Intelligent System for Automated Fish Sorting and Counting

Sébastien Cadieux[1,2], Francois Lalonde[1] and Francois Michaud[2]

[1]Systèmes d'automatismes et de mesure
Institut de recherche en électricité du Québec (IREQ)
Varennes, Québec, Canada, J3X 1S1
lalonde.francois@ireq.ca

[2]Department of Elec. & Comp. Engineering
Université de Sherbrooke
Sherbrooke (Québec, Canada) J1K 2R1
{cadieux,michaudf}@gel.usherb.ca

Abstract

This paper presents an automated system for counting fish by species. This system is to be used in fishways for monitoring and surveying fish. The system requires very few adjustments and no special installation. An infrared silhouette sensor is used to acquire the fish silhouettes. These silhouettes are then processed on a personal computer for fish counting and classification by species. The system allows the operator to select the species of interest according to the fauna of the specified river. Classification is made based on the combined results of a Bayes maximum likelihood classifier, a Learning Vector Quantification classifier and a One-Class-One-Network (OCON) neural network classifier. Through the use of specialized classifiers of different types, a robust, modular and expandable recognition system is created.

1 Introduction

Over the years, the construction of dams has lead to the gradual disappearance of migratory fish in certain rivers. To avoid this, fishways allow the fish to cross the dams and to reproduce upstream. Scientific monitoring and survey of the migrating fish passing through the fishways are however required, to confirm the usefulness of the solution. Such monitoring consists of counting, by species, the number of fish crossing the dam. However, up to now, this tedious task has mostly been done manually by capturing the fish passing through the fishway or by putting a video camera in the fishway.

A better solution would be to automatize the process. To achieve this goal, Castignolles [1] presents a method using binary images taken using video camera to recognize the species of the fish. The system has obtained good results but has low expandability, i.e., it is difficult to increase the system's ability of recognizing new species of fish. In addition, the use of a video camera has some drawbacks, since water quality and lighting conditions are critical in acquiring the images. Special installations and the use of fragile equipment are then required, and can be problematic when used in remote locations. Some other systems have been developed for fish species recognition using artificial vision [6,8], but these systems are designed for fish plants where the conditions can easily be controlled.

Our system uses the infrared fish silhouette sensor "RiverWatch". This sensor uses infrared diodes to acquire images of objects passing through the fishway. This sensor has the advantage of using low power and to work in a wide variety of water conditions, without any special installation on the fishway. The images are sent to a personal computer where the images are analyzed, classified and counted by the analysis algorithm. Classification is made based on the combined results of a Bayes maximum likelihood classifier, a Learning Vector Quantification classifier and a One-Class-One-Network (OCON) neural network classifier. Results from these classifiers are combined using a majority vote method. Through the use of specialized classifiers of different types, a robust, modular and expandable recognition system is created. Our classification system also innovates by allowing the operator to select the species of interest and by deactivating the classification of other species. This makes it possible to account for the fauna of any rivers.

This paper is organized as follows. Section 2 gives a general overview of the system and of the silhouette sensor. Section 3 covers the features extraction methods, the classifiers used and how their results are combined. Results are presented in Section 4, followed by the conclusion.

2 Fish Counting System

Represented in Figure 1, the system can be divided in two parts: the silhouette sensor and the analysis algorithm.

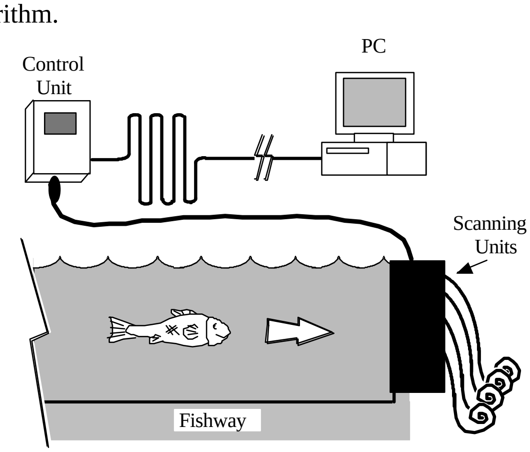

Figure 1 - Fish counting system.

2.1 Silhouette Sensor

The silhouette sensor is developed by an Icelandic company called Vaki. The sensor is made of two scanning units and a control unit. Each scanning unit has two linear sensors, each of them composed of about a hundred infrared diodes and sensing elements. As the fish swims through the scanning units, it obstructs the infrared beams and a silhouette is constructed.

Figure 2 – "RiverWatch" silhouette sensor.

Once acquired, the data are sent periodically to a personal computer using a direct link, a satellite link or a cellular phone, for analysis. The system has low power requirements (only a solar panel is required), making it possible to use the sensor in remote locations. Also, it does not need long adjustments to work in different water conditions or installations.

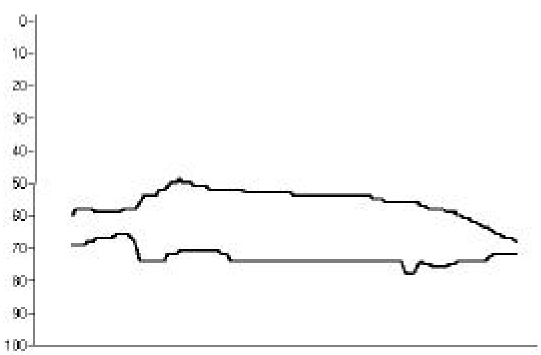

Figure 3 - Example of a Northern Pike silhouette.

Figure 3 and 4 show two examples of such images for two different species of fish (the northern pike and the Atlantic salmon).

Figure 4 - Example of an Atlantic Salmon silhouette.

3 Analysis Algorithm

The analysis algorithm classifies the silhouettes it receives and then counts the recognized fish by species. The algorithm works in three steps: feature extraction, classification, and classifier fusion.

3.1 Feature Extraction

The goal of the feature extraction step is to extract specific features from a silhouette, required for classification. Classification of two-dimensional shapes necessitates features that are invariant to position, size and orientation. Two of the most widely used features for silhouette recognition are moment invariants and Fourier boundary descriptor. Castignolles [1] also uses different shape descriptors. Our system uses several combinations of these features to get a more complete description of fish silhouettes.

Moment Invariants

The concept of moments as invariant image features was first introduced by Hu [4]. Because moments do not only account for the contours of the shapes, like some of the other methods, they have the advantage of representing very complex shapes. This type of features has been used in many applications, including fish species recognition [1]. Here is a very brief description of the moments.

For digital images, a $(p+q)$th order moment is defined by Equation (1).

$$m_{pq} = \sum_x \sum_y x^p y^q f(x,y). \tag{1}$$

The central moments of a digital image can then be expressed as shown by Equation (2):

$$\mu_{pq} = \sum_x \sum_y (x-x_c)^p (y-y_c)^q f(x,y). \tag{2}$$

where x_c and y_c are the centres of gravity. The normalized central moments are derived from these moments by using Equation (3).

$$N_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\frac{p+q}{2}+1}} \tag{3}$$

Hu [4] derived seven invariant moments using non-linear combinations of the second and third normalised central moments. Our systems uses the first seven invariant moments.

Fourier Descriptors

A common set of features used to recognize a silhouette is the Fourier transform of the contour [10]. Fourier descriptors represent the shape of the silhouette in the frequency domain. The lower frequency descriptors contain information about the general shape, and the higher frequency descriptors give indications about the smaller details. The benefits of Fourier descriptors are their robustness to scale and orientation variations.

Consider the closed contour of a silhouette composed of successive pixels. The centroidal radius function presented in Equation (4) defines the distance r_i of the contour points to the center of gravity of the silhouette.

$$r_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \quad (4)$$

Since this function is computed around the center of the silhouette, it is automatically translation invariant. The Fourier transform provides rotation invariance. We use the Discrete Fourier Transform to compute contour signature of the silhouette. The DFT magnitudes give a rotation invariant feature vector. Scale invariance is accomplished by normalizing all the DFT magnitudes by the first value of the DFT, as shown by Equation (5):

$$FD = \left[\frac{|F_1|}{|F_0|}, \frac{|F_2|}{|F_0|}, \ldots, \frac{|F_{N/2}|}{|F_0|} \right] \quad (5)$$

where N is the contour length and F_i the ith component of the Fourier spectrum.

Other Shape Descriptors

Moment invariants and Fourier descriptors do not perform well in the case of noisy images, like in our case. This is why in Castignolles [1] other shape descriptors taken from image processing are used. Our system uses nine of these features: area, perimeter, length, height, compactness, convex hull area, concavity rate, object area divided by the minimal rectangular area and the height divided by the length.

3.2 Classification

Our classification procedure uses a multi-classifier [6,9] approach to improve robustness of the classification. A multi-classifier approach consists of using several different classifiers with different properties and features. According to Suen et al. [9], this approach allows to use smaller classifiers, and their combined performance is higher than the best of the classifiers when all classifiers have similar performances.

Since our classification process has to be very robust, algorithms of different natures having different weaknesses and strength are used. In order to facilitate the addition of new species or the removal of the identification of some species based on knowledge about the fauna of a particular river (as specified by the system's operator), a modular classifier approach is also required. A modular classifier is a classifier in which each class can be trained independently and the trained parameters can be reassembled together in order to add or remove new classes. The classifiers used in our system satisfy this criterion. Modularity simplifies the recognition and makes sure the analysis algorithm is not trying to fit a silhouette of a species not possible to find in a particular river, improving also the robustness of the system.

Three types of classifiers are used: Bayes maximum likelihood classifier [2] with a quadratic discrimination rule, learning vector quantization (LVQ), and multi-layer perceptron (MLP) trained with backpropagation. The trained classifier parameters for each class are the mean and covariance matrices of the Bayes Maximum-Likelihood, the codebook of the species for the LVQ, and the sub-networks of the species for the OCON.

Bayes Maximum Likelihood Classifier

The maximum likelihood classifier is a classical parametric classifier that relies on second order statistics of Gaussian probability density model for each class [2]. The class probability is being assumed to be Gaussian the discriminant function is given in Equation (6):

$$p(x|w_i) = \frac{1}{\sqrt{2\pi^d |\Sigma_i|}} e^{\left[-\frac{1}{2}(x-m_i)^t \Sigma_i^{-1} (x-m_i) \right]} \quad (6)$$

where d is the number of class, m_i is the mean vector for class i and Σ_i the covariance matrix of class i.

Each new image is classified to its most likely class of membership. This can be achieved by classifying each silhouette to the class with the highest probability density function or the highest *a posteriori* probability of membership.

Learning Vector Quantification

Learning vector quantification process is a supervised vector quantization developed by Kohonen [5]. LVQ as the advantage of making weaker assumptions about the shapes of underlying feature vector distribution than statistical classifiers.

A vector quantization process optimally allocates M codebook reference vectors, $v_i \in R^n$, to the space of n-dimensional feature vectors, $x \in R^n$. The local point density of the allocated v_i can be used to approximate the probability density function $p(x)$ [5].

First, the codebook references v_i are initialized by random feature vectors. Each vector is assigned to a class. The number of vectors in each class must be proportional to the *a priori* probability of the class. A training algorithm is then used to optimize the code book vectors.

Let the input training vector x belong to class C_t and the closest codebook vector v_i labeled as class C_i. The codebook vector is updated by the learning rules shown in Equation (7):

$$\begin{aligned} \Delta v_i &= \alpha(x - v_i) && \text{if } C_i = C_t \\ \Delta v_i &= -\alpha(x - v_i) && \text{if } C_i \neq C_t \\ \Delta v_k &= 0 && \text{for } k \neq i \end{aligned} \tag{7}$$

with α as the learning rate. Only the closest of the vectors v_i is updated, with the direction of the correction depending on the correctness of the classification. Effectively, these codebook vectors are either pushed toward the good classification region or pulled away from zones where misclassifications occur.

After training, the nearest neighbour rule is used to classify the input test vector according to the class label of its nearest codebook vector. The Euclidean distance between the input vector x and the codebook vector v_I is used.

Neural Network

Neural networks are among one of the most commonly used classifiers for pattern recognition problems. Multi-layered perceptron (MLP) trained with backpropagation has good generalization capability and can recognize even non-linear distribution [3]. Despite their advantages, they suffer from some very serious drawbacks that make their use for certain kinds of problems very difficult. One of these problems is the lack of flexibility in the case of pattern recognition. When a new class is added, the network has to be retrained. The One-Class-One-Network (OCON) [7] version of the MLP overcomes this drawback, making the classifier modular.

In this approach, each class is modelled independently by a fully connected MLP of a fairly small size. Figure 5 shows the structure of OCON. Each of these sub-networks specializes in only one class and has only one output neuron. The highest output of all the sub-networks is chosen as the final decision.

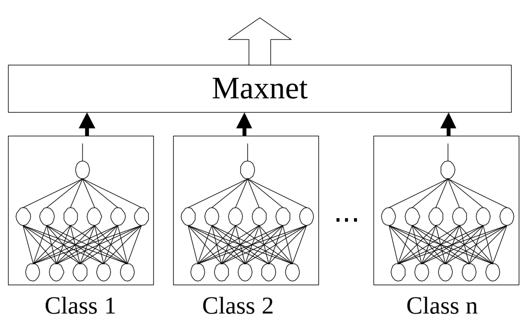

Figure 5 - Structure of OCON.

One advantage of using an OCON architecture is that by using smaller sub-networks, the training process becomes computationally less intensive compared to using a bigger network for all classes.

3.3 Classifier Fusion

As shown in Figure 6, classifications made by the classifiers must be combined to derive a final decision. The idea behind this approach is that even though each classifier has different performances, each one may not be able to identify correctly the same patterns. The fusion algorithm used is the majority vote [9] because of its simplicity, the fact that it does not need a lot of *a priori* knowledge and that it support modularity.

The majority vote is based on the following rule: classify an image as being in class i if at least two of the three classifiers makes the same decision. The images that do not satisfy this rule are rejected and submitted to the operator for manual classification/rejection.

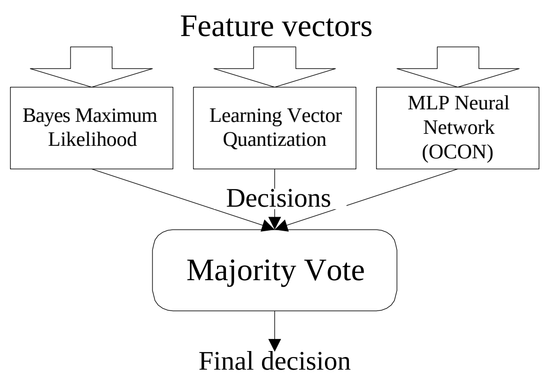

Figure 6 - Combination of classifiers.

4 Experiments

Experiments reported in this paper evaluate the usefulness of the three types of features extracted for characterizing fish silhouettes, and the performance of the three types of classifiers using different combinations of the possible features. We also validate the addition of new species to the analysis algorithm.

Real fish silhouettes are used in these experiments. These silhouettes were obtained by having five specific species of fish passing through the fishway at the Hydro-Québec's Rivière-des-Prairies dam near Montreal. Table 1 summarizes the number of silhouette images acquired. These images were then separated into a training set of 60 images and a test set made of the remaining images, for each species. Section 4.1 describes classification results obtained using all except the Largemouth Bass data. These data were used to test the ability of the analysis algorithm to classify new specifies, as presented in Section 4.2.

TABLE 1 – NUMBER OF IMAGES OF THE FIVE FISH SPECIES USED

Species	Number of images
Bowfin (*Amia calva*)	106
Copper Redhorse (*Moxostoma huddsi*)	354
Largemouth Bass (*Micropterus salmoides*)	208
Northern Pike (*Esox lucius*)	182
Walleye (*Stizostedion vitreum*)	350

Before classification, each classifier is created using the trained sets of species selected by the user. Several tests were made to find adequate configurations for each classifier. The mean vectors and covariance matrices of each of the selected species are put together to form the Bayes Maximum-Likelihood classifier. The quadratic discriminant law is used for the Bayes Maximum-Likelihood classifier. The LVQ classifier uses 15 vectors per class in the codebook, a learning rate of 0.01 and 1000 epoch of training. The OCON neural network uses 15 hidden neurones, 1200 epoch of training and learning rate of 0,01. Each sub-network is specialized in recognizing a specific species.

4.1 Classification Results

First, we investigated several sets of the features and selected the one that appears to be the most appropriate. These feature sets are:

I. the 7 moment invariants;
II. the shape descriptors;
III. the Fourier descriptor;
IV. the first two moments invariants and the shape descriptors;
V. the first three groups of features (feature set I+II+III).

Table 2 shows classification results obtained using these feature sets with each type of classifiers. Only using moment invariants features does not give good results. This may be attributed in part to the fact that all fish have the same basic shape and differ only in the details for which the moments seem to be not well adapted. For the other feature sets, the OCON classifier gives the best results on the training set and also on the test set, except for the feature set IV. The best performances for Bayes classifier is obtained using this feature set. As for Castignolles [1], this feature set gives very good results. But for LVQ and OCON, the best results are obtained using feature set V. This may be cause to the greater variety of features considered to make the classification.

TABLE 2 - CLASSIFICATION RESULTS

Feature set	Training set			Test set		
	Bayes	LVQ	OCON	Bayes	LVQ	OCON
I	39.3%	48.1%	46.9%	38.5%	39.6%	38.6%
II	64.7%	68.1%	82.8%	57.3%	59.4%	61.5%
III	37.8%	62.5%	76.6%	33.3%	52.1%	63.5%
IV	71.9%	63.4%	87.5%	69.8%	60.4%	66.7%
V	69.1%	67.8%	95.0%	57.3%	62.5%	75.0%

Table 3 shows results obtained when classifier fusion is done. The performance on the training set is lower than the best results obtained for individual classifiers. This is probably caused by the fact that individual performances of each classifier are quite different, and the majority vote method requires two identical results out of three to classify a silhouette. This confirms the indication given in [9] that performances of a multi-classifier approach increase only when all classifiers have similar performances. But while this is true for the training set it is not for the test set, and performances on the test set are more important since they validate the generalization ability of the system. Except for feature set III, the combined classification results are better than for individual classification. The best overall result is 77.88%, obtained using feature set V. For comparison, according to an informal survey, the video analysis of images by an expert has a recognition rate of around 80%. Therefore, the results of the best overall compare well with human performance.

TABLE 3 - RESULTS USING CLASSIFIER FUSION

Feature set	Multi-Classifier Fusion	
	Training set	Test set
I	48.13%	41.67%
II	75.62%	70.83%
III	69.69%	57.29%
IV	81.87%	72.96%
V	88.75%	77.88%

4.2 Adding a New Species

Using the feature set V and the complete classification process (i.e., with classifier fusion), we also evaluated the performance of the system when a new species is added, the Largemouth Bass. We wanted to see the effect on the recognition performance of the previously trained parameters, when adding parameters for new species to the classifiers. The parameters of species are the mean and covariance matrices of the Bayes Maximum-Likelihood, the codebook of the for the species for the LVQ and the sub-network of the species for the OCON. Some reduction of performance can be expected because the parameters of the already trained species were not trained with examples of the new species. For example, the trained classifiers parameters for Northern Pike have not been trained not to recognize Largemouth Bass as Northern Pike.

Two tests were done. First, a new set of classifiers (Bayes maximum-likelihood, LVQ and OCON) were trained to recognize the Largemouth Bass. Negative examples from the Copper Redhorse and the Walleye were also included to the training and test sets. The second test was the evaluation of the performance of the combination of parameters of the trained classifiers for the Largemouth Bass with the ones of the four-species classifier trained with feature set V of the section 4.1.

TABLE 4 – ADDITION OF A NEW SPECIES

Test	Multi-Classifier Fusion (V)	
	Training set	Test set
Largemouth Bass + Copper Redhorse + Walleye	81.8%	72.7%
All species	80.3%	70.8%

Table 4 reports the results. The first line presents the results of training the new species, the Largemouth Bass, and the other two species. The performance is a little less than the best result presented in Table 3. This may be due to the that because both the performance of OCON and LVQ depend on their random initialization. Therefore the performance of these to algorithms may vary a little from training to training.

The second line presents the results of assembling the trained parameters of the four species trained in section 4.1 with the parameters for the Largemouth Bass trained in the first part of this section. As discussed earlier in this section, the performance when the new parameters are added are a little reduced compared to the results in section 4.1. This is probably due to the fact the parameters of the four species trained in 4.1 were not negatively trained with the new species (Largemouth Bass). For example, they were not trained not to recognize Largemouth Bass as Northern Pike. Therefore some of the Largemouth Bass are recognized as some of the other four species.

This reduction is acceptable because the retraining of the classifiers are not necessary every time new species need to be recognized. In practice, the problem could be substantially reduced by training a classifier for new species including the most common species in the training sets. This nonetheless proves the possibility of using modular classifiers for building a flexible classification system with satisfying performance.

5 Conclusion

This paper describes an intelligent system for automated fish counting by species. Our system is based on a multi-classifier approach that fuses classification made by Bayes maximum likelihood classifier, Learning Vector Quantification classifier and One-Class-One-Network (OCON) neural network classifier, using a majority vote method. Their combined classification results in better overall performances. The approach helps increase the robustness of the system, and facilitates its configuration according to the specific fauna of the river.

In future work, we will further improve the recognition performance and the robustness of the system by using a feature selection algorithm to optimize the selection of relevant characteristics for fish classification, and by increasing the ability of the system to reject silhouettes that are not from fish.

Acknowledgment

The authors would like to thank Dr. R. Verdon, Hydro-Québec, VPTI and FCAR for making this project possible. F. Michaud is also supported by NSERC, FCAR and CFI.

References

[1] Castignolles, N., *Automatisation du comptage et de la reconnaissance des espèces dans les passes à poissons par l'analyse d'images*, Ph.D. Thesis (in French), U. Paul-Sabatier de Toulouse, 1995.

[2] Duda, R.O., Hart, P.E., *Pattern Recognition and Scene Analysis*, New York, John Wiley and Sons, 1973.

[3] Haykin, S. *Neural Networks – A Comprehensive Foundation*, New York, Prentice Hall, 1999.

[4] Hu, M.K., "Visual pattern recognition by moment invariant", *IRE Transactions on Information Theory*, vol. 1T-8, p. 179-187, 1962.

[5] Kohonen, T. *Self-Organization and Associative Memory*, 3rd Edition, Springer-Verlag, 1989.

[6] Loonis, P., *Conception d'un prototype de trieuse automatique de poissons par vision artificielle en milieu industriel*. Ph.D. Thesis (in French), Université de La Rochelle, 1996.

[7] Pandya, A., Macy, R. *Pattern Recognition with Neural Networks in C++*, Boca Raton, CRC Press, 1995.

[8] Strachan, N.J.C., Nesvada, P., Allen, A.R., "Fish species recognition by shape analysis of images", *Pattern Recognition*, vol. 23, no. 5, p. 539-544, 1990.

[9] Suen, C.Y., Xu, L., Krzyzak, A., "Methods of combining multiple classifiers and their applications to handwriting recognition", *IEEE Trans. on Sys., Man and Cyb.*, vol. 22, n° 3, p. 418-435, 1992.

[10] Zahn, C., Roskies, R.Z., "Fourier descriptors for plane closed curves", *IEEE Trans. on Computer*, vol. 21, no. 3, p. 269-281, 1972