Fish Image Reorganization Construction using Unsupervised Learning Performance

N.Chithra,
PG Scholar,
Department of Computer Science and Engineering,
Dr. Sivanthi Aditanar College of Engineering,
Tiruchendur , India.

Dr. J. Jemi Florinabel
Associate Professor,
Department of Computer Science and Engineering,
Dr. Sivanthi Aditanar College of Engineering,
Tiruchendur, India.

***Abstract*— Live fish recognition is one of the most crucial elements of fisheries survey applications where the vast amount of data is quickly acquired. Different from general scenarios, challenges to underwater image recognition are posted by poor image quality, uncontrolled objects and environment, and difficulty in acquiring representative samples. In addition, most existing feature extraction techniques are hindered from automation due to involving human supervision. Toward this end, we propose an underwater fish recognition framework that consists of a fully unsupervised feature learning technique and an error-resilient classifier. Object parts are initialized based on saliency and relaxation labelling to match object parts correctly. To exploit information from ambiguous images, the notion of partial classification is introduced to assign coarse labels by optimizing the benefit of indecision made by the classifier. Experiments show that then proposed framework achieves high accuracy on both public and self-collected underwater fish images with high uncertainty and class imbalance.**

***Index Terms*—Feature learning, fish species identification, object recognition, underwater imagery, unsupervised learning.**

I. INTRODUCTION

Fish abundance estimation which often calls for the use of bottom and mid-water trawls, is critically required for the commercially important fish populations in oceanography and fisheries science. [1].However, fish captured by trawls often do not survive, and thus trawl survey methods are inappropriate in some areas where fish stocks are severely depleted. To address these needs, we developed the Cam-trawl to conduct video-based surveys. The absence of the codend allows fish to pass unharmed to the environment after being sampled (captured by cameras[2].The captured video data provide much of the information that is typically collected from fish that are retained by traditional trawl methods.[1]Various challenges for the underwater fish / video analysis includes the variable lighting conditions and ubiquitous noise from non fish objects. Using automated image processing algorithms. for detection, segmentation, length /area and size measurements, classification, the challenges are used Underwater video processing for fish detection, tracking, and counting using monocular or stereo cameras have been investigated .There are, however, several challenges for underwater image/video analyses. First, the fast attenuation and uniformity of LED illumination make many foreground objects have relatively low contrast with the background, and fish with similar ranges from the cameras can have significantly different intensity because of the differences in angle of incidence as well as reflectivity of fish body among species.[2] These factors make segmentation of fish difficult. Second, the ubiquitous noise is created by non fish objects, such as bubbles, organic debris, and invertebrates, which can easily be mistaken a real fish. Third, the low frame rate (LFR) of capturing results in poor motion continuity and frequent entrance/exit of the field of view (FOV) for fish targets makes conventional multiple-target tracking algorithms infeasible under such circumstances.The object recognization in various aspects have been well investigated , there has some challenges to identify fish in an restricted natural habitat. Because of poor image quality , non – lateral fish views or curved body shapes ,there is high indecision existing in an data. To address the indecision issue , the partial classification approach is used.

In this paper, a multiple fish-tracking algorithm for trawl-based underwater camera systems is proposed to perform automatic fish size estimation and counting. Specifically, the contributions of this paper include: 1) a novel non-rigid part model that represents both appearance and geometric attributes of the fish body; 2) an unsupervised learning algorithm of non-rigid part model based on systematic part initialization and an expectation-maximization-like alternating optimization algorithm; 3) a novel hierarchical partial classification that successfully handles data uncertainty and class imbalance; 4) a formal approach that determines the decision criteria based on an optimization formulation. The rest of this paper is organized as follows. Section II gives a brief review of the related work. Section III describes the problem formulation. Section IV introduces the unsupervised non-rigid part model learning algorithm. Section V describes the hierarchical partial

classification method. Section VI reports the experimental results on fish species recognition, and the conclusion is given in Section VII.

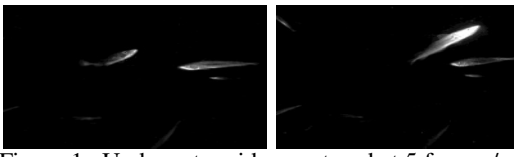

Figure 1 : Underwater video captured at 5 frames/s,

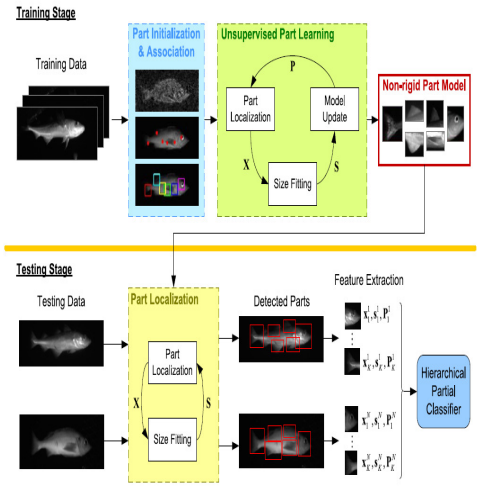

Figure2: The overall system architecture which describes the steps in precise manner.

II. FISH RECOGNITION

Live fish recognition is one of the most crucial elements in camera-based fisheries survey systems . Similar to most recognition frameworks, the successful extraction of informative features is the key to enhancing fish recognition performance[1] .Existing feature extraction techniques are divided into two categories, namely the supervised and unsupervised methods. Supervised methods represent a fish by pre-specified features that adopt common low-level image descriptors such as contour shape. For instance, Lee et al[3] used a curvature analysis approach to locate critical landmark points. The contour segments of interest were extracted based on these landmark points to achieve satisfactory shape-based species classification results.[3].In their subsequent work ,the features were further extended to include several shape descriptors such as Fourier descriptors, polygon approximation and line segments.For the dealing the vast amount of data ,we choose the phase Fourier transform (PFT) approach.

A .Part Initialization:

For the efficiency in dealing with vast data amount, we adopt the phase Fourier transform (PFT) approach. Given an image, we calculate its 2-D discrete Fourier transform, which can be expressed by

$$F(I(x,y)) = M(u,v)e^{j\emptyset(u,v)} \quad (1)$$

For part initialization, The saliency is obtained by taking the inverse Fourier transform of only the phase term,

$$S(x,y) = G_{\sigma}(x,y) * F^{-1}(e^{j\emptyset(u,v)}) \quad (2)$$

Non-maximal suppression is applied to extract local maxima from the saliency map. Here we use the object segmentation mask produced to discard salient points in the background. Finally, we choose the top K local maxima locations, each of which serves as an initial part.

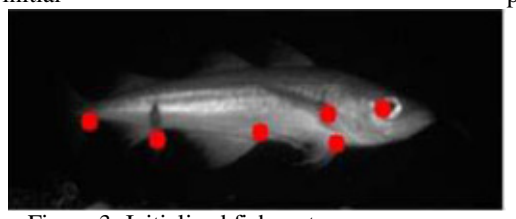

Figure 3: Initialized fish parts.

B. Part Association:

Due to the pose variation, one object part may appear in different locations in two images[2]. To ensure the correctness of learning, it is important to align the extracted points from one image to another. In the proposed method, we formulate part identification as a one-to-one association problem and apply the relaxation labelling process. Introducing the outlier notion brings in two advantages[1]. Firstly, it allows for handling substantial pose variations or partial occlusions. Moreover, it facilitates the imposition of one-to-one match constraint between two part location sets.

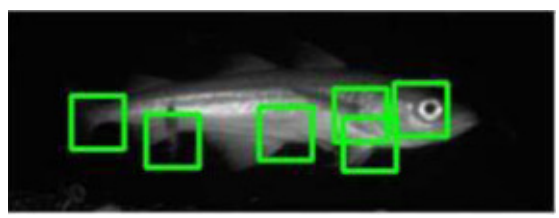

Figure 4: Associated fish parts.

III. UNSUPERVISED MODELDISCOVERY

The non-rigid object part model is learned by solving the optimization problem where the objective function J $(\mathbf{P}, \mathbf{X}, \mathbf{S})$. In order to effectively update the variables $\mathbf{P}, \mathbf{X}, \mathbf{S}$ without human assistance in the loop, For each training image the part locations and sizes are initialized based on the saliency detection and relaxation labelling procedure [1].The appearance for each part is initialized by the average value of the corresponding block over the training set.

A. Part appearance:

The goal here is to find the optimal part appearance **P** without changing location **X** and size **S**. When optimizing ,the separation cost term from is constant with fixed **X** and **S**.. Therefore the part appearance can be found by,

$$\sum_{m=1}^{N} d(P_i, \phi(I_i^m)) - \sum_{j=1}^{K,} d(P_i, P_j) \quad (3)$$

B. Part Localization:

In this step, the part features **P** and sizes **S** are given. By updating **X**, we localize the sub-region that corresponds to each part in each image.[2]. The discrimination cost term becomes a constant since **P** is fixed. The gradient estimate is iteratively calculated by using pixels within the located sub-region. The part location is updated by

$$x_i^m(t+1)=\frac{\sum_{j=1}^{n_p} k\left(z_j-x_i^m(t)\right)w_j\left(z_j-x_i^m(t)\right)}{\sum_{j=1}^{n_p} k\left(z_j-x_i^m(t)\right)w_j} \quad (4)$$

where K(.) is the kernel function, n_p is the number of pixels in the part and w_j is the sample weight at z_j,

$$w_j = -\frac{\partial}{\partial x_i^m}\left[d\left(P_i,\phi(I_i^m)\right)\sum_{j\neq i} V_{i,j}^m\right]\Big|_{x_i^m=z_a} \quad (5)$$

C. Part size selection:

The goal is to optimize the part size ,while fixing the appearance **P** and location **X**. Same as the part localization step, the discrimination cost term held constant since **P** is fixed.

$$S_i^m(t+1) = S_i^m(t)b^{r'},\ r'=\frac{\sum_{r\epsilon\Omega}\sum_{j=1}^{n_p} H(z_j,r)w(z_j)r}{\sum_{r\epsilon\Omega}\sum_{j=1}^{n_p} H(z_j,r)w(z_j)} \quad (6)$$

where Ω refers to range in the scale space centred at the current part size $S_i^m(t)$, H is the scale kernel, $n_p$is the number of pixels inside the current part size, and $w(z_a)$ is the sample weight defined . The iteration stops when$|r'|$is small enough..

V. HIERARCHICAL CLASSIFICATION

To exploit the information from uncertain data without introducing misclassification, we develop a novel technique that learns a hierarchical structure for the classes and allows for indecision for ambiguous data.[1]. A class hierarchy, i.e., a binary decision tree with one classifier at each node, is generated to determine the grouping of classes in higher levels. The grouping labels can serve as coarse categorization results when the exact class label cannot be identified. In the testing phase, the input data instance is examined by layers of classifiers, each of which gives a prediction label. If the instance falls in the indecision range at any layer, the classification procedure stops and returns an incomplete sequence of class labels. In this way, misclassifications are avoided without losing the entire information provided by uncertain data.

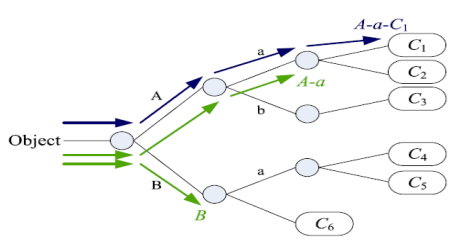

Figure 5: The class hierarchy is earned from the training data via an unsupervised clustering procedure. The method for solving in part model learning step depends on the distance metric chosen for features. In our experiments, we measure the distance metric between feature descriptors **P** and **Q** by the normalized correlation function, and the accuracy of feature learning methods are given in table I.

TABLE I ACCURACY OF FEATURE LEARNING METHODS

Method	Accuracy (%)
Template [14]	88.4
Alignment [15]	91.6
PAFF [25]	90.1
Proposed (grid, spatial-order)	86.9
Proposed (saliency, spatial-order)	89.5
Proposed (saliency, relax-label)	**93.8**
Proposed (removing pectoral fins)	87.3

VI. PERFORMANCE ANALYSIS

A visualization of some fish parts discovered by the proposed algorithm .In addition to the head and the caudal fin (tail), which are most seminal fish parts, the non-rigid part model locates the pectoral fin and classify the fish based on its characteristics. In particular, the pectoral fin part is systematically discovered by the unsupervised non-rigid part model learning. If we remove the corresponding features deliberately, the recognition performance is decreased, the performance analysis are shown in table II.

TABLE II PERFORMANCE MEASURES

Method	AP (%)	AR (%)	AC (%)
Flat SVM	88.5	76.9	95.7
PCA-Flat SVM	88.9	77.7	95.4
CART [37]	52.9	53.6	87.0
Taxonomy	87.2	76.1	95.3
BEOTR [4]	91.4	84.8	97.5
Proposed	**92.1**	**91.6**	**97.7**

VI. EXPERIMENTAL RESULT

This section presents experimental results with a datasets of top 15 species which consists of 26418 fish images, and the fish images are captured by cam-trawl system .In the image , we apply our fish segmentation algorithm to find the bounding boxes and the segmentation masks.

A) Experimental Setup – each image is scaled so that the bounding box is no larger than 200 × 200 pixels with its aspect ratio preserved. The number of parts is empirically determined as K = 6 in the experiments.

i) Pre-processing – The number of parts is empirically determined as K = 6 in the experiments Each part is initialized with a size of 48 × 48 pixels in the rescaled images.

ii) Part initialization and association- The saliency is obtained by taking the inverse Fourier transform of only the phase term Non-maximal suppression is

applied to extract local maxima from the saliency map. Then the object segmentation mask is produced by, to discard salient points in the background. Finally, we choose the top K local maxima locations, each of which serves as an initial part. Whereas due to the pose variation, one object part may appear in different locations in two images. To ensure the correctness of learning, it is important to align the extracted points from one image to another. In the proposed method, we formulate part identification as a one-to-one association problem and the relaxation labeling process is applied. The goal of part identification is to find an optimal association from candidate parts to reference parts, which is similar to the matching problem between two sets of 2-D points that undergo some non-rigid Deformation.

a b c

Figure 6a,b,c – represents the fish parts by non – rigid part learning.

iii) Feature extraction(non- rigid part model) – Followed by the part initialization and association, the part features are extracted using features of location, size, and appearances. For location x, and size s, the mean shift and scale-space mean shift algorithm is applied and hence it is solved by means of optimization function. Whereas for appearance P, the SIFT descriptors and the weighted color histogram are applied to each and every 4 pixels, and they are sampled densely and the dimensionality is reduced to 128. From that local and global features are obtained. Finally, the resultant is scaled to 48 X 48 pixel and the feature vector of an image is obtained.

iv) Hierarchical classification- The partial labels are assigned and gather information about the uncertain data. The grouping labels can serve as coarse categorization results when the exact class label cannot be identified. In the testing phase, the input data instance is examined by layers of classifiers, each of which gives a prediction label. If the instance falls in the indecision range at any layer, the classification procedure stops and returns an incomplete sequence of

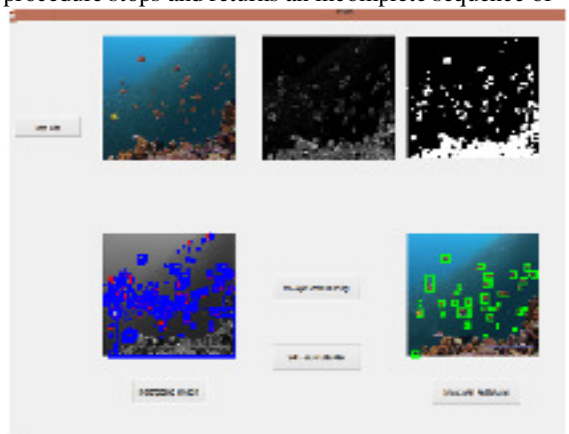

Figure7 – This represents the fish images are recognized in an green color bounding boxes.

VII. CONCLUSION

A novel framework for underwater fish recognition is proposed. The proposed framework is facilitated by unsupervised learning algorithms and thus reduces the requirement of human interference comparing to existing approaches. The non-rigid part model effectively discovers discriminative parts by adopting saliency and relaxation labelling. Fitness, separation and discrimination of parts are considered for finding meaningful representations of fish body parts in a fully unsupervised fashion. On the other hand, data uncertainty and class imbalance are two of the most common issues in practical classification applications. The proposed hierarchical partial classification successfully handled these issues by enabling coarse-to-find categorization and thus retrieving partial information from those ambiguous data which are possibly misclassified or rejected by other algorithms. We further develop a systematic optimization approach to selecting decision criteria for partial classifiers by introducing the exponential benefit function. Experimental results show a favourable performance of fish recognition on both large-scale public dataset and practical highly-uncertain dataset of live fish.Future work includes that along with the fish species, the type of fish is also identified. It is the hope that, this study may give the maximum yield from the fishery field.

REFERENCE

[1] M.-C. Chuang,J.-N. Hwang, K. Williams, and R. Tower, "Tracking live fish from low-contrast and low-frame-rate stereo videos," *IEEE(Trans. Circuits Syst. Video Technol., vol. 25, no. 1, pp. 167–179,*Jan. 2015).

[2] Chuang, Meng-Che, et al. "Automatic fish segmentation via double local thresholding for trawl-based underwater camera systems." 2011 18th IEEE International Conference on Image Processing. IEEE, 2011.

[3] Li, Xiu, et al "Accelerating fish detection and recognition by sharing CNNs with objectness learning." *OCEANS 2016-Shanghai*. IEEE, 2016

[4] Chuang, Meng-Che, et al. "Multiple fish tracking via Viterbi data association for low-frame-rate underwater camera systems." *2013 IEEE International Symposium on Circuits and Systems (ISCAS2013)*. IEEE, 2013.

[5] Chuang, Meng-Che, et al. "Underwater Fish Tracking for Moving Cameras based on Deformable Multiple Kernels." (2016).

[6] Chuang, Meng-Che, et al. "Recognizing live fish species by hierarchical partial classification based on the exponential benefit." *2014 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2014.

[7] Li, Chongyi, et al. "Underwater Image Enhancement by Dehazing With Minimum Information Loss and Histogram Distribution Prior." *IEEE Transactions on Image Processing* 2016.

[8] Y. Baram, "Partial classification: The benefit of deferred decision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 769–776,Aug. 1998.

[9] E. Gavves, B. Fernando, C. G. M. Snoek, A. W. M. Smeulders, and T. Tuytelaars, "Fine-grained categorization by alignments," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Dec. 2013, pp. 1713–1720.

[10] S. Gao, I. W.-H. Tsang, and Y. Ma, "Learning category-specific dictionary and shared dictionary for fine-grained image categorization," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 623–634, Feb. 2014.

[11] N. Zhang, R. Farrell, F. Iandola, and T. Darrell, "Deformable part descriptors for fine-grained recognition and attribute prediction," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2013, pp. 729–736.

[12] C. Göering, E. Rodner, A. Freytag, and J. Denzler, "Nonparametric part transfer for fine-grained recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2014, pp. 2489–2496.

[13] S. Singh, A. Gupta, and A. A. Efros, "Unsupervised discovery of mid-level discriminative patches," in *Proc. 12th Eur. Conf. Comput. Vis. (ECCV)*, 2012, pp. 73–86.

[14] C. Doersch, A. Gupta, and A. A. Efros, "Mid-level visual element discovery as discriminative mode seeking," in *Proc. Adv. Neural Inf. Process. Syst.*, 2013, pp. 494–502.

[15] 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 *Proc. 18th Brazilian Symp. Comput. Graph. Image Process. (SIBGRAPI)*, 2005, pp. 173–180.

[16] Z.-G. Liu, Q. Pan, G. Mercier, and J. Dezert, "A new incomplete pattern classification method based on evidential reasoning," *IEEE Trans. Cybern.*, vol. 45, no. 4, pp. 635–646, Apr. 2015.

[17] C. Lian, S. Ruan, and T. Denoeux, "An evidential classifier based on feature selection and two-step classification strategy," *Pattern Recognit.*, vol. 48, no. 7, pp. 2318–2327, 2015.

[18] D. J. Park and C. Kim, "A hybrid bags-of-feature model for sports scene classification," *J. Signal Process. Syst.*, vol. 81, no. 2, pp. 249–263, 2014.