

COBEM-2017-1676

DENSITY BASED FISH COUNTING WITH COMPUTER VISION FOR AQUARIUM AND BOX STORAGE IN NOISY ENVIRONMENTS

Thiago Henrique Gomes Lobato
Gustavo da Silva Vieira de Melo

Universidade Federal do Pará, Department of Mechanical Engineering, Belém, Brazil.
thiagohgl@hotmail.com, gmelo@ufpa.br

Abstract. *When making fishing inventory there is hardly room for optimization. The process takes too long, reaching days when the number of aquariums and boxes are very high. Sometimes fishes need to be removed from their recipient and it can be a harmful process to them, especially for some kinds of ornamental ones. To avoid this time consuming and difficulty procedure, a very imprecise counting can be done with less frequency, counting fishes in the aquariums or in the boxes by eye, which, however, can be very prone to errors because of the fish size and quantity, producing errors that are greater than 40%. To address that problem, a system of fishing counting with computer vision was proposed, using a grid image binarization together with image segmentation techniques such as color and texture segmentation. The binary image could then be filtered and the fishes be counted with the use of element density, reducing the error caused by the fish overlap. The system was able to reach a high accuracy (ca. 93%) even in noisy environments, with high ambient light and reflections. The proposed system is way faster than the conventional way, does not harm the fishes, is more precise than the bare eye and can be automated, proving the algorithm high value in fish counting.*

Keywords: *Computer Vision, Fish Counting, Image Segmentation*

1. INTRODUCTION

An automated way of counting fish is already some very wishful thing to have (Silverio, 2016)(French *et al.*, 2015) and (Labuguen and Volante, 2012), especially due to the difficulty and time consuming work that is to count them individually, taking more or less a minute and a half to count 31 fishes(Silverio, 2016) by eye and making the counting of 100 or more mostly hopeless. Some approaches used to address that problem were foreground subtraction with video footages(Silverio, 2016) and convolution neural networks(French *et al.*, 2015). The first has the problem that it needs the fishes to be in constant movement, which is not always the case. The second one has the limitation that a neural network needs many samples from the type of fish it is searching, so if one wants to acquire new type of fishes, it will demand a new training of the network every time (and a different one for each fish to increase precision). A general convolution network could possibly be trained to detect any fish with an average accuracy, but the alternative proposed here was meant to be more general and totally fish independent.

In (Labuguen and Volante, 2012) a similar approach as the one used here was done, which is image binarization. That process is independent of the of fish type (works based on intensity contrast) and works really well if there's a nice distinction between the fish and the background, as shown in (Labuguen and Volante, 2012). That good distinction, however, is not always the case, and there's also the problem of fish overlap. Thus it is necessary to use some measures to have a noisy robust system. In this work a grid binarization together with colour and texture segmentation as well as density fish estimation was used to solve those problems.

2. COMPUTER VISION ALGORITHMS

2.1 Color Segmentation

Mostly images, when digitalized, are nothing more than a intensity matrix, and depending of those intensities it possible to identify the colour of a determined pixel based in some colormap (Shapiro and Stockman, 2012), and one that it is really helpful to segmentation is the Hue-Saturation-Value or HSV (Shamik *et al.*, 2002) (Vitabile *et al.*, 2001). In this system, one is able to easily find the entries of a specific colour(Hue) and also to remove some very high intensity(value) parts, as, for example, the environment lights. In the figure 1 below it is possible to understand how the HSV system is divided.

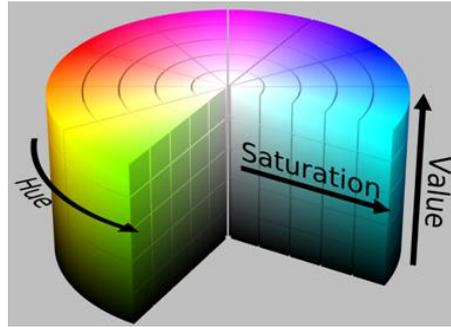


Figure 1: HSV colormap.

Source: <https://i.stack.imgur.com/3zJuP.png>, accessed in 1/10/2017.

The segmentation is very straightforward, one simply find the pixels which belong to the undesired colour and/or intensity and remove it.

2.2 Texture Segmentation

A texture segmentation can be done in several ways(Gonzalez R.C, 2003) and has already showed great results (Portes de Albuquerque *et al.*, 2003) (Andre *et al.*, 2011). in this work, however, it was used a range filter approach, where, for each pixel in the moving window $H_{n \times n}$, it would give as result for the intensity i :

$$i = \max(H_{n \times n}) - \min(H_{n \times n}) \quad (1)$$

In this way it is possible to identify abruptly changes of texture, which, afterwards, can be dilated with morphological operations in order to give appropriate blobs of the fishes.

2.3 Binarization

Binarization can be easily described as:

$$if(i_n < Treshold) : i_n = 0; else : i_n = 1 \quad (2)$$

which means that every value above a given threshold will be given the value 1, while every value below would get the value zero (Shapiro and Stockman, 2012).

Binarization can be used to image counting, where one finds the so called blobs, which are like isolated islands of one's. Those blobs usually represents the objects that we're looking for and a main goal in segmentation counting is to identify those blobs (Shapiro and Stockman, 2012).

In the figure 2 below on can see an example of a image and its blobs after binarization with inverted values:

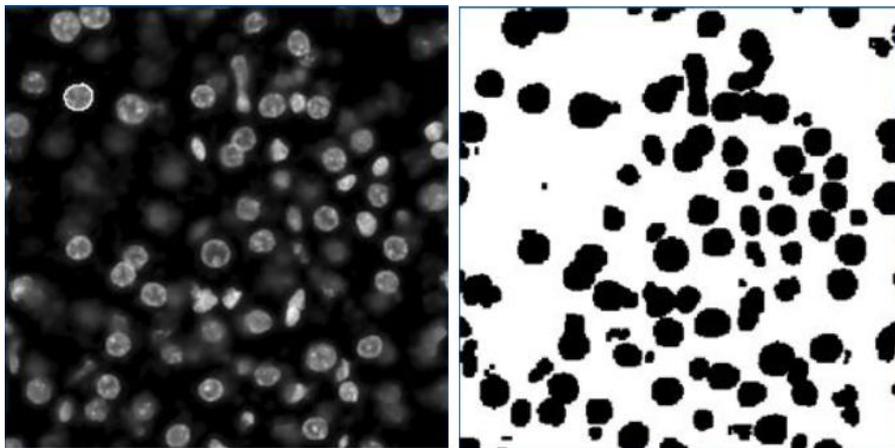


Figure 2: A image and its corresponding blobs after binarization.

Source:

[http://slideplayer.com/slide/7596540/25/images/8/Example:+Labeling+Blobs+\(here+application+is+medical+image+%E2%80%93+ce](http://slideplayer.com/slide/7596540/25/images/8/Example:+Labeling+Blobs+(here+application+is+medical+image+%E2%80%93+ce)
Accessed in 1/10/2017.

The values obtained by binarization have normally much noise and thus must be filtered with some morphological operations like Erosion and Dilatation, which can together be used to close blobs and remove noise (Shapiro and Stockman,

2012). The threshold can automatically be chosen using the Otsu method (Otsu, 1979).

3. EXPERIMENTAL PROCEDURE AND METHODOLOGY

There were two types of fish recipients, the aquariums and the boxes (basquetas), both can be seen in the figure 3: It



Figure 3: Example of aquarium and box to have the number of fishes counted.

is clear for the pictures how troublesome it is to count those fish as well as the environmental noise of the images.

The equipment and software used in the process were:

- Digital Camera Nikon Coolpix;
- Notebook Lenovo Y-50-70 i7;
- Matlab 2016b(MathWorks, 2016).

Photos were taken from 10 aquariums and boxes and then processed in the developed algorithm. The algorithm follows like this:

- 1 - Select ROI;
- 2 - Do colour segmentation in the HSV color map;
- 3 - Do range based texture segmentation;
- 4 - Use only the points that are both in the colour as well the texture range.
- 5 - Do a binarization in a grid form in the resulted masked image to find blobs, in this work, a 12x12 grid;
- 6 - Remove disproportional blobs, filter image with morphological filters and calculate the mean blob area;
- 7 - Use the mean area to make estimations of how many fish are in each blob.

This can be more easily seen in the block diagram in figure 4. The end result for the above two images can also be seen in the figure 5.

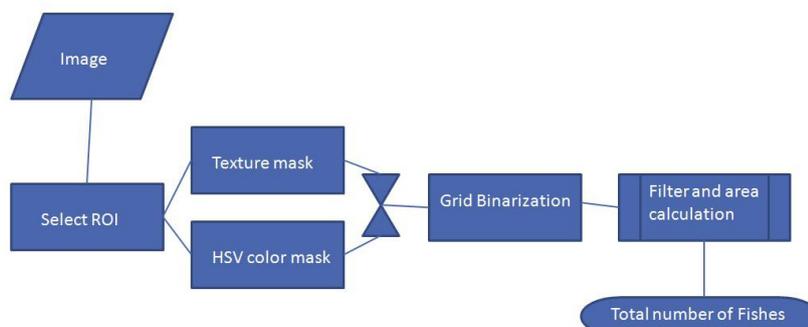


Figure 4: Process Block Diagram.

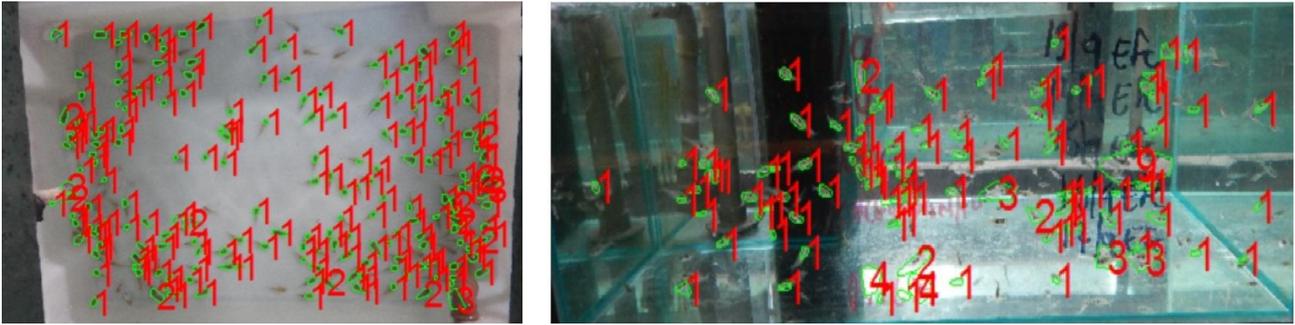


Figure 5: Left: Calculated/Real: 216/217. Right: Calculated/Real: 120/122.

The use of a grid binarization was to increase the accuracy of the blobs. This process was chosen over a full adaptive binarization as in (Bradley and Roth, 2007) because, even though it is a great method and works better than the global threshold, it has some trouble find the cleanest blobs, as can be seen in the comparison in figure 6.

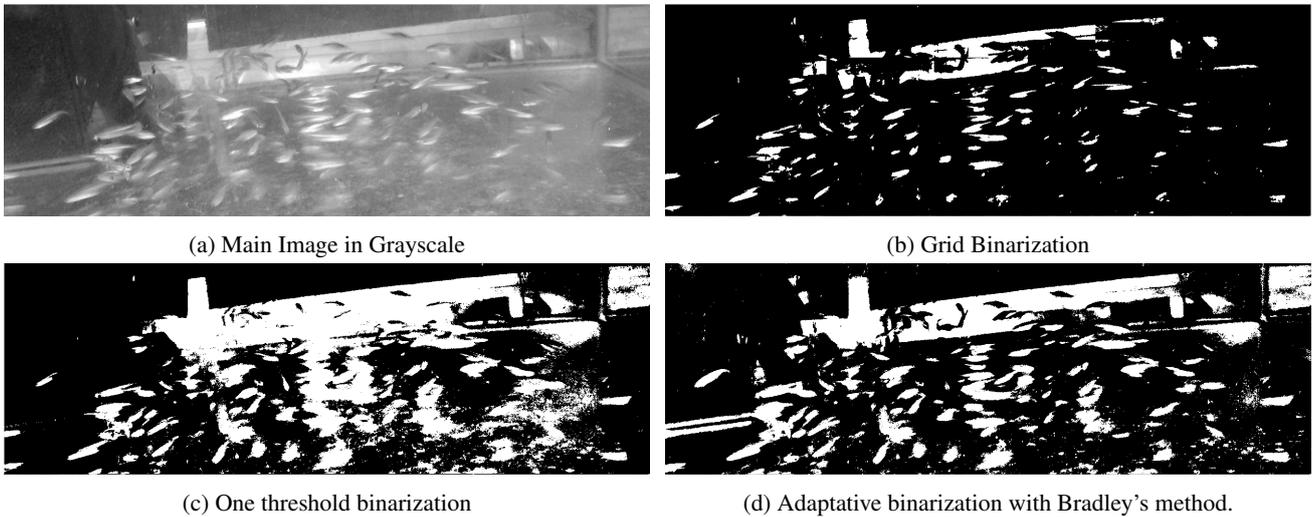


Figure 6: Comparison of different types of binarization. One can see that the proposed grid binarization can best proportionate the correct blobs and avoid most of the reflection in the aquarium, followed by the Brandley's method and the only one threshold being the worst.

To get these results in the grid binarization, however, it was used an empiric constant of 0.74 in the Otsu's threshold to all images, in order to reduce discontinuities effects in the border of each grid. The grid in the image can be seen in the figure 7.

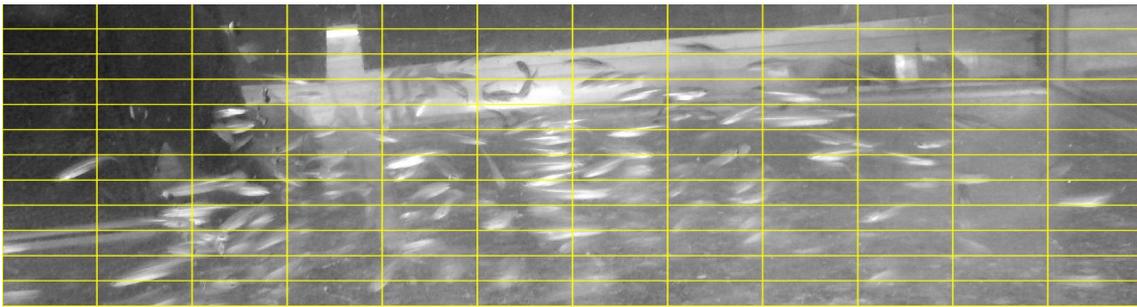
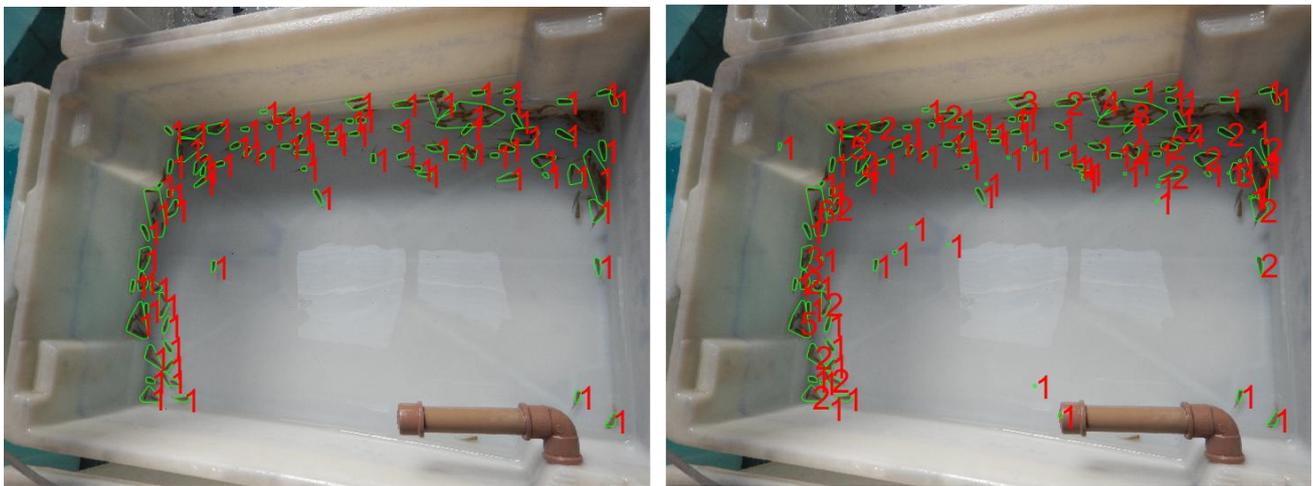


Figure 7: Image with the 12x12 grid to binarization.

The most obvious problem of a image fish counter is what to do when one fish is behind another. Well, for this problem, it was considered that it is extremely unlikely that a fish would be completely behind another one, what would leave at least part of his body in sign, making a identified blob become bigger than the others. Thus, in order to determine how many fish are in each blob, a density approach was used, dividing the area of a blob for the mean blob area and rounding the number. This approach has enormously improved the accuracy of the method, as can be seen in the figure 8 where one can compare the result of a binarization with and without the density blob consideration.

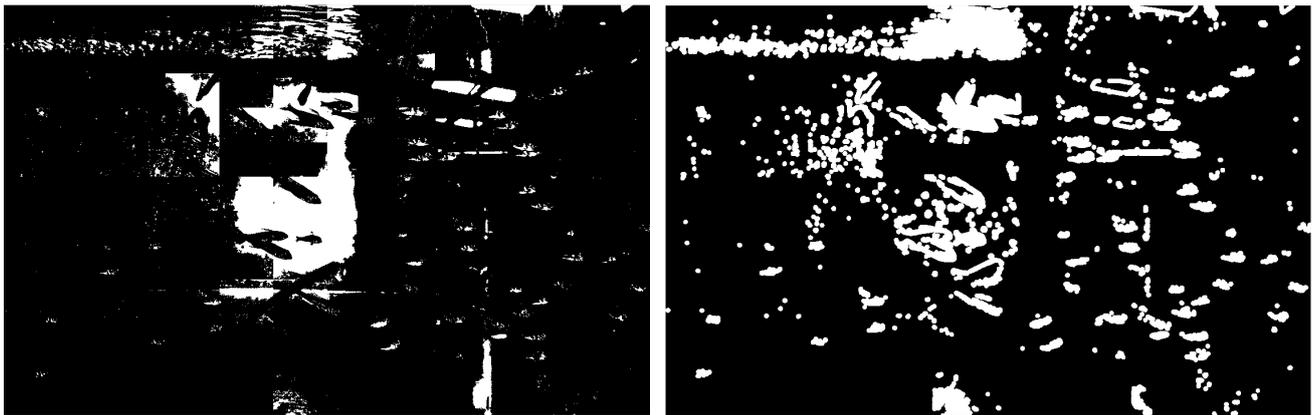


(a) No density. 83/161 fishes

(b) With density. 169/161 fishes

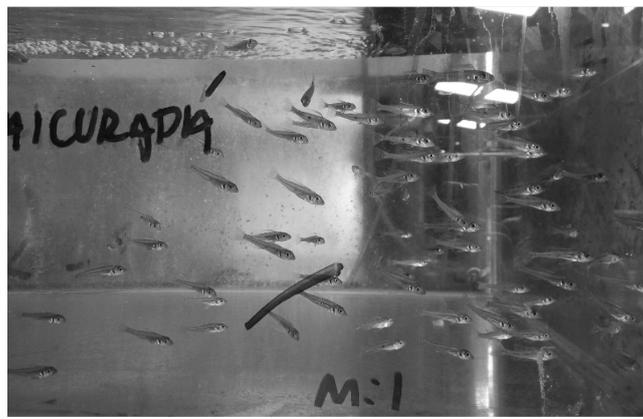
Figure 8: Comparison of the result with and without a density based approach. It is clear how the use of the density can improve the result of the analysis.

The filters of texture, colour as well as the morphological operations are specially needed for the most difficult images, in which is extremely hard to binarize the fishes due to their very high contrast. Using the filters, however, it becomes possible to have a somewhat accurate result, as seen in the figure 9.



(a) No Filter

(b) Texture, Colour and Morphological Filters



(c) Original Image

Figure 9: Comparison of the results with and without filters in a very difficulty image with a lot of intensity contrast.

4. RESULTS AND DISCUSSION

The algorithm was then tested for all 10 aquariums and boxes, the relative error can be seen in the figures 10 and 11.

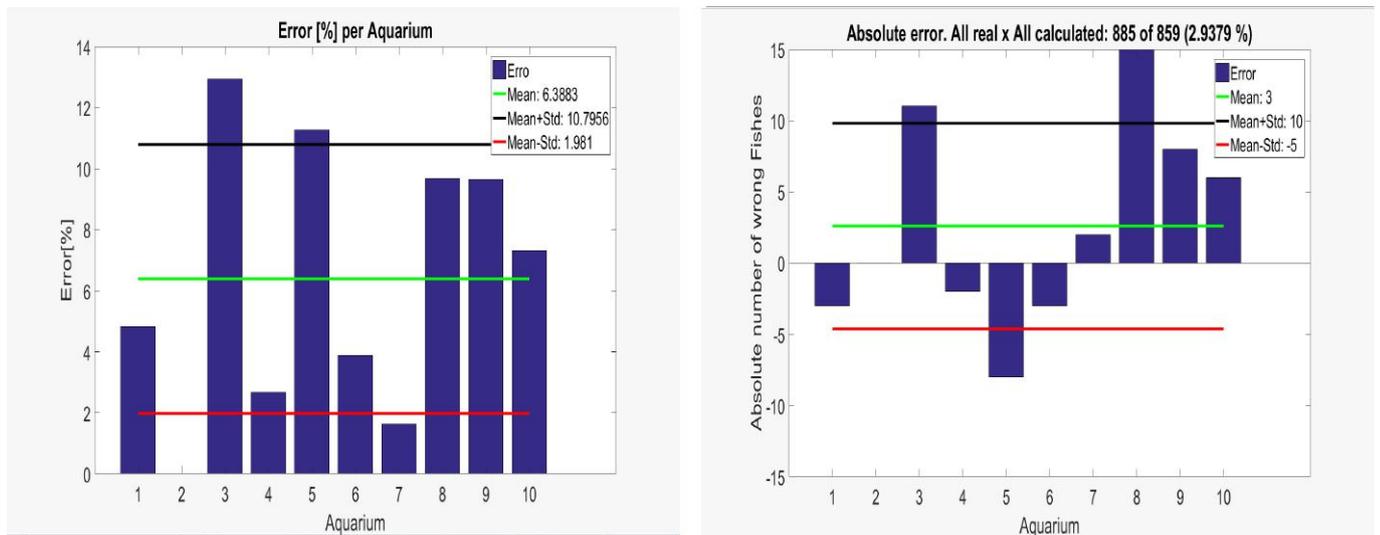


Figure 10: Percentage and absolute error for all tested aquariums.

For both cases we have an average accuracy for about 93%, which is way higher than the eye approach and way faster than the conventional. We can also analyse the global error for both cases, which is ca. 3% and 2% for the aquarium and boxes respectively, this lower error is due to the fact that the errors tend to compensate themselves in the long run,

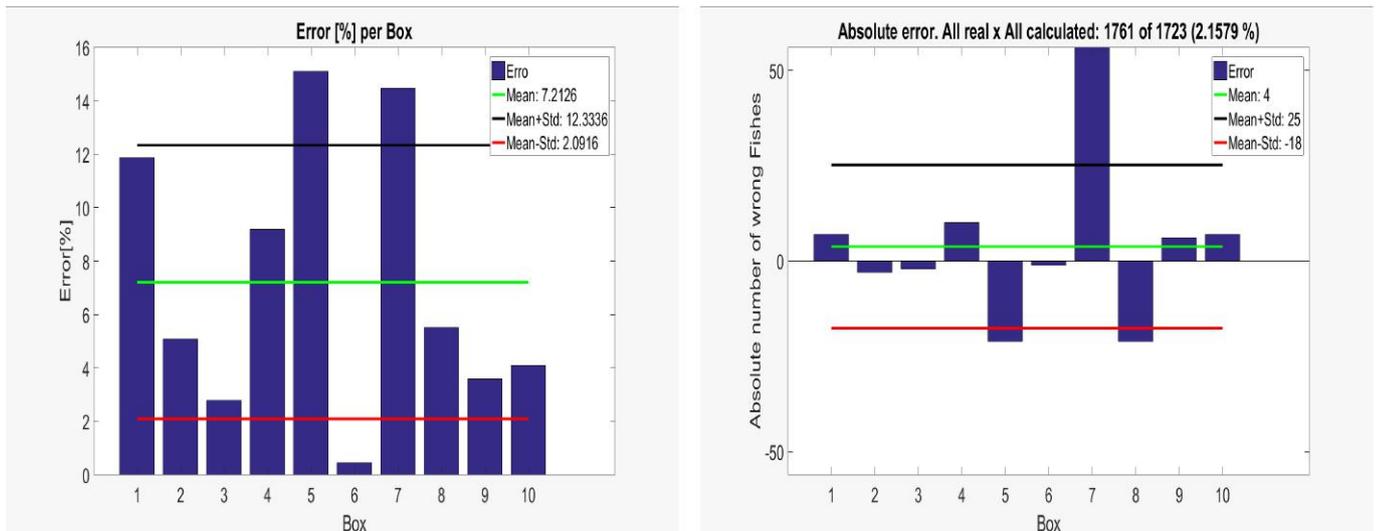


Figure 11: Percentage and absolute error for all tested boxes

supporting that the method is a very good alternative for global stocking control and a feasible option to fish control. We are also able to see that the biggest error in the model is about 15%, which is way less than the medium error in the naive approach of eye counting. The time to process each image after the filters settings is approximately 4 seconds in the hardware presented before. Assuming that to take each photo a time of 20 seconds is used, for 2000 photos (aquariums + boxes) it would take a single person 13 hours to have an accuracy of roughly 90% in the counting, while in the naive method a whole team with more than 6 persons take more than 20 hours (3 to 4 days) to count all with very low accuracy. It is also possible (and recommended) to automate the photo processes, which could leave to a real time stocking count with a almost perfect accuracy, since it could take many photos and use a robust mean.

5. CONCLUSION

We were able to see that the algorithm performs with relatively high accuracy even in very noise environments and that it is a very good and fast alternative to fish counting. The grid binarization method together with the colour and texture segmentation was able to give better results than both the global threshold approach as the adaptive Brandley's method. The density method was proved a satisfactory approach to account for overlapping fishes. The increment in the average accuracy from the naive method was of more than 30% and the time used to make the counting of the proposed method is reduced in almost five times considering the same number of workers for both cases. The possibility to automate the process and use more photos to each aquarium/box creates a enormous attractive for practical uses of the algorithm, and also opens the door for future works with videos streams or regular taken photos.

6. REFERENCES

- Andre, L.B., G.F., d.A., Francisco, A.R., Odemir M., B. and Luciano, d.F.C., 2011. "An entropy-based approach to automatic image segmentation of satellite images". *Physica A: Statistical Mechanics and its Applications*, Vol. 390, pp. 512–518.
- Bradley, D. and Roth, G., 2007. "Adaptive thresholding using integral image". *Journal of Graphics Tools*, Vol. 12, pp. 13–21.
- French, G., Fisher, M., Mackiewicz, M. and Needle, C., 2015. "Convolutional neural networks for counting fish in fisheries surveillance video".
- Gonzalez R.C, Woods R.E., E.S., 2003. *Digital Image Processing Using MATLAB*. Prentice Hall, New Jersey, 1st edition.
- Labuguen, R. and Volante, E., 2012. "Automated fish fry counting and schooling behavior analysis using computer vision". *IEEE 8th International Colloquium on Signal Processing and its applications*.
- MathWorks, 2016. *MATLAB Release 2016b The MathWorks, Inc.,*. Massachusetts, United States.
- Otsu, N., 1979. "A threshold selection method from gray-level histograms". *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 9, pp. 62–66.
- Portes de Albuquerque, M., Esquef, I., Gesualdi Mello, A. and Portes de Albuquerque, M., 2003. "Image thresholding using tsallis entropy". *Pattern Recognition Letters*, Vol. 25, pp. 1059–1065.
- Shamik, S., Gang, Q. and Pramanik, S., 2002. "Segmentation and histogram generation using the hsv color space for image retrieval". *Proc. Int. Conf. Image Process.*, Vol. 2, pp. II–589.
- Shapiro and Stockman, 2012. *Computer Vision*. Pearson, London, 1st edition.

Silverio, F., 2016. *Automatic Fish Counting in Aquariums*. Instituto Superior Tecnico, University of Lisbon, Portugal.
Vitabile, S., Pollaccia, G., Pilato, G. and Sorbello, F., 2001. "Road signs recognition using a dynamic pixel aggregation technique in the hsv color space". *Proc IEEE Image Anal Process*, pp. 572–577.

7. RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.