

Determining the appropriate feature set for fish classification tasks

M. S. Nery, A. M. Machado
Pontifícia Universidade Católica de Minas Gerais
Rua 8, 200 - Jardins, Guanhães, Brasil
msnery,alexei@pucminas.br

M. F. M. Campos, F. L. C. Pádua, R. Carceroni
Universidade Federal de Minas Gerais
Av. Antônio Carlos, 6627, Belo Horizonte, Brasil
{mario,cardeal,carceroni}@dcc.ufmg.br

J. P. Queiroz-Neto
Centro Federal de Educação Tecnológica
Av. Danilo Areosa, 138, Manaus, Brasil
pinheiro@cefetam.edu.br

Abstract

We present a novel fish classification methodology based on a robust feature selection technique. Unlike existing works for fish classification, which propose descriptors and do not analyze their individual impacts in the whole classification task, we propose a general set of features and their correspondent weights that should be used as a priori information by the classifier. In this sense, instead of studying techniques for improving the classifiers structure itself, we consider it as a "black box" and focus our research in the determination of which input information must bring a robust fish discrimination. All the experiments were performed with fish species of Rio Grande river in Minas Gerais, Brazil. This work has been developed as part of a wider research [3], which has as main goal the development of effective fish ladders for the Brazilian dams.

1 Introduction

Careless interventions of all sorts human kind has been imposing to its environment have severely compromised the existence of several living beings. It has also deeply altered the distribution and abundance of the native life in general, and specially the fish fauna has suffered great loss all over the world due mainly to water pollution and predatory fishing practices. Resource exploitation such as mining and forestry and the construction of dams for hydroelectric power generation are some of activities with widespread, and oftentimes dramatic, impact on local fish biodiversity. As a result, extensive changes in aquatic life break down reproductive isolation between some species [1, 2].



Figure 1. The observation window at the fish ladder located at UHE-Igarapava, Minas Gerais, Brazil.

As an attempt to minimize the aforementioned problems, various legal provisions have been made available in Brazil, more specifically those caused by hydroelectric impoundments. In some states construction of fish transposition mechanisms – also known as *fish ladders* – is now mandatory. These channel like structures are built to assist fishes in their upriver journey to spawning grounds (a natural phenomenon called *piracema* in Brazil) [1]. Besides the direct benefits, those fish transposition mechanisms provide adequate means for acquisition of relevant information about the migration and other relevant features of different fish species (see Figure 1), such as swimming ability, time of migration and peak flow rates.

Several design issues must be considered in the design and placement of a fish ladder. Every blockage in a river course ushers in many environmental concerns and each case is unique as far as biology is concerned. Furthermore, spawning behavior may be quite different among species within the same ecosystem. These are but a few problems that need to be further understood before a biologically correct fish ladder is built. Therefore, as an effort to understand the benefits and

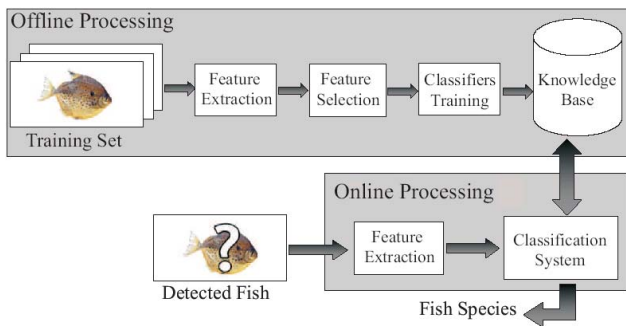


Figure 2. Block diagram of the fish classification system developed by the CTP Research Group [3].

drawbacks of fish ladders and due to the vast amount of data that needs to be collected and analyzed, we have been working on vision based systems to assist fish biologists to better understand the impacts of such mechanisms. The present work was developed as part of a wider research effort [3], and its goal was to provide efficient vision-based techniques for automatic acquisition of accurate information about the behavior of fish species that swim upstream a fish ladder. Based on our methodology for feature selection, a robust fish classification system has been developed, a block diagram of is depicted in Figure 2. With the information provided by that system, biologists may provide engineers with relevant clues which may be fundamental in the design of fish transposition mechanisms.

Object classification problem lies at the core of the task of estimating the prevalence of each fish species. In other words, the problem is to accurately identify and classify a fish according to its species. We understand that any general solution for automatic fish classification should adequately tackle at least the following:

- *Arbitrary fish size and orientation:* fish size and orientation are unknown *a priori* and can be totally arbitrary;
- *Feature variability:* some features may present large differences among different fish species;
- *Environmental changes:* variations in illumination parameters, such as power and color and water characteristics, such as turbidity, temperature, not uncommon. The environment can be either outdoor or indoor;
- *Poor image quality:* image acquisition process can be affected by noise from various sources as well as by distortions and aberrations in the optical system;

- *Segmentation failures:* due to its inherent difficulty, segmentation may become unreliable or fail completely.

The wealth of species present in most Brazilian rivers transform the identification task into an extremely hard problem. Therefore, in order to make the problem tractable, the project was to focus on only a few species which are either of significant economic impact, or those which are in fear of extinction. Our approach was to unburden the classifier as much as possible by a careful selection of meaningful features. As a step toward this goal, we present a methodology to select from a wide set of features, those which significantly improve the classifiers' performance under all but the last one of the aforementioned conditions.

It is assumed, without loss of generalization, that the input of our system is the image of a single fish. This precludes a previous segmentation step that is able to provide such images in spite of noise and the existence of occlusions in the scene. Figure 1 is a typical image acquired through the observation window at the fish ladder of the UHE (hydroelectric plant)-Igarapava, in the state of Minas Gerais, Brazil, showing some of the fishes that are classified by our system.

The vast majority of the works on fish classification reported in the literature such as [4, 5, 6] approach the problem fundamentally in the same way, that is, from the classifiers perspective. They typically use small groups of features with no previous thorough analysis of the individual impacts of each factor in the classification accuracy. Feature selection is achieved by brute force approaches where the classifier is trained and its performance is measured for each subset of N possible feature combinations. The computational cost is clearly very high, since the number of tests is the cardinality of the set defines by $\mathcal{P}N$. The contributions of those works are typically proposals of new classification systems as well as improvements on existing classification techniques. Feature selection, albeit mentioned, is not developed in any substantial manner.

In [4], an infrared silhouette sensor is used to acquire contours of fish in constrained flow. Classification is based on the combined results of three different classifiers which use invariant moments and Fourier boundary descriptors for fish silhouette recognition. Those features, however, do not perform well with noisy images – which is our case. The authors report a classification accuracy of about 78%. In [6] a fish weight automated sorter system based on background subtraction was developed. The proposed solution was designed exclusively for fish scanning processes and is strongly affected by environmental variations and noise of the image acquisition system. The technique presented in [5]

and the one described in [6] are the most closely related to our approach, but we present a more complete and thorough feature selection methodology. As it will be later demonstrated by our experimental results, even very simple classifiers might be able to achieve reasonable performance levels if the proper set of features is carefully chosen using our methodology.

The remainder of this paper is organized as follows: Sections 2 and 3 cover our feature extraction and selection methods, respectively. Experimental results are presented in Section 4, followed by the conclusions and discussion in Section 5.

2 The Feature Extraction Approach

As a first step we have set out to determine a largest set of features. For each fish species, we have computed 47 different features, which can be divided in four main groups, differentiated by the type of extracted information. Those groups and their corresponding numbers of features are:

- *Size measurements*: 4 features;
- *Shape measurements*: 19 features;
- *Color signatures*: 8 features;
- *Texture measurements*: 16 features;

The computation of the above-mentioned features was performed based on fish“masks”, which were obtained by applying morphological operations such as dilation and erosion. In what follows, we briefly describe each group of features that were analyzed in our work.

2.1 Size Measurements

This group consists of planar measurements on the fish mask *area*, the *center of mass* of the mask, its *perimeter*, its *length* and *width*. Although these features are not invariant under translation, rotation and scale, they are fundamental for computing other relevant features that will be described next.

2.2 Shape Measurements

This group is composed of the following features:

- *Rectangle fit factor*: indicates how close is the object to the *Minimum Enclosure Rectangle* (MER). As the object’s shape gets closer to a rectangle this factor tends to 1;

- *Aspect ratio*: given by the ratio between the MER’s length and width. This feature is commonly used to differentiate gaunt objects from rounded or rectangular ones;
- *Circularity*: computed by using the squared perimeter and the area of the object. This feature captures the contour complexity of the segmented object;
- *Moments*: describe the image contents (or distribution) with respect to its main axes. They are designed to capture both global and detailed geometric information about the image. Depending on the *kernel function* used [7], the estimated moment may uncover a range of useful properties, producing descriptions which can be invariant under rotation, scale, translation and orientation. In this work we compute the zeroth, first, second and third order moments for each segmented fish, which are further used to compute the seven Hu moments [8, 9].

2.3 Color Signatures

According to studies of fish biologists [10], the dorsum and ventral colorations constitute very important features that might be used to discriminate different fish species. Based on this fact, we use this information by assigning to each fish species a color signature which is composed of the average color of the dorsum and the ventral region of the fish. Both YUV and HSI color models were used, since they provide luminance and chrominance information in separated bands.

2.4 Texture Measurements

Some authors [8, 7] divide texture computation techniques into three main groups, namely, statistical techniques, syntactic methods and spectral methods. Statistical techniques are interesting alternatives due to their generally lower computational costs. Moreover, those methods do not use color information and are based on a set of energy functionals that are responsible for characterizing each type of texture.

In order to compute texture measurements we use co-occurrence matrices. A co-occurrence matrix contains the probability that a gray level configuration will come up in an arbitrary pair of neighboring pixels in the image. From these matrices we compute two descriptors: *Inertia* and *Energy* [7].

3 The Feature Selection Approach

Feature selection may be viewed as an optimization problem of exponential time complexity along several dimensions. Many feature selection algorithms have been developed and systems have been deployed for real-world applications [11]. However, there exists a gap between theoretical results and what practice reveals. The proliferation of feature selection algorithms has not yet produced a methodology that is both general and at the same time eases the burden of selecting the best features. Feature selection is at the very core of our methodology for fish classification, and it is based on the statistical concepts of *discrimination* and *uncorrelatedness* [12].

Discrimination of a feature measures how well-separated are the clusters that represent each class (fish species, in the present case), when only that feature is used for classification. More specifically: given a feature x , its *discrimination* value \hat{D} is computed by Equation 1, where j and k represent two different classes and $\hat{\mu}$ and $\hat{\sigma}$ are the estimated mean and variance of x on each class, respectively:

$$\hat{D} = \frac{|\hat{\mu}_{xj} - \hat{\mu}_{xk}|}{\sqrt{\hat{\sigma}_{xj}^2 + \hat{\sigma}_{xk}^2}} \quad (1)$$

Two features are *uncorrelated* when they are not statistically correlated. In order to evaluate the *correlation* \hat{I} between two features x and y , we use the Equation 2, where j represents a specific class (fish species), i is the number of patterns (fish examples) in that class, $\hat{\mu}$ and $\hat{\sigma}$ are the estimated means and variances, respectively:

$$\hat{C} = \frac{\frac{1}{N_j} \sum_{i=1}^{N_j} (x_{ij} - \hat{\mu}_{xj})(y_{ij} - \hat{\mu}_{yj})}{\hat{\sigma}_{xj}\hat{\sigma}_{yj}} \quad (2)$$

The possible values of this *correlation* metric falls within the range [-1,1], where a value close to +1 indicates a strong correlation, zero indicates that the two features are not correlated and a value close to -1 indicates that one feature is proportional to the complement of the other one.

The algorithm may be shortly described as a three step process: i) compute the discrimination using Equation 1 for all features and sort the resulting vector in descending order; ii) build a graph G , where nodes are features and edges represent the pairwise uncorrelated (as per Equation 2) between two features; iii) coalesce each clique in the graph to its representative node, chosen by the one with larger discriminating value. The result is the set of “optimal” features.

Algorithm 1 Feature ranking approach

```

Create an empty Vector
for each feature  $x$  do
   $Sum \leftarrow 0$ 
   $K \leftarrow 0$ 
  for each fish species  $i$  do
    for each fish species  $j$  do
      if  $i \neq j$  then
        Compute  $\hat{D}$  (Equation 1)
      end if
      if  $\hat{D} \geq T$  then
         $K \leftarrow K + 1$ 
         $Sum \leftarrow Sum + \hat{D}_{ij}$ 
      end if
    end for
  end for
   $Mean \leftarrow Sum/K$ 
   $Vector[x] \leftarrow [K, Mean]$ 
end for
Order Vector by  $K$  as primary key and by  $Mean$  as secondary key

```

In the first step, the *discrimination* value is computed for 34 features described in Section 2. They are: the seven Hu moments, aspect ratio, circularity, rectangle fit factor, eight color signatures and 16 texture measurements. These features are computed based on the 13 remaining descriptors. The output of this first step is a database in decreasing order of discrimination value. Figure 3 illustrates the discrimination and uncorrelated values contained in that database for some of the extracted features.

A slightly modified version of steps ii) and iii) above was used here. Of the 34 features, the correlation computation pruned off 6 highly correlated ones, namely, the texture’s first, third, fourth, sixth, seventh and eighth co-occurrence matrices, reducing the cardinality of the feature set to 28. Then Algorithm 1 the was used to rank the 28 remaining features in decreasing order of *discrimination*. Based on both Equation 1 and our database values, a parameter T was chosen to be 1, since a *discrimination* larger than 1 gives a strong indication that the corresponding tuple is composed by fish species that can be told apart by the feature.

Finally, Algorithm 1 enabled us to determine a meaningful set of features and a ranking responsible for indicating their corresponding *weights* (quantitative influences) in the whole classification process. Now, the next step consists in using such information for improving a classifier’s performance. In our work, we consider the classification system as a “black box” and in the next section we briefly describe the solution adopted.

4 The Classification System

In order to illustrate the robustness of our feature selection methodology, we use in this work a

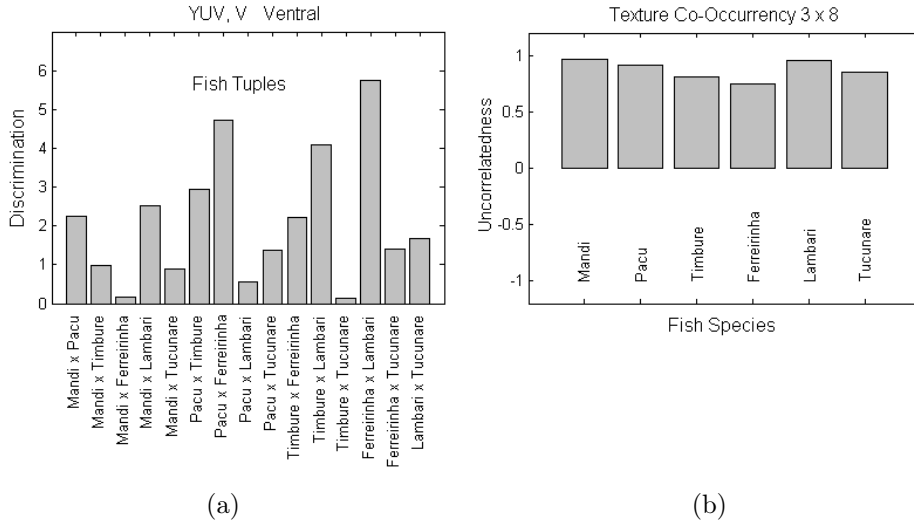


Figure 3. (a) Color signature’s discrimination of the fish’s ventral part. Color information used: V channel of the YUV color model; (b) Statistical uncorrelatedness between two co-occurrence matrix-based textures (*inertia* descriptors), being the computation of the first one based on pixels (i, j) and $(i+1, j+1)$ and the computation of the second one based on pixels (i, j) and $(i+1, j+2)$.

Bayesian classifier, whose effectiveness has already been proved in many different pattern classification problems [12, 13]. Equation 3 represents the Bayesian decision functions for Gaussian classes under the 0-1 loss function condition:

$$d_j(\mathbf{x}) = \ln P(\omega_j) - \frac{1}{2} \ln |\mathbf{C}_j| - \frac{1}{2} \left[(\mathbf{x} - \mathbf{m}_j)^T \mathbf{C}_j^{-1} (\mathbf{x} - \mathbf{m}_j) \right] \quad (3)$$

where \mathbf{x} represents an input pattern, ω_j represents a specific class ($j = 1, 2, \dots, M$, being M the number of classes) and \mathbf{m}_j and \mathbf{C}_j are the mean vector and the covariance matrix of class j , respectively.

As a Bayesian classifier makes the assumption that its numeric input attributes (features) are Gaussian random variables, we have previously analyzed the probability distribution of each one of the 28 selected features. Since we do not have *a-priori* knowledge about the actual mean and variance of our data set, we used the Lilliefors method to perform the necessary normality test of our input data. By using that technique, we identified 8 features that could not be considered as Gaussian variables, namely, the 8 co-occurrence matrix-based texture descriptors (e.g. all energy texture descriptors).

Therefore, in order to use the Bayesian classifier, we reduced our database to a final set of 20 features. The features are listed below in decreasing order of relevance to the classification process, according to the

feature selection methodology presented in Section 3. (1) color signature given by the mean of the V channel values (YUV color model) in the ventral part of the fish, (2) circularity, (3) color signature given by the mean of the U channel values (YUV color model) in the dorsum part of the fish, (4) first order Hu moment, (5) second order Hu moment, (6) color signature given by the mean of the H channel values (HSV color model) in the ventral part of the fish, (7) aspect ratio, (8) third order Hu moment, (9) sixth order Hu moment, (10) color signature given by the mean of the S channel values (HSV color model) in the ventral part of the fish, (11) color signature given by the mean of the H channel values (HSV color model) in the dorsum region of the fish, (12) fourth order Hu moment, (13) co-occurrence matrix-based texture (*inertia* descriptor with its computation based on pixels (i, j) and $(i, j+2)$), (14) color signature given by the mean of the V channel values (YUV color model) in the dorsum region of the fish, (15) color signature given by the mean of the S channel values (HSV color model) in the dorsum region of the fish, (16) co-occurrence matrix-based texture (*inertia* descriptor with its computation based on pixels (i, j) and $(i, j+1)$), (17) fifth order Hu moment, (18) color signature given by the mean of the V channel values (YUV color model) in the ventral part of the fish, (19) rectangle fit factor and (20) seventh Hu moment.

Based on the above *a-priori* information, we performed experiments with the Bayesian classifier, in order to demonstrate the effectiveness of our feature se-

lection technique for improving the classifier’s performance. The experimental results are discussed in the next section.

5 Experimental Results

In order to make the fish classification problem even more challenging and to demonstrate the robustness of the designed classification system against significant variations of the 3D fish orientation, we used fish perfectly preserved in formaldehyde solutions and, by employing auxiliary equipment, we rotated those individuals about the z axis, as illustrated in Figure 4, from -40° to 40° in 10° steps. Those rotations were performed by taking initially both right and left sides of the fish parallel to the camera’s image plane. The maximum rotation angles (40° and -40°) were chosen according to fish biologists observations about the swimming characteristic of those species.

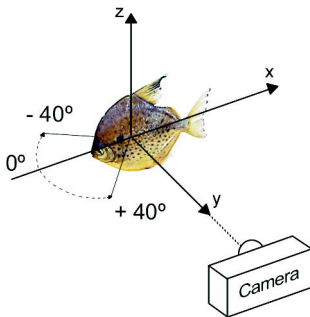


Figure 4. Fish rotation about the z axis.

In Table 1 we list the six fish species classified in this work as well as their corresponding number of fish examples and images acquired per individual. The feature selection methodology validation was performed by a straightforward technique. All 225 images were randomly divided into five subsets – S_1 through S_5 , under the constraint that each one should contain at least five images of each fish species. Then, subset S_i was used to train the classifier and experiments were conducted for each S_j , for $j \neq i$, for $i, j=1..5$. For each trial, the classifier’s accuracy is given by the total number of fishes that were correctly identified, that is, were correctly defined as belonging to its actual class. The classifiers overall accuracy was computed as the mean of the averages of each trial.

Firstly, in order to perform a careful analysis of the accuracy, efficiency and scalability of our approach, we have performed three main experiments with image sets acquired under controlled conditions in the laboratory. Those experiments are differentiated by the training and validation classes used, the feature sub-

Fish Species	individuals	images per individual
Mandi (<i>Pimelodus Fur</i>)	2	18
Pacu (<i>M. Maculatus</i>)	3	9
Tucunaré (<i>Cichla</i> Spp.)	2	18
Lambari (<i>A. Altiparanae</i>)	2	18
Timburé (<i>L. Amblyrhynchus</i>)	3	18
Ferreirinha (<i>L. Octofasciatus</i>)	2	18

Table 1. The classified fish species with their corresponding number of fish examples and images per individual.

set considered for representing a specific fish and the presentation order of each feature subset to the classifier. Finally, in order to demonstrate the applicability of our approach, we tested it with several challenging real-world images acquired at the fish ladder in the UHE-Igarapava. In the following, we present and discuss their corresponding results.

Experiment 1 As a first step, we trained the classifier with all six fish species so that it was able to identify each one of them. Figure 5 illustrates the classifier’s change in accuracy as features in decreasing order of discrimination were included in an one by one basis. The maximum classification accuracy (81%) was achieved by presenting to the classifier fishes represented by the 4 most discriminant features. The inclusion of further features decreased classification down to an accuracy of 50%. These results are in agreement with the literature [13], which shows that the average error rate is strongly related to the number of features and to the number of image samples. Also, for a specific set of image samples, the average error rate increases with the increasing number of features.

Experiment 2 In a second experiment, we trained the classifier to discriminate among four different classes: Mandi, Pacu, Ferreirainha and a fourth class composed of other fishes (*e.g. the "rest" of the world*).

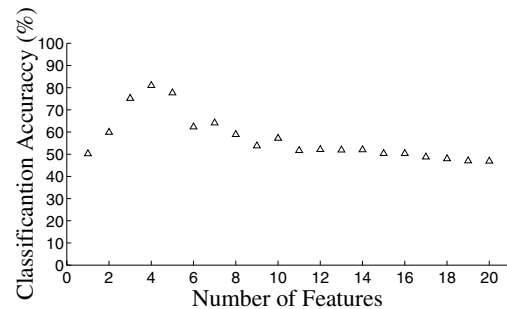


Figure 5. Experiment 1: Variation of the classification accuracy. The classifier was trained with the six fish species.

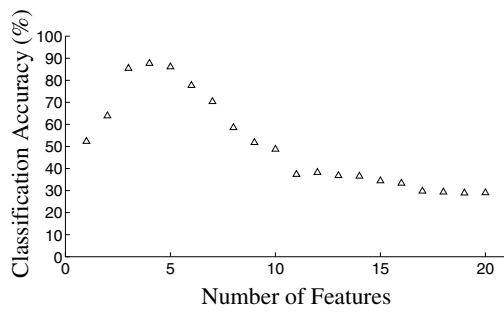


Figure 6. Experiment 2: Variation of the classification accuracy. The classifier was trained with three fish species (Mandi, Pacu and Ferririnha), while the others were considered as a fourth different class.

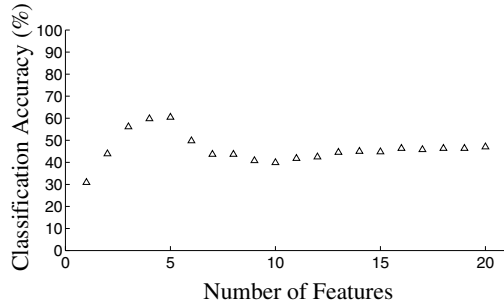


Figure 7. Experiment 3: Variation of the classification accuracy. The classifier was trained with six fish species.

As in Experiment 1, the feature subset was gradually increased. Figure 6 depicts the classifier’s accuracy, where a behavior similar to Experiment 1 was observed. Maximum accuracy (87%) was achieved by representing a fish with only four features. However, now a sharper decrease in classifier’s performance was observed when less discriminant features were added. Differently from the first experiment, the classification’s accuracy stabilized around a smaller value (30%). This behavior may be explained by the fact that three different fish species, which are described by completely distinct features, were associated to form a single class. Under this kind of training, a substantial overlap may exist among clusters which may dramatically decrease the classification’s performance.

Experiment 3 In this experiment the classifier was trained with the same six fish species used in Experiment 1, but here the features were included in increasing order of discrimination. Figure 7 depicts the classifier’s accuracy. Observe that the maximum accuracy achieved was only 61%, a value 20% smaller than the maximum obtained in the Experiment 1. This experiment together with the first one, demonstrate how our feature selection methodology can significantly improve

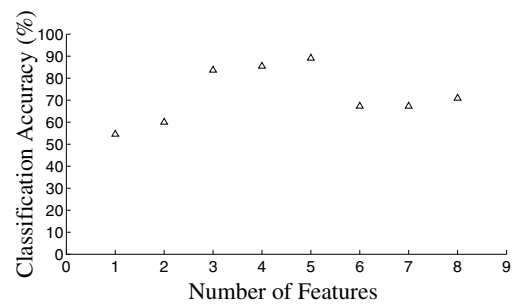


Figure 8. Experiment 4: Variation of the classification accuracy. The classifier was trained with four fish species: Mandi, Ferririnha, Piau and Tucunaré. The first two were the classification targets and the accuracy achieved was 87.3% with five features.

the classification’s performance. Using our technique the impact of each feature to the classification process as a whole is individually evaluated, so that the best features can be appropriately combined to optimize the classifier’s response.

Experiment 4 In this experiment the classifier was trained with image of actual fish species captured at the fish ladder in the UHE-Igarapava. We then applied the same methodology as in previous experiments which produced the following features, in relevance order: rectangularity, circularity, average U component of the ventral area, average U component of the dorsum area, aspect ratio, average H component of the ventral area, first and second order Hu moments. As expected, no texture features were selected by the method, since the co-occurrence matrices have shown to be very sensitive to rotation and scale changes. On the other hand, features related to the form (i.e. rectangularity and circularity), were most relevant. The classifier was trained to recognize the four main species, and based on biological and environmental criteria, only two – Mandi and Ferririnha–, were chosen by the fish biologists to be sorted out in this experiment. In order to validate the methodology other 55 images that were not presented to the classifier in the training process were used as inputs to the classifier. As far as the system developers are concerned, this experiment was completely "blind" since the contents of the data set was totally unknown to them. The automatic classification results were given to the fish biologists for verification.

The eight features selected by our methodology where used by the classifier in this experiment. Figure 8 shows the classifier results for the four trained species starting from the more to less discriminant features. The best classification was obtained with a set consisting of four or five among the most discriminant features, which was consistent with the behavior pre-

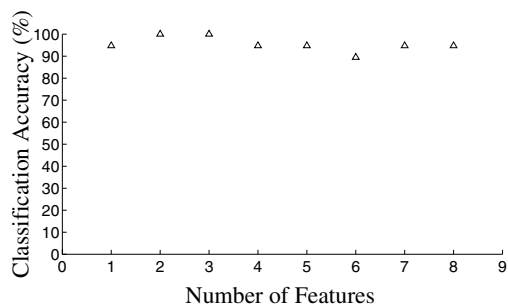


Figure 9. Experiment 4: Variation of the classification accuracy. The classifier’s analysis considers only Mandi species. Accuracy achieved was 94.7% with five features.

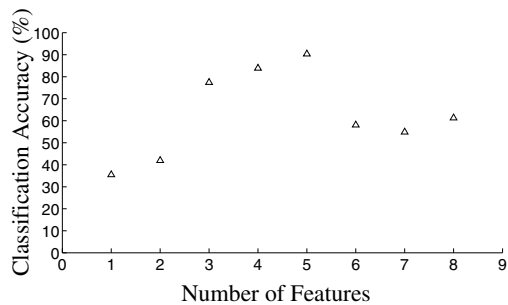


Figure 10. Experiment 4: Variation of the classification accuracy. The classifier’s analysis considers only Ferreirinha species. Accuracy achieved was 90% with five features.

sented in previous experiments. Figures 9 and 10 illustrate the analysis for each one of the target species. An accuracy of almost 90% was achieved which in this specific case was better than human results. For the Ferreirinha species, the accuracy was slightly smaller than that of the Mandi. This result may be explained by the fact that the shape of the other two trained species were more similar to the Ferreirinha than they were to the Mandi. Nevertheless, our methodology obtained very good results with real fish images and the few classification errors occurred among Ferrerinhas, Piaus and Tucunarés specially because of their shape similarity.

6 Concluding Remarks

Our experimental results suggest that our feature selection methodology can be successfully used to significantly improve the performance of fish classification systems. Unlike previous approaches which propose descriptors and do not analyze their impact in the classification task as a whole, we propose a general set of 20 features and their corresponding weights which may be used as *a priori* information by the classifier. Moreover, our work presents a novel set of features based on

color signatures associated to the dorsum and ventral average colors of a fish, which have proven to be fundamental for good classification results, including real images tests. A system for automatic fish counting, based on the results of our system was developed. The classification accuracy obtained in our methodology is so significant as in [4, 5, 6], inclusive considering all the noise problems in experiment with real images. We are currently investigating new feature selection techniques based on Factorial and Principal Component Analysis to improve both the accuracy and the robustness of our methodology.

Acknowledgments

The authors thank the support of CNPq-Brazil.

References

- [1] H. Godinho and A. Godinho. Fish Communities in Southeastern Brazilian River Basins Submitted to Hydroelectric Impoundments. *Acta Limnologica Brasiliensia*, 5:187–197, 1994.
- [2] E. Woynarovich. The Hydroelectric Power Plants and the Fish Fauna. *Verh. Internat. Verein. Limnol.*, 24:2531–2536, 1991.
- [3] C. de Transposição de Peixes. Endereço eletrônico: <http://www.ctpeixes.ufmg.br/>, 2004.
- [4] S. Cadieux, F. Lalonde, and F. Michaud. Intelligent System for Automated Fish Sorting and Counting. *Proceedings of the IEEE International Conference on Intelligent Robots and Systems*, 2000.
- [5] C. Bendall, S. Hibert, and G. Mueller. Experiments in Situ Fish Recognition Systems Using Fish Spectral and Spatial Signatures. Open File No.99-104, U.S. Depart. of the Interior - Bureau of Reclamation, 1999.
- [6] E. Dadios, A. Calderon, J. Ledesma, C. Que, and C. Sy. Knowledge Based Vision Guided Fish Weight Sorter. *Proceedings of the Philippines Computing Science Congress*, November 2000.
- [7] R. Gonzalez and R. Woods. *Processamento Digital de Imagens*. Edgard Blücher Ltda, São Paulo, 2000.
- [8] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing Analysis and Machine Vision*. PWS Publishing, Pacific Grove, California, USA, 2nd. edition, 1999.
- [9] M. K. Hu. Visual Pattern Recognition by Moment Invariants. *IRE Transactions Information Theory*, 8(2):179–187, 1962.
- [10] M. H. A. Keenleyside. *Diversity and Adaptation in Fish Behavior*. New York: Springer-Verlag, 1979.
- [11] Y. Kim, W. N. Street, F. Menczer, and G. J. Russell. *Data Mining: Opportunities and Challenges*. Idea Group Publishing, 2001.
- [12] K. R. Castleman. *Digital Image Processing*. Prentice Hall, Upper Saddle River, New Jersey, USA, 1996.
- [13] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis*. John Wiley & Sons, NY, USA, 1973.