A COMPACT UNDERWATER STEREO VISION SYSTEM FOR MEASURING FISH

César Silva [a], Ricardo Aires[a], Flávio Rodrigues[a]

[a] Fishmetrics, Lda, Horta, Portugal, cesar.silva@fishmetrics.pt

Abstract

In the context of aquaculture, fish length is a key parameter to assess fish stocks, essential for feeding regime and contributes for decisions at several production levels. However, taking length measurements is a cumbersome task that, when applied to living individuals, can induce considerable stress, increasing the risk of damage or hindering their growth. Computer vision is one of the most used non-contact tools to overcome this issue, being fast, consistent and repeatable. However, its use in aquatic environments is limited by the high cost, the difficulty of calibrating the system in underwater conditions and the complexity of implementation.

This paper proposes a low-cost easy-to-use vision system that can take measurements on live fish in aquatic conditions, without the need for a special calibration and a demanding image analysis service. The present work implemented a compact stereo vision system and developed a method that estimates the correct length of fish, based on the variation of the angle of incidence of the light rays in the water. Given some structural conditions such as a short baseline, the system is able to measure fish with an error of less than 1%. The short baseline allows to have a compact system and reduces the effect of water refraction on the 3D reconstruction. A set of experiments were performed with real fish, working robustly for a set of orientations of the fish (even when the caudal fin and snout are on different distances to the cameras).

Keywords

Underwater Stereo Vision; Fish Measuring; Fish Length; Aquaculture Technology.

1. Introduction

Fish length is a key parameter to assess fish stocks (Jardim et al, 2015), helps stock management and brings economic benefits in the context of fisheries and aquaculture. The collection of information on the size of fish is essential in aquaculture, in many stages of production (Li et al, 2020), such as (a) determining the ideal moment of capture, (b) estimating a series of parameters at the population level and studying its evolution (volume, weight, gender or fat content); (c) stock assessment during creation; (d) classification and separation of specimens by size; (e) optimization in food distribution; (f) controlling specimens' growth; (g) forecasting market value. The length estimation and the total biomass value are usually taken manually based on a reduced sample of fish or, more indirectly, based on feed consumption. Manual measurements are climate-dependent time-consuming laborious tasks, being generally invasive and stressful for fish (Antonucci et al, 2020). Moreover, measurement-based size estimates for few specimens may not be representative of the entire population (Silva et al, 2018). Inaccurate estimates of length may result in overfeeding and, consequently, higher production costs and negative environmental impact (Li et al, 2020). In recent years, fish monitoring technology, with probes and optical systems, has gradually gained importance in the aquaculture and fishing sector (Antonucci et al, 2020, Li et al, 2020). Computer vision is one of the technologies used as a noninvasive technique to estimate animal biometrics, using images to measure different body features (Hao et al, 2016).

In general, to obtain information of an object from the image is used classical image processing algorithms (e.g., algorithms for feature detection, object segmentation, clustering and/or stochastic classification) (Hartley et al, 2003) or algorithms with supervised learning, such as the popular neural networks (Monkman et al 2019, Silva et al 2020). Most efforts in the use of computer vision systems in aquaculture are associated with categorization, behavior modeling, counting, estimation of fish body size (mainly length) or body weight prediction, mostly using an underwater stereo camera module (Boutros et al, 2015, Torisawa et al., 2011, Costa et al., 2006, Risholm et al., 2022). Usually, these methods advocate the idea that a conventional calibration allows to achieve a precise underwater 3D reconstruction, provided that a large base distance between cameras is guaranteed (Boutros et al, 2015, Shortis et al, 1998). Thus, the calibration of the stereo pair is simply performed with the structure dipped in water, and no alteration or correction to the mathematical models is proposed, which deals with the refraction effects of water properly. In some cases, the refraction effect is considered (Chadebecq et al, 2020), but, since most of the projective constraints are no longer valid, the calibration and reconstruction process become quite demanding.

This paper proposes a low-cost easy-to-use vision system to obtain measurements on live fish, avoiding the need for a demanding calibration. Next section presents a brief theoretical framework of the subject and how the system was implemented to estimate the correct length of fish based on the variation of the angle of incidence of the light rays in the water. Given some structural conditions such as a short baseline (on the order of 100mm) and an ROI (Region of Interest of the FOV, measured in degrees) of less than 60º, it is shown that the system is able to measure fish with an error of less than 1%, using a modified model for the stereo pair. The short baseline allows to have a compact system and benefits the estimation algorithm. In the following sections, a set of experiments is presented both in simulated and real conditions, showing that it works robustly for several orientations of the fish (even when the caudal fin and the snout are at different distances from the cameras).

2. System and Method

2.1 Theoretical Background

The system proposed in this paper consists of a pair of cameras, located at approximately parallel axes, installed in an air chamber (A), separated from the aquatic environment (W) by a transparent planar substance (T) of glass or acrylic.

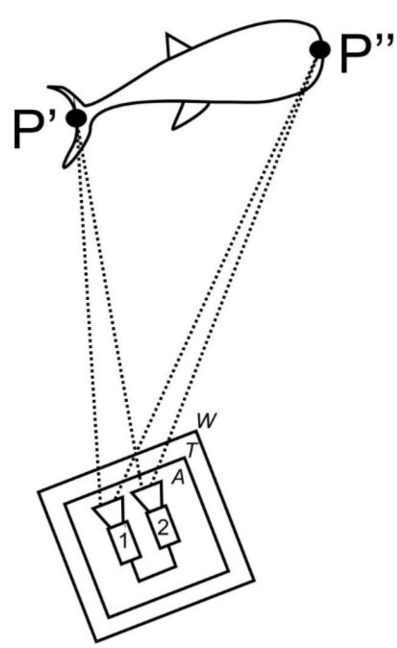

Figure 1. Underwater Stereo Vision System measuring the distance between two points (P' and P"). For both cameras 1 and 2, light rays cross three different media: Air (A), Glass or Acrylic (T) and Water (W).

The surface of the transparent substance shall be approximately perpendicular to the optical axes of the cameras, as shown in Figure 1.

Usually, the image pair taken by the said stereo camera system appears as shown in Figure 2. So, the left camera captures the left image (whose points have coordinates (x_1, y_1)), and the right camera captures the right image (whose points have coordinates (x_2,y_2)).

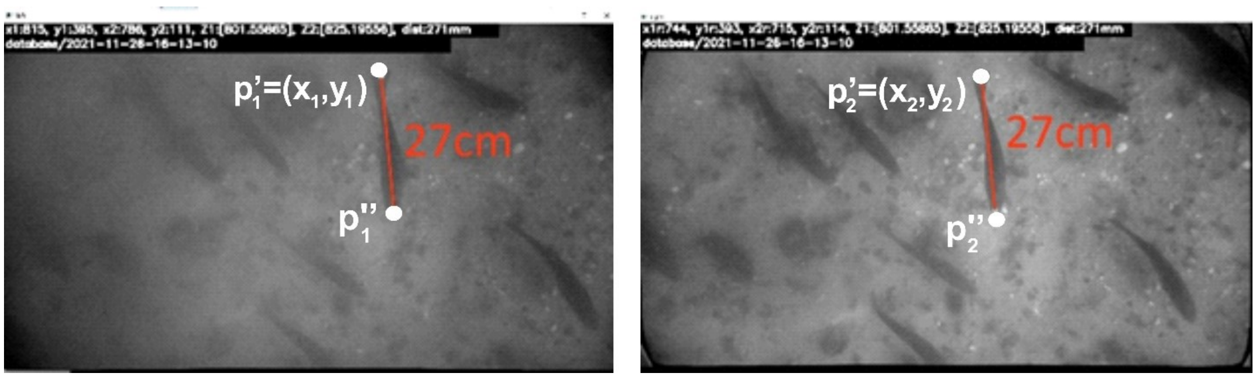

Figure 2. Images captured by camera 1 (on the left) and camera 2 (on the right). P' and P'' are projected on both images at different pixel coordinates. These differences allow to reconstruct tridimensionally P' and P''.

If the stereo pair and fish were out of water, defined as a conventional non-aquatic model (this happens whenever the fish is caught, brought on board or land), the measurements of the fish could be performed through a simple triangulation, using a standard projective model (Hartley et al., 2003). Considering an arbitrary point P' located on the fish, its coordinates X, Y and Z are projected directly on the cameras, through a unique angle of incidence. To exemplify this, Figure 3a shows an example for the coordinates X and Z, which are related to the angle of incidence a through a simple relationship of similarity between triangles. This can be described mathematically as follows (to simplify we will consider that both cameras are normalized, are equal, and have focal length equal to 1):

$$X = Z \tan a \qquad \textbf{(1)}$$

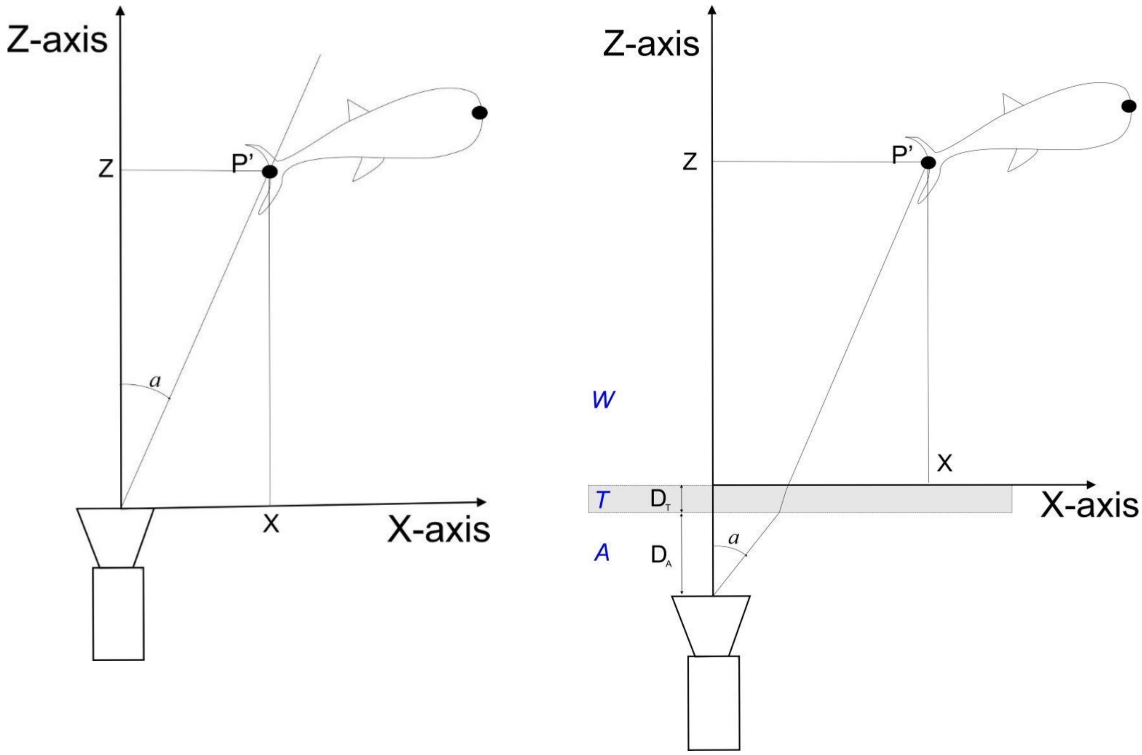

Figure 3. (a) On the left: Perspective projection of point P', considering that the light crosses a unique substance. (b) On the right: Perspective projection of the same point, considering that the light crosses three different media (A, T and W).

To achieve a complete 3D reconstruction, there are robust and well-established processes (Hartley et., 2003) to estimate Z and X (and Y for the generic case) from the triangulation of two projection rays, projected on two cameras, in the context of a conventional calibrated stereo model.

However, by forcing the light rays to cross different substances, as previously proposed in Figure 1, light is diverted at their transition points. For our example, represented in Figure 3b, there are two points of transition: between A and T and between T and W. For a given angle range, X relates to the angle of incidence a and Z through a new formula, much more complex than the formula presented in the equation (1):

$$X = D_A \tan a + D_T \, k_T \frac{\sin a}{\sqrt{1 - k_T^2 \sin^2 a}} + Z \, k_W \frac{\sin a}{\sqrt{1 - k_W^2 \sin^2 a}} \quad \textbf{(2)}$$

Where k_T and k_W are the inverse of the refraction index of the transparent medium T and the inverse of the refraction index of water, respectively.

Unfortunately, there is no explicit formula that computes exactly the value of *a* directly from X and Z (Chadebecq et al., 2020). Moreover, calibrating this stereo system with this formulation is a time-consuming and a very error-sensitive iterative process. In order to avoid a complex calibration algorithm (or large measurement errors), instead of using directly the above formula, we propose to use a set of simpler expressions. In this approach, described in the Appendix, a first estimate is produced using a prior calibration performed out of water, and, then, a transformation is applied to deal with the refraction effect; this allows to achieve an explicit computation of the solution with a negligible approximation error.

2.2 System implementation

As mentioned previously, one of the main goals of this work is to achieve a portable and compact underwater stereo system, knowing that a compact structure is much more maneuverable and robust to calibration changes. A compact configuration implies that the stereo pair exhibits a short baseline (defined as the distance between cameras), which leads to a larger error variance associated to the 3D point estimation (Boutros et al, 2015). However, this negative effect caused by a short baseline is somehow compensated by a better and more robust calibration due to the global structure compactness. Moreover, due to a reasonable view overlapping (since the cameras are closer to each other), we are able to work in a relatively narrow field of view, condition needed to make the expressions of Appendix valid, facilitating the 3D computation significantly.

The proposed system is then presented as a ten-centimeter baseline camera pair, integrated inside a solid acrylic chamber, a watertight enclosure capable of supporting deep underwater conditions, as shown in the Figure 4. Both cameras are leaning against a ten-millimetre-thick planar acrylic surface, minimizing the distance between cameras and water. Inside this acrylic enclosure, in addition to the cameras, a mini pc and connecting cables are included.

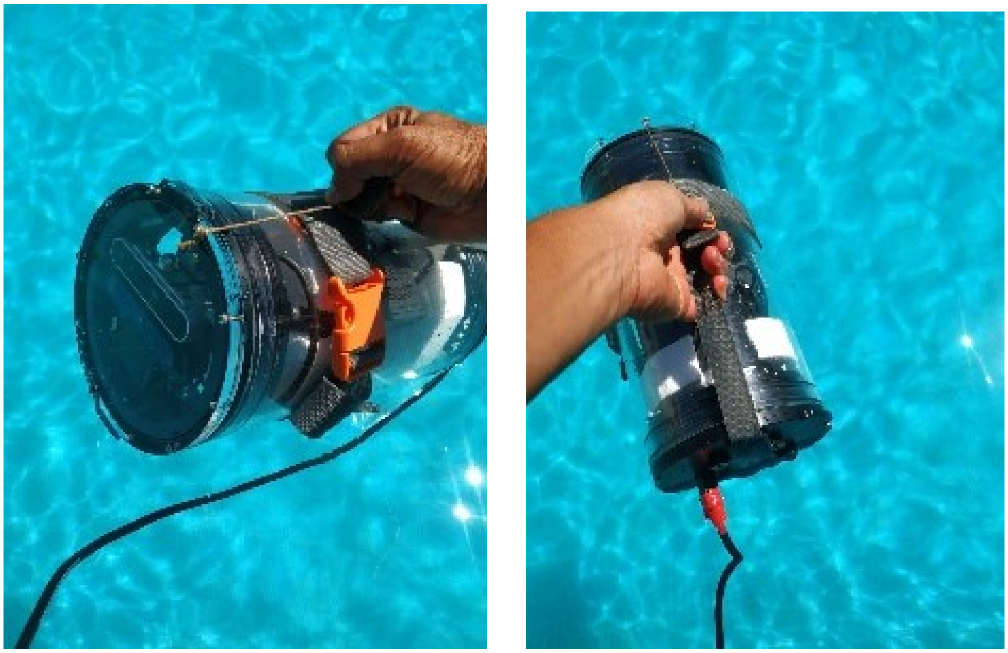

Figure 4. Two views of the developed stereo system.

The system has a power cord, although batteries could be an option. The system is able to communicate via ethernet cable or Wi-Fi (functioning when it comes to the water surface). In our implementation, the system is always placed on the surface, half-dipped in water, allowing access via Wi-Fi. This permits to visualize the images remotely in real time and capture new images when needed.

3. Results

Several experiments were carried out both in a simulated and a real environment with different levels of water turbidity. In the simulated environment, several scenarios have been tested to verify whether the theoretical model is close to the real one or not, and how it behaves within the limits of the validity range.

In a first experiment, it was experienced moving the specimen along a horizontal line between two boundaries of the image and verifying to what extent its location in the image induces an error in length estimation. The error was calculated by computing the difference between each estimate (given by the system) and the actual length of the fish, scaled by the actual length (in %).

Figure 5a shows the estimation error when the fish movement is performed at constant depth, closer to both cameras (0.8m). The observed error has an approximately parabolic shape, being higher at the boundaries of the image but never exceeds the 1% error. Figure 5b shows the estimation error when the fish is placed further away from the cameras (2m) and its path remains at this depth level. In this case, the observed error becomes negligible (less than 0.3%) for all range of positions.

In a third experiment (Figure 5c) we want to evaluate whether the system is robust to estimate the length for different poses, or, in other words, the measurement remains constant even when the fish spins on itself.

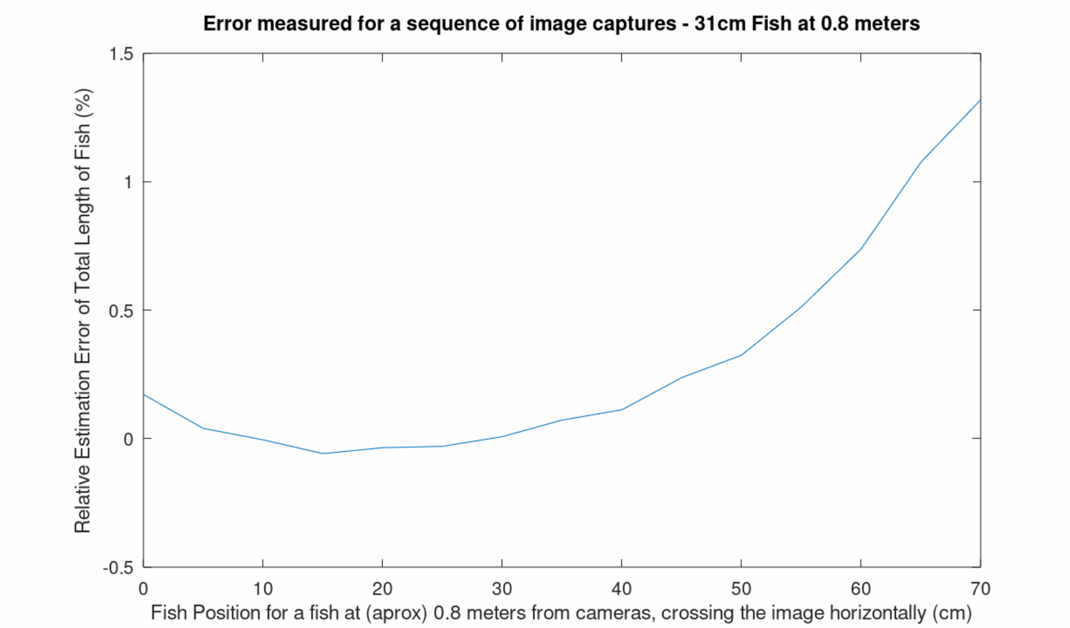

Figure 5a. Error measured for a moving specimen at 0.8 meters.

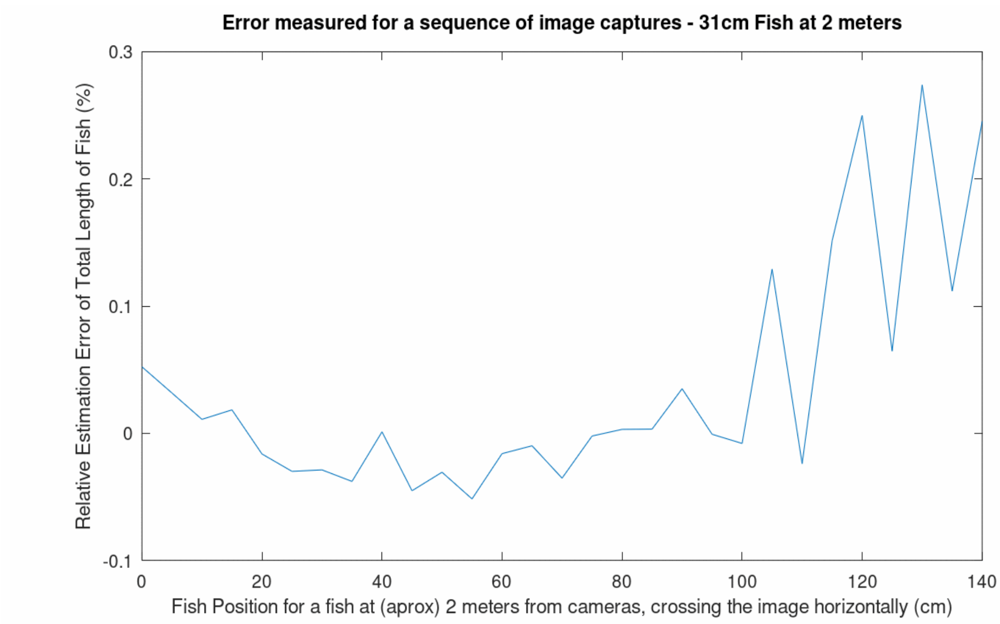

Figure 5b. Error measured for a moving specimen at 2 meters.

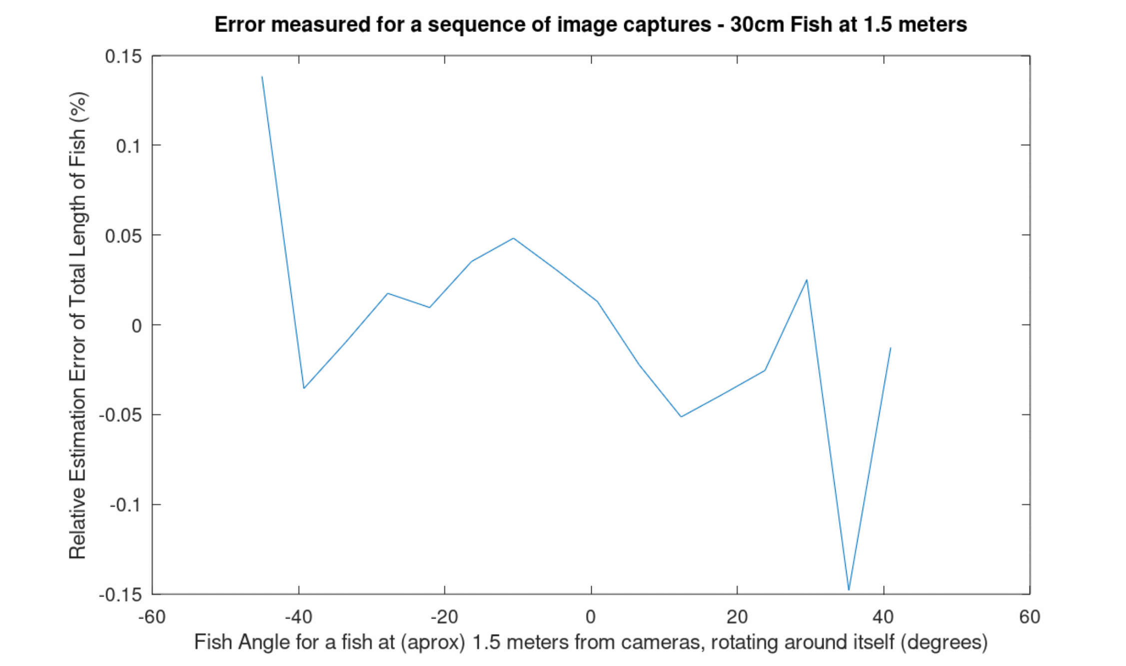

Figure 5c. Error measured for a rotating specimen at fixed position (1.5m)

It turns out that, if the specimen rotates at an angle between -40 and 40 degrees, the measurement error is negligible (less than 0.15%), so the system can handle well even when the caudal fin and the head are not at the same depth.

Several experiments were made in a real environment, namely in water with different levels of turbidity, with greater or lesser degree of concentration of fish. As shown in Figure 6, the stereo pair takes images in two different perspectives (represented by Figure 6a and Figure 6b), so that the parallax observed in each specimen allows to retrieve its depth information. In this experiment, the method is applied for specimens that are fully visible, and the results are superimposed in the Figure 6b. As expected, the fish are not all at the same depth nor are they in a horizontal pose relatively to the ground. Nevertheless, the method estimated the actual positions of the caudal fin and the snout. To better estimate the lengths, especially for fish that are folded, a method of estimating lengths by sections was implemented, by estimating two intermediate control points along the fish. In this case, the total length corresponds to the line integral evaluated along the spline defined by the selected control points and fish ends.

Figure 6a. Image captured by left camera.

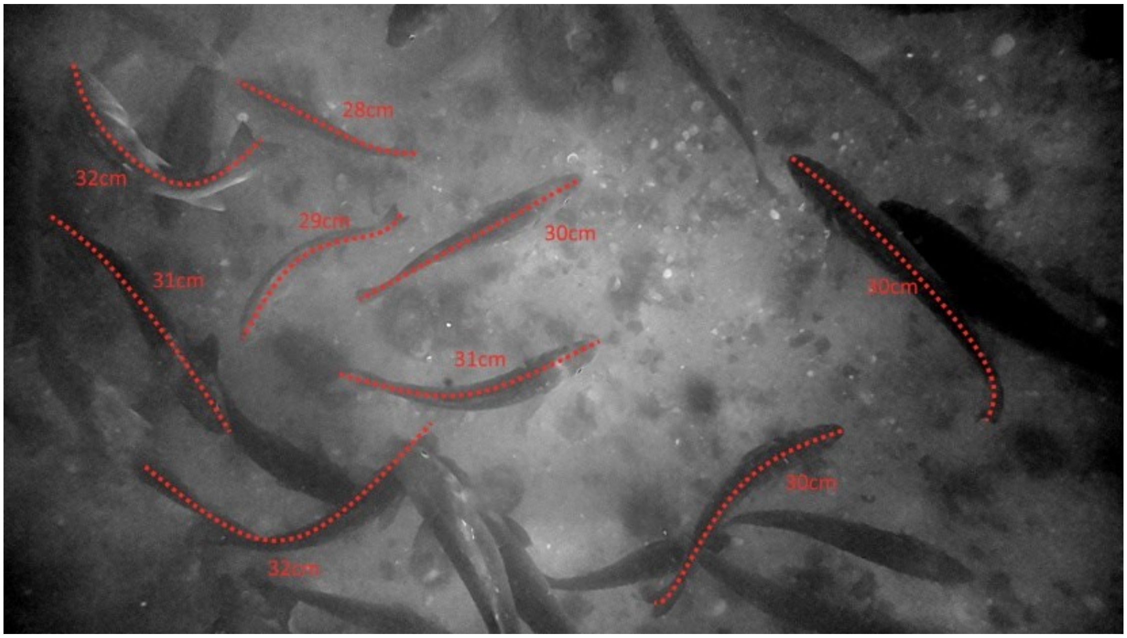

Figure 6b. Image captured by right camera, where the measures appear superimposed.

4. Conclusions

This paper proposes a low-cost easy-to-use stereo vision system for measuring live fish in an aquatic environment, without the need for a demanding calibration. The system has a short baseline (on the order of 100mm) and an ROI (Region of Interest of the FOV, measured in degrees) of 60º. It is shown that the system is able to measure fish with an error of less than 1%, using a default calibration. A set of experiments was carried out, working robustly for a set of orientations and positions of the fish. The major contributions of this work are the following: (1) it is shown that a short baseline and a narrow field of view reduces significantly the negative effect of refraction on the stereo reconstruction; (2) a new and straightforward method to compute underwater 3D reconstruction is proposed: this method produces a first estimate using a prior calibration performed out of water, and, then, applies a transformation to the first estimate to deal with the refraction effect; (3) it is shown that, under certain conditions, only the water refraction index is relevant to obtain the 3D reconstruction (or, equivalently, the refraction parameters of intermediate substances can be neglected).

5. Appendix

In this Appendix, the main mathematical expressions that rule the projection and reconstruction are presented, in the context of a underwater stereo vision system. For a projection model with narrow field of view (<60º), a new formulation was derived to replace equation (2), disregarding the Taylor terms of order greater than 3, as shown below:

$$X = D_A\left(a - \frac{a^3}{6}\right) + D_T k_T\left(a + (3k_T - 1)\frac{a^3}{6}\right) + Zk_W\left(a + (3k_W - 1)\frac{a^3}{6}\right) \tag{3}$$

Unlike what happened in equation (2), the angle *a.* can be computed as a solution of a third-degree polynomial equation (using the resolvent formula of a third-degree equation). This means that the angle calculation can be set explicitly in relation to *X* and *Z*. This approach allows to reproject (explicitly) a given 3D point on the camera images. The reprojection of the estimated 3D points on the images is especially useful to compute errors in the image space and to estimate the intrinsic and extrinsic parameters of this stereo pair of cameras.

For two parallel cameras as defined in the Figure 1, we can combine both projection equations, using its projections on image x_1, x_2 (Figure 2) and the angle of incidence a_1, a_2 . This combination can be used to estimate the three coordinates ($\tilde{X}$, $\tilde{Y}, \tilde{Z}$) of a given underwater 3D point, based on an initial estimate *(X_A, Y_A, Z_A)*, computed by considering a conventional stereo system (Hartley et al. 2003), calibrated out of water. The coordinates ($\tilde{X}$, $\tilde{Y}, \tilde{Z}$) are computed as follows:

$$\tilde{X} = X_A \frac{x_1 - x_2}{cosm(a_1)x_1 - cosm(a_2)\,x_2} \tag{4}$$

$$\tilde{Y} = Y_A \frac{x_1 - x_2}{cosm(a_1)x_1 - cosm(a_2)\,x_2} \tag{5}$$

$$\tilde{Z} = \frac{1}{k_w} Z_A \frac{x_1 - x_2}{cosm(a_1)x_1 - cosm(a_2)\,x_2} - D_A - D_T \tag{6}$$

where:

$$cosm(a) \;=\; \frac{\cos a}{\sqrt{1 - k_W^2 \sin^2 a}}$$

Notice that, in this formulation, only the refraction index of water is relevant for the reconstruction computation. The refraction index of the intermediate medium has a negligible influence on the estimation results.

The formulation presented above was used throughout this work to reconstruct underwater points from matched points observed on the stereo pair of images. The reconstruction uses a standard stereo calibration (Zhang, Z. 1999) and a few parameters of the underwater structure, such as the refraction index of every substance involved and the corresponding thickness. This approach allows to simplify both the reconstruction process and the calibration needs, without neglecting, however, the effect of refraction on the system.

6. ACKNOWLEDGMENTS

We thank the MARE Marine and Environmental Sciences Centre and IMAR Institute of Marine Research, University of the Azores, Portugal, for providing conditions to carry out the experiments in a real environment. The authors work for Research Department of Fishmetrics, and their work was developed in the context of the project MEDUSA (AÇORES-01-0247-FEDER-000023), partially funded by the Research Program ACORES2020, with the participation of Azorean Funds, Portuguese Funds, and European Funds.

7. REFERENCES

Hartley, R., Zisserman, A. 2003. *Multiple View Geometry in Computer Vision*. New York, NY, USA: Cambridge University Press. ISBN: 0521540518

Jardim E., Azevedo M., Brites N. 2015. *Harvest control rules for data limited stocks using length-based reference points and survey biomass indices*. Fisheries Research, 171 : 12 –19 .

Antonucci, F., Costa, C. 2020. *Precision aquaculture: a short review on engineering innovations*. Aquaculture Int 28, 41–57 (2020). https://doi.org/10.1007/s10499-019-00443.

Silva C., Gaspar T., Aires R., Ramalho M., Reis D., Diogo H., Canha Â. et al. 2018. *An exhaustive sampling approach for fish data collection, using an automatic measuring system.* In Proceedings of the 9th International Fisheries Observer and Monitoring Conference. Ed. by Kennelly S. J., Borges L.. Vigo, Spain. ISBN: 978-0-9924930-7-3. 395 pp.

Silva C, Aires R, Rodrigues F, 2020. Automatic fish measurement using a camera and a 3D sensor applied to a long-term experiment, ICES Journal of Marine Science, Volume 77, Issue 7-8, December 2020, Pages 3050–3057, https://doi.org/10.1093/icesjms/fsaa190

Li, D., Hao, Y. & Duan, Y. 2020. *Nonintrusive methods for biomass estimation in aquaculture with emphasis on fish: A review.* Rev. Aquac. 12, 1390–1411.

Hao, M., Yu, H. & Li, D. 2016. *The measurement of fish size by machine vision—A review.* IFIP Adv. Inf. Commun. Technol. 479, 15–32.

Monkman, G. G., Hyder, K., Kaiser, M. J. & Vidal, F. P. 2019. *Using machine vision to estimate fish length from images using regional convolutional neural networks*. Methods Ecol. Evol. 10, 2045–2056.

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*. Limnology and Oceanography: Methods. https://doi.org/10.1002/lom3.10020.

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.* Aquat. Living Resour. **24**, 107–112.

Costa, C., Loy, A., Cataudella, S., Davis, D. & Scardi, M. 2006. *Extracting fish size using dual underwater cameras*. Aquac. Eng. 35, 218–227.

Risholm P., Ahmed Mohammed, Trine Kirkhus, Sigmund Clausen, Leonid Vasilyev, Ole Folkedal, Øistein Johnsen, Karl Henrik Haugholt, Jens Thielemann, 2022. *Automatic length estimation of free-swimming fish using an underwater 3D range-gated camera*, Aquacultural Engineering, Volume 97.

Shortis, M., Harvey, E. (1998). Design and calibration of an underwater stereo-video system for the monitoring of marine fauna populations. Int. Arch. Photogramm. Remote Sensing. 32.

Chadebecq, F., Vasconcelos, F., Lacher, R. et al. 2020. *Refractive Two-View Reconstruction for Underwater 3D Vision*. Int J Comput Vis 128, 1101–1117 (2020). https://doi.org/10.1007/s11263-019-01218-9

Zhang, Z. 1999. *Flexible Camera Calibration by Viewing a Plane from Unknown Orientations*. In: IEEE kim.L.Boyer.Sudeep.Sarkar, eds.IEEE International Conference on Computer Vision, Greece. Washington: IEEE Computer Society: 666-673P.