Automated Identification of Individual Great White Sharks from Unrestricted Fin Imagery

Benjamin Hughes benjamin.hughes@bristol.ac.uk Tilo Burghardt tilo@cs.bris.ac.uk

Introduction. Recognising individuals repeatedly over time is a basic requirement for field-based ecology and related life sciences [5, 6]. In this paper we propose a visual identification approach for great white shark fins as outlined in Figure 1, one that is applicable to unconstrained fin imagery and fully automates the pipeline from feature extraction to matching of identities. We pose the associated vision task as a fine-grained, multi-instance classification problem for flexible, smooth and partly occluded object parts. To the best of our knowledge, the proposed system is the first fully automated contour-based animal biometrics system.

Contour Stroke Object Model. A key technical contribution of the paper is a contour stroke model trained for fin detection. It combines a partitioning of ultrametric contour maps (UCM; generated via [3, 4]) with normal descriptors, dense local features and a random-forest regressor. UCM region boundaries are partitioned at keypoints detected as local maxima of the Difference of Gaussian (DoG) function $D(u, \sigma)$ detailed by Zhang et al. [10]. We generate fin candidates as contour strokes by sampling the region contour between every permutation of keypoint pairs. Each stroke is then described by a 160 dimensional feature vector consisting of two components: the first is a bag of opponentSIFT [9] visual words, the second describes boundary shape using a histogram of boundary normals. A random forest regressor is finally trained to predict the fin-like quality of contour stroke candidates. Figure 2 illustrates fins detected by this strategy. In common with the segmentation stage of Arandjelovic and Zisserman's 'Bag of Boundaries' pipeline [2], our contour stroke model separates objects in natural images and against cluttered backgrounds, but differs in that it exploits species characteristic shape encoded in the open contours, to specifically and explicitly detect fin parts.

Biometric Encoding: [1, 2] perform smooth object recognition using semi-local boundary normal and HoG+occupancy representations, respectively. These allow efficient and robust matching, but their encoding of inter-class variance will always be sub-maximal. Meanwhile, the semi-automated fin recognition system DARWIN [8] uses a global 2D Cartesian representation. This encoding maximises inter-class variance, but is sensitive to partial occlusions and detection errors. By contrast we utilise semi-local *and* global shape descriptions in a vector-based combinatorial encoding strategy, one that enables efficient and accurate individual recognition while being robust to noisy, partially occluded input.

As with generating contour strokes, we combinatorially sample fin boundaries between pairs of keypoints detected using $D(u, \sigma)$, reparameterised for fine-grained feature detection. We then scale-normalise the resulting contour subsections and encode their shape over various spectral scales. Denoting the set of descriptors for a query object at scale *j* as D_Q^j , for each query descriptor $d_i^j \in D_Q^j$, we find the two nearest neighbours $(NN_C(d_i^j), NN_{\bar{C}}(d_i^j))$ where *C* is the class of the nearest neighbour and \bar{C} is the set of all other classes. Subsection information is then combined using a multi-scale decision rule based on the Local Naive Bayes Nearest Neighbour (LNBNN) classifier [7]:

$$\hat{C} = \arg\max_{C} \sum_{j=1}^{\nu} w_{j} \cdot \sum_{i=1}^{|D_{Q}^{j}|} ||d_{i}^{j} - NN_{\bar{C}}(d_{i}^{j})||^{2} - ||d_{i}^{j} - NN_{C}(d_{i}^{j})||^{2}$$
(1)

Results: We benchmark system performance using a dataset presenting 85 individuals and consisting of 2456 images (one labelled reference image per shark), exhibiting significant variability in waterline and splash occlusion, viewpoint, orientation and scale (see various Figures). Overall, 72% of queries are correctly identified with a mAP of 0.79. Discounting detection errors and partial occlusions, this rises to 82% accuracy with a mAP of 0.84. Fin detection performance depends on the quality threshold *t* for detected fin candidates, as measured by *F*-scores. For unoccluded instances we obtain: AP=0.98 for t > 0.7 and AP=0.85 for t > 0.9.

Conclusion: We have combined a contour stroke model with a spectral and combinatorial contour encoder. The method shows promising individual shark fin identification performance. We conclude applicability at accuracy levels ready to assist human identification without labelling.

Department of Computer Science University of Bristol Bristol, UK



Figure 1: SYSTEM OVERVIEW: The system performs a coarse and a fine-grained recognition task. The first is to simultaneously segment and detect shark fins, and the second is to recognise individuals.



Figure 2: FIN DETECTION AS OPEN CONTOUR STROKES: Segmentation algorithms [5] (left images) alone often fail to detect fins. Our stroke model returns the fin section more robustly as shown in solid white.

- [1] Ognjen Arandjelovic. Object matching using boundary descriptors. BMVC 2012.
- [2] Relja Arandjelovic and Andrew Zisserman. Smooth object retrieval using a bag of boundaries. ICCV 2011.
- [3] Pablo Arbeláez et al. Multiscale combinatorial grouping. CVPR 2014.
- [4] Piotr Dollár and C Lawrence Zitnick. Structured forests for fast edge detection. ICCV 2013.
- [5] Hjalmar S K\u00fchl and Tilo Burghardt. Animal biometrics: quantifying and detecting phenotypic appearance. Trends in ecology & evolution 2013.
- [6] AD Marshall and SJ Pierce. The use and abuse of photographic identification in sharks and rays. Journal of Fish Biology 2012.
- [7] Sancho McCann and David G Lowe. Local naive bayes nearest neighbor for image classification. CVPR 2012.
- [8] R Stanley. Darwin: Identifying dolphins from dorsal fin images. Eckerd College 1995.
- [9] Koen EA Van De Sande et al. Evaluating color descriptors for object and scene recognition. PAMI 2010.
- [10] Xiaohong Zhang et al. Robust image corner detection based on scale evolution difference of planar curves. Pattern Recognition Letters 2009.



Figure 3: INDIVIDUAL IDENTIFICATION EXAMPLES: left images are queries and right ones are predicted individuals. Coloured lines indicate ten sections contributing most evidence for the matched individual.