Error Index (FEI), based on the quadratic distance between model points and target edge points, is computed to analyse the accuracy of the fitting. FEI takes values in the [0. .10] range, where FEI = 0 denotes a perfect fit between the segmented blob and the geometric model. Fittings with high values (FEI>3) are discarded. An example of image segmentation and blobs filtering with many animals swimming at the same time is shown in Fig. 5. For further details on the segmentation, filtering and model fitting procedures see (Atienza-Vanacloig et al., 2016).

In the first row, sequence of processes performed automatically in our proposal. In the second row, intermediate results of each step in

3.3. Stereoscopic vision system calibration and stereo correspondence

The results for left and right videos, obtained separately in Section 3.2, are merged to calculate fish length. The image plane information is transformed to 3D measurements using the calibration parameters of the stereoscopic vision system and 3D triangulation.

Images for calibration were acquired in a tank containing seawater at IEO (Spanish Oceanographic Institute) facilities in Mazarrón (Spain). A 1.40 × 1.10 m checkerboard pattern was guided from −45° to 45° with respect to the optical axis and moved between 1 and 10 m away from the cameras. The MATLAB® Stereo Calibration Application based on (Heikkila and Silven, 1997) and (Zhang, 2000) was used to estimate the calibration parameters. The diagonal length of the checkerboard pattern was computed for 5018 samples in the stereo images to analyse our calibration accuracy in terms of proportional error between true and measured lengths, resulting in a mean error of 0.3%, similar to other results reported in the literature (Shortis, 2015).

With a proper calibration of the stereoscopic vision system, the following epipolar geometry restriction can be used: given two characteristic points of the fish model in one image, the matching points in the other image must lie on the epipolar line defined by the calibration parameters. In our case, samples are discarded if the stereo correspondence is not met for the first and last model mid-body points, that is, if the distance from the points to the epipolar lines is greater than a threshold of 10 pixels. The adoption of this value is based on extensive experience with underwater stereo-image measurements and current image resolution. Such distances may occur because the individuals identified in the left and right videos do not correspond to the same fish or because left, right or both model fittings are not accurate enough.

3.4. Fish length measurements using stereoscopic vision

Snout Fork Length (SFL) is the fish length generally used in the literature and the fish length measured at harvesting. However, as mentioned before, the caudal fins cannot be modelled due to their great variability, so SFL cannot be directly measured in the model. To overcome this issue, we defined Snout Peduncle Length (SPL), as the Euclidean distance between the 3D coordinates of the snout tip and the caudal peduncle vertebra, which corresponds to the last vertebra of **M1** and the prior-to-the-last vertebra in **M2**, **M3** and **M4**. Then, we deduced the relation between SFL and SPL from a dataset of experimental semiautomatic measurements. For this purpose, 1000 samples from the video frames were selected with the following requirements: the mid-body points form a straight line (no bending) and the tail fork is clearly identifiable and aligned with the vertebral column, as shown in Fig. 6a. For these samples, SPL and SFL were computed using a semiautomatic procedure: the snout tip, the peduncle point and the tail fork point are manually marked in the left image using the mouse pointer, whereas the corresponding right image points are also marked, but with the aid of the epipolar lines resulting from the stereovision calibration. SPL and SFL are computed using the stereovision calibration parameters as the 3D Euclidean distances between snout-peduncle and snout-fork, respectively. A polynomial fitting was then applied, resulting in an SPL-SFL linear relation, as shown in Eq. (6) and Fig. 6b. Note that this relation has been deduced from the stock under study, that is, adult Bluefin Tuna with SFL between 1.60 and 2.60 m. Hence, it is species specific and based on a finite range of lengths.

$$SFL = 1.0879\ SPL + 0.050964 \quad (6)$$

Fish are deformable due to the swimming motion and, consequently, measurements taken from a single frame may not be reliable (Shortis et al., 2016). Two main options are used in the literature to reduce the effect of swimming motion on length measurement: (i) take measurements in all frames and deduce straight body length from a sinusoid-like pattern (Shortis et al., 2016); (ii) account for body bending by adding contiguous linear segments (Williams and Lauffenburger, 2016). In our case, the swimming length problem is

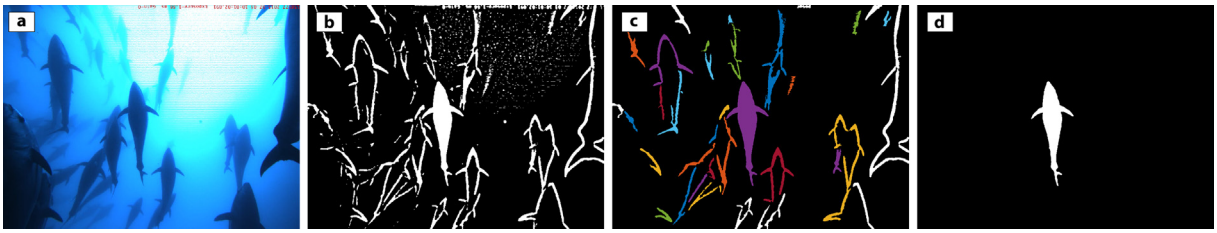

Fig. 5. Segmentation and blobs filtering. (a) original image, (b) segmented image using local thresholding and morphological operations, (c) blobs labelling and characterization and area filter, (d) aspect ratio and pixel density filters.

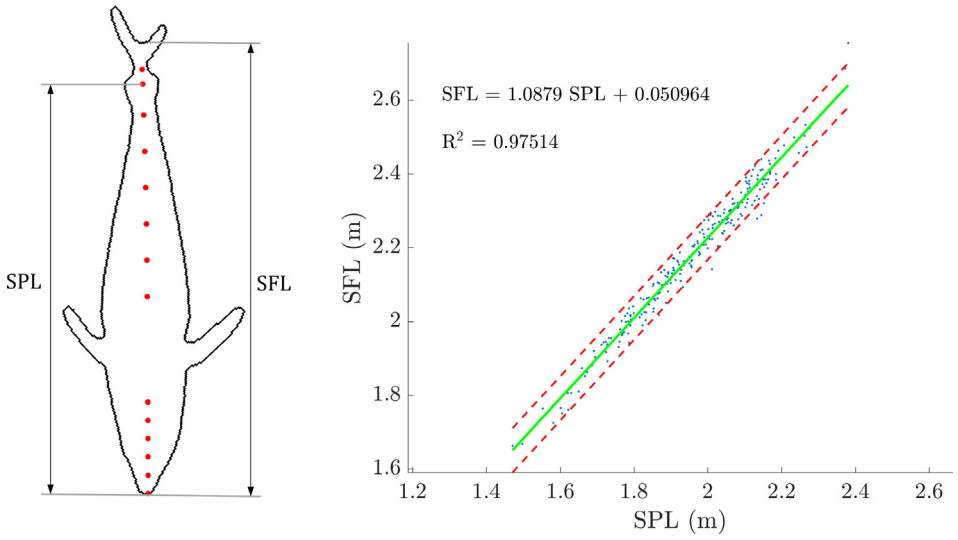

Fig. 6. (a) Sample selected to study the relation between Snout Peduncle Length (SPL) and Snout Fork Length (SFL). (b) Polynomial fitting for the SPL-SFL relation. The green solid line is the linear fitting and the red dashed lines are the 95% confidence interval. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

resolved using the tuna model bending angle θ, by identifying as valid samples the ones whose mid-body points form a straight line and discarding the others.

4. Results

For a quantitative evaluation of the **M1**, **M2**, **M3** and **M4** models, the results are presented in this section from three points of view: (i) Quantity of successful fittings achieved by each model, which allows us to assess the degree of adaptation of each model to the blobs and, therefore, its capacity to adapt to the tuna body while swimming. (ii) Computing time, analysed to compare tuna models, but taking into account that both the code and the algorithms could be optimized to speed up automatic measurements. (iii) Accuracy of the SFL measurements. In addition, four different methods of initializing the model parameters are studied, using the feature extraction algorithms detailed in Section 2.5: (a) fast template matching, (b) fast template matching and snout tip feature extractor, (c) fast template matching and caudal peduncle feature extractor, (d) fast template matching and both snout tip and peduncle feature extractors.

4.1. Datasets and ground truths

Two datasets and ground truths were configured from a stereo video of 135 min duration, acquired with the stereoscopic vision system described in Section 3.1. The video was recorded in real conditions on an adult Atlantic Bluefin Tuna growing cage in the Mediterranean Sea containing 316 fish.

The first dataset (DS1) is composed of 1000 samples, selected from the video frames but avoiding difficult cases caused by overlapping or bad lighting. These samples are measured whit the aforementioned semiautomatic procedure to configure the ground truth (GT1) and compared to automatic measurements using the different tuna models. To configure the ground truth (GT1), samples have been measured semiautomatically by three different operators to obtain a mean value. The procedure is repeated for the samples with discrepancies between operators greater than 0.5% in fish length. Relative error between measurements is defined for the different tuna models, and calculated for each sample as stated in Eq. (7), where SFL_a and SFL_m are the automatic and semiautomatic SFL, respectively.

$$e_r(\%) = \frac{SFL_a - SFL_s}{SFL_s} \cdot 100 \tag{7}$$

A second dataset (DS2) comprises all the 97,341 pairs of frames from the stereo video. The whole video is processed and SFL measurements are extracted using the different tuna models. Real SFL measurements of the tuna stock collected by Grup Balfegó at harvesting configure the second ground truth (GT2). It is important to highlight that this stock corresponds to the same cage where the video was acquired because this increases the veracity of the results. SFL frequency histograms are built to compare the automatic SFL measurements from DS2 with real SFL measurements of GT2.

The DS1 and DS2 results are detailed in Sections 4.2 and 4.3, respectively.

4.2. Automatic versus semiautomatic measurements

Each sample of DS1 is automatically measured with the different models (**M1**, **M2**, **M3** and **M4**) and with four different methods of initializing the model parameters (a-d). Fig. 7 shows the results from different perspectives.

The first subplot in Fig. 7 represents the relative errors (Eq. (7)) between automatic and semiautomatic measurements of GT1 in a box plot. For each box, the central rectangle represents the interquartile range or IQR, which comprises 50% of the samples, from 25th to 75th percentiles. A segment inside the rectangle shows the median error and whiskers above and below the box comprise 90% of the samples, from 5th to 95th percentile. Error margins of 3%, 5% and 7% are displayed with parallel horizontal dotted lines in green to provide a visual guide. In comparison with fast template matching initialization (a), relative errors decrease when either the snout tip feature extractor (b) or the caudal peduncle feature extractor (c) is used, and greatly decrease when both feature extractors are used (d). To such an extent, that 90% of the samples for **M1**.d are bounded in a 7% error margin, for **M2**.d they are bounded in a 5% error margin, and for **M3**.d and **M4**.d they are bounded in a 3% error margin. Note that relative errors decrease for **M2**, **M3** and **M4** with respect to **M1**, independently of model initialization.

The second subplot in Fig. 7 shows the number of samples successfully measured automatically. Note that a sample is considered successfully measured if these two conditions are satisfied: good FEI and stereo correspondence, as defined in Sections 3.2 and 3.3, respectively. When the snout tip feature extractor is used (b), the number of samples increases only with **M4**, whereas it increases for all the models when the caudal peduncle feature extractor is used (c). More automatic measurements are obtained when both feature extractors are used (d), with 749, 842, 853 and 950 automatic measurements out of 1000 samples for **M1**, **M2**, **M3** and **M4**, respectively.

The third subplot in Fig. 7 represents the mean and standard deviation of computing time per fish measurement. Note that computing

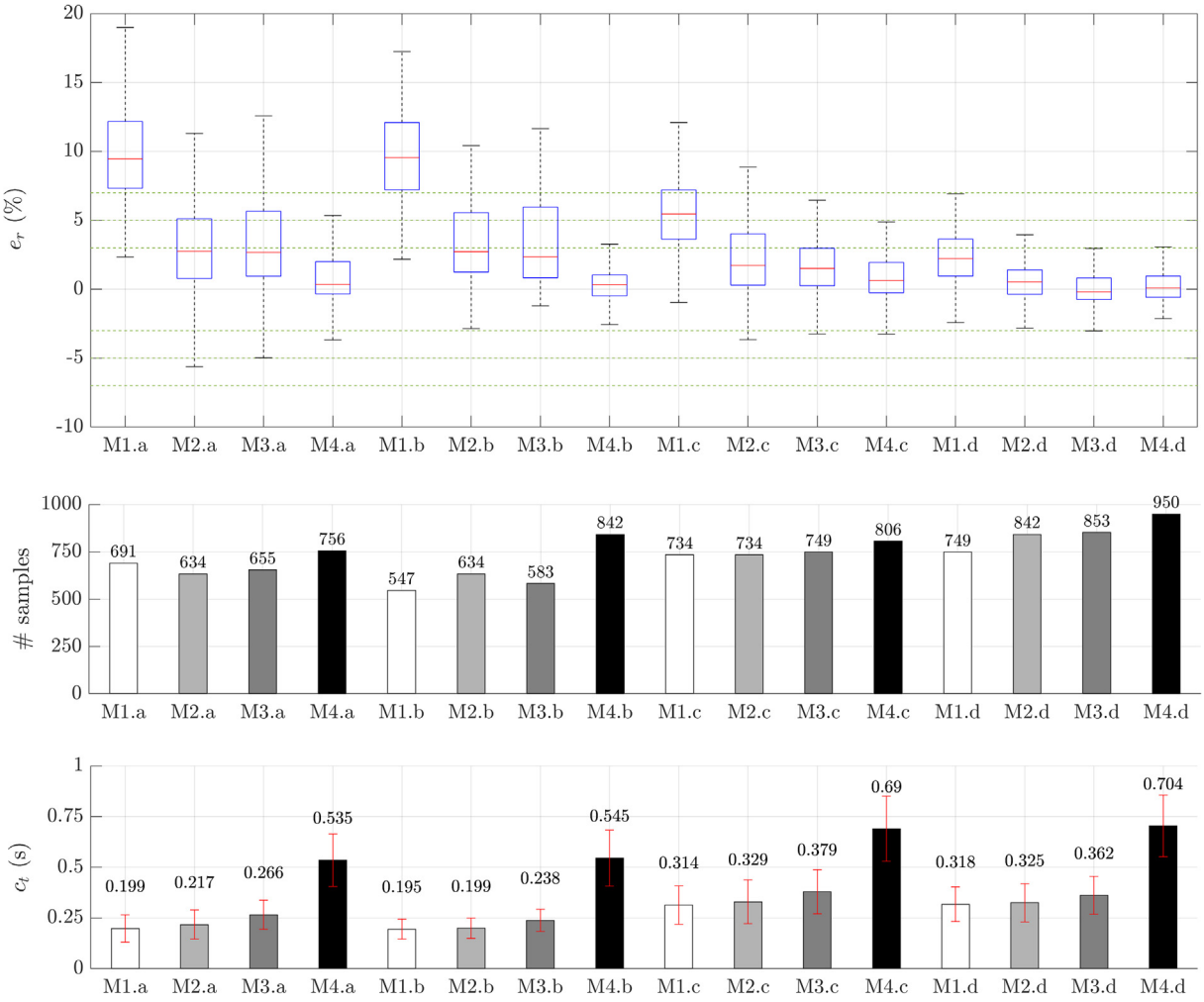

Fig. 7. Automatic versus semiautomatic measurements for the different tuna models (**M1**, **M2**, **M3** and **M4**) and feature extractors (a, b, c and d). At the top, box plot of individual relative errors (the central rectangle represents the interquartile range or IQR, which comprises 50% of the samples, the segment inside the rectangle shows median error and whiskers above and below the box comprise 90% of the samples). Error margins of 3%, 5% and 7% are displayed with parallel horizontal dotted lines in green to provide a visual guide. In the middle, number of samples automatically measured. At the bottom, mean (bar and value) and standard deviation (whiskers) of computing time per fish measurement. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

time for **M4** is almost twice the time spent by **M1**, **M2** and **M3** and computing time increases by approximately 0.1 s when the caudal peduncle feature extractor is used (c and d), to give computing times of 0.318, 0.325, 0.362 and 0.704 s for **M1**, **M2**, **M3** and **M4**, respectively.

4.3. Automatic measurements versus real data from harvests

DS2 is processed to compute automatic measurements with the different tuna models, which are compared to real SFL measurements of GT2, collected at harvesting by Grup Balfegó. Model parameters are initialized using fast template matching and both the snout tip feature extractor and the caudal peduncle feature extractor (d). This initialization method accomplishes the most accurate measurements for all models, as demonstrated in Section 4.2.

Fig. 8 shows the normalized SFL frequency histograms. Differences in $\overline{\text{SFL}}$ between harvests and automatic measurements were examined with analysis of variance tests. Since the two groups have unequal sample sizes and homoscedasticity (homogeneity of variance) cannot be ensured, Welch's ANOVA test (Welch, 1951) is used, as recommended in (Rasch et al., 2011) and (McDonald, 2014). Differences in SFL frequency distributions are analysed with the Kolmogorov-Smirnov test (Massey, 1951).

As Table 2 indicates, the tests for $\overline{\text{SFL}}$ and SFL distribution frequency give p-values higher than the 5% significance level for models **M2**, **M3** and **M4**. With these models there is no statistically significant difference between GT2 and DS2 automatic measurements. However, note that the results obtained with **M1**, that is, 3 cm (1.4%) error in mean length, can be admissible in some applications. For example, for transfer operations, (ICCAT, 2015) establishes that fish length estimations using stereoscopic camera systems must be lower than ± 5%. Moreover, Table 2 shows the number of successful model fittings for left and right videos, the number of automatic measurements and computing time, which covers the entire automatic measurement process (as outlined in Fig. 4). It can be seen that using model **M4**, automatic measurements increase by 2536 samples (60.8%, 19 samples per minute of recording) with respect to **M2** and by 1895 (39.4%, 14 samples per minute of recording) with respect to **M3**, at the expense of increasing computing time by more than 6 h (57%) and 5.5 h (47%), respectively.

5. Discussion

The results show the enhanced capabilities of the new models, from fish discrimination (**M1**) to accurate length estimation (**M3** and **M4**), by

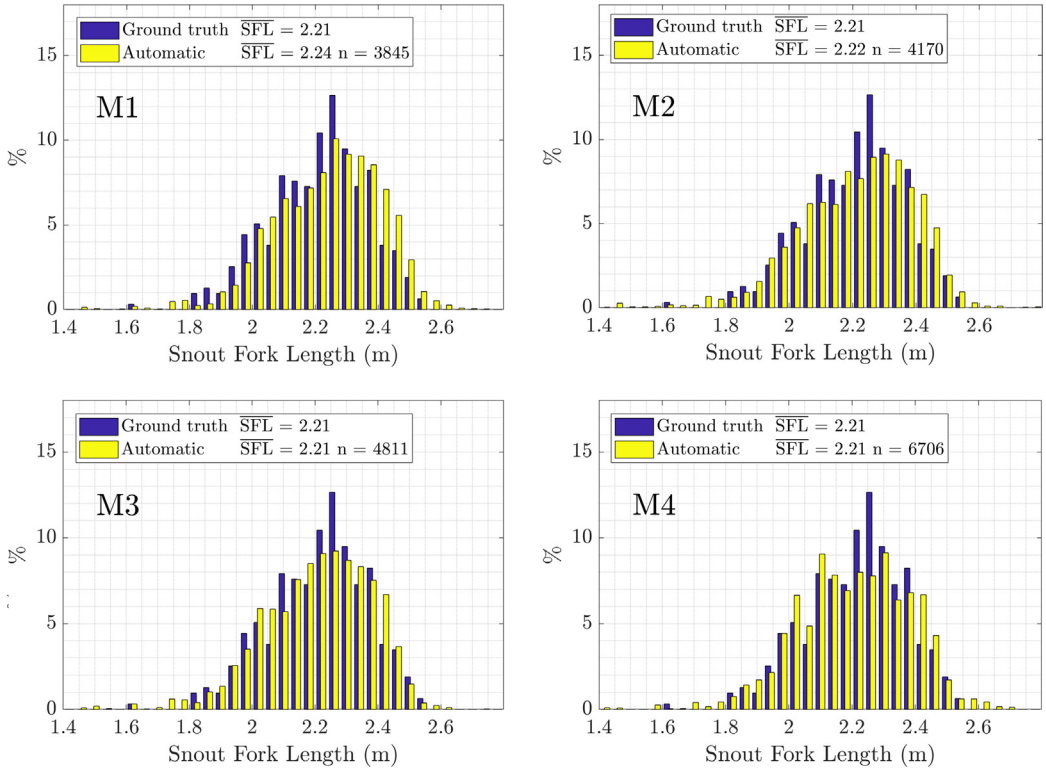

Fig. 8. Automatic measurements versus real data from harvests. Normalized SFL frequency histograms. Automatic measurements in light-yellow and ground truth in dark-blue. $\overline{SFL}$, mean SFL; n, number of automatic measurements. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Automatic system measurements and ground truth statistical comparison for the different tuna models.

	M1	**M2**	**M3**	**M4**
# successful fitting in left video	41,014	48,017	48,312	57,167
# successful fitting in right video	43,962	51,059	51,236	60,750
# automatic measurements	3845	4170	4811	6706
# automatic measurements per minute of recording	28	31	36	50
Computing time	11 h 04′ 12″	11 h 10′ 54″	11 h 56′ 57″	17 h 37′ 03″
Welch's ANOVA test p-value	0.0012	0.5127	0.5192	0.5278
Kolmogorov-Smirnov test p-value	0.0017	0.1081	0.1503	0.1544

combining improved geometric model definitions and feature extraction algorithms to initialize the model parameters.

Firstly, the algorithms developed for snout tip and caudal peduncle features extraction allow an initialization of model parameters, which leads to enhanced results in terms of error in fish length estimation and number of automatic measurements, at the expense of increasing computing time. Secondly, the improved geometric models consider the following main modifications with respect to the original model (**M1**): the fish silhouette does not include pectoral fins and the caudal peduncle is modelled with a higher density of mid-body points. The models that consider these two modifications (**M2**, **M3** and **M4**) obtain smaller errors in fish length estimation than the original model (**M1**). The capabilities of **M3** model have been increased compared to **M2**, to include also a width parameter that allows better fit to the fish silhouette (smaller FEI) when both snout tip and caudal peduncle features extractors are used. **M4** model improves the fit to the fish silhouette by redesigning the caudal peduncle and model width.

For all models, the most accurate length estimations are obtained when both snout tip and caudal peduncle features extractors are used (**M1**.d, **M2**.d, **M3**.d and **M4**.d). Focusing on the differences between tuna models in this case, the error in fish length estimation and the number of automatic measurements improve from **M1** to **M4**, at the expense of higher computing times. **M1**.d can be discarded, because worse length estimation is obtained, with no advantage in either number of samples or computing time, and **M2**.d can also be discarded, because **M3**.d has very similar but slightly better results. In conclusion, **M3**.d and **M4**.d have the highest number of samples and most accurate fish length estimations, and choosing between them would depend on the application. To extract a higher number of samples with very low errors **M4**.d would be preferred, but for faster automatic measurements, with slightly higher errors and lower number of samples, **M3**.d would be used. In both cases, very accurate results are obtained, as proved in the comparison with ground truth data from harvests.

Fish length information is an important indicator of the health of wild fish stocks and for predicting biomass using length-weight relations (Lines et al., 2001), (Martinez-de Dios et al., 2003). The most common mathematical model between fish length (L) and mass (W) is $W = aL^b$, where a and b are empirically characterized species and strain-dependent parameters (Zion, 2012). The total biomass of a fish stock is commonly determined by obtaining the mean length of a statistically representative number of fish and counting the number of fish (Costa et al., 2009), (Shafait et al., 2017). Recent studies attempt to show that biomass can be estimated more accurately if fish measurements in dimensions other than length, like width and depth, are available (Aguado-Gimenez and Garcia-Garcia, 2005), (Harvey et al., 2003). Nevertheless, as stated in (Harvey et al., 2003), measuring the width of a fish is relatively subjective due to the lack of defined points in the fish silhouette. We are currently working on finding an adequate way of estimating fish width, using the width parameter and mid-body points position in the **M3** and **M4** tuna models, and generating a ground truth dataset that includes width measurements at harvesting.

6. Conclusions and further work

The proposed procedure might be a significant contribution towards a commercial system for fully automatic fish sizing using stereoscopic vision. The proposed geometric models are able to provide high number of samples and accurate measurements of the fish length in adult tunas. The completely automatic process is the main difference of this work with respect to other studies with similar goals, such as (Lines et al., 2001), (Harvey et al., 2003), (Letessier et al., 2015), (Williams and Lauffenburger, 2016) and (Shafait et al., 2017). The results demonstrate highly accurate SFL estimations and validate the automatic procedure. The system could be used to track the growth of fish trough time by scheduling a recording plan and could be integrated in an autonomous monitoring system, whose computing performance and energy requirements should be dimensioned to allow recording and analysis of statistically representative amount of measurements. The automatic system has been used for fish sizing on adult Atlantic Bluefin Tuna, but the procedure could be applied to other species, adapting segmentation parameters, blobs filtering criteria and geometric model.

As further work, we are working on computing fish width from the geometric models and studying different possibilities for fish tracking, such as Kalman and particle filters (Morais et al., 2005). We plan to improve some aspects of the models to provide accurate 3D measurements of bent fish and to fit the tuna silhouette from other views in addition to the ventral one. Moreover, we want to combine this computer vision procedure with acoustic information to estimate biomass in more complex situations, such as wild environments and transfers from tow to grow-out cages.

Acknowledgements

This work was supported by funding from ACUSTUNA project ref. CTM2015-70446-R (MINECO/ERDF, EU). This project has been possible thanks to the collaboration of IEO (Spanish Oceanographic Institute). We acknowledge the assistance provided by the Spanish company Grup Balfegó S.L. in supplying boats and divers to acquire underwater video in the Mediterranean Sea.

References

Aguado-Gimenez, F., Garcia-Garcia, B., 2005. Growth, food intake and feed conversion rates in captive Atlantic bluefin tuna (*Thunnus thynnus* Linnaeus, 1758) under fattening conditions. Aquaculture Res. 36, 610–614. http://dx.doi.org/10.1111/j.1365-2109.2005.01210.x.

Atienza-Vanacloig, V., Andreu-García, G., López-García, F., Valiente-González, J.M., Puig-Pons, V., 2016. Vision-based discrimination of tuna individuals in grow-out cages through a fish bending model. Comput. Electron. Agric. 130, 142–150. http://dx.doi.org/10.1016/j.compag.2016.10.009.

Ballard, D.H., 1981. Generalizing the Hough transform to detect arbitrary shapes. Pattern Recog. 13, 111–122. http://dx.doi.org/10.1016/0031-3203(81)90009-1.

Boutros, N., Shortis, M.R., Harvey, E.S., 2015. A comparison of calibration methods and system configurations of underwater stereo-video systems for applications in marine ecology. Limnol. Oceanography: Methods 13, 224–236. http://dx.doi.org/10.1002/lom3.10020.

Costa, C., Loy, A., Cataudella, S., Davis, D., Scardi, M., 2006. Extracting fish size using dual underwater cameras. Aquacultural Eng. 35, 218–227. http://dx.doi.org/10.1016/j.aquaeng.2006.02.003.

Costa, C., Scardi, M., Vitalini, V., Cataudella, S., 2009. A dual camera system for counting and sizing Northern Bluefin Tuna (*Thunnus thynnus*; Linnaeus, 1758) stock, during transfer to aquaculture cages, with a semi automatic Artificial Neural Network tool. Aquaculture 291, 161–167. http://dx.doi.org/10.1016/j.aquaculture.2009.02.013.

Dunbrack, R.L., 2006. In situ measurement of fish body length using perspective-based remote stereo-video. Fish. Res. 82, 327–331. http://dx.doi.org/10.1016/j.fishres.2006.08.017.

Espinosa, V., Soliveres, E., Cebrecos, A., Puig, V., Sainz-Pardo, S., de la Gándara, F., 2011. Growing monitoring in sea cages : Ts measurements issues. In: Proceedings of the 34th Scandinavian Symposium on Physical Acoustics, Geilo, Norway, 30 January – 2 February, 2011.

Hao, M., Yu, H., Li, D., 2016. Computer and Computing Technologies in Agriculture IX. IFIP International Federation for Information Processing 2016 Published by Springer International Publishing AG 2016 478, 15–32. http://dx.doi.org/10.1007/978-3-319-48357-3.

Hao, M., Yu, H., Li, D., 2015. The Measurement of fish size by machine vision - a review. In: Li, D., Li, Z. (Eds.), Computer and Computing Technologies in Agriculture IX - 9th IFIP WG 5.14 International Conference, CCTA 2015, Beijing, China, September 27–30, 2015, Revised Selected Papers, Part II, IFIP Advances in Information and Communication Technology. pp. 15–32. http://dx.doi.org/10.1007/978-3-319-48354-2_2.

Harvey, E., Butler, J.J., McLean, D.L., Shand, J., 2012. Contrasting habitat use of diurnal and nocturnal fish assemblages in temperate Western Australia. J. Exp. Marine Biol. Ecol. 426, 78–86. http://dx.doi.org/10.1016/j.jembe.2012.05.019.

Harvey, E., Cappo, M., Shortis, M., Robson, S., Buchanan, J., Speare, P., 2003. The accuracy and precision of underwater measurements of length and maximum body depth of southern bluefin tuna (*Thunnus maccoyii*) with a stereo-video camera system. Fish. Res. 63, 315–326. http://dx.doi.org/10.1016/S0165-7836(03)00080-8.

Heikkila, J., Silven, O., 1997. A four-step camera calibration procedure with implicit image correction. In: Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97), CVPR '97. IEEE Computer Society, Washington, DC, USA, p. 1106.

Hong, H., Yang, X., You, Z., Cheng, F., 2014. Aquacultural engineering visual quality detection of aquatic products using machine vision. Aquacultural Eng. 63, 62–71. http://dx.doi.org/10.1016/j.aquaeng.2014.10.003.

Hu, J., Li, D., Duan, Q., Han, Y., Chen, G., Si, X., 2012. Fish species classification by color, texture and multi-class support vector machine using computer vision. Comput. Electron. Agric. 88, 133–140. http://dx.doi.org/10.1016/j.compag.2012.07.008.

Huang, P.X., Boom, B.J., Fisher, R.B., 2013. Underwater live fish recognition using a balance-guaranteed optimized tree. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 7724 LNCS, 422–433. http://dx.doi.org/10.1007/978-3-642-37331-2_32.

ICCAT, 2015. Recommendation by ICCAT amending the recommendation 13–07 by ICCAT to establish a multi-annual recovery plan for Bluefin Tuna in the eastern Atlantic and Mediterranean. Rec [14-04]. In: 2015 Compendium Management Recommendations and Resolutions Adopted by ICCAT for Conservation of Atlantic Tunas and Tuna-like Species. pp. 47–82.

Kloser, R.J., Ryan, T.E., Macaulay, G.J., Lewis, M.E., 2011. In situ measurements of target strength with optical and model verification: a case study for blue grenadier, *Macruronus novaezelandiae*. ICES J. Marine Sci. 68, 1986–1995. http://dx.doi.org/10.1093/icesjms/fsr127.

Langlois, T.J., Harvey, E.S., Meeuwig, J.J., 2012. Strong direct and inconsistent indirect effects of fishing found using stereo-video: testing indicators from fisheries closures. Ecol. Indicators 23, 524–534. http://dx.doi.org/10.1016/j.ecolind.2012.04.030.

Letessier, T.B., Juhel, J.-B., Vigliola, L., Meeuwig, J.J., 2015. Low-cost small action cameras in stereo generates accurate underwater measurements of fish. J. Exp. Marine Biol. Ecol. 466, 120–126. http://dx.doi.org/10.1016/j.jembe.2015.02.013.

Lines, J.A., Tillett, R.D., Ross, L.G., Chan, D., Hockaday, S., McFarlane, N.J.B., 2001. An automatic image-based system for estimating the mass of free-swimming fish. Comput. Electron. Agric. 31, 151–168. http://dx.doi.org/10.1016/S0168-1699(00)00181-2.

Mallet, D., Pelletier, D., 2014. Underwater video techniques for observing coastal marine biodiversity: a review of sixty years of publications (1952–2012). Fish. Res. 154, 44–62. http://dx.doi.org/10.1016/j.fishres.2014.01.019.

Martinez-de Dios, J.R., Serna, C., Ollero, A., 2003. Computer vision and robotics techniques in fish farms. Robotica 21, 233–243. http://dx.doi.org/10.1017/S0263574702004733.

Massey, F.J., 1951. The Kolmogorov-Smirnov Test for goodness of fit. J. Am. Statis. Assoc. 46, 68–78.

McDonald, J.H., 2014. Handbook of Biological Statistics, third ed. Sparky House Publishing, Baltimore, Maryland.

McLaren, B.W., Langlois, T.J., Harvey, E.S., Shortland-Jones, H., Stevens, R., 2015. A small no-take marine sanctuary provides consistent protection for small-bodied by-catch species, but not for large-bodied, high-risk species. J. Exp. Marine Biol. Ecol. 471, 153–163. http://dx.doi.org/10.1016/j.jembe.2015.06.002.

Morais, E.F., Campos, M.F.M., Padua, F.L.C., Carceroni, R.L., 2005. Particle filter-based predictive tracking for robust fish counting. In: XVIII Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI'05). IEEE, pp. 367–374. http://dx.doi.org/10.1109/SIBGRAPI.2005.36.

Petrou, M., Petrou, C., 2011. Image segmentation and edge detection. In: Image Processing: The Fundamentals. John Wiley & Sons Ltd, Chichester, UK, pp. 527–668. http://dx.doi.org/10.1002/9781119994398.ch6.

Phillips, K., Rodriguez, V.B., Harvey, E., Ellis, D., Seager, J., Begg, G., Hender, J., 2009. Assessing the operational feasibility of stereo-video and evaluating monitoring options for the Southern Bluefin Tuna Fishery ranch sector. Fish. Res. Devel. Corp. Bureau Rural Sci. (Australia).

Puig, V., Espinosa, V., Soliveres, E., Ortega, A., Belmonte, A., de la Gándara, F., 2012. Biomass estimation of bluefin tuna in sea cages by the combined use of acoustic and optical techniques. Collect. Vol. Sci. Pap. ICCAT 68, 284–290.

Rahim, M., Shafry, M., Rehman, A., Kumoi, R., Abdullah, N., 2012. A new approach in measuring fish length using fish length from digital images (FiLeDI) framework. Int. J. Phys. Sci. 7, 607–618. http://dx.doi.org/10.5897/IJPS11.1581.

Rasch, D., Kubinger, K.D., Moder, K., 2011. The two-sample t test: pre-testing its assumptions does not pay off. Statis. Papers 52, 219–231. http://dx.doi.org/10.1007/s00362-009-0224-x.

Rosen, S., Jörgensen, T., Hammersland-White, D., Holst, J.C., Grant, J., 2013. DeepVision: a stereo camera system provides highly accurate counts and lengths of fish passing inside a trawl. Canad. J. Fish. Aquatic Sci. 70, 1456–1467. http://dx.doi.org/10.1139/cjfas-2013-0124.

Ruff, B.P., Marchant, J.A., Frost, A.R., 1995. Fish sizing and monitoring using a stereo image analysis system applied to fish farming. Aquacultural Eng. 14, 155–173. http://dx.doi.org/10.1016/0144-8609(94)P4433-C.

Saberioon, M., Gholizadeh, A., Cisar, P., Pautsina, A., Urban, J., 2017. Application of machine vision systems in aquaculture with emphasis on fish: state-of-the-art and key issues. Rev. Aquaculture 9, 369–397. http://dx.doi.org/10.1111/raq.12143.

Santana-Garcon, J., Newman, S.J., Harvey, E.S., 2014. Development and validation of a mid-water baited stereo-video technique for investigating pelagic fish assemblages. J. Exp. Marine Biol. Ecol. 452, 82–90. http://dx.doi.org/10.1016/j.jembe.2013.12.009.

Sawada, K., Takahashi, H., Abe, K., Ichii, T., Watanabe, K., Takao, Y., 2009. Target-strength, length, and tilt-angle measurements of Pacific saury (*Cololabis saira*) and Japanese anchovy (*Engraulis japonicus*) using an acoustic-optical system. ICES J. Marine Sci. 66, 1212–1218. http://dx.doi.org/10.1093/icesjms/fsp079.

Seiler, J., Williams, A., Barrett, N., 2012. Assessing size, abundance and habitat preferences of the Ocean Perch Helicolenus percoides using a AUV-borne stereo camera system. Fish. Res. 129, 64–72. http://dx.doi.org/10.1016/j.fishres.2012.06.011.

Shafait, F., Harvey, E.S., Shortis, M.R., Mian, A., Ravanbakhsh, M., Seager, J.W., Culverhouse, P.F., Cline, D.E., Edgington, D.R., 2017. Towards automating underwater measurement of fish length: a comparison of semi-automatic and manual stereo– video measurements. ICES J. Marine Sci. 10–1093. http://dx.doi.org/10.1093/icesjms/fsx007.

Shieh, A.C.R., Petrell, R.J., 1998. Measurement of fish size in atlantic salmon (*Salmo salar* l.) cages using stereographic video techniques. Aquacultural Eng. 17, 29–43. http://dx.doi.org/10.1016/S0144-8609(97)00012-5.

Shortis, M., 2015. Calibration techniques for accurate measurements by underwater camera systems. Sensors 15, 30810–30827. http://dx.doi.org/10.3390/s151229831.

Shortis, M.R., Ravanbakhsh, M., Shafait, F., Mian, A., 2016. Progress in the automated identification, measurement, and counting of fish in underwater image sequences. Marine Technol. Soc. J. 50, 4–16. http://dx.doi.org/10.4031/MTSJ.50.1.1.

Spampinato, C., Giordano, D., Di Salvo, R., Chen-Burger, Y.-H.J., Fisher, R.B., Nadarajan, G., 2010. Automatic fish classification for underwater species behavior understanding. In: Proceedings of the First ACM International Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams, ARTEMIS '10. ACM, New York, NY, USA, pp. 45–50. http://dx.doi.org/10.1145/1877868.1877881.

Tillett, R., Mcfarlane, N., Lines, J., 2000. Estimating dimensions of free-swimming fish using 3D point distribution models. Comput. Vision Image Understanding 79, 123–141. http://dx.doi.org/10.1006/cviu.2000.0847.

Torisawa, S., Kadota, M., Komeyama, K., Suzuki, K., Takagi, T., 2011. A digital stereo-video camera system for three-dimensional monitoring of free-swimming Pacific bluefin tuna, *Thunnus orientalis*, cultured in a net cage. Aquatic Living Res. 24, 107–112. http://dx.doi.org/10.1051/alr/2011133.

Wakefield, C.B., Lewis, P.D., Coutts, T.B., Fairclough, D.V., Langlois, T.J., 2013. Fish assemblages associated with natural and anthropogenically-modified habitats in a marine embayment: comparison of baited videos and opera-house traps. PLoS One 8, e59959. http://dx.doi.org/10.1371/journal.pone.0059959.

Watson, D., Anderson, M., Kendrick, G., Nardi, K., Harvey, E., 2009. Effects of protection from fishing on the lengths of targeted and non-targeted fish species at the Houtman Abrolhos Islands, Western Australia. Marine Ecol. Prog. Series 384, 241–249. http://dx.doi.org/10.3354/meps08009.

Welch, B.L., 1951. On the comparison of several mean values: an alternative approach. Biometrika 38, 330–336. http://dx.doi.org/10.1093/biomet/38.3-4.330.

White, D.J., Svellingen, C., Strachan, N.J.C., 2006. Automated measurement of species and length of fish by computer vision. Fish. Res. 80, 203–210. http://dx.doi.org/10.1016/j.fishres.2006.04.009.

Williams, K., Lauffenburger, N., 2016. Automated measurements of fish within a trawl using stereo images from a Camera-Trawl device (CamTrawl). Methods Oceanography 17, 138–152. http://dx.doi.org/10.1016/j.mio.2016.09.008.

Willis, T.J., Babcock, R.C., 2000. A baited underwater video system for the determination of relative density of carnivorous reef fish. Marine Freshwater Res. 51, 755. http://dx.doi.org/10.1071/MF00010.

Zhang, Z., 2000. A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Machine Intelligence 22, 1330–1334.

Zintzen, V., Anderson, M.J., Roberts, C.D., Harvey, E.S., Stewart, A.L., Struthers, C.D., 2012. Diversity and composition of demersal fishes along a depth gradient assessed by baited remote underwater stereo-video. PLoS One 7, e48522. http://dx.doi.org/10.1371/journal.pone.0048522.

Zion, B., 2012. The use of computer vision technologies in aquaculture - a review. Comput. Electron. Agric. 88, 125–132. http://dx.doi.org/10.1016/j.compag.2012.07.010.

Zion, B., Alchanatis, V., Ostrovsky, V., Barki, A., Karplus, I., 2007. Real-time underwater sorting of edible fish species. Comput. Electron. Agric. 56, 34–45. http://dx.doi.org/10.1016/j.compag.2006.12.007.