

Research Article

Normalization of Active Appearance Models for Fish Species Identification

Charles-Henri Quivy and Itsuo Kumazawa

Department of Information Processing, Tokyo Institute of Technology, Nagatsuta-Cho, Midori-Ku, Yokohama, Kanagawa 226-850, Japan

Correspondence should be addressed to Charles-Henri Quivy, charleshenri.quivy@gmail.com

Received 12 January 2011; Accepted 7 February 2011

Academic Editors: C.-M. Kuo and L. Shen

Copyright © 2011 C.-H. Quivy and I. Kumazawa. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In recent years, automatic visual coral reef monitoring has been proposed to solve the demerits of manual monitoring techniques. This paper proposes a novel method to reduce the computational cost of the standard Active Appearance Model (AAM) for automatic fish species identification by using an original multiclass AAM. The main novelty is the normalization of species-specific AAMs using techniques tailored to meet with fish species identification. Shape models associated to species-specific AAMs are automatically normalized by means of linear interpolations and manual correspondences between shapes of different species. It leads to a Unified Active Appearance Model built from species that present characteristic texture patterns. Experiments are carried out on images of fish of four different families. The technique provides correct classification rates up to 92% on 5 species and 84.5% on 12 species and is more than 4 times faster than the standard AAM on 12 species.

1. Introduction

The evaluation of the impact of human activity on the environment is a recent concern. Nevertheless, coral reefs have been monitored all over the world during decades. This early interest in coral reefs is explained by three main factors: reefs are fragile and shelter rich animal communities studied by biologists, they are a revenue stream for tourism industry and supply food for local communities.

Standard monitoring techniques such as belt transect or rapid visual census [1] are based on statistical tables of fish density. Data collection is performed by trained volunteers at regular time intervals. Manual coral reef studies are difficult to conduct for diving is required and the results highly depend on the observers' ability. Hence, automatic data collection has been proposed, such as the Ecogrid project [2] that implemented underwater scene-streaming to biologists' desktop. However, despite the fact that the data acquisition difficulties have been partly overcome with remote sensors, it is still necessary to process the data automatically. Similarly to manual methods, it is required to

build statistical tables of fish concentration, which involves fast fish species identification techniques. Different methods have been proposed to perform live fish identification [3–5]. Nonetheless, most of the techniques are based on shape descriptors, which are not sufficient for separating species presenting similar shape and shape variations.

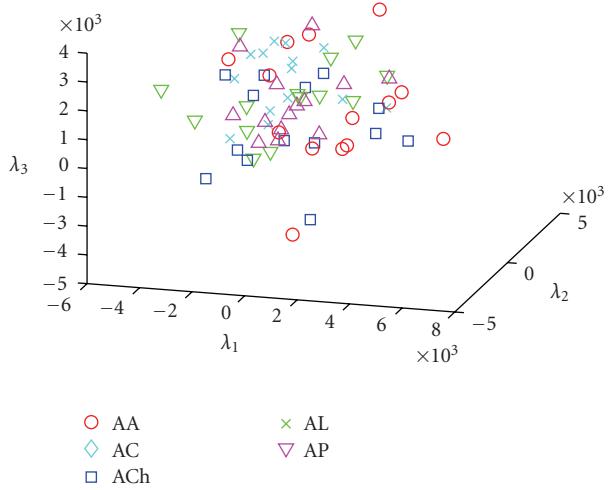
The Active Appearance Model (AAM) is a parameterize model of shape and appearance variations proposed for face recognition [6]. It is generally used for texture sampling. In that case, the sampled texture is a shape-free texture: it does not depend on the shape of the target. This is why this paper focuses on the AAM algorithm: regarding fish species identification, it is required to sample fish texture precisely because there are numerous similar species. The AAM is a statistical model built over an image data set. Each image of the data set includes an object of interest (face, bone) under the same conditions (angle, scale). It is required to manually annotate each object at specific feature points of its texture pattern, such as the eye pupils or eyebrows right and left limit in the case of face models (feature points determined by the operator). In the case of fish species identification, it

TABLE 1: Comparison table between the existing approaches for fish species identification.

Authors	Application	Inputs	Environment	Species	Recognition rate	Year
Zion et al. [7]	Polyculture	Moment Invariants	Constrained	3	91%	2000
Cadieux et al. [3]	Monitoring	Moment Invariants	Unconstrained	5	77%	2000
Storbeck and Daan [8]	Food Industry	Width/height	Constrained	6	95%	2001
Akg [4]	Monitoring	Texture	Unconstrained	9	76%	2003
Nery et al. [5]	Monitoring	Geometric, texture, color features	Unconstrained	6	81%	2005
White et al. [9]	Food Industry	Shape + Color	Constrained	3	99.8%	2006

TABLE 2: Presentation of the target families and species. Each column indicates the species that belong to the family presented in the first row (example: Chaetodon Auriga, Acanthurus Lineatus, Amphiprion Latezonatus).

Amphiprion	Acanthurus	Chaetodon	Pomacanthus
Akindynos (AA)	Achilles (AAC)	Auriga (CAu)	Annularis (PA)
Clarkii (AC)	Leucosternon (ALe)	Ephippium (CE)	Imperator (PI)
Chrysopterus (ACh)	Lineatus (ALi)	Lineolatus (CL)	Maculosus (PM)
Latezonatus (AL)	Sohal (AS)	Ocellicaudus (COc)	Semicirculatus (PS)
Polymnus (AP)	Tennenti (ATe)	Oxycephalus (COx)	Sexstriatus (PSe)
		Trifascialis (CTr)	Xanthometopon (PX)

FIGURE 1: Projection of the vectors of points sampled along the contours of Amphiprion's species shapes onto the associated three dimensional eigenspace. AA, AC, ACh, AL, and AP indicate species name's abbreviations. Each point's coordinates correspond to a parameter λ_X , (1) that represents a contour.

is possible to create species-specific AAMs, because texture patterns are similar amongst the same species. This approach performs well; however, its computation cost turns to be too high regarding the amount of data to process. Using an AAM for each family and a hierarchical approach is a way to speed up the identification. Nevertheless, labeling images of fish of different species is very difficult, because it is impossible to manually make correspondences between different texture patterns. For this reason, this paper proposes a novel method to reduce the computational cost of the standard AAM by

normalizing a set of models automatically (based on existing techniques).

The paper is structured as follows. Section 2 reviews related studies and presents the shortcomings of the previously proposed methods. Section 3 presents the Active Appearance Model and its application to fish species identification. It includes the AAM theory and the evaluation of species-specific AAMs for fish species identification. Section 4 contains the normalization algorithms, and the evaluation of this normalization technique against the results obtained in Section 3. Both Sections 3 and 4 include discussion parts regarding the merits and demerits of the evaluated methods. Lastly, Section 5 presents the conclusion of this paper.

2. Related Works for Fish Species Identification

Fish species classification has been investigated for about twenty years. There are two main application areas for fish species identification: commercial purpose and environment monitoring purpose.

Zion et al. [7] proposed a computer vision system that aims to sort fish that grow together in polyculture according to species shape and size. The authors propose to set a feeding station separated from the growing pool and to capture fish shapes when the targets are going from one tank to another one. Therefore, fish profiles and tail end shape, enhanced with particular lighting conditions, can be used for classification. Shapes are represented with the Moment Invariant (MI) algorithm, and identification rates reach more than 90% for the three interest species. The authors state that the tail end shapes are necessary to reach such results. Although the method performs well, it is impossible to apply in unconstrained situations.

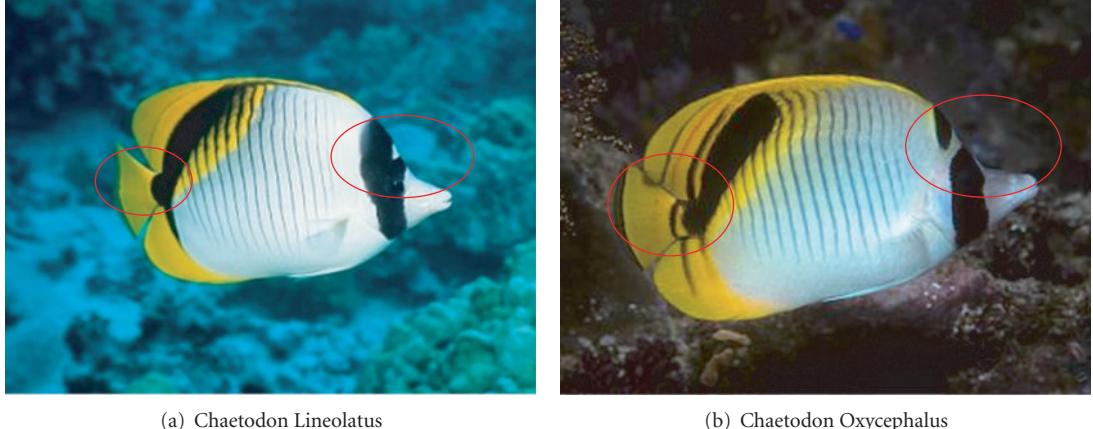


FIGURE 2: *Chaetodon Oxycephalus* and *Chaetodon Lineolatus* present similar texture patterns. The differences are observable in the ellipses. In the experiments, the two species are merged together.

TABLE 3: Correct identification rates of the Experiment 1 for species of two families. The first row of each table indicates the target species (arbitrarily selected amongst the species presented in Table 2).

Pose parameters	Species					
	AAc	ALe	ALi	ATe	AS	total
Optimal	100%	100%	65%	85%	85%	87%
Random	95%	50%	65%	60%	65%	67%
	AA	AC	ACh	AL	AP	Total
Optimal	67%	87%	93%	100%	93%	88%
Random	0.07%	47%	80%	93%	87%	63%

Storbeck and Daan [8] proposed a system that recognizes the species of fish seen on a conveyor belt. A neural network is trained with widths and heights at specific locations of samples. The height of fish is captured using the distortion of a light line produced by a laser on the fish. Height and width information is encoded as a pseudoimage. The recognition rate is about 95% for 6 species.

White et al. [9] also proposed a way to sort species of fish transported along a conveyor and got up to 99.8% of correct classification for 7 species. The method first detects a target, then increases the contrast of the image to extract edges, and finally sorts the fish from shape and color information as an input of a Canonical Discriminant Analysis.

Regarding the above commercial applications, shape information is often used because the illumination conditions, as well as the background, make the edge detection possible and precise. On the other hand, a large part of monitoring systems is based on many other kind of features.

Cadieux et al. [3] proposed to count fish present in fishways by species. Shape information is captured with a silhouette sensor instead of a standard camera. A multiclassifier composed with a Bayes maximum likelihood classifier,

a Learning Vector Quantification classifier, and a back-propagation neural network is applied and reaches a correct classification rate of 77% on 5 species by using Moment Invariant Features, Fourier descriptors, and geometric features. This application is not comparable to the previous and next studies because it uses a specific sensor.

Semani et al. [10] proposed to characterize fish among 12 species present on an aquarium basin. After capturing geometric, photometric, and texture features as well as Moments of Hu and motion, 18 of 38 features are selected using clustering operations, which represents a compression rate of about 52.

Nery et al. [5] analyzed the impact of different types of features on classification results of 6 species. The authors propose to first extract features (size, shape aspect ratio, circularity and moments, color signature, texture inertia and energy, etc.) and then to determine the ones that are meaningful for the classification process. The system identifies the 6 species with a correct classification rate of 81% by using only 4 features.

Despite the fact that the results of shape-based methods reach sufficient recognition rates, texture-based approach has also been proposed for fish classification [4]. The standard eigenface algorithm is improved with a target angle estimation technique, and the recognition accuracy is 76% on 9 species. Five images per species are retained for the data set, and a leave-one-out technique evaluation is adopted.

Although it is possible to identify fish species from shape information, there are two drawbacks. First, segmentation failures make the contour extraction step uncertain (still images and images extracted from videos) [11]. Secondly, fish belonging to the same family or genus often present the same shape but different texture patterns. Figure 1 presents the three-dimensional eigenspace built from shapes of fishes belonging to the same family (15 shapes per species). It is impossible to classify the species from these representations because no clusters are present in the space.

Regarding color information, calibrated cameras, or the use of the same camera in the training and testing

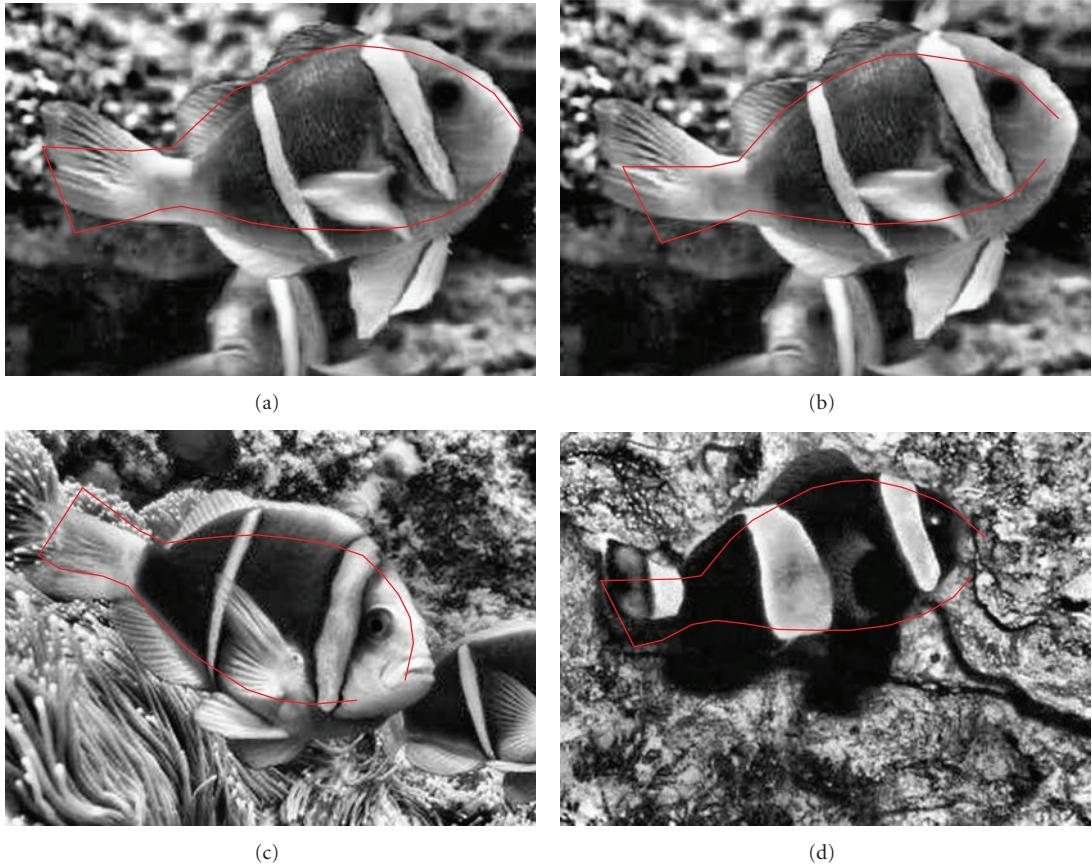


FIGURE 3: Illustration of the shape of the sampling window corresponding to the average of contours of fishes belonging to the *Amphiprion* species data sets (red curves). The two top images present two different pose parameters of the sampling window on the same image (the scale of the (a) sampling window is larger than the scale of the (b) sampling window), while the two bottom images present the pose parameters of the sampling window on two other images of the data set.

TABLE 4: Correct identification rates of the Experiment 2. The fitting is constrained in the pose parameters space given the constraints stated in the introduction of Section 3.2. The first row of each table presents the species' names and the family names are shown in the first column. Due to their similarities in shape and texture, the COx and CL fishes are merged together, as explained in Figure 2. The number of objective function evaluations is proportional to the number of target species.

Chaetodon	COc	COx	CAu	CE	CTr	CL	COx/CL	Total
Identification rate	85%	90%	95%	100%	90%	80%	97.5%	90%
Number of objective function evaluations	1554	1454	1479	1406	1535	1476	—	—
Acanthurus	AAc	ALe	ALi	ATe	AS	—	—	Total
Identification rate	95%	90%	60%	90%	85%	—	—	86%
Number of objective function evaluations	1512	1479	1565	1444	1559	—	—	—

procedures, are required. Hence, in the general case, as well as in this study, color information is not used. Another point to focus on is the lack of common data sets. For this reason, the results of previous works can only be approximately compared. In this study, data sets are created from images of various sources and various resolutions to give a fair evaluation.

An appearance-based method is proposed in order to overcome the two drawbacks of previous works. Moreover, given the above reasons, it is better to use appearance-based

techniques on still images as the texture is important for fish species identification.

3. Active Appearance Models

This study focuses on fish species identification based on texture information, because many species present similar shapes but different texture patterns. In order to identify targets from texture, it is necessary to sample it precisely. Indeed, as presented in Figure 2, some few species present

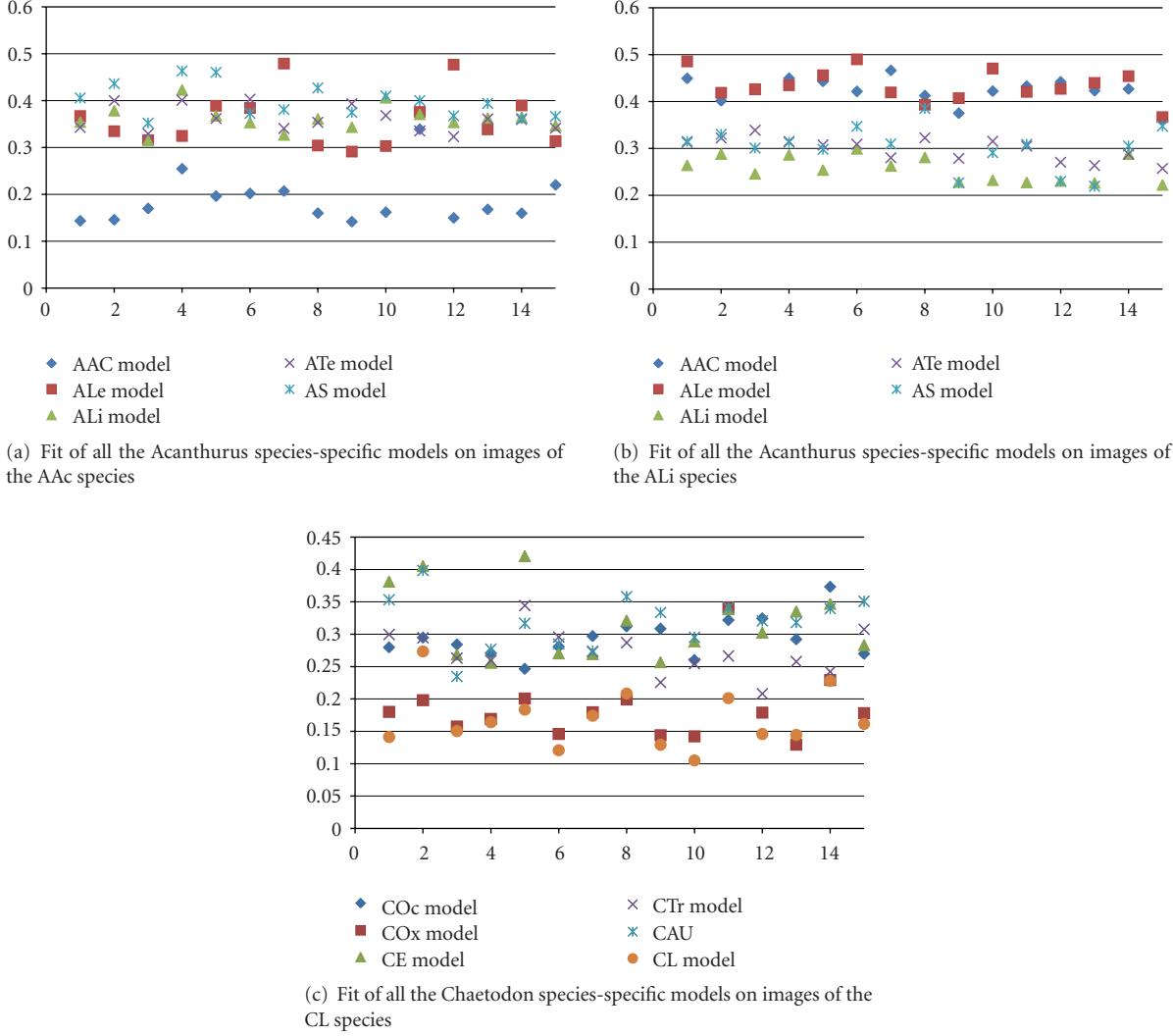


FIGURE 4: The x -axis represents image IDs of species specific data set (15 images) and the y -axis represents the normalized error between the model texture and the sampled texture (4). The difference of errors between the use of ALi and AS models on one hand, and COx and CL models on the other hand is due to appearance similarities between ALi and AS fishes and COx and CL fishes.

very similar texture patterns. Eventually, to deal with shape changes, the use of a morphable sampling window that adapts itself to the target's shape is suitable.

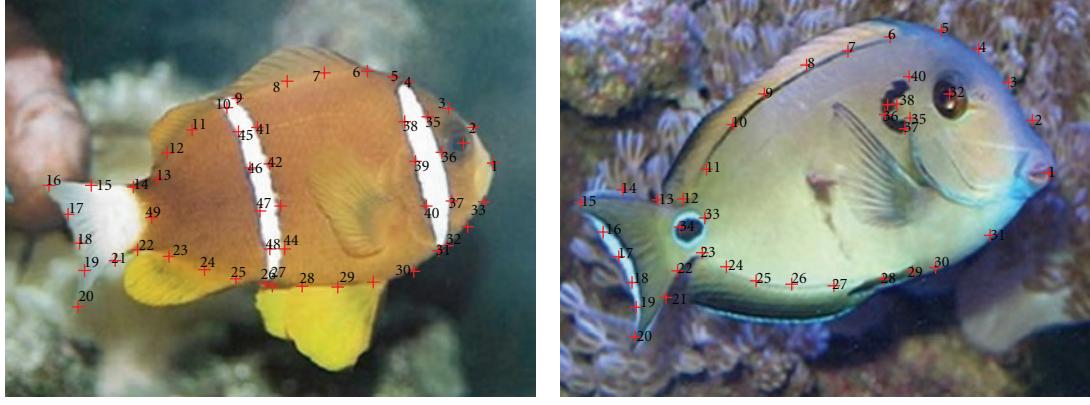
The AAM is stated to be useful if the targets have to be classified using shape or texture information, if the targets have well-defined shapes, and if the position of the target is known (in the frame coordinates). On the contrary, it is not appropriate for objects with widely varying shapes [12]. Since these conditions meet the requirements stated above, the AAM stands as the basis of this study.

3.1. Theory. The AAM is a linear morphable model based on shape and texture information [6]. This section briefly explains the 3-step process to build the model.

First, given a training data set composed of N images, M feature points (landmarks) are manually selected for each image. Each feature point, characterized by its coordinates (x_j, y_j) , represents the same physical position on the fish

(mouth, thin, tail, stripes) on all images. The feature points selection has to be precise, which justifies why it is usually performed manually, despite the existence of methods that propose automatic landmarks selection [13]. A shape vector X_i is defined as the concatenation of the coordinates of those M feature points resulting from the i th image of the training set: $X_i = (x_1 \ x_2 \ \dots \ x_M \ y_1 \ y_2 \ \dots \ y_M)'$, where X_i is the shape vector defined from the i th labeled image of the training set and $\{x_j, y_j\}$ are the coordinates of the j th feature point in the image frame. A shape vector set is created from the training set of labeled images: $X = \{X_1, X_2, \dots, X_N\}$. A compact shape model is computed by applying Principal Component Analysis on the normalized shape data set (Procrustes Analysis [6]). The shape model follows the equation

$$X_G = X_0 + \Phi_X \cdot \lambda_X, \quad (1)$$



(a) Feature points of the Amphiprion Akindynos

(b) Feature points of the Acanthurus Tennenti

FIGURE 5: Illustration of the feature points for two different species. Computing a shape model from shape representations that have different numbers of feature points is impossible. For this reason, normalization of shapes is necessary to build an AAM from different species.

TABLE 5: Correct identification rates of the Experiment 3, with the objective function computed about 600 times for each image. Unified Active Appearance Models built from species belonging to the same family, keeping 5 variation modes. The identification rates are indicated for each species from which the UAAM is built. The COx/CL result, associated to the CL result, shows the confusion between the CL and the COx species.

Amphiprion Model	AA	AC	ACh	AL	AP	—	—	Total
Identification rate	84%	100%	80%	93%	80%	—	—	84%
Acanthurus Model	AAc	ALe	ALi	ATe	AS	—	—	Total
Identification rate	100%	100%	73%	67%	100%	—	—	88%
Chaetodon Model	CAu	CE	CL	COc	COx	CTr	COx/CL	Total
Identification rate	73%	100%	53%	93%	80%	93%	100%	92%
Pomacanthus Model	PA	PI	PM	PS	PSe	PX	—	Total
Identification rate	100%	87%	100%	87%	73%	60%	—	84%

where X_G is the generated shape according to the parameter λ_X , X_0 is the mean shape computed over the normalized shape data set, λ_X is the parameter that controls shape generation, and Φ_X is the matrix that describes shape variation modes (A “mode” is a principal component direction). Column vectors of Φ_X correspond to the principal components of the normalized shape data set X .

Then, the same principle is applied to the textures bordered by the manually selected shapes contours. For each image i of the training set, the texture that lies inside the manually selected shape contour is warped so that the feature points of the shape X_i match the feature points of the mean shape X_0 , and is sampled. A texture vector T_i is defined as the concatenation of the values of the intensity of pixels lying inside the warped shape. In the same way as the shape processing, texture vectors form a texture vector set: $T = \{T_1, T_2, \dots, T_N\}$ (N number of images). PCA is applied to the normalized texture data set to create a compact texture model:

$$T = T_0 + \Phi_T \cdot \lambda_T, \quad (2)$$

where T_0 is the mean texture computed over the normalized texture data set, λ_T is the parameter that controls texture generation, and Φ_T is the matrix that describes texture

variation modes. Column vectors of Φ_T correspond to the principal components of the normalized texture data set T .

Finally, another PCA is computed on the normalized concatenated shape and texture parameters:

$$C = \Phi_C \cdot \lambda_C, \quad (3)$$

where C represents combined shape and texture, λ_C is the parameter that controls combined generation of shapes and textures, and Φ_C is the matrix that describes combined variation modes. It is possible to write $[X, T] = f(\lambda_C)$, where X and T are shape and texture vectors, respectively, and f is the function defined from the training data set.

In this paper, an *instance* refers to the shape and texture vectors resulting from the AAM, an *appearance parameter* refers to the parameter that controls shape and texture generation, and *pose* refers to the position, scale, and orientation of an instance in the image frame (four to six dimensions). *Fitting* an AAM to an unseen image consists in generating a shape and a texture that are as close as possible to the target’s shape and texture. *External feature points* designate points that belong to the manually selected shapes’ contours, while *internal feature points* designate feature points that are inside the selected shapes’ contours. *Species-specific-AAM* is an AAM built from images of fishes belonging to the same species.

```

(1) for each species do
(2)   Normalize shapes (Procrustes Analysis)
(3)   Compute the mean shape
(4)   Compute a Delaunay triangulation on the mean shape
(5)   Re-sample the external mean shape
(6)   for each new external feature point do
(7)     Find the triangle in which lies the current new external feature point
(8)     for each shape of the current specie do
(9)       Compute the position of the new external feature point in the corresponding triangle
               of the current shape using an affine interpolation
(10)    end for
(11)   end for
(12) end for

```

ALGORITHM 1: Computation of the normalized shapes external feature points, basic approach. The term “species” designates the species manually chosen for the creation of the UAAM.

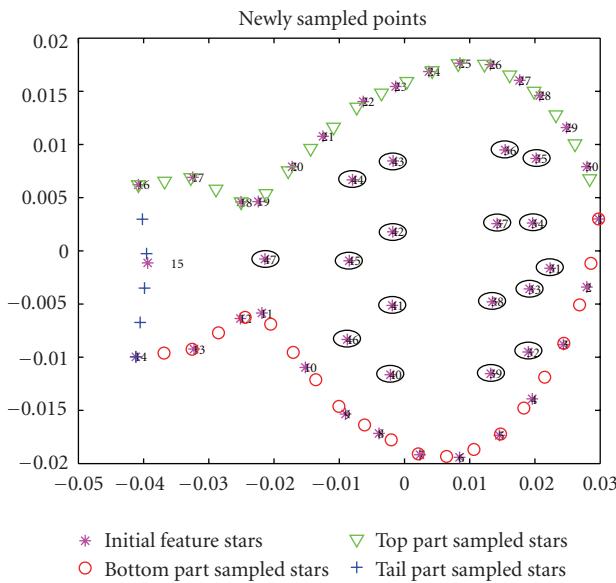


FIGURE 6: This figure illustrates Algorithm 1. Stars represent the initial feature points for the considered shape. Triangles, crosses, and circles represent the top, tail, and bottom parts of the newly sampled external feature points, respectively, while ellipses represent the internal feature points.

Fitting an AAM to an input image requires minimizing the objective function defined as

$$E = \sqrt{\sum_{i=1}^p (T_{\text{Model}}[i] - T_{\text{Sampled}}[i])^2}, \quad (4)$$

where p is the number of pixels of the model texture, T_{Model} is the texture generated using the AAM, and T_{Sampled} is the warp of the texture that lies inside the AAM instance to the AAM mean shape X_0 (1).

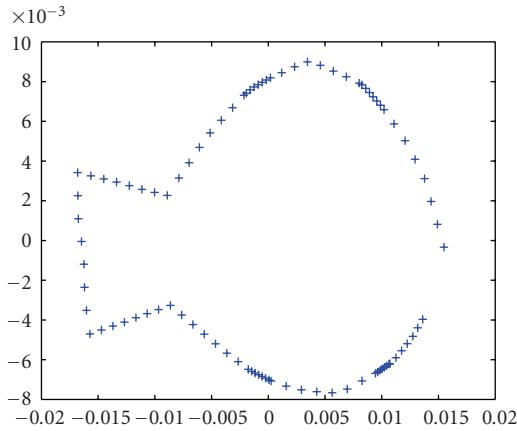
In order to compute the objective function, it is necessary to generate sampling windows using the AAM and to warp the texture that lies inside the sampling windows toward the mean shape X_0 of the AAM (1).

In the case of standard AAMs, the inverse compositional algorithm [14] is one of the ways to minimize the objective function, as well as the Nelder-Mead Simplex (NMS) algorithm [15] that has been proposed for the AAM fitting in [16]. The NMS algorithm requires less memory and has better generalization properties than regression techniques and is stated to perform better than the Regression Matrix method [16]. Thus, the NMS algorithm is used for AAM fitting in this study.

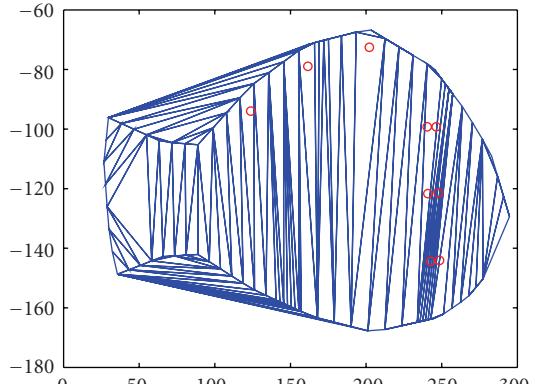
3.2. Evaluation of the AAM for Fish Species Identification. The AAM-based fish species identification results are compared to the best result in fish species identification (although the data sets are different), that is, 81% for 6 species (unconstrained environment, Table 1) and to the results obtained using a nonmorphable sampling window.

All the experiments of this paper follow the same experimental conditions. A data set of 15 images is built up for each species. No juvenile fishes are visible on the data set images, and fishes are visible sideways. For each species, the 15 images are collected from different sources and present various resolutions. Images are converted to grayscale after a histogram equalization preprocessing step. A Leave-One-Out cross validation is used for the evaluation of the algorithms. The identification rates are either presented for each species or as the average of the identification rates for each target species. In all experiments, only four pose parameters are used: translation, scale, and in-plane rotation. Pose parameters are initialized randomly from the optimal position given the following constraints: Translation: ± 10 pixels, Scale: $\pm 10\%$, and Rotation: ± 10 degrees. A Linear Discriminant Analysis (LDA) is conducted in the texture eigenspace (2) to identify the species, except for Experiment 2. Table 2 presents the different target families and species.

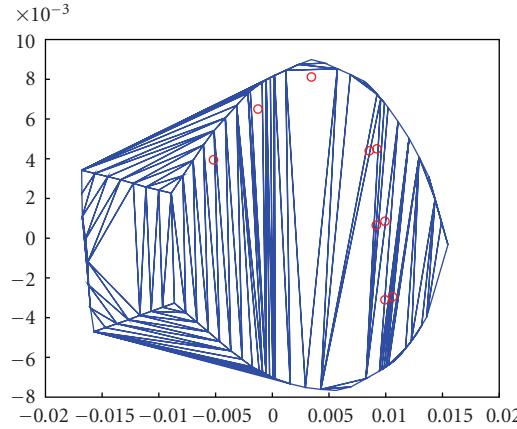
3.2.1. Experiment 1: Evaluation of a Nonmorphable Sampling Window. This experiment aims to evaluate texture sampling using a nonmorphable sampling windows, in comparison with the morphable sampling window of the AAM. Instead



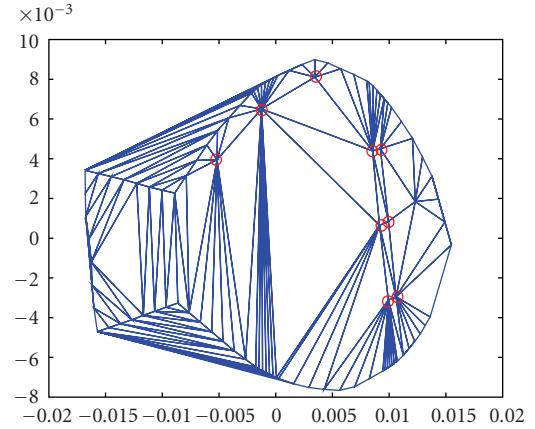
(a) External mean shape computed over all the species, Algorithm 2 step 1



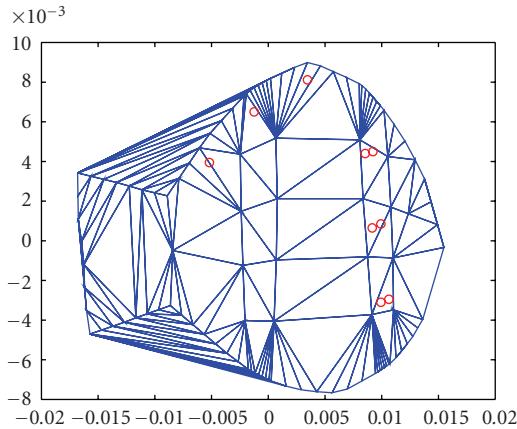
(b) Computation of the Delaunay triangulation on the mean contour of the temporary shapes for the first species, Algorithm 2 step 4. The points represent the internal points of the first species



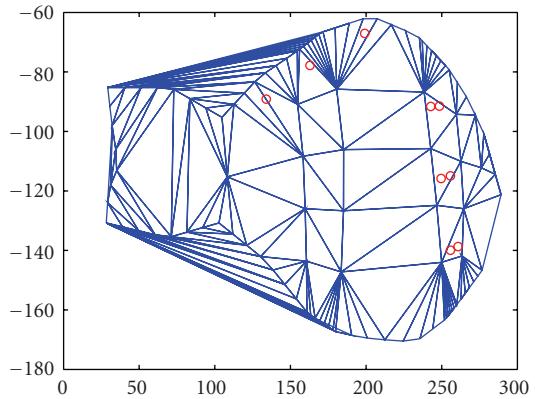
(c) Computation of the position of internal feature points of the first species in the Unified mean shape, Algorithm 2 step 5



(d) First species frame, Algorithm 2 step 9. The points represent internal points of the first species in the Unified Mean shape. It corresponds to the first species frame



(e) The points represent internal points of the first species in the second species frame, Algorithm 2 step 17



(f) The points represent internal points of the first species in a shape of the second species, Algorithm 2 step 19

FIGURE 7: Illustration of the steps of Algorithm 2.

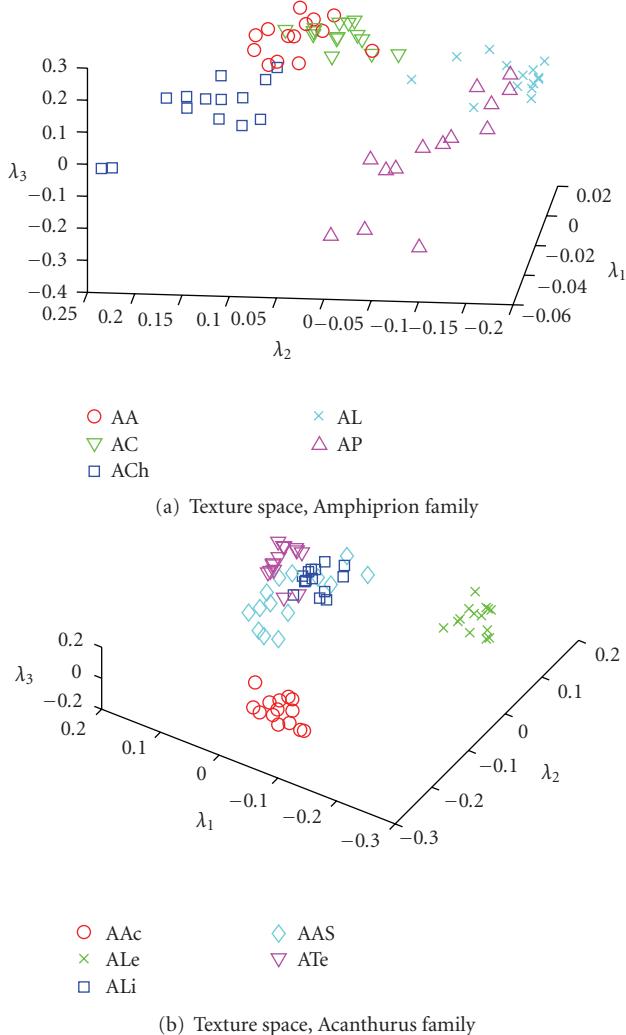


FIGURE 8: Projection of textures in their associated eigenspace for two families. Each symbol represents a species, and each point corresponds to a parameter λ_T , (2). Clusters make the convergence of AAM difficult.

of a rectangle window, the average shape of contours of fishes is used as the sampling window. The influence of pose parameters on the classification results is evaluated by using two sets of parameters: the best pose parameters, computed from the Procrustes analysis [6] and random pose parameters computed as explained above.

Figure 3 presents some examples of pose parameters for the same sampling window on different images, and Table 3 presents the results of the experiment based on the Acanthurus and the Amphiprion families. It shows that a nonmorphable sampling windows is not robust to pose parameters. In real applications, this technique may lead to poor results due to target segmentation failures [11].

3.2.2. Experiment 2: Evaluation of the Active Appearance Model. This experiment aims to validate the use of AAMs for fish species identification. Two subdatasets are built up from manually selected species (One dataset for the species

TABLE 6: Four situations are evaluated in order to compare the robustness to pose variations of the UAAM technique. Columns represent each pose parameters, and rows correspond to the evaluated cases. Random pose parameters are generated from the optimal pose parameters given the proposed constraints.

	Translation, x	Translation, y	Rotation	Scale
Case 1	0 pixels	0 pixels	0 degree	0 degree
Case 2	2.5 pixels	2.5 pixels	2.5 degree	2.5 degree
Case 3	5 pixels	5 pixels	5 degree	5 degree
Case 4	10 pixels	10 pixels	10 degree	10 degree

belonging to the Chaetodon family and one dataset for the species belonging to the Acanthurus family). For each subdataset, species-specific-AAMs are computed and fitted to unseen images using the NMS algorithm, as explained in part 2. The model with the best fit indicates the species. This approach can be considered as a brute-force (exhaustive) search. The maximum number of iterations for the Nelder-Mead simplex algorithm is set to 100.

Figure 4 presents the fitting results of species-specific-AAMs on images of the two subdatasets. The y -axis represents the normalized error between the target texture and the model texture after the fitting (error computed using the objective function defined in (4)). A good fitting is represented by a low error, as shown in the three subfigures. For the three species (AAC, ALi and CL), the lowest errors mainly correspond to the fitting of the corresponding model. Table 4 presents the results of fittings constrained in pose (given the constraints stated above) for the two subdatasets. The brute-force approach based on the AAM outperforms the previous studies results for applications in unconstrained environments. It also outperforms by 20% the results obtained using a nonmorphable sampling window for the Acanthurus family. Unconstrained fitting in the pose parameters space has been evaluated but lead to high misclassification rates in the case of uniform textures, which is particularly present in the Acanthurus family.

3.3. Discussion. In this section, fish species identification based on texture is evaluated by two experiments. The AAM is robust to segmentation failures and gives correct identification rates higher than those of previous works.

Regarding the computational cost, all the experiments are conducted using Matlab R2009b, and texture warps (Section 3.1) are performed using the OpenGL.NET Tao library. Achieving real-time identification with Matlab is not possible.

On the other hand, warps represent 90% of the computational time of the objective function, and one warp takes 0.02 ms in C language on a GeForce 8800 GTX (50 feature points, texture composed of 3000 pixels). According Table 4, it is required to compute an average of 1500 times the objective function for fish species identification. Thus, the computational time is estimated to about 30 ms (1500 by 0.022 if the NMS algorithm computation time is neglected) to identify 6 species only.

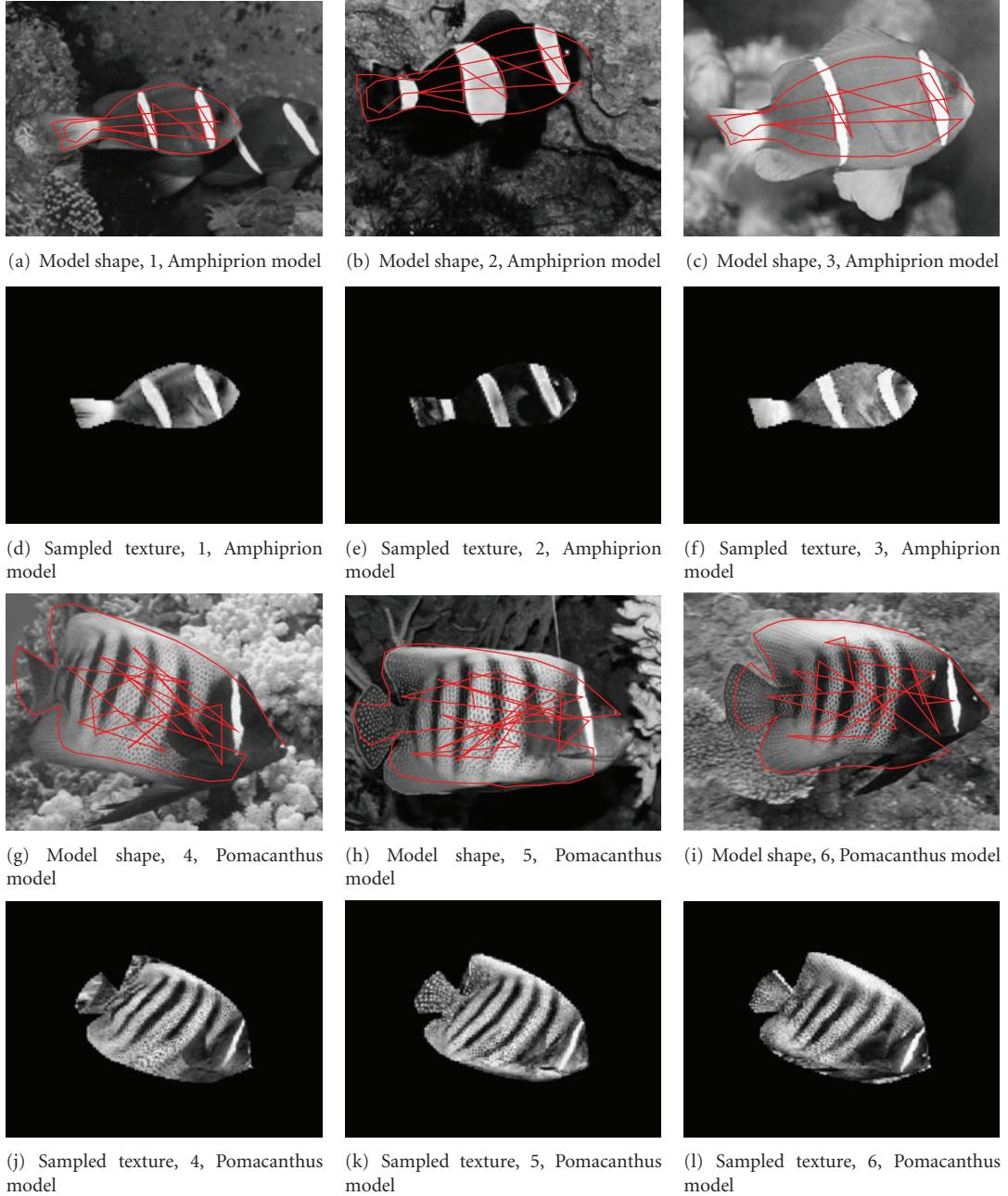


FIGURE 9: Illustration of the fitting of two UAAMs and the associated sampled textures. The sampled texture patterns are the same although the targets shape change (Pomacanthus model results) or the targets species differ (Amphiprion model results).

The two requirements for automatic monitoring are the processing speed, because the data is streamed in real time, and the identification precision. Although the AAM meets the accuracy requirements, it is necessary to speed up the identification process without decreasing the qualitative performance of the method. The next section presents a method that aims to speed up the identification process without sacrificing the accuracy, by replacing the set of species-specific AAMs with a unified model built from the target species data sets.

4. Normalization of Active Appearance Models

In order to reduce the computational cost of the fish species identification based on the AAM, the set of species-specific AAMs should be replaced with a unified AAM. However, it is impossible to use the AAM for objects that do not present similar texture patterns. Indeed, the AAM algorithm requires computing a shape model based on a manual selection of feature points. The feature points are determined depending on the texture pattern, which means that the number of

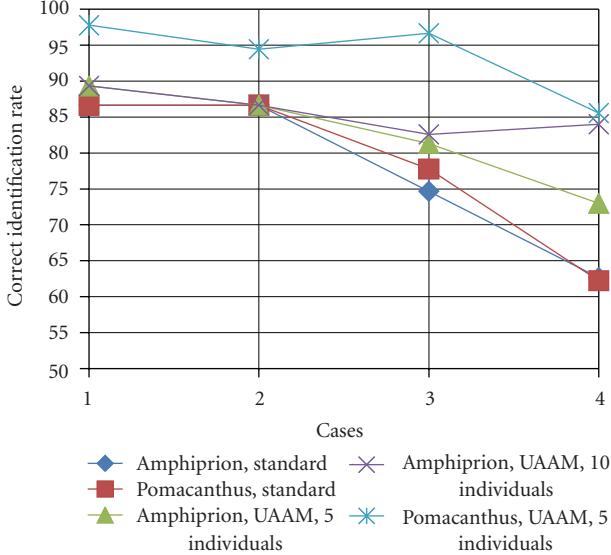


FIGURE 10: The x -axis represents the 4 pose initialization conditions, and the y -axis represents the correct identification rate. The standard (nonmorphable) sampling method is not robust to pose variations. In case of segmentation failures, the UAAM still provides high correct identification rates (20% greater than the standard technique). However, it requires more GA individuals for the fitting procedure.

feature points, as well as their positions, differs for each species, as presented in Figure 5. It is related to the “missing features” problem (the “missing features” problem appears when some texture patterns are not visible on all images of the data set, as presented in Figure 5). This problem has already been addressed in [17], to deal with “missing features, occlusion, substantial spatial rearrangement of features” by introducing the concept of layered AAM (Each feature corresponds to a layer, and layers occluding each others). However, this technique was developed to meet the requirements that are different from those of this study. The purpose of this study is to reduce the computational cost of the brute-force approach (presented in Section 3) that proved to outperform previous works. This section presents an original technique that normalizes species-specific-AAM shape models together and that leads to the creation of the Unified Active Appearance Model (UAAM).

4.1. Theory. The creation of the UAAM follows the same steps as the creation of a standard AAM. Nevertheless, since the UAAM is built over different species, it is required to normalize the species-specific shape data sets so that all shapes have the same vector length while conserving the initial properties of the data set.

4.1.1. Normalization of Shapes. After the manual selection of the species from which the UAAM is created, the following algorithms are applied. Algorithm 1 provides sampling of the external feature points for all the shapes of the selected species. Figure 6 presents an example of newly sampled external feature points for one shape. Algorithm 2 consists

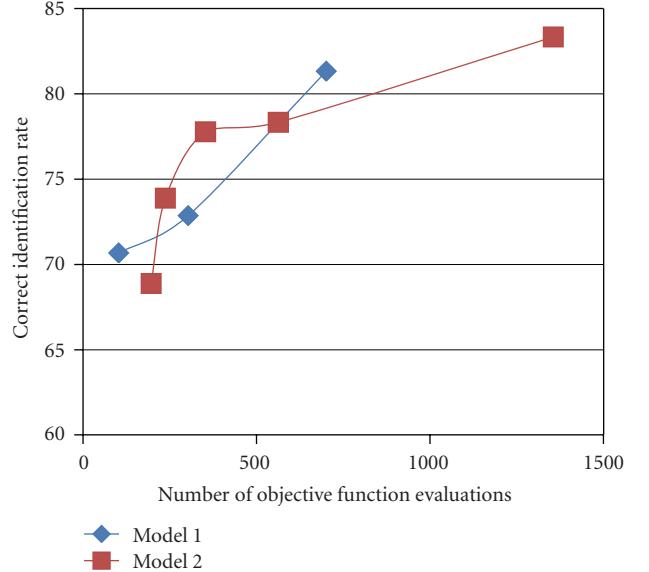


FIGURE 11: The x -axis represents the number of times the objective function is computed by the optimization algorithm and the y -axis represents the correct identification rate, computed on all the species on which each model is built. A large number of objective function evaluations produce satisfying identification rate but make the identification process slow.

in sampling the position of internal feature points for all the shapes of the selected species, as illustrated in Figure 7.

Algorithms 1 and 2 normalize the length of shape vectors among different species by adding *virtual* feature points which coordinates linearly depend on the manually selected feature points. Thus, the number of feature points increases with the number of species, which increases the time required for each objective function computation during the fitting procedure.

Two refinements are proposed to reduce the number of feature points of the UAAM shape model (1) without influencing the texture model (2). First, correspondences of feature points between species are manually set up before the execution of Algorithms 1 and 2: each feature point or group of feature points are labeled regarding the texture pattern they belong to. If two or more species present common texture patterns (Example of the “eye” feature point visible on the two shape representations of Figure 5), then the corresponding *virtual* feature points are added only to the shapes of the species that do not present the point (instead of ending up with groups of feature points that represent the same texture pattern). Secondly, shapes are down-sampled after the execution of Algorithms 1 and 2. The number of feature points varies from 50 to 100 depending on the number of species and the computational speed requirements.

4.1.2. Fitting Procedure. The UAAM is a multi-class model created through the normalization of species-specific AAMs. Figure 8 illustrates the broken-up low-dimensional representation of textures (each texture represented by a parameter

```

(1) Define the Unified mean shape as the mean shape computed over all the species using new external shapes (Algorithm 1)
(2) for species  $i = 1$  to  $N$  do { $N$ : number of species}
(3)   Define temporary shapes by adding the original shapes' internal feature points to newly sampled external feature points
       of the current species  $i$ 
(4)   Compute the Delaunay triangulation on the mean computed over the temporary shapes' external feature points
(5)   for  $p = 1$  to  $P$  do { $p$ : internal feature point of species  $i$ ,  $P$ : number of internal feature points for the species  $i$ }
(6)     Find in which triangle of the step 4's triangulation  $p$  lies
(7)     Compute the position of  $p$  in the corresponding triangle of the Unified mean shape (step 1) using an affine interpolation
(8)   end for
(9)   Define the current species Frame as the Unified mean shape and the species specific internal points expressed in the Unified
       mean shape (step 5)
(10) end for
(11) for species  $i = 1$  to  $N$  do { $N$ : number of species}
(12)   for  $k = 1$  to  $M$  { $M$ : number of shapes for the current species  $i$ }
(13)     Add the original internal feature points of the shape  $k$  to the newly sampled external feature points of shape  $k$ 
           (Algorithm 1)
(14)   end for
(15)   for species  $j = 1$  to  $N$ ,  $i \neq j$  do { $N$ : number of species}
(16)     for  $p = 1$  to  $P$  do { $p$ : species  $i$  internal feature point,  $P$ : number of internal feature points for the species  $i$ }
(17)       Find in which triangle of the species  $j$  frame  $p$  lies
(18)       for  $k = 1$  to  $M$  do { $M$ : number of shapes for the current species  $j$ }
(19)         Compute the position of  $p$  in the corresponding triangle of the newly sampled external feature points of shape  $k$ 
           using an affine interpolation
(20)       end for
(21)     end for
(22)   end for
(23) end for

```

ALGORITHM 2: Computation of the normalized shapes internal points, basic approach. The term “species” designates the species manually chosen for the creation of the UAAM.

TABLE 7: Correct identification rates of the Experiment 5, with the objective function computed about 650 times for each image. The identification rates are indicated for each species from which the UAAM is built. The Unified Active Appearance Models built from species belonging to 2 families: Amphiprion and Acanthurus (Model 1), and Chaetodon and Pomacanthus (Model 2), keeping 5 variation modes.

Model 1	AA	AC	ACh	AL	AP	—
Identification rate	60%	93.3%	46.7%	86.7%	80%	—
Model 1	AAc	ALe	ALi	ATe	AS	—
Identification rate	100%	100%	73.3%	73.3%	100%	—
Model 2	CAu	CE	COc	COx	CL	CTr
Identification rate	73.3%	100%	53.3%	53.3%	86.7%	86.7%
Model 2	PA	PI	PM	PS	PSe	PX
Identification rate	100%	80%	86.7%	66.7%	60%	93.3%

λ_T in (2)) in the case of the UAAM. Because of the presence of clusters in the texture space (usually, one cluster per species), the use of optimization methods based on gradient is difficult [18]. The NMS algorithm, used for the AAM fitting in the previous section, has good exploitation properties [19] but lacks in the capabilities in exploration,

contrarily to the GA-based methods. Regarding the UAAM, exploration of the search space during the fitting procedure is fundamental because of the presence of clusters, but a precise fitting requires a correct exploitation. Thus, a hybrid GA optimization method that combines the NMS and the GA approaches [19] is employed in this paper. Individuals of the GA are defined as simplex. One or more flips of the NMS algorithm are applied at each GA generation, while crossovers and mutation operators are applied on vectors of the simplexes. This method is stated to be more efficient than the GA regarding the number of function evaluations. Vectors of the simplexes are defined as the concatenation of an appearance parameter λ_C and a pose parameter (translation, rotation, scale), and the function to be minimized as the (4).

4.2. Evaluation of the UAAM. The species used for the construction of UAAMs are selected arbitrarily amongst the species presented in Table 2 given shape considerations. In this paper, four models are built from species of the same family (e.g., one UAAM is built from the 5 species of the Amphiprion family: Akindynos, Clarkii, Chrysopterus, Latezonatus, and Polymnus), while two models are built from species of 2 different families that present close shape variations: species of the Amphiprion and the Acanthurus families and species of the Chaetodon and the Pomacanthus families. For all the experiments based on the UAAM, the models are fit to images that show species belonging to the

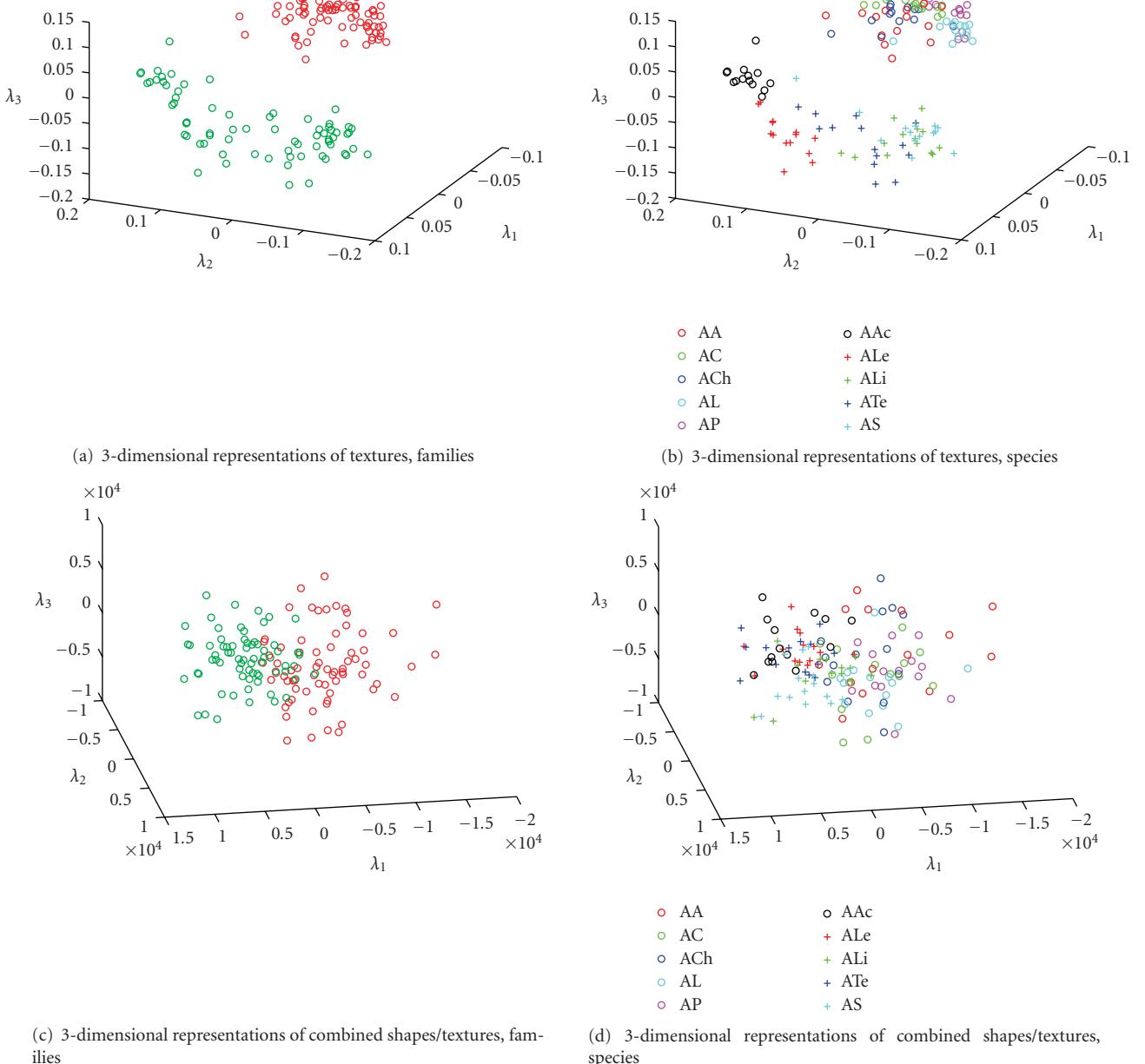


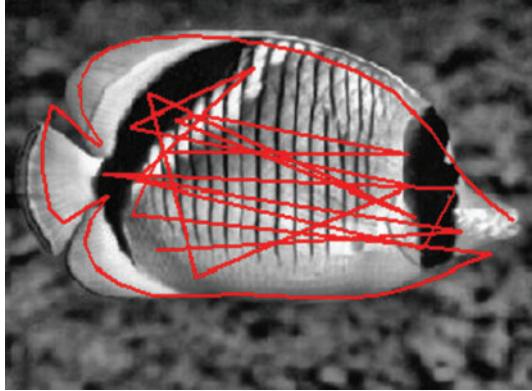
FIGURE 12: Illustration of the 3-dimensional textures and combined shapes and textures spaces of the Model 1. Each circle or cross represents a texture vector or both a shape and a texture vector, computed as explained in Section 3.1. In the family figures, red circles represent the *Amphiprion* family, whereas green points represent the *Acanthurus* family. Separability between family clusters is better in the texture space than in the combined shape and texture space, which justifies why the LDA is conducted in the texture space.

UAAM training data set. (i.e., the *Amphiprion* model is evaluated using images of *Akindynos*, *Clarkii*, *Chrysopterus*, *Latezonatus*, and *Pomnus*.)

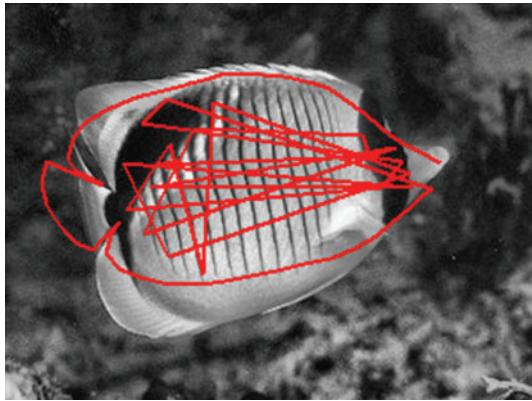
The UAAM evaluation is conducted given the experimental conditions stated in Section 3. Regarding the initialization of the hybrid GA, the k -mean clustering algorithm is applied to the low-dimensional representation of the training data set, λ_C (3). The number of clusters is initially defined as the number of GA individuals (i.e., the number of simplexes), and each individual is initialized from appearance

parameters belonging to the same cluster (Vectors that constitute simplexes all come from the same cluster). The optimization is conducted using 10 individuals and 6 GA epochs.

4.2.1. Experiment 3: Evaluation of the UAAM. The purpose of this experiment is to validate the UAAM and to confront the normalization technique to the brute force approach of Experiment 2. Figure 9 presents six fitting results for two UAAMs, the *Amphiprion* model and the *Pomacanthus* model.



(a) Correct fitting of the UAAM



(b) Failure in the fitting of the UAAM

FIGURE 13: Illustration of the fitting of a UAAM on two *Chaetodon Lineolatus* images. The failure in the UAAM fitting may lead to species misclassifications.

Table 5 represents the classification results of four different families. Comparing to the brute-force approach, the normalization of species-specific AAMs brings a speed up of a factor greater than 2 in the case of the *Chaetodon* and *Acanthurus* families. Moreover, the correct identification rates obtained from the UAAM are comparable to the brute-force approach rates.

4.2.2. Experiment 4: Robustness to Pose Variations. As explained in the previous section, sampling textures using nonmorphable windows is not robust to segmentation failures. This experiment aims to evaluate the robustness of the UAAM against that of the standard texture sampling. Segmentation failures result in estimation errors for the four pose parameters at the same time: translation, rotation, and scale. For this reason, instead of evaluating the robustness to translation, rotation and scale separately, the 4 arbitrary situations presented in Table 6 are evaluated. Results are presented in Figure 10. The nonmorphable sampling window technique evaluation follows the same principle as the Experiment 1 using the case constraints, while the case constraints are used to initialize the UAAM fittings. The experiment is conducted on the *Amphiprion* and *Pomacanthus* families. The species belonging to the

TABLE 8: Confusion matrices for the UAAM built from the *Amphiprion* and the *Acanthurus* species. There are 15 images for each species. Rows represent the target species, and columns represent the identified species. There are 6 species per family (horizontal and vertical delimitations). The identification of families requires very few objective function evaluations.

(a) Objective function evaluated 100 times

	AA	AC	ACh	AL	AP	AAC	ALe	Ali	Ate	AS
AA	5	5	5							
AC		11	1	1	2					
ACh	3	3	8			1				
AL				12						
AP				3	1	11				
AAC							14			1
ALe		2						13		
ALi								12		3
Ate								5	6	4
AS								1		14

(b) Objective function evaluated 700 times

	AA	AC	ACh	AL	AP	AAC	ALe	ALi	Ate	AS
AA	9	1	5							
AC	1		14							
ACh	5	3	7							
AL		1		13						1
AP				3		12				
AAC								15		
ALe									15	
ALi									11	4
Ate								3	11	1
AS										15

Pomacanthus family present various texture patterns, which explains why the identification rates are very high in the Case 1 for both methods, although the UAAM outperforms the nonmorphable sampling window. On the contrary, species belonging to the *Amphiprion* family all present two stripes of variable width. Some of these species are very similar, which justifies that the identification rates are lower than the rates for the *Pomacanthus* species. However, models of both families outperform by about 20% the standard sampling for the Case 4.

4.2.3. Experiment 5: Hierarchical Approach. The purpose of this experiment is to evaluate the impact of the number of species on the classification results. The previous experiments are conducted using UAAM created from species that belong to the same family, that is, that share shape variations. In this experiment, the *Amphiprion* and *Acanthurus* models are merged together (Model 1), as well as the *Chaetodon* and *Pomacanthus* models (Model 2).

Table 7 presents the classification results on 10 and 12 species. Model 1 reaches 81.3% of correct identification on 10 species, while Model 2 reaches 78.3% on 12 species (about 700 objective function evaluations), which outperforms all

TABLE 9: Comparison between a hierarchical technique and a standard method. Using two models, one for the family identification and one for the species identification, outperforms the nonhierarchical method by about 5 percent.

	Identification rate	Number of objective function evaluations
10 species, Model 1	81.3%	600~700
10 species, Model 1, hierarchical	85.1%	100 (Model 1) + 600 (Family Specific Models)
12 species, Model 2	78.3%	600~700
12 species, Model 2, hierarchical	84.5%	230 (Model 2) + 600 (Family Specific Models)

TABLE 10: Summary of the performance of the three evaluated sampling methods on two families.

	Nonmorphable sampling window	AAM, brute force	UAAM
Acanthurus (5 species)	67%	86.6%	88%
Chaetodon (6 species)	76%	90%	92%

the previous works and confirms the need of a normalization technique (the brute force approach would require more than (1500 evaluations/5 Species)* 10 species = 3000 objective function evaluations).

Figure 11 illustrates the relation between the identification rate and the number of objective function evaluations (i.e., related to the exploration and the exploitation properties of the hybrid GA algorithm). Given only 100 function evaluations, the algorithm yields 70% of correct classification rate. However, as presented in Table 8, families are correctly identified. Hence, using a UAAM built from species of 2 families for family identification and a UAAM built from species of the same family for species identification (hierarchical approach) is evaluated. Results are presented in Table 9. The hierarchical approach surpasses by about 5 percent the non-hierarchical approach and achieves more than 81% of correct identifications for 10 and 12 species, against 81% on 6 species for the best related work.

Figure 12 presents the 3-dimensional texture and combined shape/texture spaces of the Model 1 and illustrates why the hierarchical approach provides good results. In the texture space, clusters corresponding to species are visible, and clusters corresponding to families are completely separated. It is not the case in the combined shape/texture space, where clusters corresponding to species are not clearly visible and clusters corresponding to families are very close. Since the LDA is performed in the texture space, 100 objective function iterations are sufficient to start converging to the correct species cluster and to converge to the correct family cluster. This figure also explains why the classification is performed in the texture space instead of the combined shape and texture spaces.

4.3. Discussion. The fitting of the UAAM is a global optimization problem, in which the main disadvantage is the occurrence of local minima. UAAM that does not use texture pattern correspondences (Algorithms 1 and 2) presents the same texture patterns at different locations, which in turn

generates artificial local minima. For instance, in the case of the *Amphiprion* family, all the species of the data set show two vertical stripes. The orientation of one of the stripe is a discriminant feature for species identification. If the UAAM is not based on manual correspondences between species, the orientation of that stripe is captured by the texture model, which leads to artificial local minima, and to misclassification. On the contrary, if it is captured by the shape model, such problem can be avoided.

The proposed normalization algorithm has a merit to simplify the building procedure of an AAM for many species. It is also proved in this paper that it surpasses other fish species identification algorithms, although the data sets are different (Table 10).

Regarding the computational time, given Tables 4 and 5, only 600 objective function evaluations are required for the chaetodon model, against about 1500 evaluations in the case of the brute-force approach. Moreover, the UAAM provides identification rates comparable to the brute-force approach. By considering that the computational cost mainly depends on the number of objective function computations, the UAAM is about 4 times faster than the brute-force approach for 12 species. On the other hand, the main limitation of the AAM normalization stands in the fitting of the UAAM. It is more prone to fall into local error minima than a standard AAM (as illustrated in Figure 13), which leads to misclassifications. Furthermore, it is still necessary to manually label images for the construction of species-specific AAMs, as there is no effective technique for automatic feature points selection.

5. Conclusion

This paper brings out the Active Appearance Model (AAM) for fish species identification. The AAM is evaluated and compared to existing methods and proved to surpass related works in terms of accuracy. Because of the high computation cost of the brute-force approach evaluated in the first section of this paper, an original AAM normalization algorithm is proposed to speed up the identification procedure. The Unified Active Appearance Model that results from the proposed method is more than four times faster than the brute-force approach, while getting comparable identification rates. It yields 84.7% of correct identifications on 10 species. The future work is to focus on increasing the generalization properties of the Unified Active Appearance Model by taking advantage of physical properties of fish. The extension of the method to common species is also of interest, as well as

the use of shape information to speed up the fitting of the Unified Active Appearance Model.

Acknowledgments

The present work was financially supported by a Japanese ministry of Education Scholarship. The authors would like to thank Dr. Renaud Seguier for his assistance and his vital encouragements.

References

- [1] J. Hill and C. Wilkinson, *Methods for Ecological Monitoring of Coral Reefs*, Australian Institute of Marine Science (AIMS), Townsville, Australia, 2004, <http://www.icran.org/>.
- [2] Ecogrid, February 2010, <http://ecogrid.nchc.org.tw/sites.php?site=kt>.
- [3] S. Cadieux, F. Michaud, and F. Lalonde, "Intelligent system for automated fish sorting and counting," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '00)*, pp. 1279–1284, November 2000.
- [4] C. B. Akg, "Abstract automatic fish classification from under-water images," 2003.
- [5] M. S. Nery, A. M. Machado, M. F. M. Campos, F. L. C. Pádua, R. Carceroni, and J. P. Queiroz-Neto, "Determining the appropriate feature set for fish classification tasks," in *Proceedings of the 18th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI '05)*, pp. 173–180, October 2005.
- [6] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681–685, 2001.
- [7] B. Zion, A. Shklyar, and I. Karplus, "In-vivo fish sorting by computer vision," *Aquacultural Engineering*, vol. 22, no. 3, pp. 165–179, 2000.
- [8] F. Storbeck and B. Daan, "Fish species recognition using computer vision and a neural network," *Fisheries Research*, vol. 51, no. 1, pp. 11–15, 2001.
- [9] D. J. White, C. Svelingen, and N. J. C. Strachan, "Automated measurement of species and length of fish by computer vision," *Fisheries Research*, vol. 80, no. 2-3, pp. 203–210, 2006.
- [10] D. Semani, C. Saint-Jean, C. Frélicot, T. Bouwmans, and P. Courtellemont, "Alive fishes species characterization from video sequences," in *Proceedings of the Structural, Syntactic and Statistical Pattern Recognition (SSPR/SPR '02)*, pp. 689–698, 2002.
- [11] M. K. S. Alsmadi, K. B. Omar, S. A. Noah, and I. Almarashdah, "Fish recognition based on the combination between robust feature selection, image segmentation and geometrical parameter techniques using Artificial Neural Network and Decision Tree," *International Journal of Computer Science and Information Security*, vol. 6, no. 2, pp. 215–221, 2009.
- [12] T. Cootes, C. Taylor, and M. M. Pt, "Statistical models of appearance for computer vision," Tech. Rep., 2004.
- [13] S. Baker, I. Matthews, and J. Schneider, "Automatic construction of active appearance models as an image coding problem," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 10, pp. 1380–1384, 2004.
- [14] J. Matthews and S. Baker, "Active appearance models revisited," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 135–164, 2004.
- [15] J. A. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965.
- [16] Y. Aidarous, S. L. Gallou, and R. Seguier, "Simplex optimisation initialized by Gaussian Mixture for active appearance models," in *Proceedings of the 9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications (DICTA '07)*, pp. 79–84, December 2007.
- [17] E. Jones and S. Soatto, "Layered active appearance models," in *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV '05)*, pp. 1097–1102, October 2005.
- [18] A. U. Batur and M. H. Hayes, "Adaptive active appearance models," *IEEE Transactions on Image Processing*, vol. 14, no. 11, pp. 1707–1721, 2005.
- [19] N. Durand and J. Marc Alliot, "A combined nelder-mead simplex and genetic algorithm," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '99)*, pp. 1–7, 1999.

