# Automatic Fish Counting in Aquariums

Francisco Silvério, Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal

***Abstract*—In this project we propose a computer vision method, based on background subtraction, to estimate the number of zebrafish inside a tank. We addressed questions related to the best choice of parameters to run the algorithm, namely the threshold blob area for fish detection and the reference area from which a blob area in a threshed frame may be considered as one or multiple fish. Empirical results obtained after several tests show that the method can successfully estimate, within a margin of error, the number of zebrafish (fries or adults) inside fish tanks proving that adaptive background subtraction is extremely effective for blob isolation and fish counting.**

***Index Terms*—computer vision, zebrafish counting, background subtraction, Hu moments, image processing.**

## I. Introduction

Zebrafish (*danio rerio*) is a small freshwater fish that is widely used as an animal model in biomedical research with origins in specific locations across the globe. Research laboratories around the world require a huge number of individuals to perform a great variety of experiments. Those fish are breed and maintained in big fish facilities managing hundreds to thousands of fish tanks. Usually these tanks are standardized containers (for instance, from 3 to 8 l) which may host several dozens of animals each. Obtaining an up to date count of the total number of animals in a fish facility is an essential task, performed by human technicians who manually extract animals with the help of small fish nets. This manual counting process requires a significant amount of time and is error prone. Moreover, handling animals for counting induces significant stress, with all the harmful consequences that may cause to the animals and, consequently, affect the scientific experiments they are involved in. Finding a noninvasive automatic procedure to obtain the precise number of zebrafish in facilities tanks, avoiding all the disadvantages of manual counting, is a long sought goal of fish facilities managers.

Fish counting automation may be approached using computer vision. In fact, today there are many examples of complex applications making use of computer vision techniques [1] such as:

- Optical character recognition: reading handwriting and automatic number plate recognition
- Machine inspection: measure tolerances on aircraft wings or inspect steel casting with X-ray vision
- 3D model building: 3D models from aerial photographs
- Medical Imaging: perform long-term studies of brain morphology
- Face detection: to be used in image searching
- Visual authentication: grant people permission for accessing buildings based on morphology features
- People tracking: monitor passenger motion in airports.

Some of the previous applications make use of techniques such as optical flow and background subtraction which have particular interest in this project. Optical flow is highly related to pixel motion and its variability between frames [2]. Furthermore , it gives a reliable estimation of displacement between different frames. Background subtraction is specially relevant when, for example, the need for isolating moving regions in a sequence of images arises. In this fish counting project, since images are two-dimensional, the main difficulties to overcome are: regions where fish overlap, mirroring effect, i.e., fish reflections on the tanks sides and, mainly, fish shoaling.

### A. Environment

Fish facilities usually use zebrafish of the following genotypes: wild type zebrafish from strains AB and TU, mutant Nacre zebrafish and transgenic fish from multiple lines with TU and Nacre background. It is important to state that to the different genotypes correspond different phenotypes.

From this point, we will assume that fish are divided into four different categories with different genotype and age: AB/TU fries (30-day old), AB/TU adults (90-day old), Nacre fries (30-day old) and Nacre adults (90-day old). The different categories can be seen from Figure 1 to 4. This type of fish is extremely sensitive when changes in its environment occur, thus, inflicting too much stress while handling may, in some cases, lead to death.

Figure 1. Nacre zebrafish fries.

Figure 2. AB/TU zebrafish fries.

Fish growth rate depends on factors such as fish density in a tank and feeding. In this project, fish have an average length (fork lenght) of 1.8$\pm$0.16 cm (fries) and 2.5$\pm$0.15 cm (90-day old adults) and, as adults, should not be longer than 3 cm. At the fish facility, fish are maintained in standard sized tanks with, for instance, 3.5L , as it can be seen in Figure 5, where water temperature is around 28$^o$C (27.94$\pm$0.05) and the external temperature is approximately 25$^o$C.

Figure 3. Adult Nacre zebrafish. [3]

Figure 4. Adult AB zebrafish. [3]

Figure 5. Zebrafish tank type. [4]

Tanks are stored, side by side, in appropriate housing systems, where the water recycling system makes sure that the water in aquariums is constantly being renewed and the automatic feeding system provides the correct periodic feeding for each tank. Figure 6 shows an example of a housing system where that is verified.

Figure 6. Housing system at the fish facility

In our project, the number of fish per tank may vary from less than a dozen to a maximum number of 35. At ”day zero” fish eggs are placed inside tanks after microscopic larvae counting. Figures 7 and 8 show the physical features of a zebrafish larva at day zero for the two genotypes. After 30

Figure 7. Nacre zebrafish larva.

Figure 8. AB zebrafish larva.

days, a manual counting is performed to analyse how many

| Fish Number | Time [s] |
|---|---|
| 12 | 28 |
| 16 | 40 |
| 18 | 35 |
| 19 | 48 |
| 23 | 58 |
| 23 | 44 |
| 26 | 55 |
| 26 | 78 |
| 27 | 67 |
| 27 | 91 |
| 28 | 53 |
| 28 | 56 |
| 28 | 69 |
| 28 | 52 |
| 29 | 43 |
| 31 | 82 |

Table I
COUNTING TIMES OBTAINED IN 16 DIFFERENT COUNTS FOR ZEBRAFISH FRIES PERFORMED BY AN ACCREDITED TECHNICIAN.

fish survived the first month allowing the extraction of data to calculate the mortality ratio. This procedure is repeated on the third month when zebrafish are adults.

To have an idea about the time that the manual counting process usually takes, we can analyse Table I. By analysing the table, we can conclude that there is no relationship between the number of fish and the time spent to perform the task. In those counts, the lowest value obtained was 28 seconds for 12 fish in a tank. These values can be used as a reference to compare to the time performance of our algorithm that will be done in Section IV.

## II. RELATED WORK

Nowadays there are many different computer vision techniques used in object counting. Since the purpose of this work is to develop a fish counting system, all of the following approaches represent different solutions related to this specific problem. Typically, these systems follow common patterns such as:

- Capturing images (normally video acquisition)
- Extracting features (such as colors)
- Detecting blobs
- Provide a counting based on the blobs detected.

For instance, Y.H. Toh et al. [5] present a method of counting feeder fish through image processing techniques. A video of a school of fish is acquired and analysed to output a counting number. At some point filters are applied to background and noise of the image with already identified blobs. The number of fish contained in one image is estimated taking into consideration the individual area of the blobs in one frame and the average number of fish over all frames is calculated. It is relevant to take into account that the setup provided made possible to have a background very easily distinguished from the fish. The tank in the experiment has, apparently, low-level water (which reduces the chance of having overlapping fish), there are no referred restrictions regarding luminosity intensity and it is assumed that the fish are most likely to appear isolated. Considering all the specifications it is possible to count correctly schools of 5, 10, 15 and 50 fish. This approach is not suitable for our project since the tanks that were used in our videos had significant differences from the ones used in Toh et al. experiments such as water depth and volume.

Khanfar et al. [6] suggest the development of algorithms which allow the recognition of fish in the images and track the locations of individual fish from frame to frame. This work takes as an assumption that the majority of the fish tend to move against an approximately stationary background and so, that motion, can be useful in detection. The background is adjusted according to images with no fish. This is used to calculate a histogram for the pixel amplitude used to define thresholds to be set and then isolate regions containing fish. Once it is possible to obtain those regions, an algorithm of edge detection and region growing are used allowing an accurate counting of the number of fish. Several frames provide information for tracking: given an identified fish in a well defined location, it is expectable that, in case of a merging at the next frame in a very close location (for instance another well identified fish), the total area of the merging equals the sum of the area of the fish in it. It is also considered that the motion pattern of fish at a frame may be used to determine if regions are likely to merge in the subsequent frame. The tracking algorithm requires information such as the location of fish, intensity of the region, length, width and area of the fish. With this information it is possible to calculate an Euclidean distance used to associate one region in a frame to a region in the next frame. The procedure enabled the tracking of regions from one frame to the following and then split any new merged regions considering former positions coordinates and that a region travels with constant velocity. This project was relevant in the development of ours since, after analysing this work, we adopted the idea that when blobs merge, the total area of the merging equals the sum of the area of fish in it.

In Fabic et al. [7] work, canny edge detection algorithm is used combined with a coral-blackening background process. Video is recorded in a coral reef environment, which represent a much more complex background than, for instance, the one used by Y.H. Toh et al. The main difference introduced by the project, comparing to the aforementioned references, is that after the detection of each blob, the Zernike moment [8] of every individual blob is calculated having in consideration a standard predefined fish template (depending on the type of fish used in the experiment). Zernike moments allow the mapping of an image using complex Zernike polynomials. The orthogonality between the polynomials makes possible to represent the properties of an image with no redundancy or information overlapping between distinct moments. Despite being dependent on the scaling and translation of the object, given a region of interest, Zernike moments have the advantages of having magnitudes independent of the rotation angle of an object. Hence, they are used to describe shape characteristics of the objects. Thus, in order to identify different fish species and count them in every frame, a set of orthogonal Zernike moments is chosen and applied due to their rotational, translational and scale invariant properties. This work had particular interest in ours since it considers the creation of fish templates with features for each type of fish. This was used as a motivation in our work since we decided to do something similar with fish areas in the different fish categories.

Significant differences regarding the setup environment control are presented by Spampinato et al. [9] where the videos are recorded in open sea and reflect the changes, for instance, in the luminosity and water flowing (background variation). Inherent to the image processing tasks are sub-processing systems consisting of texture and colour analysis, fish detection and fish tracking. The analysis of the statistical moments of the grey-level histogram is the chosen approach to describe mathematically the image texture (e.g. brightness and smoothness). In colour analysis hue, saturation and pixels values are compared to predefined threshold to decide which color one region has in a frame. A moving average algorithm consisting on frame subtractions is used to provide fish detection analysing moving pixels and is made between the background image and the current frame. It is claimed that this particular algorithm has the advantage of giving a good balance between results accuracy and total processing time. On the other hand, in scenes with no static background false positives arose and had to be removed using adaptive gaussian mixture model [10] which modelled each background pixel and classified them based on the likelihood between the real value of the pixel and the one assigned by the Gaussian Mixture Model. Combining the two algorithms, it becomes possible to generate an approximation of the number of fish in a frame by applying a connected component labelling algorithm. Finally the tracking system uses an algorithm based on the matching of blob features and on a pixel histogram matching. This approach motivated the usage of gaussian mixture models in the background subtraction algorithm in our project.

Considering the setup features used in our project, Boom et al. [11] work may have special relevance due to the applied techniques, video environment and results. For instance, regarding fish detection, Gaussian Mixture Model (as presented by Spampinato et al. [9]) which allows dealing with multimodal backgrounds and Adaptive Poisson Mixture Model variant are implemented, as mixture-based algorithms. This procedure models, for each pixel, the distribution of the intensity values typically contained in the background image. However, the computational processing costs tend to increase exponentially as more models are added. Adaptive Poisson Mixture Model is used to handle illumination variations. It is also stated that another algorithm is developed to specifically deal with sudden illumination changes where the reflectance component (static element) of each frame is separated from its illumination component (which is a parameter that varies depending on the light conditions) and the new background model is then computed as a temporal median of these two components. Each pixel has a list of its 20 most recent intensity values and if the value of a pixel in a new frame matches with a high number of values on the list, the pixel is considered background in the new frame. The ViBe algorithm is the responsible for the described verification. All of the above methods are combined and employed by a trained classifier which has no great interest for our project since our algorithm does not include machine learning techniques. Filtering is applied to remove noise and isolate blobs. A postprocessing detection module is used to filter bad detections and reduce false positives through the analysis of each blob by verifying if its shape, texture, motion, structure and segmentation match to expected values from correctly identified fish. Tracking is based on covariance-based models where the template of a fish is represented as the covariance matrix of a set of feature vectors computed for each pixel of the object. Pixel's (x,y) coordinates, RGB values, hue value and the mean of a grayscale histogram are included in each vector. Covariance matrices are compared to decide model similarity using Förstner's distance. The tracking algorithm is connected to the fish detection since a new fish is only tracked if there is the indication that a new one is detected. Tracked fish are located in the scene considering that the search window is based on the fish' speed and direction in the previous frame. Thus, it is possible to calculate candidate regions and compute their covariance matrix. In the end, a new location for a fish is set to the region which, according to Förstner's distance, is most similar to the fish model.

Our project presents different environment and features (such as fish size) comparing to the ones presented in the aforementioned references. Nevertheless, as already stated, the references were used as a motivation to some decisions we made in our project such as the use of background subtraction with gaussian mixture models and the acquisition of the typical area to create a template for each fish category. Combining all these informations we decided to develop an approach specifically for our application.

## III. Implementation

### A. Danio Recording Setup

This work includes the development of a full prototype of a counting device and it can be found in Figure 9. Videos were recorded at Champalimaud Centre for the Unknown Fish Facility and were processed offline. Fish were raised and maintained at the Champalimaud Fish Platform according to Martins et al. (2016) and manipulated by staff accredited for animal experimentation by the Portuguese Veterinary Agency (DGAV)[12].

The container was designed to achieve a parallelepiped structure, made of acrylic, and covered, inside, in blue musgami paper (waterproof). Moreover, given the different physical patterns exhibited by the zebrafish, it was thought that a clean blue environment could bring good image contrast and quality, representing an advantage for video processing. In this way the background subtraction algorithm can rapidly stabilize and, more effectively, allow the detection of fish in the foreground. There are several holes carved on the bottom of the container which are intended to fix each aquarium to it making sure that the distance to the camera is always constant and it does not influence the algorithm output. Blue LEDs are fixed to a moving piece that was designed to fit on the top of each tank, at exactly the same place, that maintains the same light intensity per area in each aquarium since its distance to the tank is always constant.

In this setup, it can be seen in 9, that there is specific hardware selected for both video recordig and user interaction.

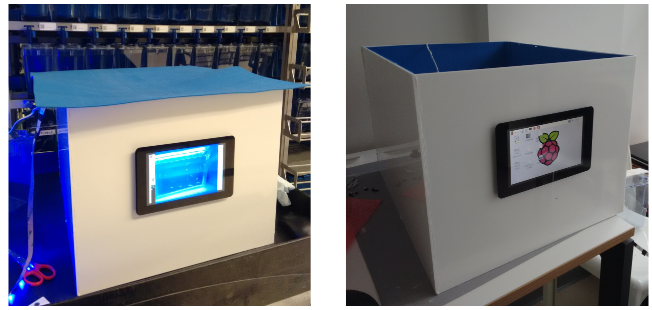
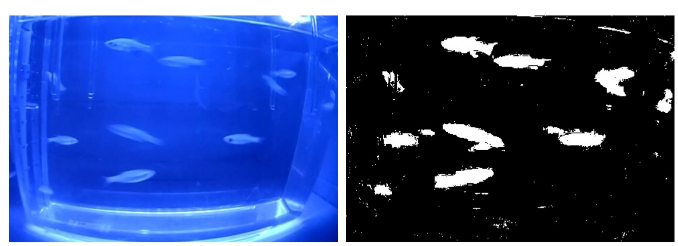


Figure 9. Zebrafish Recording Setup (left: on at the Fish Facility, right: in standby mode).

In this way, to record the videos, a Raspberry Pi 2, Model B [13] is used with an integrated camera [14]. On the front side of the container there is a fixed touch screen which transforms the container into an interactive setup for the user, representing an all-in-one recording system. An example of a frame obtained by the Danio Recording Setup may be found in Figure 10 a).

### B. Estimating the Number of the Fish

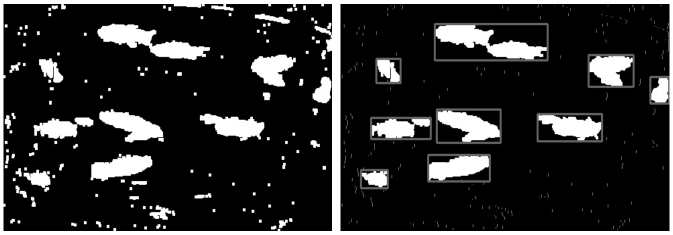
*1) Background subtraction with GMMs:* Regarding video properties, it is relevant to state that each frame has a size of 640x480 and is acquired at a frame rate of 25 Hz. Each frame is initially cropped so that the tank can be the only region of interest in it. Then, an algorithm of background subtraction is applied to the frame with the tank. This background subtraction uses a GMM background subtraction scheme [10] [15] that automatically selects specific components, for each pixel in the image. This makes the differentiation between what is considered as background (blue environment) and foreground (fish moving between successive frames) possible. In Figure 10 b) we can found an example of a frame after applying background subtraction. This technique allows adaptability during frame variation to guarantee that the background is consistently subtracted and only foreground variations are able to be identified. The resulted black and white image, where black represents the background and white the fish (blobs from now on), is noise filtered through dilation and selection of blobs above an area threshold to remove the tiny white blobs and obtain a clean background subtracted frame to analyse.



a) b)

Figure 10. Example of frame acquired with the Danio Recording Setup and cropped at the beginning of the algorithm, a) and after backgroung subtraction, b).

*2) Blob Counting:* After the dilation and noise filtering, Figure 11 a) and b), a routine for blob contours detection occurs as well as the calculation of each blob pixel area (inside the contour). In case a blob has an area larger than



a) b)

Figure 11. Example of frame after background subtraction and dilation, a) and the result after noise filtering, b).

a given threshold (different for fries and adult fish), it may represent fish or multiple fish together in a big blob. Each detected blob contour larger than the threshold is counted representing the first counting approximation. Afterwards, the decision whether a blob represents a fish or multiple fish is made taking into consideration testing of videos previously performed (different from the ones used to extract results from the counting algorithm). These videos contain only one fish (fry, or adult, in each genotype) and are used to register the average area (reference area from now on) that the fish assumes during a five minute video (7500 frames). Then, each blob area is divided by the reference area multiplied by two (the reference is two fish overlapping) giving a correction number for the counting done so far. Different reference areas, bigger or lower than the calculated in the tests, are applied as the number of fish increases in the tank. In fact, after testing it was possible to build intervals with typical average areas that specific amounts of fish represent and, this way, if a frame has a total blob area in one of this intervals, a specific reference area is used. It is understandable that, for instance, if the number of fish is 35 (maximum number of fish per tank in this project), a blob with the same size as a case where there are 25 fish may represent more fish. This justifies why the reference area used to divide the blob area by is lower than twice the reference area calculated in the tests. The reference area is decreased by 100 pixel for each area interval. In cases where shoaling at the bottom of the tank is detected (by checking if the total blob area in the bottom of the tank is equal or higher than 90% of the total blob area detected in the frame) the reference area used is typically half of the reference area given by the tests because higher compensation needs to be done.

*3) Mirrored Counting Compensation:* In the recorded videos two areas containing fish mirroring can be detected: the right side of the tank and the upper part (water surface). Near these regions fish are reflected like in a mirror which may lead to count some fish twice. Thus, the method used for mirror compensation is calculating each Hu moment for each blob based on Ming-Kuei Hu [16] work that presents a theory of two-dimensional moment invariants for planar geometric figures. There, absolute orthogonal invariants are derived which are used in our project to get the pattern identification of similar shape independent of size, position and orientation. In this way and since a reflected fish is

| | **Fish Number** | | | | | |
|---|---|---|---|---|---|---|
| **Fish Category** | **5** | **7** | **10** | **12** | **15** | **17** |
| AB/TU Fries | 4% | 13% | 7% | 10% | 4% | 9% |
| Nacre Fries | 14% | 12% | 15% | 6% | 8% | 6% |
| AB/TU Adults | 0% | 6% | 5% | 3% | 9% | 11% |
| Nacre Adults | 21% | 13% | 18% | 12% | 29% | 6% |

Table II
AVERAGE FISH COUNT ERROR FROM 5 TO 17 FISH IN THE DIFFERENT CATEGORIES (20 SAMPLES PER CATEGORY) OBTAINED IN THE FINAL SOLUTION.

similar in size and has opposite orientation comparing to the original fish, it is possible to identify which fish is producing a reflection and to compensate this counting by subtracting the number of mirrored fish to the overall counting that was made until this step.

## IV. Results and Discussion

### A. Results - Final Solution

To begin with, it is important to state that videos with fish were recorded at Champalimaud Centre for the Unknown Fish Facility and were processed offline. Fish were raised and maintained at the Champalimaud Fish Platform according to Martins et al. (2016) and manipulated by staff accredited for animal experimentation by the Portuguese Veterinary Agency (DGAV)[12].

For each of the categories described at the beginning of subsection I-A, we collected videos from tanks containing schools of 13 different sizes: 5, 7, 10, 12, 15, 17, 20, 22, 25, 27, 30, 32 and 35 fish except AB/TU fries as previously stated.

For each different category and school size we collected 20 videos of 40 seconds each, thus totalling 200 videos for AB/TU fries and 260 for each of the remaining categories.

The average error resulting from the counts can be found in Tables II and III where acceptable average errors ($\leq 15\%$) are indicated in green and not acceptable in red. Let us recall that the estimated number of fish detected in each frame of a video is denoted by: $f_m^n$, $n = 1, ..., N$ and $m = 1, ..., M$, where N and M are the number of videos and number of frames per video, respectively. The number of fish is given by: $\hat{F}_n = median(f_1^n, f_2^n, ..., f_M^n)$. The error in a single video is expressed by:

$$e_n = \frac{|\hat{F}_n - K_n|}{K_n} \tag{1}$$

where $K_n$ represents the real number of fish, which is known *a priori* and the average error for $N$ videos is calculated using

$$\bar{e} = \frac{\Sigma_{n=1}^{N} e_n}{N}. \tag{2}$$

where in this case $N = 20$ for each entry of Tables II and III.

Regarding fries' category, we can see that the highest error observed was 15%, for Nacre fries in a 10 fish tank. For the

| | Fish Number | | | | | | |
|---|---|---|---|---|---|---|---|
| **Fish Category** | **20** | **22** | **25** | **27** | **30** | **32** | **35** |
| AB/TU Fries | 7% | 9% | 4% | 3% | | | |
| Nacre Fries | 11% | 3% | 4% | 5% | 1% | 8% | 4% |
| AB/TU Adults | 14% | 3% | 5% | 24% | 15% | 12% | 29% |
| Nacre Adults | 9% | 20% | 9% | 11% | 7% | 3% | 4% |

Table III
AVERAGE FISH COUNT ERROR FROM 20 TO 35 FISH IN THE DIFFERENT CATEGORIES (20 SAMPLES PER CATEGORY) OBTAINED IN THE FINAL SOLUTION.

remaining schools of fish acceptable errors were obtained. Since fries are significantly smaller than adults, overlapping is less frequent and cases where the error is higher (than 15%) correspond to partial shoaling. Partial shoaling in a video sample may be defined as the occurrence of shoaling only in portions of a video in such way that there are enough frames where shoaling is not detected leading to an estimate in the acceptable margin.

### B. Shoaling

In the adult fish category, there were certain fish quantities which did not meet the acceptable margin. For instance, in the videos with 15 Nacre adult fish the average error is 29% which means that, in average, approximately 5 fish were not detected in those videos which is far from the real value. This is due to the significant high number of videos where shoaling occurred. As previously mentioned, fish shoal particularly in the bottom of the tank and it is extremely difficult for the algorithm to output correct counts when this behaviour is verified. This occurs for videos with few or many fish when fish are closely together in multiple layers behind each other. Hence, shoaling justifies the average errors at Tables II and III that are higher than the 15% margin.

In Figure 12 representing 22 adult fish, we can see that for videos where shoaling occurs, the percentage of total blob area detected in each frame is significantly lower than the case in Figure 13 where shoaling does not occur. Thus, it is understandable that if there is less blob area than it should in a frame, it will lead to unacceptable results.

### C. Algorithm Convergence

We can see examples of how the fish count per frame varies within individual videos for fish in the different categories from Figures 14 to 17. In those figures, the real fish value inside the tank is represented by the green line and $\pm 15\%$ of the real fish number correspond to the red lines. It can be verified that the algorithm outputs a number within the error margin in early frames and maintains the error margin until 1000 frames. Nacre zebrafish tend to reflect more the light, resulting in bigger blobs for single fish, comparing to the AB/TU category, since these are darker, which explains the multiple peaks in Figure 17. This peaks are not so evident

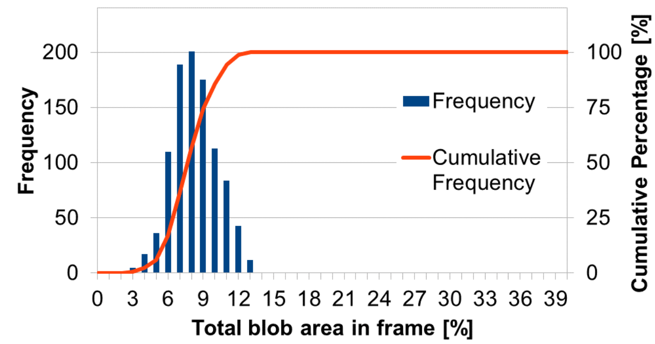



Figure 12. Frequency and cumulative frequency histogram for a video with 22 fish shoaling.

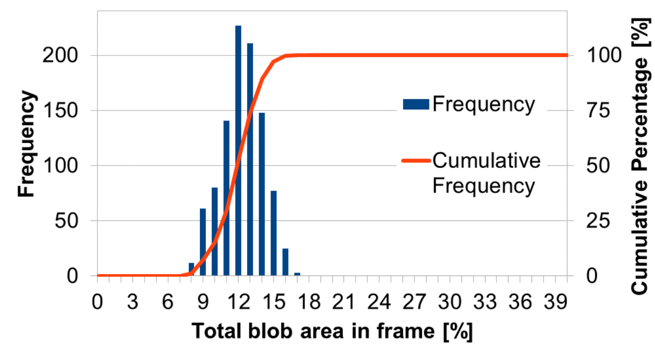



Figure 13. Frequency and cumulative frequency histogram for a video with 22 fish without shoaling.

in Nacre fries, Figure 16, due to their small dimensions which, in the ends, compensates the reflection.

From Figure 18 to 21 we can analyse the variation of the variance for each school of fish in the different categories.

Regarding Nacre, Figure 18 and AB/TU fries, Figure 19, categories, we can easily verify that the variance values are very low. This means that the algorithm presented results which were very close to the average error in the 20 different video samples for each school number. Taking into consideration that all the values in Tables II and III for fries are within the acceptable margin and the error variances are very low, we can conclude that the algorithm performs well in this category. Moreover, if we analyse the regression lines in red, we can verify that the line slope is positive. This demonstrates that the error's variance tends to increase with the with the increase of the number of fish inside a tank as expected.

In Nacre and AB/TU adult categories, there are some error's variance values that are significantly greater than the ones in fries categories. This is due to the fact that adult fish have bigger size which makes overlapping occur more often leading more frequently to situations where shoaling and partial shoaling occur. Hence, in some trials (27 AB/TU adult fish, for instance), these behaviours were more intense than in others leading, for example, to high error values or even no error in some video samples. These variations lead to the increase of the variance value. There were also cases in which shoaling occurred similarly in all the video samples, as in 35 Nacre adult category, leading to an average error very close

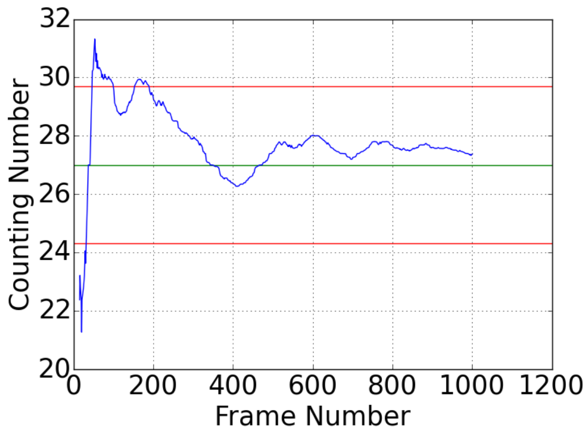



Figure 14. Counting number for each frame for 27 AB/TU zebrafish fries in a 40-second video (1000 frames). The red lines represent the acceptable error margin and the green line is the real fish number in the tank).

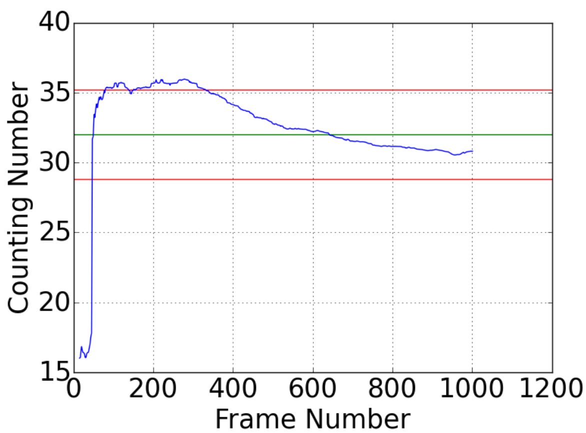



Figure 15. Counting number for each frame for 32 adult AB/TU zebrafish in a 40-second video (1000 frames).

to the (wrong) counting values obtaining, consequently, low variance. It was once more possible, through the analysis of the regression lines, to verify that the slope is positive representing the same tendency observed in fries categories (variance tends to increase with the increase of the number of fish inside the tank).

### D. Mirroring Compensation

In order to emphasize the importance of the fish mirroring compensation (i.e., subtracting blobs representing fish reflections that were previously counted as fish), we present, in this section, the results of the algorithm, over the same samples as in the previous section, without the calculation of Hu moments (mirroring identification). The counting results may be found in Tables IV and V.

When performing fish counting with the final version of the algorithm, we can obtain an estimate of the counting equal, above or below the real number of fish inside the tank. If the estimate, in a frame, is lower than the real value of fish, the overall counting estimate value tends to decrease and the error

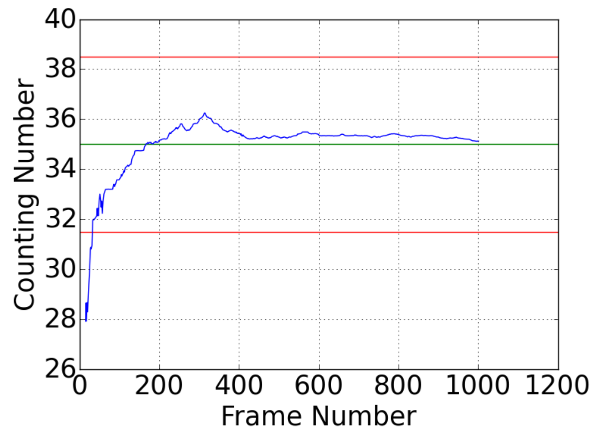



Figure 16. Counting number for each frame for 35 Nacre zebrafish fries in a 40-second video (1000 frames).

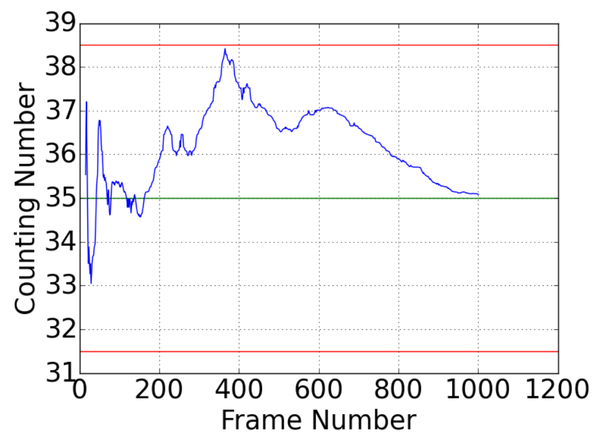



Figure 17. Counting number for each frame for 35 adult Nacre zebrafish in a 40-second video (1000 frames).

| | Fish Number | | | | | |
|---|---|---|---|---|---|---|
| **Fish Category** | **5** | **7** | **10** | **12** | **15** | **17** |
| AB/TU Fries | 4% | 16% | 9% | 16% | 11% | 12% |
| Nacre Fries | 14% | 11% | 17% | 8% | 6% | 11% |
| AB/TU Adults | 0% | 13% | 17% | 10% | 12% | 18% |
| Nacre Adults | 21% | 13% | 18% | 17% | 29% | 18% |

Table IV
AVERAGE FISH COUNT ERROR FROM 5 TO 17 FISH IN THE DIFFERENT CATEGORIES WITHOUT MIRRORING COMPENSATION (20 SAMPLES PER CATEGORY).

of the estimate increases when we compensate the blobs that represent fish mirroring. On the other hand, if before mirroring compensation the estimate of the counting is greater than the real value, when we subtract the mirroring counts, the value of the estimate tends to decrease and converge to the real number of fish inside the tank. Having this into consideration, we can see in Tables IV and V how the error varies, when comparing to Tables II and III, if the mirroring compensation

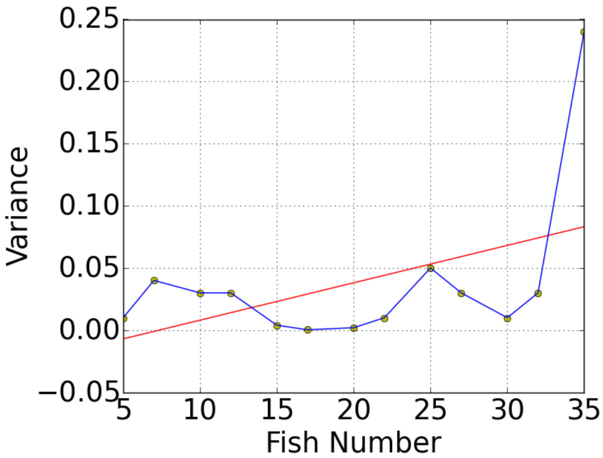



Figure 18. Plotted variance values and linear regression for each school of Nacre fries.

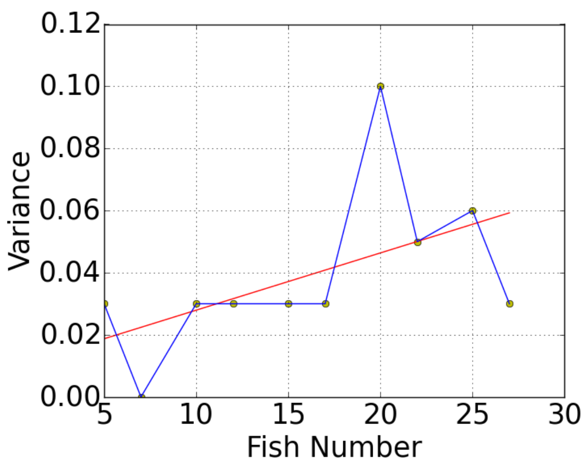



Figure 19. Plotted variance values and linear regression for each school of AB/TU fries.

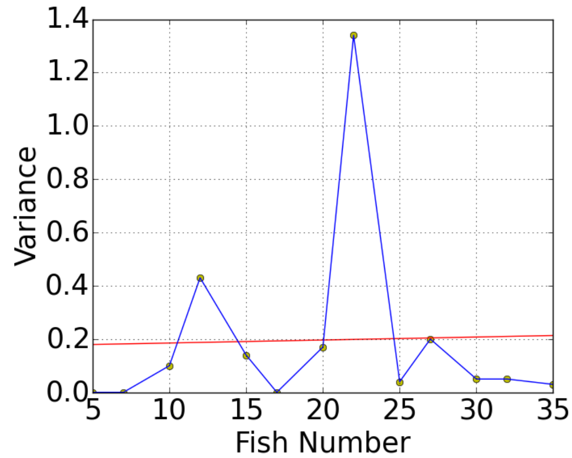



Figure 20. Plotted variance values and linear regression for each school of Nacre adults.

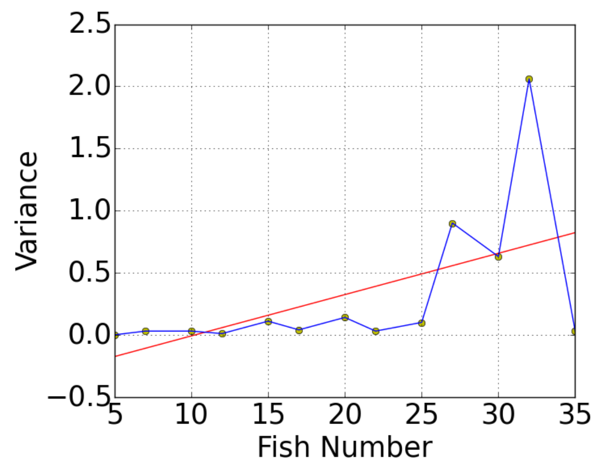



Figure 21. Plotted variance values and linear regression for each school of AB/TU adults.

| | **Fish Number** | | | | | | |
|---|---|---|---|---|---|---|---|
| **Fish Category** | **20** | **22** | **25** | **27** | **30** | **32** | **35** |
| AB/TU Fries | 18% | 13% | 9% | 7% | | | |
| Nacre Fries | 13% | 7% | 10% | 12% | 9% | 12% | 5% |
| AB/TU Adults | 21% | 9% | 11% | 24% | 17% | 13% | 29% |
| Nacre Adults | 16% | 20% | 13% | 12% | 15% | 9% | 7% |

Table V
AVERAGE FISH COUNT ERROR FROM 20 TO 35 FISH IN THE DIFFERENT CATEGORIES WITHOUT MIRRORING COMPENSATION (20 SAMPLES PER CATEGORY).

was not performed. In fact, we can easily verify that the last two tables we have many more situations where the 15% error margin was not achieved. For instance, for 20 AB/TU adults we obtained 14% error margin in Table III, when mirroring compensation is done and in Table V we can see that the error increases to 21% representing a value outside the margin.

We can also verify that, for instance in 7 and 15 Nacre fries, the error margin is lower in Table IV. This means that before compensation, the count estimate was below the real number of fish. Only in this situation can the average error in Tables IV and V be lower than in Tables II and III.
It is also important to verify that the cases which we identified as shoaling in Table III, for instance 27 AB/TU adults and 22 Nacre adults, present the same value in Table V whether compensation occurs or not. This is due to the fact that the blobs that represent the shoaling occurred mostly at the bottom of the tank where no reflections are detected. Moreover, since blobs that results from shoaling present, generally, very distinct shape characteristics among each other, the first condition in Hu moments comparison (shape features in the first orthogonal invariant) is not verified leading to a situation where reflections can not be identified.

### *E. Algorithm Time Performance*

In this section we present examples of execution time of the algorithm in videos with different quantities of zebrafish.

It is possible to find those values in Table VI where times for adults, Table VI a) and fries, Table VI b), are represented.

| Fish Number | Execution Time[s] |
|---|---|
| 5 | 18,2 |
| 7 | 26,4 |
| 10 | 25,5 |
| 12 | 24,4 |
| 15 | 25,2 |
| 17 | 29,5 |
| 20 | 32,4 |
| 22 | 31,9 |
| 25 | 36,6 |
| 27 | 37,8 |
| 30 | 38,1 |
| 32 | 41,3 |
| 35 | 44,9 |

a)

| Fish Number | Execution Time[s] |
|---|---|
| 5 | 14,1 |
| 7 | 15,2 |
| 10 | 17,4 |
| 12 | 17,7 |
| 15 | 17,9 |
| 17 | 19,5 |
| 20 | 23,3 |
| 22 | 23,7 |
| 25 | 24,7 |
| 27 | 28,6 |
| 30 | 29,3 |
| 32 | 28,1 |
| 35 | 28,7 |

b)

Table VI
EXAMPLES OF THE AMOUNT OF TIME (IN SECONDS) TO RUN THE ALGORITHM IN VIDEOS WITH DIFFERENT NUMBER OF NACRE ZEBRAFISH ADULTS, A) AND FRIES, B).

It is important to state that this examples were obtained in a single 40-second video, for each number of fish, using an Intel Core i7-4500U CPU @ 1.8GHz - 2.4GHz processor and 8GB of RAM.

We can easily verify that this offline video processing never takes more than approximately 45 seconds. That value is verified in the (Nacre) adults category, as expected, since adult fish are represented by bigger blobs which demands more time to process.

It is interesting to compare these values to the ones present in Table I where manual counting times for zebrafish fries are presented. Fries' counting is more time consuming comparing to adults counting due to the fish' size. If we compare the two tables, we can see that in Table I most of the counting times exceed or is approximately equal to 30 seconds and the decrease in the number of fish does not always traduce the decrease in the amount of time. Using the information in both tables, we can, finally, conclude that the time our algorithm takes to run a 40-second video (1000 frames) is acceptable taking into consideration the manual counting time for fries. Even though adult fish tend to be easier to count manually than fries, we can also consider the execution time for adults as acceptable since the longer our algorithm takes to output an estimate is around 45 seconds for 35 fish (only 15 seconds more than the maximum obtained for the fries).

Other processing was made using an Atom where a 40-second video took approximately 80 seconds to be analysed and a Raspberry Pi outputted an estimate after approximately 2 minutes.

Finally, it is important to refer that with this solution there is no need to manipulate the fish, then, less stress will be induced due to manual handling.

## V. CONCLUSION AND FUTURE WORK

After the development of this project, we were able to present a noninvasive technique for zebrafish count in fish facility tanks. It was possible to develop and deliver the production of a full recording setup prototype as well as the counting software for fish number estimation.

Regarding the recording prototype, we were able to produce a solution that allows the recording of videos always at the same conditions (for instance constant luminosity and distance to the camera) which is very important to guarantee the repeatability of the process.

We were able to implement computer vision techniques and mathematical tools for mirroring compensation during frames analysis, in the algorithm developed in this project, in order to output an estimate of the number of fish in a tank. At a first stage of this project, including an optical flow technique in the algorithm was considered. However, after realizing that fish are not detected when their motion is low and the fact that optical flow implicates more calculations for each frame and, consequently, more processing time, we decided to drop this idea. Hence, since the algorithm has in its base the counting of blobs resulting from background subtraction, information regarding the average area that one fish, for the different categories, represented in 7500 frames was collected. Using this information as a reference, we were able to build area intervals for different fish quantities and use that area to develop a fish overlapping compensation method. For the two different phenotypes, Nacre and AB/TU zebrafish reflect light differently since the latter are physically darker than the first which lead to the usage of different parameters for blob overlapping compensation.

Regarding performance, the algorithm demonstrated very good results particularly in fish fries categories, where the average error was always in the acceptable margin defined at the beginning of the project and the values for the variances of the error in the different video samples were very low.

In some cases, particularly in videos with adult fish, shoaling occurred and affected significantly the algorithm performance. In fact, the cases where higher errors were verified correspond to samples where shoaling occurred during almost the entire video sample or partial shoaling could be observed in a large number of frames. Due to the significant difference in size between fries and adults, overlapping, thus shoaling, tend to occur more often when comparing to the fries categories. As a matter of fact, all the results above the error margin were verified in the adults category. Nevertheless, we could conclude that in cases where shoaling does not occur, this algorithm does also demonstrate acceptable results for adult zebrafish.

After analysis of graphs such as in Figure 14, we could observe that after a few hundreds of frames the counting estimate was already inside the acceptable error margin which indicates that we may not need 1000 frames (40 seconds) to output an acceptable estimate.

Considering an approach with only one camera as in our project and since this algorithm does already identify shoaling behaviour, frames where shoaling occurred could be skipped and not used in fish estimation. In this way we would expect to have lower error since partial shoaling and shoaling would not affect the counting. However, the disadvantage of this process would be the extra waiting time until we had enough frames without shoaling to provide an acceptable estimate.

Another relevant information that could be used to reduce the error in adults counts is previous counting records. Since at this Fish Facility zebrafish are not generally added to tanks as time passes, we could take advantage of the good results obtained in fries categories and use the counting information obtained after the first 30 days (fries) to limit the maximum number of fish that are likely to be in the tank after that time (90 days, adults). However, this would be useful to obtain an upper limit but would not, obviously, solve the shoaling issue.

This algorithm could also be implemented with two cameras: one as used in this project, and another recording on top of the tank. In this way, we could use the shoaling behaviour identification and the frame obtained from the top camera to give more accurate counting because we would have information regarding the depth of the shoal.

Another interesting and useful approach would be to use background subtraction with features detection and extraction algorithms to identify, in each tank:

- how many males or females exist
- in tanks with more than one phenotype understand how many fish exist
- study the relationship between the biomass in each tank (which would implicate fish weighing) and the average blob areas of fish in videos.

It would also be extremely useful to:

- design an implementation of the algorithm using multi-threading techniques to reduce execution times in devices such as the Raspberry Pi
- measure the average time in adult zebrafish manual counting
- perform experiences in order to evaluate the manual error counting
- perform an in depth study to evaluate, quantify and compare the stress caused to fish during manual counting and automatic counting.

## References


[1] R. Szeliski, *Computer vision: algorithms and applications*, ser. Texts in computer science. London: Springer, 2011, page 5.

[2] G. Farnebäck, "Two-frame motion estimation based on polynomial expansion," *Proceedings of the 13th Scandinavian Conference on Image Analysis*, pp. 363–370, 2003.

[3] K. Stoletov and R. Klemke, "Catch of the day: zebrafish as a human cancer model," *Oncogene*, vol. 27, no. 33, p. 4511, Jul. 2008, accessed: 2016-04-09. [Online]. Available: http://www.nature.com/doifinder/10.1038/onc.2008.95

[4] Tecniplast, "Tecniplast ZebraFish Tank," accessed: 2016-04-13. [Online]. Available: http://www.scanbur.com/fileadmin/News/Educational_tour/ZebTEC_PPT.pdf

[5] Y. Toh, T. Ng, and B. Liew, "Automated fish counting using image processing," in *Computational Intelligence and Software Engineering, 2009. CiSE 2009. International Conference on*. IEEE, 2009, pp. 1–5.

[6] H. Khanfar, D. Charalampidis, E. Yoerger, J. Ioup, and C. Thompson, "Automatic Fish Counting in Underwater Video," *66th Gulf and Caribbean Fisheries Institute*, pp. 1–9, Nov. 2013.

[7] J. N. Fabic, I. E. Turla, J. A. Capacillo, L. T. David, and P. C. Naval, "Fish population estimation and species classification from underwater video sequences using blob counting and shape analysis." IEEE, Mar. 2013, pp. 1–6.

[8] H. A. Qader, A. R. Ramli, and S. Al-haddad, "Fingerprint recognition using zernike moments," 2006.

[9] C. Spampinato, Y.-H. Chen-Burger, G. Nadarajan, and R. B. Fisher, "Detecting, Tracking and Counting Fish in Low Quality Unconstrained Underwater Videos," *Proc. 3rd Int. Conf. on Computer Vision Theory and Applications (VISAPP)*, pp. 514–519, 2008.

[10] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction." IEEE, 2004, pp. 28–31 Vol.2, accessed: 2016-02-16. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1333992

[11] B. J. Boom, J. He, S. Palazzo, P. X. Huang, C. Beyan, H.-M. Chou, F.-P. Lin, C. Spampinato, and R. B. Fisher, "A research tool for long-term and continuous analysis of fish assemblage in coral-reefs using underwater camera footage," *Ecological Informatics*, vol. 23, pp. 83–97, 2014.

[12] M. S, J. Monteiro, M. Vito, D. Weintraub, J. Almeida, and A. Certal, "Toward an Integrated Zebrafish Health Management Program Supporting Cancer and Neuroscience Research," 2016.

[13] Pi Foundation, "Raspberry Pi 2 Model B," 2015, accessed: 2016-03-06. [Online]. Available: https://www.adafruit.com/pdfs/raspberrypi2modelb.pdf

[14] Pi Foundation, "Raspberry Pi Camera," 2013, accessed: 2016-04-13. [Online]. Available: https://www.raspberrypi.org/documentation/hardware/camera.md

[15] Z. Zivkovic and F. van der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," *Pattern Recognition Letters*, vol. 27, no. 7, pp. 773–780, May 2006, accessed: 2016-02-16. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0167865505003521

[16] Ming-Kuei Hu, "Visual pattern recognition by moment invariants," *IEEE Transactions on Information Theory*, vol. 8, no. 2, pp. 179–187, Feb. 1962, accessed: 2016-02-10. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1057692