Optical Fish Classification Using Statistics of Parts

Mohcine Boudhane[1,2], Benayad Nsiri[1] and Hamza Toulni[1]
[1]University Hassan II, Faculty of Sciences, Ainchock B.P 5366 Maarif 20000, Casablanca, Morocco
[2]Faculty of Computer Science and Electrical Engineering, University of Applied Sciences, Grenzstr. 5, 24149 Kiel, Germany
Email: {mrboudhane@gmail.com, benayad.nsiri@enst-bretagne.fr, hamza.toulni07@etude.univcasa.ma

***Abstract*—In this paper, We consider the problem of fish classification underwater. We described an algorithm based on robust fish decomposition and feature extraction. Our goal is to develop a system that detects and recognizes many kinds of fish in images and video sequences including various structures in underwater imagery. In order to get pertinent features, many classifiers is defined. These classifiers are responsible to extract parts of interest from the fish body. Each part forms a sub-classifier and represents various local properties of the fish. A Bayesian artificial neural network (ANN) is used to classify each part. Probabilities are given by each ANN and was combined and recalculated in order to give a decision. Experimental results shows that the proposed approach give a good accuracy to make distinguish between different fish species.**

***Keywords*—Underwater fish recognition, Artificial neural networks, Image/video analysis, Image and signal processing.**

I. INTRODUCTION

Computer vision is the science of the processing, analysis , and interpretation of visual images. It includes many techniques which are useful in their own right, as well as image processing and pattern recognition. It can help also to detect significant events, track and analyze the content of thousands frames. The typical functions which are found in many computer vision systems are image acquisition, preprocessing, feature extraction, detection or segmentation and analysis.

Actually, The most detection and monitoring systems underwater are based on cameras and the exploitation of the image data. Underwater vision is an important issue in ocean engineering. Computer vision techniques can also help biologists observe marine ecosystems where the manual annotation is too expensive. Underwater vehicles are used to survey the ocean floor, generally with optical sensors for their capability of remote sensing in recent years. Different from the common images, underwater optical images suffer from poor visibility due to the medium scattering and light distortion [1]. There are many challenges for object detection and classification using camera because of the constraints of water visibility, non-uniform color distribution under water, occlusion, changing appearance patterns of both the object and the scene [2] $\cdots$

Because of the challenging environmental conditions and the differential light dissemination, most of traditional computer vision methods cannot be applied directly in underwater images. Recent works try to enhance the underwater image quality, and to reduce the level of noise, in order to detect and localize the objects in it successfully. Some researchers in [3], [4], [5] propose filter based methods for reduction of undesirable noise. In [6] and [7] wavelet based methods are proposed.

In this work, we propose a method for detection, tracking and recognition for fish in underwater environment using camera. Our goal is to offer the submarine biologist to explore the underwater environment and analyze the behavior of different fish species without prior knowledge of the environment. In which, We investigate the recognition task of more fish species in a more complex and fundamentally challenging natural environment.

The next sections in this paper are organized as follows. Section II describes the system overview. In Section III, experimental results and comparison with other methods are shown. Section IV conclude this work.

II. SYSTEM OVERVIEW

The study of underwater species is an important topic to marine biologists and environmental experts. Optical fish sequences analysis is a popular approach to study marine life. To determine the size and distribution, or to study the behavior of species, the researchers need to locate them in images and (or) videos, but this can be very time consuming when processing a large volume of data. An robust fish recognition is required in this subject. The purpose of this paper is to identify fish in the video or image sequences, by analyzing and exploring various visual information. The proposed system steps is divided into six steps (Figure 1).

A. Training data processing

We want to classify some fish species and recall our agreement that any given fish is either included in the vector of fish categories **c**. Lets define a (probabilistic) variable that describes the state of nature:

$$c = \{c_1, \cdots c_N\} \text{ with } i \in [1, N]. \quad (1)$$

with

- c_i represent the i-th category.
- N is a number of fish species.
- $\sum_{i=1}^{N} P(c_i) = 1;$

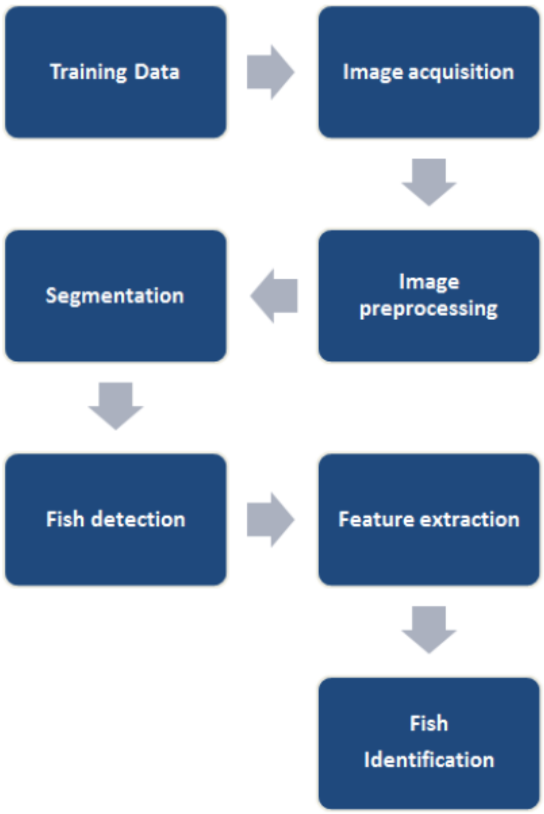

Figure1 : System architecture.

This step is an investigation in each species. It observes all the necessary information (features in traing data) that can make the distinguish between them. A feature is some attribute of the object that is considered important in describing and recognizing the object in relation to other objects. Color, size, shape, texture, head, number of fins, ... are some commonly used features. The goal in this stage to find relevant information in many different features that deal with large quantities of information. Figure 2 shows a feature-model matching that demonstrate the parts of interest of our model. In the end of this step, the probabilities inter-species is processed, and the probabilities apriori have been generated.

All object recognition systems use models either explicitly or implicitly and employ feature detectors based on these object models. An object recognition system must select appropriate tools and techniques for the steps discussed above. Many factors must be considered in the selection of appropriate methods for a particular application. The central issues that should be considered in designing an object recognition system are (Figure 1):

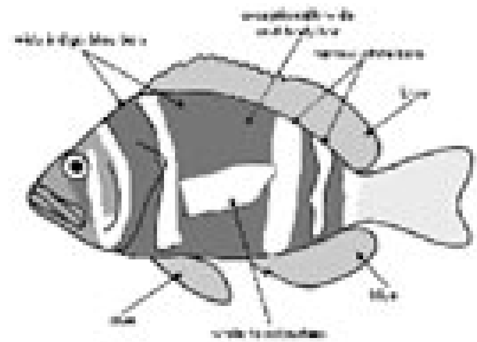

Figure 2 : Feature-model matching.

- Data acquisition
- Preprocessing
- Segmentation and detection
- Feature extraction
- Fish recognition

B. *Data acquisition*

The underwater platform is designed to detect and record optical signals over an extended deployment. The system is able to take videos or image sequences in the input. Figure 3 shows some examples of raw scenes.

Figure 3 : Example of scenes used in the classification.

C. *Preprocessing*

Frame analysis needs preprocessing whatever the noise occurred in the capture. Thus, this noise should be eliminated. The aim of this part is to enhance the visual appearance of the image, in order to improve the manipulation of raw frames. When we look on the underwater images we can find two major problems : smoothness and noise. For that, we use a preprocessing method cited in [8]. The method model the underwater environment as overlaps of two processes. The first process is considered as a Poisson distribution, while the second one is considered as a Gaussian mixture. The resulting distribution is called Poisson-Gaussian mixture (PGM). The distribution is defined as follows

$$p(z|\mathbf{y}) = \sum_{k=0}^{\infty} \left(\frac{x^k}{k!} e^{-x} \sum_{m=1}^{M} \alpha_m \cdot f(\mathbf{y}, \mathbf{C}_m, \boldsymbol{\mu}_m) \right),$$

where λ is a strictly positive real number, z is an observation, x resp(y) is the realization of the Poisson distribution (resp of the Gaussian distribution f), m is the number of Gaussians, C_i is the covariance matrix of the i^{th} Gaussian, μ_i and α_i are the i^{th} mean and mixture coefficient, respectively.

D. Segmentation

In this stage we use so called "Mean shift" algorithm. Mean Shift is a powerful and versatile non parametric iterative algorithm that we will use it in the segmentation of the images. Mean Shift was introduced in Fukunaga and Hostetler [9] and has been extended to be applicable in many fields in Computer Vision. It is a powerful non-parametric iterative algorithm that can be used in many purposes (segmentation, clustering, $\cdots$). Mean-shift associates these segments with the nearby pixels of the dataset probability density function. As results, Meanshift segment images on different regions.

E. Features extraction

The aim of this step is to extract as much as possible information from input data. These features are expected to characterize different proprieties of structures and objects in each data source. After feature extraction a large amount of important information are obtained.

1) Challenges: : Automatic fish recognition is a difficult undertaking. In over 30 years of research in computer vision, progress has been limited. The main challenge is the amount of variation in visual appearance. An feature detector must cope with both the variation within the object category and with the diversity of visual imagery that exists underwater. For example, fish vary in size, shape, color, and in small details such as the head shape, texture, and position/number of fins, tail $\cdots$

2) Formulation: : After observing fish in the database, We proceed by feature extraction by part (by region) of the fish pattern. We split the image into 3 parts. the first part represent the head, second represent the body, and the third represent the tail. where the size of the body is twice the size of the head (resp. tail). The categorization of this decomposition is not derived from a same point of view. As results, feature vectors (of fish parts) is not identical. These categories are later used as the class labels in the classification problem. Figure 4 shows an example of fish parts extracted.

For each fish species, we have computed more than 50 different features, which are divided in five groups. Those groups and their corresponding numbers of features are:

Size (8 features)	**e.g:** *Area, length.*
Shape (20 features)	**e.g:** *Moment, aspect ratio.*
Color (9 features)	**e.g:** *luminance, chrominance.*
Texture (16 features)	**e.g:** *Inertia, energy.*
Particular features	**e.g:** Depend of each part of fish.

Figure 4 : Fish decomposition.

Region feature extractors process square image neighborhoods and represent its central pixel by the resulting feature vector. This is useful to account for local spatial information and structure in images. We can applied in texture classification and distinguishing edges. For each part of fish, They are some common feature (size, color, area, shape, texture) and some special features. Given input fish, we form a set of parts, each consisting of a group of variables that are statistically dependent. We then treat these parts as statistically independent. With this assumption, our classifier takes the following form:

$$P_{fish} = \prod_{i=1}^{M} P(part = i) \tag{2}$$

where

- P_{fish} is a global probability of the whole fish.
- $P(part = i)$ is the probability of the $i-th$ part.
- M is a number of fish parts.

F. Fish identification

Its performed using Artificial Neural network ANN [10]. ANN is not that different from our magnificent biological system in their most basic definitions. They are composed of input channels, units and output channels. An artificial network defines the interaction of several units. A single unit (also called neuron, or cell) has one or more real valued edges as input, an integration function that process the inputs and produce a value passed to an activation function which basically computes the output of the unit. This neural network is formed in three layers, called the input layer, hidden layer, and output layer:

- **Input layer:** represent a feature vector for selected part.
- **Hidden layer:** feature combination .
- **Output layer:** include the fish categories.

Each layer consists of one or more nodes, represented in this diagram by the small circles. The lines between the nodes indicate the flow of information from one node to the next (Figure. 5).

- They take as input n real values $I_1, I_2, ..., I_n$. These values represent feature vector of part of fish.
- The integration function computes the sum $\sum_{i=1}^{n} I_i w_i$ where w_i is the input associated edge weight $w_1, w_2, ..., w_n$.
- The activation function is a boolean function which tests the inequality $\sum_{i=1}^{n} x_i w_i \geq a$, where a is the threshold.

III. CLASSIFICATION RESULT:

In this section, our method is compared with conventional methods. The system is implemented on a standard PC (8GHz). Different images of different sizes have been used in the experiment.

As we say before, the classification is done by parts (segments). Segment analysis was employed to identify fish part based on robust feature extraction. In order to perform segment analysis, the feature vector of each segment were used as sub classifier. First, Bayesian ANN method was performed in the cluster to get an classification results of the cluster. Second, the sub-results was combined and recalculated to give a global decision. Third, The resultant probability was verified and final decision is given. The corresponding operating point curve for the test data is shown in Figure. 5

The objective is to apply the sub-segment data fusion approach in decision level identity fusion in order to combine multiple neural networks. The aggregated neural network model can be represented as

$$y_t^* = w_1 y_1 + w_2 y_2 + \cdots + w_m y_m. \tag{3}$$

where t is the aggregated single network model prediction and w_k is the weight for combining the k-th network.

In order to implement the Bayesian inference data fusion techniques in the combining of multiple neural networks. It can be taken from the basic formula of Bayesian Combinations. In the Bayesian inference process, supposed we have H_1, H_2, $\cdots$,H$_i$ representing hypotheses that can explain an event E (an observation) then the posterior probability of hypothesis H_i being true given the evidence E, $P(H_i|E)$, can be defined as:

$$P(E|H_i) = \frac{P(H_i|E)P(H_i)}{\sum_{i=1}^{N} P(H_i|E)P(H_i)} \tag{4}$$

where $P(H_i)$ is the prior probability of hypothesis H_i being true and $P(E|H_i)$ is the probability of observing evidence E given that H_i is true.

For example we take the case of the *"tail"* part, and we treat the feature *"shape of the tail"*. Let fr_1 defines this feature. We consider the case of the probability of fish tail to be forked knowing that the category is Salmon. In that situation, the probability in (4) will be as follows (in (5) and (6)).

$$P(\text{Head forked}|Salmon) = \frac{P(Salmon|\text{Head forked})P(\text{Head forked})}{P(Salmon)} \tag{5}$$

after simplification, the probability will be as

$$P(\text{Head forked}|Salmon) = \frac{P(Salmon|\text{Head forked})P(\text{Head forked})}{\sum_{i=1}^{N} P(c_i|\text{Head forked})P(c_i)} \tag{6}$$

where c_i defines the i-th category.

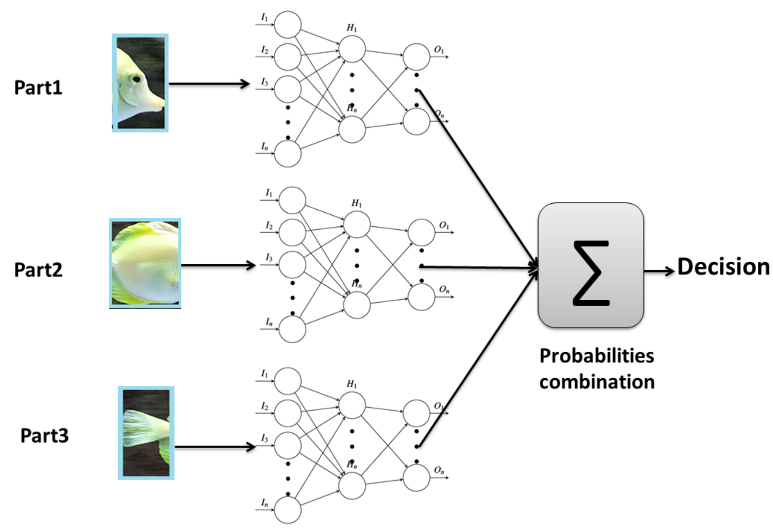

Figure 5 : The proposed fish recognition model.

In this paper, six fish sequences are randomly picked from each species and altogether there are several fish images. When running the classification algorithm, a set of fish categories could be recognized. A total Accuracy, which is defined as the percentage of correctly classified fish. Usually, Accuracy is represented as a real value between 0 and 1.

$$\text{Accuracy} = \frac{\text{Correct decision}}{\text{Total decision}} \tag{7}$$

The first evaluation only evaluates the performance of our fish detection system, which consists of two classifiers. When evaluating each fish class separately, the classification performance for each category could be concluded:

i	c_i	Classification Accuracy
1	Salmon	99.4%
2	Blowfish	95.5%

Table1: Classification accuracy

The individual class recall/precision is shown in Table 1. The proposed approaches achieve a good accuracy in fish distinguish. The accuracy of 95.5% and the matching speed of 0.15 seconds for several test sample images have been obtained. It is shown in Figure 6.

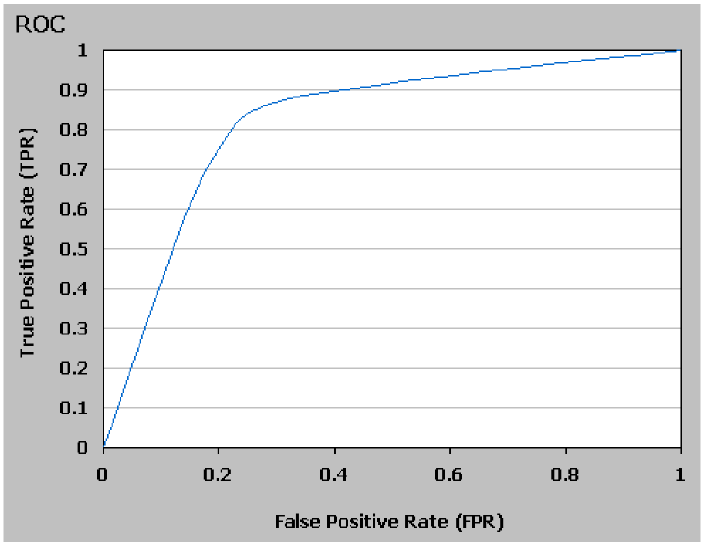

Figure 6: ROC diagram of our classification system.

IV. Conclusion

In this paper, we introduce an object recognition system in underwater environment. Our fish recognition system is based on features extraction of fish parts. Particular features for these parts are employed in the recognition system for efficiency. For the recognition system to be used in the smart environment, the system should be robust under the following environmental factors: varying luminance; scale change caused by varying distances between the camera and the target; image change caused by varying object orientation. To minimize the effects caused by the factors we adopted a Poisson-Gaussian distribution for denoising and the enhancement of raw images. Besides, a Gaussian color model is introduced to lessen the illumination effects. Neural networks are used to generate the posterior probabilities of each parts. Experimental results show that the proposed system gives a good recognition performance in recognition accuracy.

References

[1] J.S Jaffe ”Underwater Optical Imaging: The Past, the Present, and the Prospects” Oceanic Engineering, IEEE Journal of (Volume:40 , Issue:3),(2015), pp.683-700.

[2] Donna M. Kocak, Frank M. Caimi ”The Current Art of Underwater Imaging - With a Glimpse of the Past and Vision of the Future” Marine Technology Society Journal. 39(3):5-26.

[3] C. Chang, J. Hsiao, and C. Hsieh, ”An adaptive median filter for image denoising,” in Proc. IEEE IITA ’08, Qingdao, China, 2008, pp. 346-50.

[4] C. Wang and J. Zhang, ”Image denoising via clustering-based sparse representation over Wiener and Gaussian filters,” in Proc. IEEE S-CET ’12, Qingdao, China, 2012, pp. 1-4.

[5] A. Nath, ”Image denoising algorithms: A comparative study of different filtration approaches used in image restoration,” in Proc. IEEE CSNT ’13, Gwalior, India, 2013, pp. 157-163.

[6] P. Prabhakar and P. Kumar, ”Underwater image denoising using adaptive wavelet subband thresholding,” in Proc. IEEE ICSIP ’10, Chennai, India, 2010, pp. 322-327.

[7] S. Feifei, Z. Xuemeng, and W. Guoyu, ”An approach for underwater image denoising via wavelet decomposition and high-pass filter,” in Proc. IEEE ICICTA ’11, Shenzhen, China, 2011, pp. 417-420.

[8] M. Boudhane, S. Badri-Hoeher, B. Nsiri ”Optical fish estimation and detection in noisy environment” IEEE Oceans - St. John’s, 2014. pp1-6.

[9] K. Fukunaga, L. Hostetler ”The estimation of the gradient of a density function, with applications in pattern recognition” IEEE Trans. Information Theory, vol. 21, no. 1, pp32-40, 1975.

[10] E.A Wan, ”Neural network classification: a Bayesian interpretation” IEEE Transactions on Neural Networks (1990), (Volume:1, Issue:4), pp303-305.