Chapter 5

Automatic Fish Segmentation and Recognition for Trawl-Based Cameras

Meng-Che Chuang
University of Washington, USA

Jenq-Neng Hwang
University of Washington, USA

Kresimir Williams
National Oceanic and Atmospheric Administration, USA

ABSTRACT

Camera-based fish abundance estimation with the aid of visual analysis techniques has drawn increasing attention. Live fish segmentation and recognition in open aquatic habitats, however, suffers from fast light attenuation, ubiquitous noise and non-lateral views of fish. In this chapter, an automatic live fish segmentation and recognition framework for trawl-based cameras is proposed. To mitigate the illumination issues, double local thresholding method is integrated with histogram backprojection to produce an accurate shape of fish segmentation. For recognition, a hierarchical partial classification is learned so that the coarse-to-fine categorization stops at any level where ambiguity exists. Attributes from important fish anatomical parts are focused to generate discriminative feature descriptors. Experiments on mid-water image sets show that the proposed framework achieves up to 93% of accuracy on live fish recognition based on automatic and robust segmentation results.

INTRODUCTION

Fish abundance estimation (Hankin and Reeves, 1988), which often calls for the use of bottom and mid-water trawls, is critically required for the conservation and management of commercially important fish populations. To improve the quality of surveys, we developed the Cam-trawl (Williams, Towler and Wilson, 2010) to capture visual data (images and/or videos) of live fish. The absence of the codend

DOI: 10.4018/978-1-4666-9435-4.ch005
allows fish to pass unharmed to the environment after being sampled by cameras. The captured visual data provide much of the information that is typically collected from fish that are retained by traditional trawl methods.

Camera-based sampling for fish abundance estimates, however, generates vast amounts of data, which present challenges to data analyses. These challenges can be reduced by using image/video processing and computer vision techniques for automated object localization, tracking, size estimation and recognition. A successful development of these algorithms will greatly ease one of the most onerous steps in camera-based sampling. To address these needs, we have developed algorithms that successfully analyze the collected data by performing fish segmentation (Chuang, Hwang, Williams and Towler, 2011), length measurement (Chuang, Hwang, Williams and Towler, 2014), counting and tracking (Chuang, Hwang, Williams and Towler, 2013) and of species identification (Chuang, Hwang and Williams, 2014; Chuang, Hwang, Kuo, Shan and Williams, 2014). These developments allow for monitoring the amount and species composition of fish schools, and thus provides a mean of assessing the status of fish stocks and the ecosystem.

There are several challenges in developing image processing or computer vision techniques for analyses of underwater imagery. The fast attenuation and non-uniformity of artificial illumination (e.g. by LED strobes) 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. In addition, the ubiquitous noise is created by non-fish objects such as bubbles, organic debris and invertebrates, which can easily be mistaken as real fish. These factors make localization of fish difficult.

On the other hand, while object recognition in various contexts has been well investigated, there exist fundamental challenges to identifying fish in an unconstrained natural habitat. For freely-swimming fish, there is a high uncertainty existing in many of the data because of poor image quality, non-lateral fish views or curved body shapes. Critical information in these data may be lost or comes with large measurement error. Even without uncertainty, fish share a strong visual similarity among species. Common features for image classification are hence not discriminative since they represent merely the global appearance of an object.

In this chapter, a hierarchical partial classification based on the novel exponential benefit function for recognizing live fish images is proposed to address all the aforementioned issues. Specifically, we 1) adopt double local thresholding and histogram backprojection techniques to produce an accurate fish segmentation for underwater video data with non-uniform illumination and low contrast; 2) build a class hierarchy by unsupervised learning and then introduce partial classification to allow assignments of incomplete but high-level labels; 3) define the exponential benefit to evaluate partial classifiers, and hence formulate the selection of decision criteria as an optimization problem; 4) learn a fish classifier by using part-aware features to identify visually-similar fish species. Experiments show that the proposed system achieves a favorable recognition accuracy on underwater fish images with high uncertainty and class imbalance.

BACKGROUND

The conservation and management of fished stocks requires, among other things, animal abundance estimates. Fisheries scientists often use bottom and midwater trawls to estimate fish abundance when