Identifying Individual Clown Fish

Xiao Liu
School of Informatics
University of Edinburgh

Robert B. Fisher
School of Informatics
University of Edinburgh

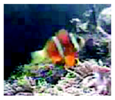
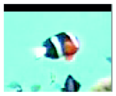
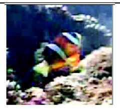
***Abstract*—This paper presents a clustering approach to estimating the number of individuals present in a set of *Amphiprion clarkii* (clownfish) observed live from a Taiwanese coral reef. Previous approaches to individual identification have addressed species where there are many individuals with some distinctive pattern and few observations of each. By contrast, fish observations from a fixed camera tend to produce many observations of a few individuals, whose differences in appearance are more quantitative than qualitative. The paper describes a k-means clustering approach to estimating the number of individuals. Although there is no absolute ground truth (*e.g.* from tagging), our data analysis shows 6 or 7 individuals were present in a collection of 785 observations.**

I. INTRODUCTION

Marine ecologists are interested in estimates of fish species abundance, *e.g.* to monitor population changes as a consequence of water quality, water temperature, and fish capture, etc. Automatic systems are being developed to detect and track undersea fish [1], but many fish species are resident or semi-resident, meaning that individuals do not move far from a home territory. One consequence is that an individual is likely to be observed repeatedly if fixed cameras are used. Therefore, estimating abundance requires estimating the number of individuals rather than the number of observations. This paper presents a clustering approach to estimating the number of individuals present in a set of observations of *Amphiprion clarkii* (clownfish) [9]. **The key contribution of the paper is a method for estimating the number of individuals when observations are very similar and there are many observations of the same individual.**

Markerless methods for identifying individuals have concentrated on external visual patterns [11] such as spot patterns on penguin chests [3], zebra stripe patterns [3], [5], snake stripe patterns [11], [2], and the fin shape and coloration patterns on sharks [12] and humpback whales [10]. By contrast, fish observations from a fixed camera tend to produce many observations of a few individuals. Therefore, one key issue is how to partition the set of observations into subsets associated with individuals, which is the topic of this paper. What makes this problem difficult is there are no markings that are distinctive to individuals, unlike many of the examples mentioned above. Here, the individuals are distinguished by slight variations in the color or positioning of features. Thus, the problem considered here is closer to the problem of unconstrained human face recognition, where identification is based on subtle differences, and the target can be seen from many unpredictable 3D viewpoints (partially constrained by a fish's tendency to largely swim upright).

The data used here was based on a subset of coral reef fish video captured *in vivo*, and computationally detected, tracked and identified by species by the Fish4Knowledge project from undersea cameras that observed some coral reefs in Taiwan [1]. The project used a subset of 785 manually validated *Amphiprion clarkii* observations. Some details of the observations are given in the table below. The data came from 7 videos that differ by date, time, or camera. Given the appearance of individual fish, the behavior of the species (tend to stay within 1-3 meters of home location), the separation of the cameras (7 m), and the temporal separation of the subsets, we are reasonably confident that these are 6 or 7 different individuals.

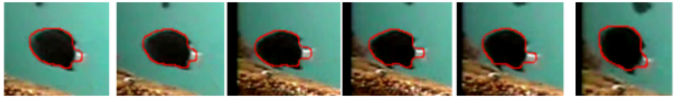
Fish				
Count	112	110	132	124
Camera	2	2	2	3
Date	8/11/2010	25/9/2010	22/8/2010	20/8/2010
Fish				
Count	84	122	101	
Camera	3	2	2	
Date	8/11/2010	25/9/2010	24/11/2010	

The main stages of the method presented here are: 1) Extract and select distinctive features from fish images (Sec II), 2) Group individuals using clustering methods and estimate individual numbers (Sec III), and 3) Evaluate experimental performance (Sec IV).

II. FISH DESCRIPTION AND FEATURE SELECTION

K-means will be used to cluster similar appearing individuals, so a set of discriminative feature values is needed. This section describes these features, and how a useful feature subset was selected. Some preprocessing [8] was required before the features were calculated. The detection algorithm created a bounding box around the fish, and the actual boundary was found using Grab-cut.

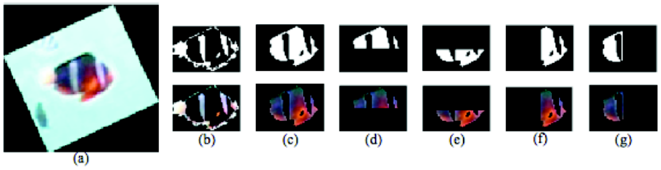
Examples of the detected fish and boundaries are shown here:



The resulting closed Grabcut boundary was used to create a binary mask for the whole fish. The isolated fish was then re-oriented computationally to face horizontally with the head facing to the right. The main steps of the re-orientation algorithm are: 1) estimate elongation axis, assuming that this is the fish main head-to-tail axis, 2) rotate the detected fish so this is horizontal, 3) identify the head based on boundary curvature, and 4) flipping the image if necessary. An example of an original detection, the binary mask from the Grabcut output and the re-oriented fish are seen here:

(a) The original image (b) The binary image (c) Orientation result

After segmentation, a large number of descriptive properties are computed for each fish. Some of the features[9] are based on the ratio of region or stripe widths, or the ratio of properties computed from the top/bottom or front/back of the fish. The stripes are found by color thresholding, and the shape splitting is based about the center of mass. An example of the regions produced by this method is seen here (showing stripes, body, body top, body bottom, body front, body rear):



The properties extracted are from a number of classes, based on color, region shape and interior texture from the segmented fish (but no data from the background was used except for that arising from segmentation errors). The 2624 features developed by [8] were used, plus 467 additional features specific to the individual identification task were developed, as described below. The full list of features is given in the Appendix. All features are z-normalized (subtracting the mean and dividing by the standard deviation) after outlier removal. Four families of new features were developed.

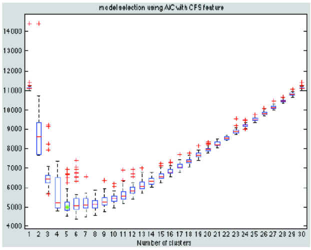
1) Color ratios. Because the observed color depends on the illumination on the fish and the color loss through the water, a set of properties based on the ratios of colors was used. The use of a ratio of data from approximately the same position, *e.g.* upper and lower body pixels on the fish, cancels the multiplicative effect of the illumination and water absorption caused color loss that affects both sets of pixels to leave a value that is largely a function of the albedo and camera sensitivity. The ratios used were: between 1) the RGB color of the non-stripe area to the RGB color of the stripe area, 2) a similar ratio of the colors from the top and bottom halves of the fish (split on the horizontal through the center of mass), 3) a similar ratio of the colors from the front and back halves of the fish (split on the vertical through the center of mass). All these were encoded from RGB histograms of the ratios. Three similar ratios were computed from each of the average RGB values from each of the two regions.
2) The percentage of the fish midline that is stripe pixels (after the tail area is removed) is calculated from the binary image seen at the left above.
3) The percentage of body area that is contained in the white stripes is another feature calculated from the binary image seen at the left above (again without the tail region).
4) The ratio of the area of the head to the tail (found by detecting the concavities at the tail/body interface).

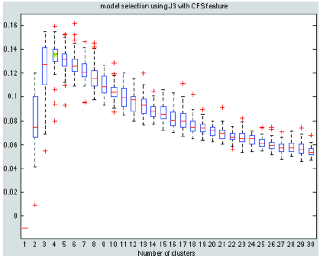
Huang's [8] 13 feature families and these 6 new feature families [9] together totaled 3091 descriptor values for each fish. Given the high dimensionality of the descriptors, feature selection was used. Four methods were considered: Principal Component Analysis on the full dataset: Principal Component Analysis on each of the 19 families, Spectral Feature Selection, and Correlation-based feature selection [7], from which the latter had the best performance. This approach chooses features that are highly correlated with the class labels (from training data), but are uncorrelated with each other. A sequential feature selection stage [4] follows generation of a set of candidate features, resulting in 50 selected features.

III. FISH CLUSTERING AND NUMBER OF INDIVIDUALS

The project assumed that individual fish would have appearance descriptors that clustered well, and there would be an unknown number of distinct clusters corresponding to different individuals. Hence, an unsupervised clustering method was used, in this case K-means with the Euclidean distance. Other clustering approaches were considered but K-means was simple and effective.

The key issue with K-means is choosing the best value of K, in that the true number of individuals was unknown. We explored the use of 5 model selection criteria: Akaike and Bayesian information criteria, the Dunns index, the Silhouette index and the J_3 cluster scatter metric. The most common value of K was used (or larger if a tie). Some typical results on a subset of fish observations containing individuals 1-4 can be seen in the figure below, which shows the value of the index vertically *versus* the number of features. From the plots we can see that the best balue for K is near $K = 5$. Five random trials were performed with different subsets of fish and with K varying from 2 to 30. The left plot shows the Akaike performance and the right plot shows the J_3 performance. The individual trajectories were allocated to separate 5-fold cross-validation subsets, so there was no contamination of known highly correlated samples between folds. Thus the number of samples in each fold was approximately equal but varied according to the number of detections in each track. The information criteria were used to select the value of K from the training data, which was then used by the test data.





The correlation-based feature selection method was chosen based on the goodness of its cluster validation measures [6] (over the training data using fish1-fish4). Five measures were considered (Entropy, Mutual Information, Cluster Purity, Rand Index and F-measure), and the correlation-based method had the best scores over all feature selection methods.

The confusion matrix for the fish1-fish4, $K = 5$ case is:

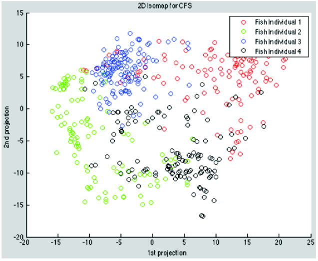
	fish1	fish2	fish3	fish4
cluster 1	21	0	0	1
cluster 2	0	26	0	0
cluster 3	5	0	20	0
cluster 4	0	0	0	25
cluster 5	0	1	6	1

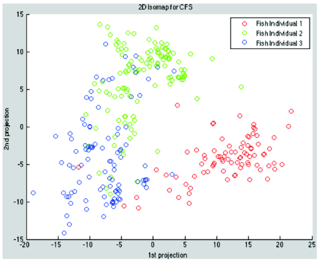
Assuming that the fish1-fish4 subsets actually represented 4 individuals, the smallest (5^{th}) cluster contained only 8 of the 106 test images, as compared to an average of 23 individuals for the other 4 clusters. We conclude that cluster 5 is not from a distinct individual. 87% of the fish were correctly clustered. We conclude that the selected features and clustering method is effective.

IV. EXPERIMENTS

Based on the methods proposed in the previous section, the two main experiments evaluated performance when using 3 new fish detection and tracking subsets (fish5-fish7) and performance when using all 7 fish subsets (which does include the training 4 subsets used for feature and K selection, but was the largest collection we had). As stated previously, we did not know the true identity of each fish. Because we used detections and trackings from different times, the locations are more than 3m separated (the typical home range for *Amphiprion clarkii* and the species is territorial), it is highly likely that there was only one individual observed in each camera and scene location and time. What is less certain is whether the individual observed at one time block is different from an individual observed in a later time block, Many of the time blocks are only 1-3 months apart, and clownfish life expectance is on the order of 6+ years Based on the dates, potentially individuals {1,2,3,6,7} and {4,5} could be the same individual, however, based on the appearances shown above, this is unlikely.

Before clustering, we used Isomap to visualize the samples in 2D, for the fish1-fish4 and fish5-fish7 sets as seen here:





There are clear groupings of fish in this data projection. Using all 7 fish shows some obvious clustering as well, but there is also considerable overlap due to limited dimensionality of the projection.

Model selection for the 3 fish subset proposed K=3 for the 3 fish and K=8 for the 7 fish subsets. K-means clustering was performed as described above (100 iterations), using the descriptive features previously selected when using the fish1-fish4 subset, with approximately 50% of data for training and the remainder for testing.

The confusion matrix for the fish5-fish7 case is:

	fish5	fish6	fish7
cluster 1	43	11	6
cluster 2	6	49	0
cluster 3	0	0	39

Clearly, the clusters are highly correlated with our hypothesized fish identification. To assess the scalability and separability, we repeated the clustering with all 7 fish subsets. Although there is a little feature selection contamination between training with the first 4 and then testing including these, poor performance here would demonstrate a scalability problem, which was not seen. The confusion matrix for the fish1-fish7 case is:

	f1	f2	f3	f4	f5	f6	f7
c1	45	0	0	5	0	43	6
c2	0	42	0	0	12	0	0
c3	12	0	64	1	2	15	4
c4	0	0	0	59	0	0	0
c5	0	0	0	0	30	0	0
c6	0	0	0	0	0	1	13
c7	1	6	0	0	1	0	24
c8	1	0	0	6	0	0	0

From the clustering results, it is probable that there are only 6 individuals in the dataset and that fish1 and fish6 are the same individual (cluster 1). Visual comparison of fish1 and fish6 does not expose any obvious differences.

V. CONCLUSIONS

The proposed combination of fish-specific features, reduced feature dimensionality by correlation-based feature selection, and k-means clustering produced reasonably plausible clusters.

Although we do not have certain ground truth here, the fact that property clustering produced clusters that mapped closely onto the *a priori* date and location subsets suggests that the analysis is close to correct.

The clownfish detections are all valid (manually checked), and the trajectories were also manually examined. So, in theory, all fish in the same trajectory are the same individual, and so should all be in the same cluster. One analysis that could be done was to look at the consistency of cluster ids assigned to fish in a given trajectory. Additional subsets of observations taken from more distant times periods could also be done to improve the confidence in the method, although, like with human face recognition, as the number of individuals increases, the likelihood of confusing individuals also increases.

As for technical improvements, one could explore additional discriminating properties to increase separation of clusters, improve the color ratio histogram analysis, identify which type of features contributed most to the clssification, incorporating trajectory information into the clustering process itself, alternative clustering methods, a weighting between the features (currently z-normalized, but all features are treated equally using the Euclidean distance), how many individuals could be included before the classification probability deteriorates, and a more stable and consistent model selection algorithm (the current use of voting between the 5 algorithms compensates for the considerable variability in the proposed K-means parameter K).

ACKNOWLEDGMENTS

The research was part of the Fish4Knowledge project, funded by the European Union Seventh Framework Programme and the Taiwan National Science Council.

REFERENCES

[1] B. J. Boom, J. He, S. Palazzo, P. X. Huang, H.-M. Chou, F.-P. Lin, C. Spampinato, R. B. Fisher; A research tool for long-term and continuous analysis of fish assemblage in coral-reefs using underwater camera footage. Ecological Informatics, 2014.

[2] A. Branzan Albu, G. Wiebe, P. Govindarajulu, C. Engelstoft, K. Ovaska; Towards Automatic Model-Based Identification of Individual Sharp-Tailed Snakes from Natural Body Markings. Proc. Int. Workshop on Visual observation and analysis of animal and insect behavior, 2008.

[3] T. Burghardt, N. Campbell; Generic phase curl localisation for an individual identification of Turing-patterned animals. Proc. Int. Workshop on Visual observation and analysis of animal and insect behavior, 2010.

[4] S. Derksen, H. J. Keselman. Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. British Journal of Mathematical and Statistical Psychology, 45(2):265 282, 1992.

[5] G. Foster, H. Krijger, and S. Bangay. Zebra fingerprints: towards a computer-aided identification system for individual zebra. African Journal of Ecology, 45(2):225.227, 2007.

[6] G. Gan, C. Ma, J. Wu. Data clustering: theory, algorithms, and applications, volume 20. Siam, 2007.

[7] M. A. Hall, L. A Smith. Feature selection for machine learning: Comparing a correlation-based filter approach to the wrapper. Proc. FLAIRS Conference, pages 235-239, 1999.

[8] P. X Huang, B. J Boom, R. B. Fisher. Hierarchical classification for live fish recognition. Machine Vision and Applications, 26(1), p. 89-102, 2015.

[9] X. Liu; Identifying Individual Clown Fish. MSc Dissertation, Informatics, Univ. of Edinburgh, 2013.

[10] E. Ranguelova, M.J. Huiskes, E.J. Paumels; Towards computer-assisted photo-identification of humpback whales, ICIP04(III: 1727-1730).

[11] C. W. Speed, M. G. Meekan, and C. J. A. Bradshaw. Spot the match - wildlife photo-identification using information theory. Frontiers in Zoology, 4(2):1.11, 2007.

[12] A. M. Van Tienhoven, J. E. Den Hartog, R. A. Reijns, and V. M. Peddemors. A computer-aided program for pattern matching of natural marks on the spotted raggedtooth shark carcharias taurus. Applied Ecology, 44:273.280, 2007.

APPENDIX

Features generated from the detected fish

	Size	Name	Item
1	510	Normalized RG Color hist	Head, Tail, Top, Bottom, Whole
2	255	H hist in HSV	Head, Tail, Top, Bottom, Whole
3	110	Normalized RG Color (re-hist)	Head, Tail, Top, Bottom, Whole
4	55	H hist in HSV (re-hist)	Head, Tail, Top, Bottom, Whole
5	1	CS ratio	
6	1	CS half tail area ratio	
7	12	Density	RGB
8	720	Co-occurrence matrix	Correlation, Energy, Homogeneity InvDiffMoment, ClusterShade ClusterProminen, MaxProbiliry Autocorrelation
9	42	Moment Invariants	Head, Tail, Top, Bottom, Whole, Half-head, Half-tail
10	680	HOG	Level 0, Level 1, Level 2, Level 3
11	15	Fourier	
12	160	Gabor	Head, Tail, Top, Bottom, Whole
13	63	Affine Moment Invariants	Head, Tail, Top, Bottom, Whole, Half-head, Half-tail
14	2	head and tail Area ratio	Head, Tail
15	153	color ratio	Chromatic body and white stripes
16	153	color ratio	Top and bottom parts of the chromatic body
17	153	color ratio	Front and posterior parts of the chromatic body
18	5	stripes length ratio	
19	1	stripes area ratio	