Ecological Informatics 23 (2014) 13–21

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Real-world underwater fish recognition and identification, using sparse representation

Yi-Hao Hsiao [a,b,*], Chaur-Chin Chen [c], Sun-In Lin [b], Fang-Pang Lin [b,**]

[a] Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan
[b] Data Computing Division, National Center for High-Performance Computing, Hsinchu, Taiwan
[c] Institute of Information Systems and Applications, National Tsing Hua University, Hsinchu, Taiwan

ARTICLE INFO

Article history:
Received 30 November 2012
Received in revised form 27 September 2013
Accepted 11 October 2013
Available online 24 October 2013

Keywords:
Fish recognition
Sparse representation classification
Compressive sensing
Fish identification

ABSTRACT

In this paper we describe how a distributed real-time underwater video observational system, developed and operated in southern Taiwan, can be used for visual environmental monitoring of a coral reef ecosystem. The method makes use of an innovative fish recognition and identification technique for real-world automatic underwater observation. Our research demonstrated that advanced fish recognition and identification techniques can be used to study fish populations and to identify species of fish that appear for the first time in particular areas of interest. The observational system subsequently accumulates massive tera-scale video data that can be used for long-term studies on coral reef fish. The system has the capacity for efficient and accurate recognition of fishes from the video dataset, which is recorded in a setting of biological abundance in a coral reef ecosystem. A simple and effective preprocess for fish detection from the video data has been developed, in which multiple bounding–surrounding boxes are introduced to discriminate between swimming fish and other moving objects, such as moving sea anemones and drifting water plants. Additional data, including images of various features from a number of fish species, taken at various angles and illumination conditions, can form the basis for a fish-category database. A maximum probability, partial ranking method, based on sparse representation-based classification (SRC-MP), is proposed for real-world fish recognition and identification. Eigenfaces and Fisherfaces are used to extract feature data, by means of the fish-category database. Two parameters — feature space dimension and partial ranking value — are used to optimize the solutions, in which the recognition and identification rates can respectively achieve 81.8% and 96%. Experimental results show that the proposed approach is robust and highly accurate for the use of fish recognition and identification of real-world underwater observational video data.

1. Introduction

Direct observation of plants and animals in their natural environment can be an important aspect of marine ecosystem research. Underwater environments constantly change and long term data sets are often required to delineate the complex processes of ecological succession. It is, however, difficult to sustain long term real time observation, mostly because of the inaccessibility of the marine environment (Edgington et al., 2007). Collection of long term data relies mainly on well-designed monitoring plans. Fish observations traditionally involve scuba diving for a period of time, to record videos and to collect physical, chemical and biological data. When using such methods, the collection of sufficient data for statistical analysis often requires a long-term, labor-intensive effort. Recent advances in information and communication technology have enabled the development of a distributed real-time underwater video observation system that can be used for long term observation of coral reef ecosystems. In the present study we made use of such techniques for underwater observations along the southern tropical coast of Taiwan (Shiau et al., 2010). We were able to broadcast the video data in real-time via the Internet. The data collection was archived to form a resource base for further analyzing. The system that we developed also has the capacity for continuous collection of tera-scale video data. Since the task of accessing and employing such large amounts of video data is challenging — particularly when having to cope with heavy-duty production cycles — we developed a high-end tera-scale data storage system, making use of a cluster of computers that served as a computational node and a gateway for accessing back-end supercomputers. The computers were further virtualized, forming a group of virtual machines (VM), to allow for the flexible use of computer resources (Smith and Nair, 2005). A VM is a software implementation of a machine that executes programs, in a similar way to a physical machine. A customized web-based interface is then developed as a

* Correspondence to: Y.-H. Hsiao, Department of Computer Science, National Tsing Hua University, No. 7, R&D 6th Rd., Hsinchu Science Park, Hsinchu City 30076, Taiwan. Tel.: +886 35776085x363.
** Correspondence to: F.P. Lin, Department of Computer Science, National Tsing Hua University, No. 7, R&D 6th Rd., Hsinchu Science Park, Hsinchu City 30076, Taiwan. Tel.: +886 35776085x360.
E-mail addresses: ihow363@gmail.com (Y.-H. Hsiao), fplin@nchc.narl.org.tw (F.-P. Lin).

1574-9541/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ecoinf.2013.10.002

Fig. 1. The video streams and location of our underwater observational site.

common portal, to access all computational resources. This system enables researchers to efficiently use these massive video datasets.

Considerable research has recently been focused on the analysis of video images obtained from video monitoring systems. Detecting and retrieving moving objects are necessary processes for applications based on monitoring video images. Numerous algorithms for land objec have been proposed (Javed and Shah, 2002). Background subtraction is a popular approach that can be used to efficiently capture the complete shapes of detected objects (Mahadevan and Vasconcelos, 2008). A real-time algorithm that combines temporal differencing with temporal filtered optical flow enables the detection of salient motion, moving in a consistent direction in complex environments (Tian and Hampapur, 2005a). Nevertheless, in uncontrolled conditions — for example, in real-life underwater monitoring systems — the system remains a challenge (Tian and Hampapur, 2005b). The real-world environment is usually unconstrained. For example, in our study, the coral reef ecosystem consists of a complicated reef environment as background and an abundance of bio-activities, such as fishes swimming and/or hiding, algae growing, sea anemones moving, and water plants drifting. This variety of movement complicates the recording of such activities. The accurate detection of fish and the effective discrimination of moving fish from other moving objects (as described above) therefore become a daunting challenge. Our underwater observation system, which has been developed and continuously operated for many years, produces tera-scale video data. This presents a problem, as the system has become too large to be processed and is, therefore, no longer suitable for scientific research. To improve the accuracy and performance during pre-processing, we have developed a simple system that makes use of multiple bounding–surrounding boxes which can integrate several existing state-of-the-art tracking algorithms, to efficiently remove irrelevant information (non-fish objects) from our tera-scale video data. The system has the capacity for a rapid detection of multiple-species fish images with variable viewing angles, shapes, and illumination. The acquired fish images are subsequently used to construct a fish-category database.

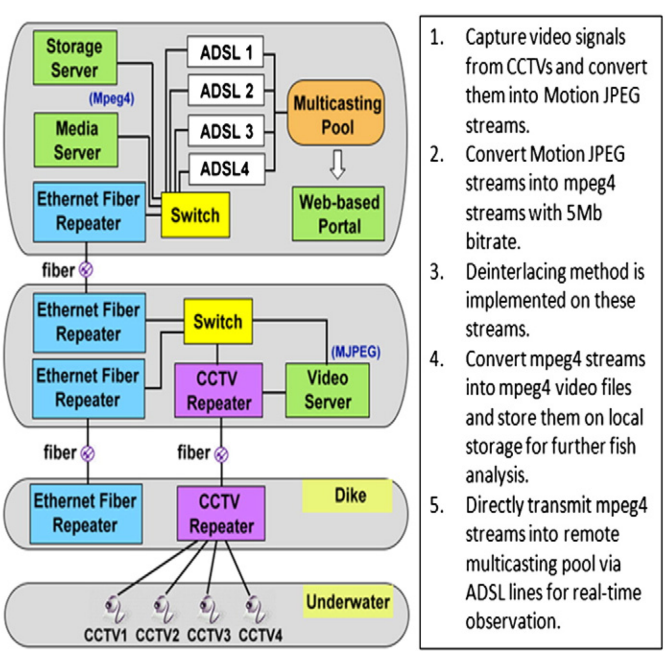

Fig. 2. The architecture of the underwater real-time video observational system.

The implementation of the fish detection method leads to the construction of a fish category database, based on the tera-scale underwater video dataset, which forms the basis of our real-world fish recognition and identification system. Object recognition is a popular topic in computer vision applications. Many publications have reported on various object recognition methods, in which the research focused primarily on projection-based methods, such as Principal Component Analysis (PCA) (Turk and Pentland, 1991) and Linear Discriminant Analysis (LDA) (Chen et al., 2008; Mika et al., 1999). These projection-based methods have been used and evaluated, for the purpose of fish recognition (Matai et al., 2012). These methods have the capacity to rapidly reconstruct object images as well as simultaneously extract image features. Nevertheless, there are some problems associated with the linear dimensionality reduction algorithm, in that the projections are a linear combination of all the original features, while all weighting coefficients in the linear combination, known as loadings, are typically non-zero. Fortunately, a proven compressive sensing theorem (Candes and Wakin, 2008), representing a novel sampling technique to find sparse solutions of underdetermined linear systems, can overcome this