Robotica (2003) volume 21, pp. 233–243. © 2003 Cambridge University Press
DOI: 10.1017/S0263574702004733 Printed in the United Kingdom

Computer vision and robotics techniques in fish farms

J. R. Martinez-de Dios, C. Serna and A. Ollero

Depto. Ingeniería de Sistemas y Automática, Universidad de Sevilla,Camino de los Descubrimientos s/n, 41092, Sevilla (Spain). Email: {jdedios, cmar, aollero}@cartuja.us.es

(Received in Final Form: November 9, 2002)

SUMMARY

This paper presents new low-cost systems for the automation of some fish farm operations. Particularly, computer vision is applied to non-contact fish weight estimation. Stereo vision systems with synchronised convergent cameras are employed to perform fish 3-D segmentation in tanks and sea cages. Several pre-processing algorithms are applied to compensate for illumination local variations. The approach applied for fish 3-D segmentation consists in detecting in both images certain fish features. Once these points have been detected and validated in both images, the fish are 3-D segmented by applying stereo vision matching considerations. Fish weight is estimated by using simple length-weight relations well known in the aquaculture domain. The paper also briefly describes robotics systems for fish feeding and underwater pond cleaning, which can be also used to implement the above mentioned computer vision techniques for the fish estimation.

KEYWORDS: Computer vision; Fish farm automation; Aquaculture; Image processing; Cost-oriented automation; Mobile robotics systems.

1. INTRODUCTION

Aquaculture has experienced a dramatic growth in the last years, with a growing rate of near 15% per year. According to the UN Food and Agriculture Organisation (FAO), aquaculture produces about 20 million of fish food tons per year in about 200,000 ponds or cages. This amount represents about 20% of the world fisheries requirement. However, the fish-farm production technology is clearly underdeveloped, still very far from the state of the art in factories of manufactured products or even from many food production processes.

The work presented in this paper is part of a project aiming to improve the production processes in fish farms using new perception, control and automation technologies. The project involves all the phases in the fish growing process including sizing, quality control, grading and the control of the variables involved in the growing. Thus, a distributed control system has been developed and implemented to control all the relevant variables involved in the fish farm process.[1]

Biomass estimation is of great interest in aquaculture with a direct impact on the optimisation of the production management, process automation and monitoring and quality control.

Traditional techniques to estimate the fish biomass in ponds under thermal cover, in large earthen ponds and sea cages involve manual sampling and weighting. However, the minimisation of the fish handling operation is highly desirable to reduce fishes stress and increase the quality of the product. Then, special attention is paid to the estimation of the number and size distribution of the fish using non-manual and non-intrusive procedures.

Sensors and techniques to estimate the biomass are still very scarce. There are some devices such as submersible frames that count and estimate the size of the passing fish using optical techniques (biomass counters or biomass estimators). Automatic systems using computer vision for monitoring the fish passage through special devices in rivers have also been presented.[2,3] It should be noted that most fish species are very reluctant to pass through artificial devices, which increases their level of stress. Furthermore, measurements of the fish passing through these devices could not be representative of the whole fish population in the cage.

In the last years, techniques using underwater video monitoring have also been proposed by several authors in the aquaculture domain.[4–6] However, many of these systems still operate with a low level of automation and require intense activity from operators.

Automatic fish segmentation and measurement requires the computation of 3-D variables, such as the distance from the fish observed, fish orientation and size. Thus, 3-D computer vision techniques should be applied. However, the cost of industrial 3-D computer vision systems is still high for many fish farms and cannot be easily applied due to difficulties in fish-farm environment.

Moreover, the hardware and software computational requirements of the system have been minimised by avoiding complex 3-D fish models. The proposed systems use simplified models based on fish features, and existing relations between the fish length and its weight. This approach has significant advantages for its simplicity, modularity and flexibility to adapt to different fish species. Besides, it needs lower computer requirements than other segmentation methods based on 3-D models or active contours.

This paper presents several low cost systems for fish farm automation. First, the paper describes stereo vision systems that provide automatically fish weight estimations (including fish average weight and weight histogram) in tanks and in sea cages. Prototypes of these systems have been tested in several fish farms including ponds with thermal cover in Ayamonte (Spanish province of Huelva)

and sea cages in the Spanish Mediterranean coast of Valencia and Murcia. The paper also introduces some other low-cost systems for fish-farm automation including a robotic system for fish feeding and an underwater robot for autonomous pond cleaning. The stereo vision systems are ready to be implemented in these robotic systems.

The paper is organised as follows: Section 2 presents the general description of the stereo vision systems; Section 3 presents the calibration and stereo matching procedures; Section 4 describes the techniques employed to estimate fish weight by using computer vision techniques; Section 5 presents some experimental results. Section 6 presents some low-cost robotic applications in fish farms. Sections 7 and 8 refer to the Conclusions and Acknowledgements, respectively.

2. DESCRIPTION OF THE SYSTEM

Figure 1 shows a general scheme of the stereo systems for fish weight estimation presented in this paper. Two stereo systems have been developed: underwater and on the water surface.

One of the main disadvantages of using non-underwater cameras is the high number of reflections on the water surface.[7] These solar reflections often cause camera saturation, which avoids the correct performance of the image-processing algorithms. These reflections can be compensated in indoor scenarios (ponds with thermal covers and inside buildings) by applying artificial lighting. However, in outdoor environments they are very difficult to compensate. Thus, an underwater stereo system is used to estimate the weight of fish in sea cages, and a system over the water surface is used for indoor tanks. The first one is employed for fish at adult state of growth and, the second is used for fish nursery (intermediate stage of growth).

The underwater system incorporates cameras with methacrylate-made housings. It also has floats and lead ballast units to maintain the stability of the structure.

Both stereo systems employ two convergent monochrome cameras. The cameras have been synchronised in order to reduce errors in the stereo system and calibration. The images from both cameras are captured by a RGB frame-grabber configured in such a way that the left image is introduced via the red component of the color image and the right view through the green one. The result is a RGB color image with an empty blue component and the two images of the stereo system in the red and green components.

Both stereo systems should be calibrated in order to compute fish length estimations in the real world coordinates. Section 3 describes the calibration method and the stereo matching technique employed.

Sea cages and tanks are highly unstructured scenarios from an image-processing point of view. These images suffer from effects such as multiple object overlapping, high level of noise and presence of unwanted objects such as bubbles. Illumination irregularities were shown as one of the most harmful effects for the automatic image analysis, both for underwater and non-underwater images. Other specific drawbacks and limitations are imposed by the particular nature of fish, such as reflections on fish scales and high speed of motion. It should be noticed that the use of artificial means to mitigate these drawbacks presents limitations due to fish nature and behavior. For the underwater stereo vision system, a white metallic panel to be part of the image background was incorporated to increase the fish contrast in the images. Artificial lighting

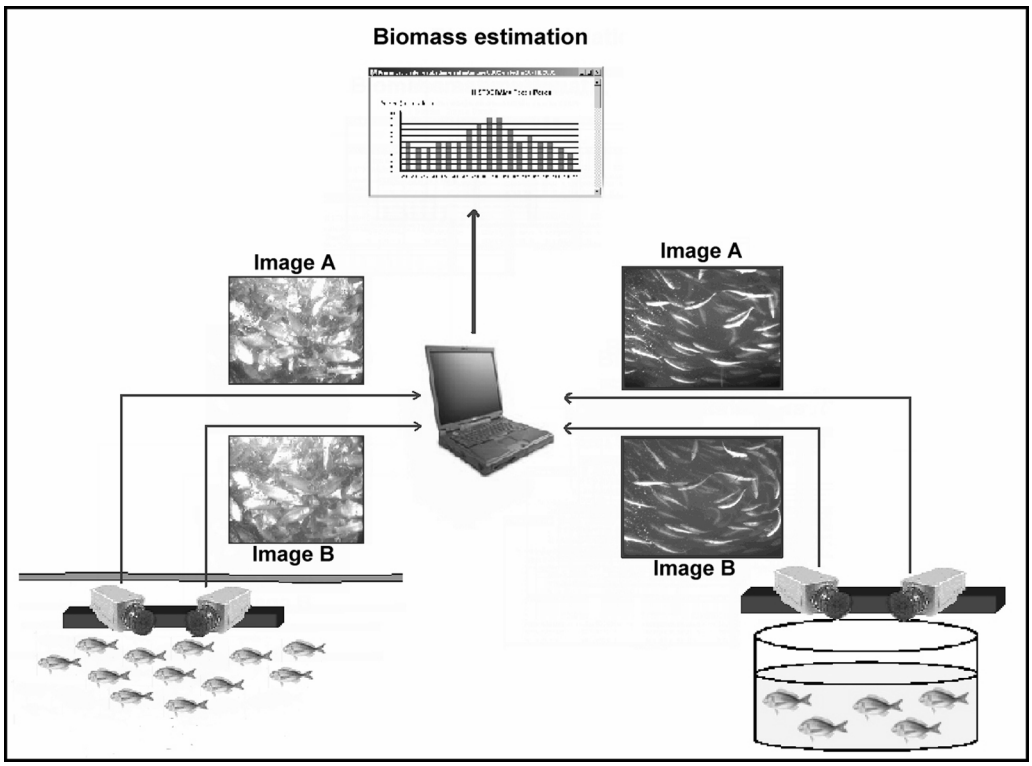


Fig. 1. General scheme of the stereo system for fish weight estimation.

https://doi.org/10.1017/S0263574702004733 Published online by Cambridge University Press

was also incorporated to compensate for illumination variations. Both ideas were soon discarded since the increments in fish stress and their extreme reluctance to approach the system. All these drawbacks make fish automatic segmentation a very complex problem.

In order to cope with these problems fish segmentation is performed at three levels: image pre-processing; image processing in both independent images; and stereo processing. The approach used consists in detecting certain fish points of interest, and then associating the points of interest of the same fish in real-world co-ordinates according to information on the fish specie and application knowledge. Other segmentation methods based on fish 3-D models have been investigated. However, the complexity of these procedures involves significant computer requirements that are not always justified for the application.

The first step consists in applying noise reduction filters and illumination compensation techniques. The second and third steps of processing consists in applying a set of methods in order to detect several fish points of interest, such as caudal fins and other features. The detection of these points is performed independently in both images. Then, stereo information is employed to match these points and locate them in the real world. The following step consists in associating the points of interest of the same fish by considering criteria based on geometrical aspects of the fish species.

Once the fish have been segmented and their size has been estimated, the weight is measured by using length-weight relations. The validation tests obtain weight errors lower than 5% for the underwater system and 4% for the system with cameras over the water surface.

3. CALIBRATION AND STEREO MATCHING

The application requires a stereo vision perception system to obtain measurements in the real world of objects located at different distances from the cameras. Figure 2a shows a simplified scheme of the stereo vision system with convergent cameras. The optical axes of both cameras have been disposed to converge at a distance of 1 meter, which has been determined, by means of extensive experimentation, as the most appropriated distance for the application. The distance between cameras has been selected to obtain the required accuracy of the stereo system.

The method proposed in Ayache[8] has been used to calibrate the stereo system. Although the execution of the calibration algorithm requires at least six non-coplanar calibration points, in practice, dozens of points are used to increase precision. In this application the aluminium-made calibration grid has two planes with a total of one hundred points distributed at different heights (see Figure 2b). The calibration grid measured by using high-precision machinery achieved an accuracy lower than 0.1 mm.

3.1. Calibration

The first step of image calibration consists in computing the perspective matrix for each camera. Assuming a simple pinhole camera model, consider that $I(u,v)$ is the intersection of the projection of a point in the space $P(x,y,z)$ on the image plane. The transformation between I and P can be modelled as a linear transformation T in projective co-ordinates:

$$I^* = \begin{pmatrix} U \\ V \\ S \end{pmatrix} = T \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}, \tag{1}$$

where $I^*(U,V,S)$ is $I(u,v)$ expressed in projective co-ordinates and T is the perspective matrix. In the general case, $S \neq 0$ and the co-ordinates of I can be computed as $I=[u\ v]^t=[U/S\ V/S]^t$. Each point on the image $I(u,v)$ corresponds to a point in the real world P, which originates two linear equations:

$$\left.\begin{aligned} P^t t_1 + t_{14} - u(P^t t_3 + 1) = 0 \\ P^t t_2 + t_{24} - v(P^t t_3 + 1) = 0 \end{aligned}\right\}, \tag{2}$$

where $t_{i,j}$ represents the (i, j) of T, t_i is the i-th column of T. T can be determined by solving the set of equations by applying simple least-squares methods.[8] Although the computation of T only requires six non-coplanar points, all the points of the calibration grid are considered to increase the precision. Stereo applications require the computation of the perspective matrix for both cameras.

3.2. Stereo matching

Assuming that the cameras are denoted as *camera 1* and *camera 2*, the problem of stereo matching, which consists in

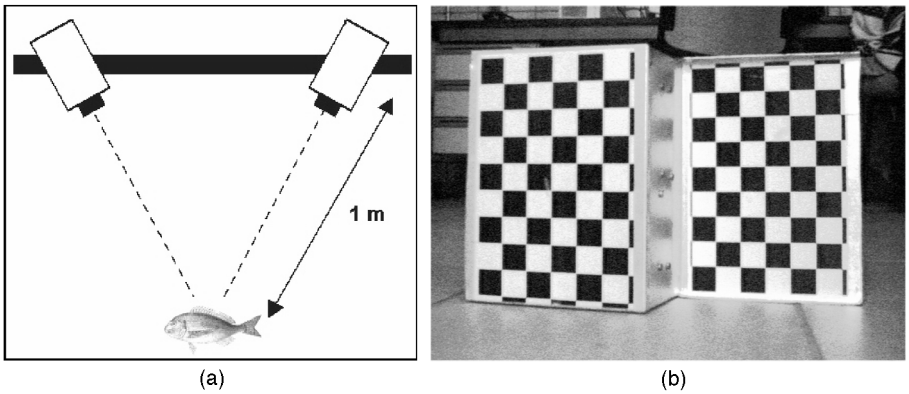


Fig. 2. (a) Scheme of the stereo vision system, (b) calibration grid used for the stereo calibration.

https://doi.org/10.1017/S0263574702004733 Published online by Cambridge University Press

identifying in the image from *camera 2* any point in the image from *camera 1*, can be divided in computing the epipolar line, and matching the point on *image 1* with its corresponding point on the epipolar line in *image 2*.

According to the Ayache method, the epipoles and the perspective matrices can be obtained during the calibration process. Thus, the determination of the epipolar line of a point of *image 1* in *image 2* only requires the computation of its direction vector, which is defined by the following expression:[8]

$$\begin{pmatrix} \Delta u^2 \\ \Delta v^2 \end{pmatrix} = M^{21} \cdot \begin{pmatrix} u^1 \\ v^1 \\ 1 \end{pmatrix}, \tag{3}$$

where M^{21} is defined as:

$$M^{21} = \begin{pmatrix} S_{E_2} & 0 & -U_{E_2} \\ 0 & S_{E_2} & -V_{E_2} \end{pmatrix} \cdot \hat{T}^2 \cdot [t_2^1 \times t_3^1 \;\; t_3^1 \times t_1^1 \;\; t_1^1 \times t_2^1], \tag{4}$$

where $\hat{T}^2$ is the 3×3 submatrix extracted from T^2 (perspective matrix of *camera 2*) by eliminating the last column, t_i^1 is the vector composed of the first three elements of the i-th row of T^1, T^l and T^2 are the perspective matrices for *cameras 1* and *2*, and $(U_{E_2} \;\; V_{E_2} \;\; S_{E_2})^t$ are the projective co-ordinates of the epipole of *image 2*.

The matching of the point in *image 1* on the epipolar line in *image 2* is solved by maximising a similarity index based on criteria such as cross-correlation functions and distance to the epipolar line, similarity in the local illumination conditions, and geometrical characteristics, such as size and orientation, of the object in which the point is located. The normalised cross-correlation function is often employed to avoid the differences in the illumination conditions in both images:

$$Corr_{A,B}(a,b) = \frac{\sum_x \sum_y I^1(x,y) I^2(x-a, y-b)}{\sqrt{\sum_x \sum_y [I^1(x,y)]^2} \sqrt{\sum_x \sum_y [I^2(x,y)]^2}} \tag{5}$$

3.3. 3-D reconstruction

The last step of the algorithm computes the co-ordinates in the real world of the points identified in both images. The reconstruction of a point P observed in the image from *camera 1* as (u^1, v^1) and in image from *camera 2* as (u^2, v^2) can be obtained by solving the following set of equations:

$$\left.\begin{aligned} (t_1^1 - u^1 t_3^1)^t P + t_{14}^1 - u^1 t_{34}^1 = 0 \\ (t_2^1 - v^1 t_3^1)^t P + t_{24}^1 - v^1 t_{34}^1 = 0 \\ (t_1^2 - u^2 t_3^2)^t P + t_{14}^2 - u^2 t_{34}^2 = 0 \\ (t_2^2 - v^2 t_3^2)^t P + t_{24}^2 - v^2 t_{34}^2 = 0 \end{aligned}\right\}, \tag{6}$$

where $t_{i,j}^2$ represents the (i, j) component of T^2. This set of equations can be easily solved applying simple least square methods. The result is a vector P composed by the co-ordinates in the real world of the observed point referred to the same co-ordinate system chosen for points of the calibration grid.

Several calibration validation tests have been performed with errors in the estimation of the distance in the real world lower than 0.7%.

4. FISH WEIGHT ESTIMATION BY USING COMPUTER VISION

3-D fish segmentation is carried out in several steps: The first one is responsible for noise filtering and for illumination compensation. Then, certain fish points of interest are detected in both independent images. Finally, the fish are segmented by associating the points of interest of the same fish, and their weight is estimated by using length-weight relations.

4.1. Pre-processing

Irregularities in illumination conditions were shown as one of the most harmful effects for the automatic image analysis, both in underwater images in sea cages and non-underwater images in nursery tanks. In indoor nursery tanks, these reflections were avoided by incorporating external lamps placed with an angle of approximately 45 degrees between the light lamps and the water surface. Artificial lightning cannot be easily used in cages in the open sea.

The main disadvantage of underwater sea-cages images illumination is due to the wide range of local illumination variations, as can be observed in Figure 3. It can be noticed that the illumination at the upper part of the images is usually very light in the Mediterranean coast due to intense sunlight. The illumination at the lower part is usually considerably dark due to sunlight attenuation and fish shades.

Image-processing techniques based on global illumination correction such as histogram equalisation fail when

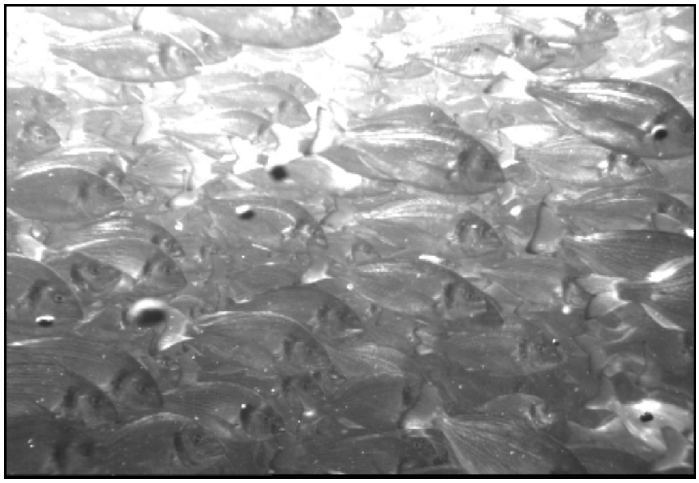


Fig. 3. Underwater image showing severe local variations of illumination.

https://doi.org/10.1017/S0263574702004733 Published online by Cambridge University Press

considering the local variations in the images. Local illumination correction methods based on square-block neighbourhoods provide unsatisfactory results since they do not fit the illuminations characteristics of this application. Methods based on modelling illumination as the low-frequency components of images[9] do not consider the sudden changes in the distribution of local illumination. The method for illumination correction employed aims to model the illumination variations and compensate them to generate an image with uniform intensity and contrast levels.

4.1.1. Illumination model of underwater images. The illumination variations in underwater sea-cages images were mainly due to fish shades, reflections on scales and sunlight attenuation. Although it is very difficult to model the detailed impact of these effects in the illumination, a qualitative model is of high utility to develop illumination compensation methods.

Image illumination is traditionally considered as the addition of three components: *ambient illumination*, *diffuse reflections* and *specular reflections*:

$$I = I_{ambient} + I_{diffuse-reflection} + I_{specular-reflection} \tag{7}$$

These three components suffer from water scattering and attenuation that depend on the distance covered by the light ray from the water surface to the observer. Ambient illumination models the reflection of ambient light, which arrives at the object after being bounced in multiple reflections from the objects of the scene. Its global effect is usually approximated by the following expression:[10]

$$I_{ambient} = \alpha_a I_a \delta_{att}, \tag{8}$$

where I_a is the intensity of the ambient light, α_a is the coefficient of ambient illumination reflection of the object surface and δ_{att} is the light attenuation factor.

Diffuse reflection originates from the scattering of light from a punctual light source (in this case the sun) on rough object surfaces. According to Lambert's law, diffuse reflection intensity is described by the following expression:

$$I_{diffuse-reflection} = \alpha_r I_p \cos(q) \delta_{att}, \tag{9}$$

where α_r is the coefficient of reflection of the object surface, I_p is the intensity of punctual light source and, q is the angle between the incident ray and the normal of the object surface.

Specular reflection originates when the light from a punctual source has stronger reflection in the line of view of the observer. Several specular reflection models have been proposed, including empirical-based models such as Phong[11] and physical-based ones such as Blinn.[12] According to Phong model, specular reflection can be modelled by:

$$I_{specular-reflection} = I_p \delta_{att} W(q) \cos(\beta), \tag{10}$$

where $W(q)$ is the reflectance angular distribution function that depends on q and on the material (fish scales) and, β is the angle between the reflected ray and the line of sight of the observer.

At the upper part of the image, ambient illumination and diffuse reflections have considerably high contributions, and specular reflections on fish scales are very intense and occur very frequently. Thus, the upper part of the images has high intensity levels and medium-high local contrast values. At the lower part of the images, ambient illumination and diffuse reflections weaken, and the probability of specular reflection decreases due to the fish shades. The intensity of specular reflections also decreases due to water attenuation and scattering. The lower parts of the images have low intensity levels and low local contrast. The central part corresponds to the transition between both situations and suffers from extreme local illumination variations mainly due to specular reflections and fishes shades. The central part of the images has average intensity levels and very high local contrast values.

Assuming vertical cameras, the sunlight attenuation effect involves different illumination features in each row in the image, which can be noticed in the row-pixel distribution (from now on, row histograms). Figures 4a and 4b show the vertical distribution of the mean intensity and contrast in the image shown in Figure 3. This distribution validates the qualitative illumination model proposed.

The method applied aims to compensate for this vertical distribution by transforming the histogram distribution of each row so that all the rows in the resulting image have the same histogram features.

4.1.2. Illumination compensation techniques. Many techniques such as *histogram equalisation* and *histogram specification*[13] perform illumination compensation by applying histogram transformation functions that depend directly on the histogram of the original images. However, the application of these techniques to all the rows for both images needs considerable computational requirements. In order to reduce the computational cost, the technique implemented is simple *direct grey-level transformation*, in which only some features from the histogram of the image are considered to design the transformation function. Two main types of transformation functions were considered: linear and non-linear functions.

Let $f(x, y)$ be an image of L grey levels with N rows and M columns. Illumination in grey-level images is often characterised by intensity and contrast values. The mean intensity and contrast of the r-th row of $f(x, y)$ can be defined as shown in the following linear expression:

$$MI_r = \frac{1}{M} \sum_{y=0}^{M-1} f(r,y), \qquad C_r = n^r_{max} - n^r_{min}, \tag{11}$$

where n^r_{max} and n^r_{min} are, respectively, the maximum and minimum intensity levels in the r-th row of $f(x, y)$. Consider that MI_{ref} and C_{ref} are the desired mean intensity and contrast value for the row grey-level distributions. Thus, the histogram transformation function for row r depends on MI_r and C_r (features from row r in the original image) and on the

https://doi.org/10.1017/S0263574702004733 Published online by Cambridge University Press

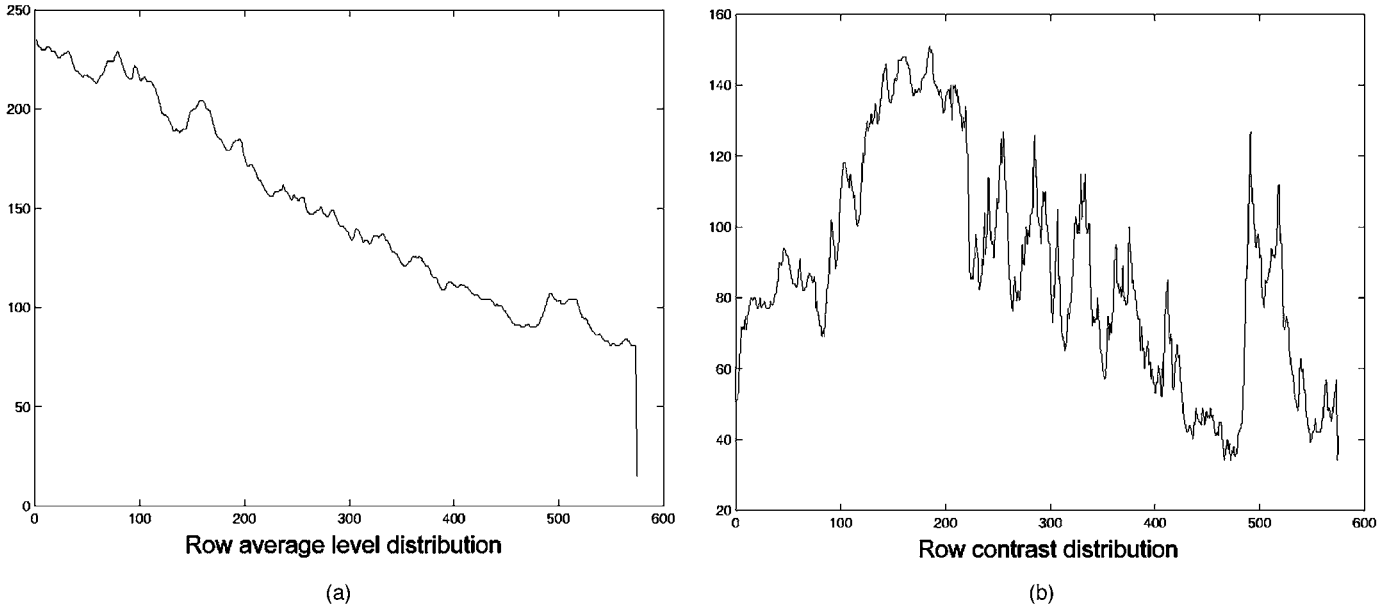


Fig. 4. Distribution of row average intensity level (a) and row contrast values (b).

desired features (MI_{ref} and C_{ref}) and can be expressed by the following expression:

$$f^*(r,y)=\frac{C_{ref}}{C_r}(f(r,y)-MI_r)+MI_{ref} \tag{12}$$

Thus, the row histograms at the upper part of the image are shifted (decreasing the mean intensity) and contracted (decreasing the contrast) and, at the lower part of the image they are shifted (increasing the mean intensity) and stretched (increasing the contrast). The application of Eq. (12) for all the rows cancels the vertical distribution of illumination, achieving images with the same mean intensity (MI_{ref}) and contrast (C_{ref}) for all the rows. Figure 5 shows the result of the application of grey-level transformation to image in Figure 3 with $MI_{ref}=120$ and $C_{ref}=130$.

A considerable number of non-linear histogram transformation functions for contrast enhance have been described in the literature, including polynomial[14] and sigmoidal[15] functions. Although non-linear functions allow a richer variety of histogram transformations, linear functions showed adequate performance with a lower computational cost.

Fig. 5. Transformed image with $MI_{ref}=130$ and $C_{ref}=80$.

4.2. Detection of points of interest

Once the image illumination has been compensated, the next step consists in the detection of fish points of interest or features. In nursery tanks fish are segmented by detecting the frontal and rear fish points and a certain number of intermediate points. The segmentation is carried out by applying region-growing techniques (see Figure 6).

However, this technique did not work well in underwater images due to the difficulty of automatic detection of the frontal and rear points because low contrast in these images. It should be noted that, in general, the most critical fish features depend on the fish species. In this paper, caudal fins have been used for the segmentation of gilthead sea breams.

4.2.1. Caudal fins detection. Figure 7 shows the images of six caudal fins extracted from the several images. One can observe the wide variety of caudal fin features and conditions. The grey-level intensities of caudal fins strongly depend on the illumination conditions. Besides, the position, size and angle of the caudal fins are not invariant and depend on the fish considered. According to the nature of

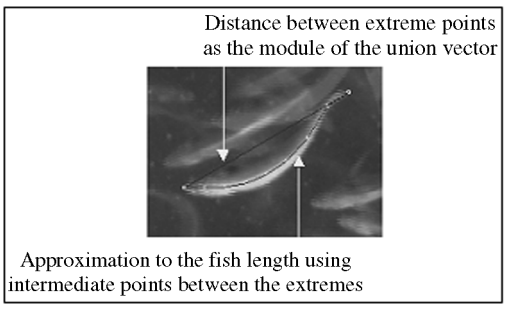


Fig. 6. Detection of frontal and rear fish points and computation of fish length.

https://doi.org/10.1017/S0263574702004733 Published online by Cambridge University Press

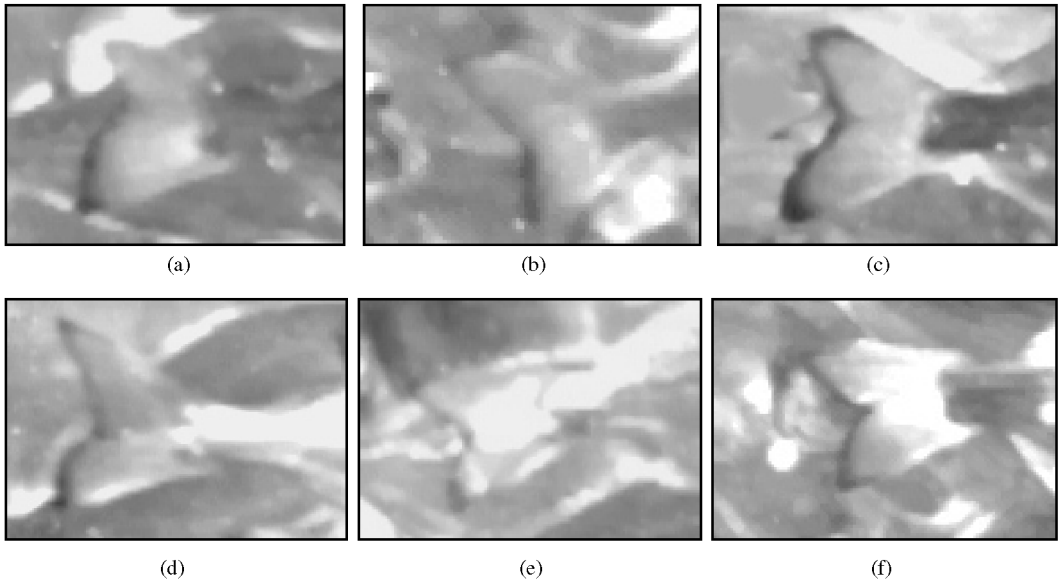


Fig. 7. Six zoomed images of caudal fins extracted from Fig. 5.

the problem the method to be employed for caudal fin detection should have the following properties:

- **illumination invariant**,
- **scale invariant**, in order to detect caudal fins from fishes at different stages of growth and at different distances from the cameras,
- **rotation invariant**, in order to detect caudal fins of fishes with different angle and to adapt the non-rigid nature of the fin,
- **high selectivity**, in order to avoid false caudal-fin detections.

The approach proposed in this paper for caudal-fin detection aims to detect the caudal fin angles by applying template-based matching (which is a *scale-invariant* operation) on a high-pass description of the images.

The application of the high-pass filter has two main objectives. The first one is to enhance the contrast of caudal fin (in order to **increase selectivity**), which in many cases is considerable low. The second objective is to increase the **illumination independence**. The illumination correction technique described in Section 4.1.2 is designed to compensate mainly the sunlight attenuation, and it does not eliminate the local illumination variations originated by fish shades. It should be noticed that high-pass filters have well-known illumination rejection properties, since illumination often corresponds to low-frequency components of the image.[9]

The selection of the high-pass filter is of high relevance to reduce the number of false detections in the performance of the overall system. Many operators were tested including the Roberts, Sobel, Prewitt and Laplacian operators. The experiments carried out revealed that emboss filter[16] achieved the most appropriate performance. Figure 8 shows the images resulting from the filtering of the images shown in Figure 7. The border originated by the caudal fins appear

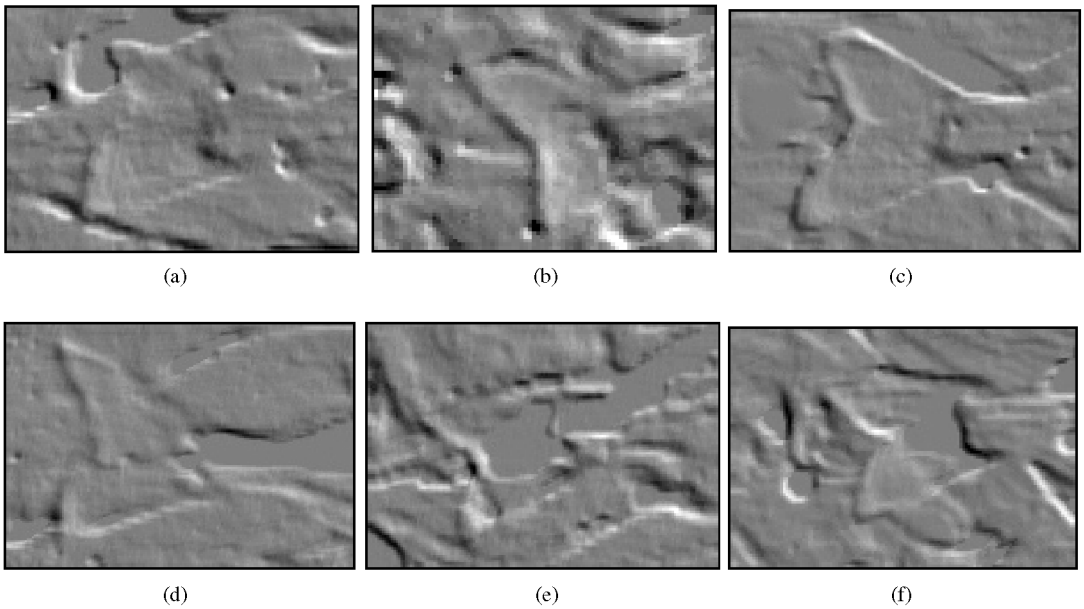


Fig. 8. Images resulting from the high-pass filtering of the images shown in Fig 7c, d and f.

https://doi.org/10.1017/S0263574702004733 Published online by Cambridge University Press

on the filtered image as transitions between pixels whose intensity levels are lower and higher than the mean level of the image.

A set of templates has been designed to match the above-mentioned transitions. The selection of templates with different caudal fin angles improves the rotation-invariance property of the method. The templates selected have low size in order to reduce the computational cost, to increase the local properties of the operator and, improve the resolution in the caudal-fin detection and stereo matching. The method consist in selecting the higher value resulting of the application of the templates on the filtered image $f^*(x, y)$:

$$f_O(x,y)=\max_i\left(\sum_{xt}\sum_{yt} template_i(xt, yt) f^*(x+xt, y+yt)\right), \quad (13)$$

where $template_i(xt, yt)$ for $i=\{1, \ldots NT\}$ represents the i-th template and NT is the total number of templates considered. The local maxima points of the resulting image $f_O(x,y)$ are considered caudal fins. Local aspects, such as level of the local maxima, can be considered to discard false detections. It should be noticed that this method also provides estimations of the caudal-fin angles, which can be useful to estimate the fish orientation. However, the flexibility of caudal-fin limits the reliability of the angle estimations.

4.3. 3-D segmentation methods for stereo images

Once the fish points of interest have been detected, the points of interest of the same fish should be associated in order to estimate fish length. The first step is the stereo matching of the points of interest detected in both images. This task is carried out according to the method described in Section 3.2. Various similarity criteria are taken into account, including cross-correlation function, distance to the epipolar line, similarity in local illumination conditions and geometrical aspects of the fish point of interest such as size and orientation. Once the points of interest have been matched, it is possible to reconstruct their location in real world co-ordinates.

Then, the points of interest of the same fish should be associated in 3-D co-ordinates by applying criteria that mainly depend on the fish species (such as fish geometry and orientation). For instance, gilthead sea breams usually move parallel to water surface, i.e. the angle between the fish axis (defined between caudal fin and mouth) and water surface is near to zero. Potential errors in the fish points association give rise to incoherent or divergent length and weight estimations, which can be eliminated by post-processing filters.

Once the fish points of interest have been associated, the fish length in real-world co-ordinates (L) can be simply computed. The weight is estimated by using length-weight relations, well known in the domain due to their relevance for production management of fish farms.[17] One of the most used relations is $W=a \cdot L^b$ (see Figure 9), where a and b are two coefficients that depend on factors such as water conditions and fish feeding. For gilthead sea breams these coefficients are commonly taken as $a=0.02263$ and $b=3.032$.

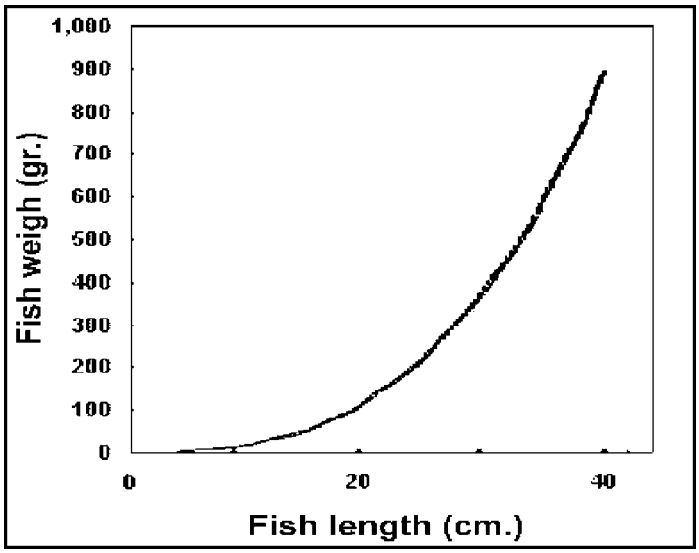


Fig. 9. Representation of the length-weight relation.

5. EXPERIMENTAL RESULTS

This Section describes the operation of the systems and presents some experimental results. The experiments of the underwater stereo system were conducted in several cages of gilthead sea breams in the open sea on the Spanish Mediterranean coast of Valencia and Murcia. The experiments of the non-underwater system were carried out in a fish farm in the Spanish province of Huelva.

Figures 10a and 10b show two synchronised images captured by the underwater stereo system in a sea cage on the Spanish Mediterranean coast of Valencia. Figures 10c and 10d show the corresponding images after the application of the illumination correction method with $MI_{ref}=120$ and $C_{ref}=130$.

Figures 11a and 11b shows the results of caudal-fin detection in both images. These images contain some false detections due to spurious fin-shaped objects and illumination effects. Figures 11c and 11d show the resulting images after caudal fins stereo matching. In this example 4 fins have been matched and located in real world co-ordinates. It should be noticed that the stereo matching process is sufficient to discard many of the above-mentioned false detections. Figure 12 shows the images resulting from the association of the points from the same fish. Three fishes have been segmented in these images.

The system is also capable of providing fish length distribution histograms such as that shown in Figure 13, which corresponds to the results in Table I obtained in the experiments performed in a nursery tank with the cameras over the water surface. In this experiment a total of 122 fishes were correctly segmented. The experiments carried out show weight errors lower that 4% for the system with the cameras over the surface in the nursery tanks, and lower than 5% for the underwater system in sea cages.

https://doi.org/10.1017/S0263574702004733 Published online by Cambridge University Press

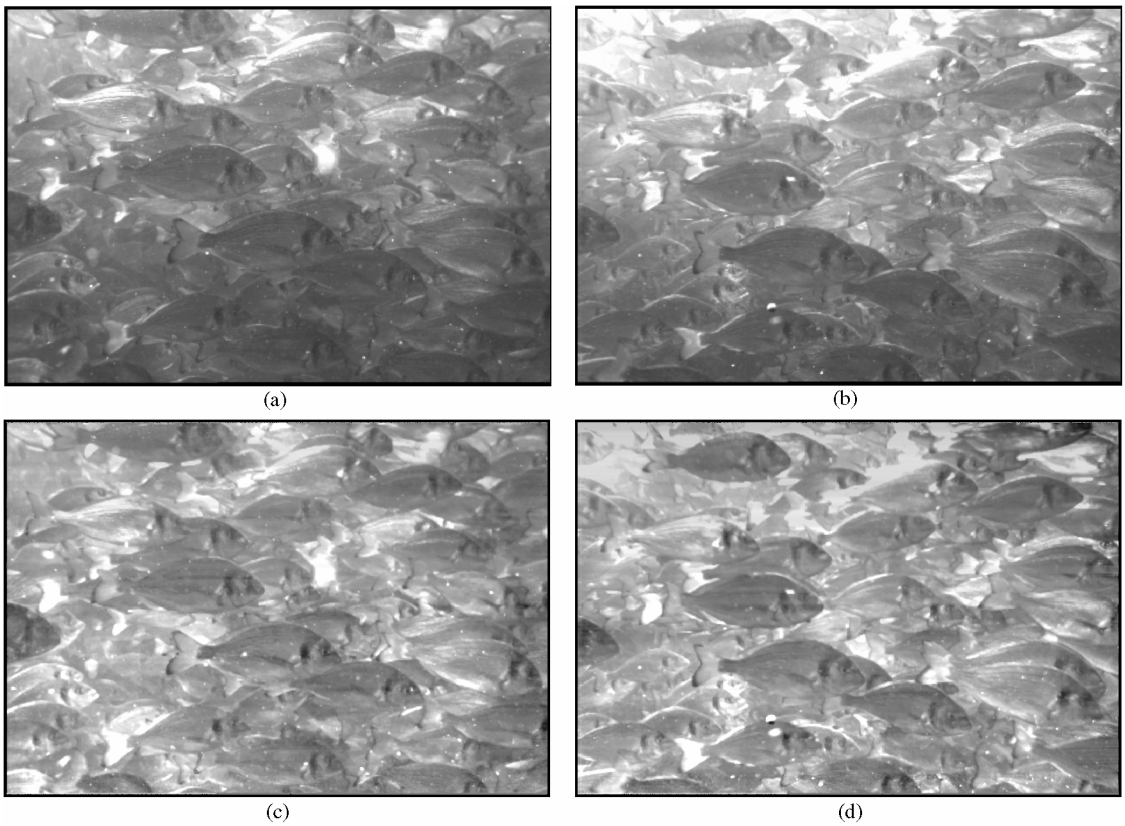


Fig. 10. (a) and (b) show two synchronised images captured by the underwater stereo system; (c) and (d) corresponding images after the application of the illumination correction method with MI_{ref}=130 and C_{ref}=130.

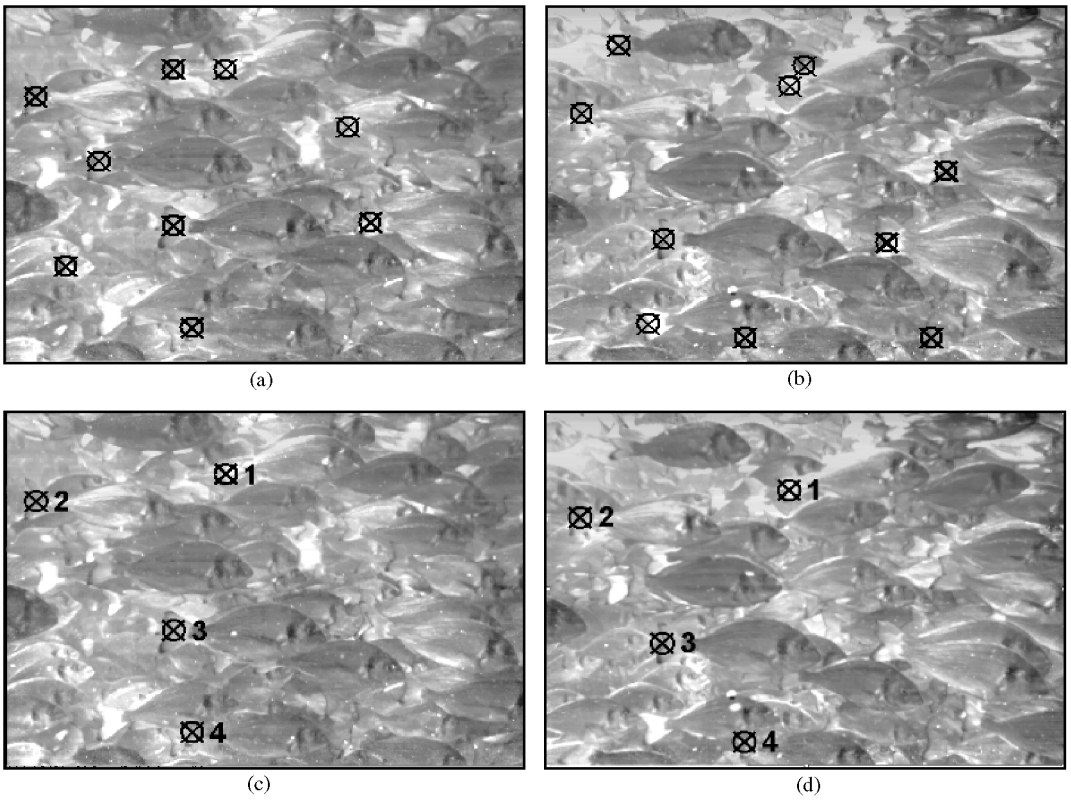
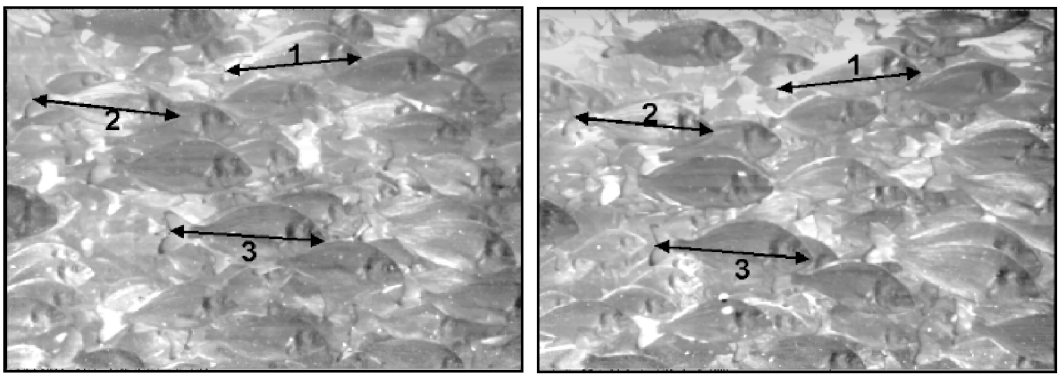


Fig. 11. (a) and (b) show the caudal fins detected in both independent images; (c) and (d) show the resulting images with the fins validated after the stereo matching.

https://doi.org/10.1017/S0263574702004733 Published online by Cambridge University Press



(a) (b)

Fig. 12. (a) and (b) show the resulting images after mouth-caudal fin association.

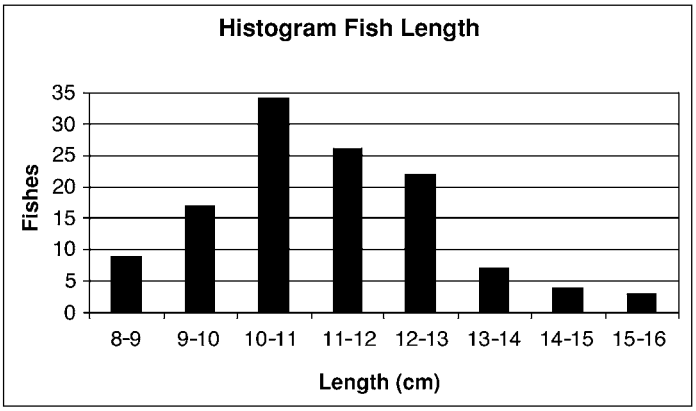


Fig. 13. Histogram of the fish length measured in the tank.

6. ROBOTIC APPLICATIONS

In addition to the above presented method for the automation of the fish weighting process by avoiding fish contact, some other new low cost systems for fish farm automation has been designed and implemented. In the following two robotics systems will be shortly described.

The first one is a system for automatic fish feeding. The system consists of an auto-demand component and a mobile robot for food transportation and supplying. Auto-demand is activated by fish contacts on a small sensor installed inside the tanks. Several criteria, such as frequency of contacts, are employed to discard false food demands due to water turbulence.

On the other hand, several mobile robot designs were considered in the project, including mobile wheeled platforms, which can navigate autonomously by using position estimation techniques based on triangulation with a number of radio beacons distributed in the fish farm that can be detected by the robot. The objective of this robot was not only feeding but also to obtain several measures of the tank, including the application of the computer vision techniques, described in the previous Section, with the pair of stereo cameras mounted on a simple arm attached to the robot which can be placed over the tank surface. However, due to reliability and cost, the final solution was a robot that moves along a rail at the top of the tanks. The robot takes the food from the food box and distributes it in several tanks according to fish demands. The stereo cameras and lights can be easily mounted in this mobile system to estimate the fish weight distribution in all the tanks as described above.

Table I. Experimental Statistics.

Item	Value
Number of fish processed	122
Average fish length (cm.)	11.14
Standard deviation of fish length (cm.)	1.58
Average fish weight (g.)	19.06
Standard deviation of fish weight (g.)	8.40

The second system is a low-cost autonomous underwater robot for pond cleaning which has also been applied in open-air fish farms with ponds. The trajectory of the robot in the pond is controlled by using visual feedback techniques. The robot is equipped with a float that is employed to indicate its position and orientation in the pond. A camera installed in a tower registers images of the float and computes the position and the orientation of the robot by applying image-processing techniques. The position and orientation of the robot are used for the automatic control of its trajectory. This robot could be also used as a platform to install the underwater system for the optical estimation of weights presented in the above sections.

Finally, it should be pointed out that the stereo vision system for fish weighting can also be integrated in a SCADA system which has been developed and implemented for the monitoring and operation of a nursery distributed control system. Thus, it would be possible to implement advanced fish-feeding strategies based on the weight distribution. Moreover, the use of this information is very valuable for production control.

7. CONCLUSIONS

Automation is an important need of modern fish farms, which are increasing their production and quality requirements very significantly. However, cost automation plays an important role and some technologies that could be applied in industrial automation are not suitable here.

Image processing is a very valuable non-intrusive technique for the biomass estimation in fish farms since it avoids the manual handling of the fish required for the sampling and weighting of the fish population in tanks and cages. However, the application of these techniques involves several constraints to the reliability and cost of the systems.

https://doi.org/10.1017/S0263574702004733 Published online by Cambridge University Press

In this paper a low-cost stereo system for the estimation of fish biomass has been presented. Two different designs have been developed for the two types of ponds: sea cages and tanks. An underwater option has been chosen for sea cages and ponds in open-air fish farms, while a technique with the cameras over the surface and controlled lights have been selected for indoor tanks.

The paper summarises the most relevant aspects of the systems, including image processing methods and calibration and stereo matching procedures. Instead of complex 3-D modelling of the fish, the systems apply simple techniques based on the estimation of the fish length and length-weight relations commonly used in the domain.

The paper describes the experiments conducted with gilthead sea breams in sea cages (in fish farms at the east coast of Spain) and nursery tanks in a fish farm in Ayamonte (south-west Spain). The fish weight distribution resulting of the proposed systems achieved errors of 4% (in nursery tanks) and 5% (in sea cages), that were validated by means of traditional methods.

These stereo perception systems could be applied in fish farms with low cost requirements. Furthermore, several robotic systems have been developed to automate various tasks including an automatic auto-demand food distribution robotic system and an autonomous pond-cleaning robot. A robotic system based on auto-demand allows one to design a feed process that adapts automatically to the state of growth of the fish and their food demands. The cleaning of the ponds is carried out by a robot controlled by visual feedback techniques.

The presented techniques could be integrated in a distributed system for the whole fish farm control, improving the quality of the product and decreasing the production cost.

ACKNOWLEDGEMENTS

The work described in this paper has been performed in the project "Intelligent systems for monitoring and control in fish farms", 1FD97-1416, partially funded by the CICYT and in a project funded by the company Step Control. The authors acknowledge the help and implementation activities developed by Step Control and particularly by F. Malet. They also acknowledge the assistance and help for the sea cage experiment of Luis Martín, as well as the access to the fish farm and the nursery information provided by J.M. Rodriguez and D. García. The research and development activities of A. Jiménez in the visual robot position estimation and control are also acknowledged.

References

1. F. Malet, A. Ollero and M. Orte, "Sistema para Automatización y Control en Piscifactorías: La solución implementada en Acuinova", *Automática e Instrumentación* **No. 305**, 82–86 (2000).
2. N. Castignolles, M. Cattoen and M. Larinier, "Automatic system for monitoring fish passage at dams", *SPIE, Application of Digital Image Processing XVII* (1994) **Vol. 2298**, 419–429.
3. N. Castignolles, M. Cattoen and M. Larinier, "Identification and counting of live fish by image analysis", *SPIE, Image and Video Processing II* (1994) **Vol. 2182**, 200–209.
4. M. Foster, R. Petrell, M.R. Ito and R. Ward, "Detection and counting of uneaten food pellets in a sea cage using image analysis", In: *Aquacultural Engineering* **14**(3), 251–269 (1995).
5. R.J. Petrell, X. Shi, R.K. Ward, A. Naiburg, and C.I. Savage, "Determining fish size and swimming speed in cages and tanks using video techniques", *Aquacultural Engineering* **16**(1,2), 63–84 (1997).
6. K.P. Ang and R.J. Petrell, "Control of feed dispensation in seacages using underwater video monitoring: effects on growth and food conversion", *Aquacultural Engineering* **16**, 45–62 (1997).
7. C. Serna and A. Ollero, "A stereo vision system for the estimation of biomass in fish farms", *Proceedings of the 6th IFAC Symposium*, Berlin, Germany (8,9 October, 2001) pp. 185–191.
8. N. Ayache, *Artificial Vision for Mobile Robots: stereo vision and multisensory perception* (Cambridge, Mass., MIT Press, 1991).
9. P. Lancaster and Q. Salkauskas, *Curve and Surface Fitting, An Introduction* (New York, Academic Press, 1981).
10. R. Hall, *Illumination and Color in Computer Generated Imagery* (Springer-Verlag, 1989).
11. Phong Bui-Tuong, "Illumination for computer generated images", *Comm. ACM* (1975) **Vol. 18** pp. 311–317.
12. J.F. Blinn, "Models of Light Reflection for Computer Synthesized Pictures", *Computer Graphics* **11**(2), 192–198 (1977).
13. R. Gonzalez and R. Woods, *Digital Image Processing* (Addison-Wesley, 1993).
14. G. Ramponi, "Contrast enhancement in images via the product of linear filters", *Signal Processing* **77**, 349–353 (1999).
15. K.R. Castleman, *Digital Image Processing* (Prentice Hall, New Jersey, 1996).
16. C. Watkins, A. Sadun and S. Marenka, *Modern Image Processing* (Academic Press, Inc, 1993).
17. Fishbase [online]. from http://www.fishbase.org.