

# Real-Time Fish Observation and Fish Category Database Construction

Yi-Haur Shiau

Data Computing Division  
National Center for High-Performance Computing  
Hsinchu, Taiwan

Fang-Pang Lin

Data Computing Division  
National Center for High-Performance Computing  
Hsinchu, Taiwan

Sun-In Lin

Data Computing Division  
National Center for High-Performance Computing  
Hsinchu, Taiwan

Chaur-Chin Chen

Department of Computer Science  
National Tsing Hua University  
Hsinchu, Taiwan

**Abstract**—This paper proposes a distributed real-time video stream system for underwater fish observation in the real world. The system, based on a three-tier architecture, includes capture devices unit, stream processor unit, and display devices unit. It supports variety of capture source devices, such as HDV, DV, WebCam, TV Card, Capture Card, and video compression formats, such as WMV, FLV/SWF, MJPEG, MPEG-2/4. The system has been demonstrated in Taiwan for long-term underwater fish observation. CCTV cameras and high-definition cameras are deployed on our system. Video compression methods and image processing methods are implemented to reduce network transfer flow and data storage space. Marine ecologists and end users can browse these real-time video streams via the Internet to understand the ecological changes immediately. These video data is preserved to form a resource base for marine ecologists. Based on the video data, fish detection is implemented. However, it is complicated in the unconstrained underwater environment, due to the water flow causes the water plants sway severely. In this paper, a bounding-surrounding boxes method is proposed to overcome the problem. It efficiently classifies moving fish as the foreground objects and the swaying water plants as the background objects. It enables to remove the irrelevant information (without fish) to reduce the massive amount of video data. Moreover, fish tracking is implemented to acquire multiple species of fish images with varied angles, sizes, shapes, and illumination to construct a fish category database.

**Keywords**—Real-time streaming; Fish observation; Fish detection; Distributed architecture.

## I. INTRODUCTION

Video stream over the Internet is a hot research topic recently. It can broadcast live events from a server, over the Internet, to end users. In recent years, owing to the advance of video stream technology and the booming of network bandwidth, live video stream is getting more and more popular. In this paper, we develop a distributed unmanned underwater video stream system for the long-term fish observation [1, 2, 3]. CCTV and high-definition cameras are set up as test cases that are installed on the Southern-most coast of Taiwan. Presently,

real-time video streams are accessible online via the Internet broadcasting. Worldwide marine ecologists and end users can now perform comparative studies between sites and attempt to understand the behavior of fish. The system facilitates marine ecologists to closely observe the ecosystem of fish, and understand the immediate phenomena of the underwater environment. It enables to enhance the public's awareness of the marine conservation. [4].

Although many applications for object detection and tracking have been proposed, application in uncontrolled conditions, i.e. in real-life underwater systems, remains a challenge [5]. Fish detection and tracking is complicated by the variability of the underwater environment. The water plants may be regarded as foreground objects as result of the severe sway from interference of the water flow, which is able to result in the complexities and difficulties to discriminate moving fish and swaying water plants. In this paper, we propose a bounding-surrounding boxes method, which effectively achieves the purpose that classifies moving fish as the foreground objects and swaying water plants as the background objects. Then, we implement the object tracking method for multiple species of fish from the stored video data to acquire fish images with varied angles, sizes, shapes, and illumination. Furthermore, we construct a fish category database by using image resizing method to let all of fish images with the same resolution.

This paper is organized as follows: Section 2 describes the distributed underwater observation system architecture details. Fish category database construction method is presented in Section 3. Section 4 shows the implemental results and the conclusion is drawn in Section 5.

## II. DISTRIBUTED REAL-TIME UNDERWATER VIDEO STREAM SYSTEM

### A. Distributed System Architecture

In this paper, a distributed real-time underwater video stream system is developed. The system is loose coupling and

three-tier architectures that includes capture devices unit, stream processor unit, and display devices unit. Figure 1 illustrates the distributed video stream system architecture and stream pipeline.

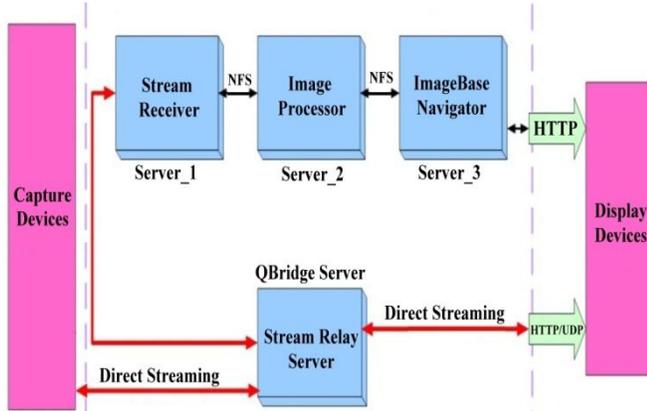


Figure 1. Architecture blocks and stream pipeline.

The left part of Figure 1 is capture devices unit. It receives signals from multiple capturing devices, such as HDV, DV, Webcam, TV Card, and automatically identifies formats of the signals. Identification of signal formats, video information and conversion is implemented by modified the functions of VideoLAN Client (VLC) [6] and FFMPEG.

The received signal can be converted to multiple video formats, such as MJPEG, MPEG-2/4, SWF/FLV, WMV, and multiple bit-rates for different bandwidths [7]. The benefit is it doesn't have to bind the specific hardware devices and the video formats. Table 1 shows the relationship between video formats and the corresponded video displayers.

TABLE I. THE RELATIONSHIP BETWEEN THE VIDEO FORMAT AND THE VIDEO PLAYER

Video format	Video player
WMV	Window media player
FLV/SWF	Wimpy player
MJPEG	Axis plug-in
MPEG-2/4	VLC media player

The center part of Figure 1 is stream processor unit. This unit is in charge of post-processing of the video stream and two modes are supported. One is direct streaming to display devices and there is a stream relay server that bridges the video stream between in unicast and multicast. The other is the video stream is stored for further implementing image processing methods, such as object detection and tracking.

The right part of Figure 1 is display devices unit. It supports multiple display devices handy to end users. Two protocols, HTTP and UDP, are adapted to transmit streaming to display devices. Figure 2 shows multiple display devices, include web-based interface, mobile display interface, Google Earth based interface, and a 4x3 Tiled Display Wall (TDW), a versatile, large, and high-resolution display system that was constructed

by National Center for High-Performance Computing (NCHC) [8].

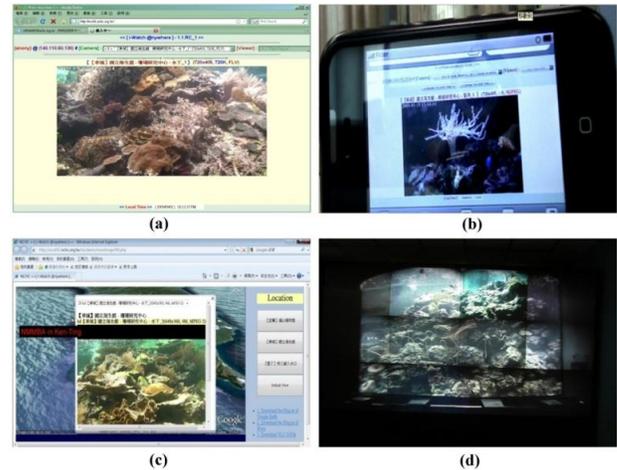


Figure 2. Multiple display interfaces (a) web-based, (b) mobile, (c) Google Earth based, and (d) TDW.

### B. Video Compression Methods

The quantities of acquired raw data of these video streams can approach 1-2 gigabyte per hour. The massive amount of video data is huge for network transfer and storage space. Two video compression concepts are implemented to reduce the network transfer flow and data capacity. After receiving signals, a compress method that converts video streams to different bit-rates is implemented for decreasing network traffic. Meanwhile, the raw video data is converted to a variety of compressive video formats such as MJPEG, MPEG-2/4, SWF/FLV, and WMV. Figure 3 shows the comparison with/without using video compression methods. The top image shows the native MPEG-2 bit-rate and the bottom image shows the compressed MPEG-2 bit-rate. The compressed ratio advances to about 25 times.



Figure 3. The bandwidth with/without video compression methods.

### III. FISH CATEGORY DATABASE CONSTRUCTION

For the stored video data, background subtraction [9, 10, 11, 12], foreground segmentation and object tracking methods are implemented for fish detection and tracking. In this paper,

Gaussian Mixture Matrix (GMM) method is adopted for background subtraction [13]. The highest color histogram similarity and the shortest distance are used for feature extraction to track the foreground objects. Figure 4(a) shows the background model and the current frame is illustrated in Figure 4(b). Figure 4(c) illustrates the foreground objects and Figure 4(d) shows the bounding boxes of these foreground objects.

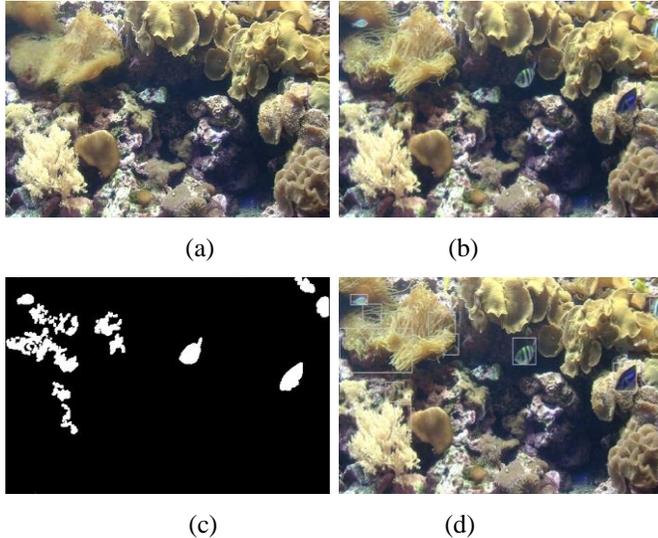


Figure 4. (a) The background model (b) the current frame (c) the foreground objects (d) the bounding boxes of foreground objects.

#### A. Bounding-Surrounding Boxes Method

The underwater environment in the real world is unconstrained, owing to the interference of the water plants sway severely. It raises the difficulty and complexity to discriminate moving fish and swaying water plants. However, the water plants always sway in a fixed field, but fish can free move to anywhere. Based on the concept, we propose a bounding-surrounding boxes method to discriminate fish as the foreground objects and water plants as the background objects. The foreground object is circumscribed by its bounding box with width  $w_1$  and height  $h_1$ . Let  $(c_x, c_y)$  be the center point of the bounding box and the upper-left point is  $(c_x - 0.5 * w_1, c_y - 0.5 * h_1)$ . Then, the surrounding box is set to T times the size of the bounding box with the same center point. Let  $B_t$  and  $S_t$  be the bounding box and surrounding box observed at time t. The location of  $S_t$  is fixed in the image, and the location of bounding box of the object is observed in a period of time  $\tau$ . If the location of the bounding box from time t to time  $t + \tau$  is always inside the range of  $S_t$ , the object is classified as a non-fish object (water plants). It is not only identified as a background object, but also eliminated from the tracked object. On the other hand, if the location of the bounding box has left the range of  $S_t$ , the object is classified as a foreground object (fish). The detecting results are shown in Figure 5. The yellow box represents the fixed surrounding box of the object. The red box in Figure 5(a) represents the object is classified as “fish”, and the blue box in Figure 5(b) represents the objects is classified as “non-fish” object.

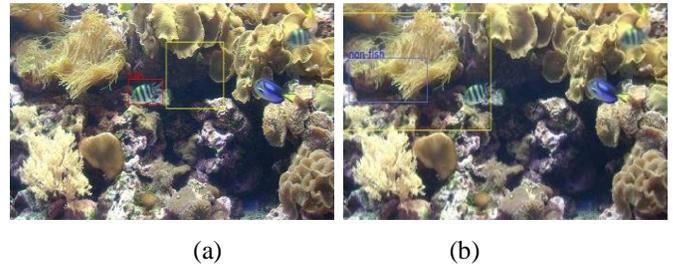


Figure 5. (a) The object (red box) is classified as fish (b) the object (blue box) is classified as non-fish (water plant).

#### B. Image Resizing Method

In order to for further implementing fish recognition and verification, the image size of each detected fish must be identical. In this case, the width and height of the resized image is the maximum width  $W_{Max}$  and height  $H_{Max}$  of all the images of multiple species of fish. In order to avoid the deformation of the images, we resized the fish images by pasting the original image to a  $W_{Max} \times H_{Max}$  black image, furthermore, the center of the black image was aligned. The result of an image resizing is shown in Figure 6.

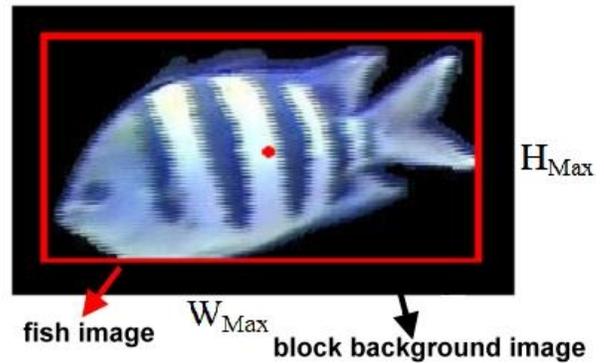


Figure 6. The result using image resizing method.

After fish detection is implemented using our proposed method, we can only record the video data that contains fish and remove the irrelevant information (without fish) to reduce the stored data volume. We also acquire multiple species of fish images with varied angles, sizes, shapes, and illumination. For each species of fish, we select some images that are almost different to construct a fish category database in the real world.

#### IV. IMPLEMENT RESULTS

Some cases are implemented to test the above-mentioned distributed real-time underwater video stream system. It is briefly introduced as followed.

##### A. Underwater Fish Observation With CCTV Cameras

Kenting, located on the southernmost tip of Taiwan, is famous for its abundant marine resources. Setting up cameras here can help marine ecologists observe fish behavior and the hydrological environment. Figure 7 illustrates the distributed architecture of the real-time underwater video stream system in Kenting.

It includes four underwater CCTV cameras with the resolution of 640x480 pixels and a sensor of water temperature and pressure. A video server that converts analog video signals into digital video streams is installed within a steel casing located on the dike. It delivers video streams to a media server via optical network with wireless network as a back-up. The stream data is transferred back to NCHC's multicasting pool, located 300km North in Hsinchu, through four ADSL lines. Figure 8 illustrates the real-time underwater video streams that the left side is the four CCTV cameras and the right side is the water temperature. These long-term continuous recordings can help marine ecologists in elucidating the ecological processes, and the real-time underwater observation system is able to enhance public's awareness of marine conservation.

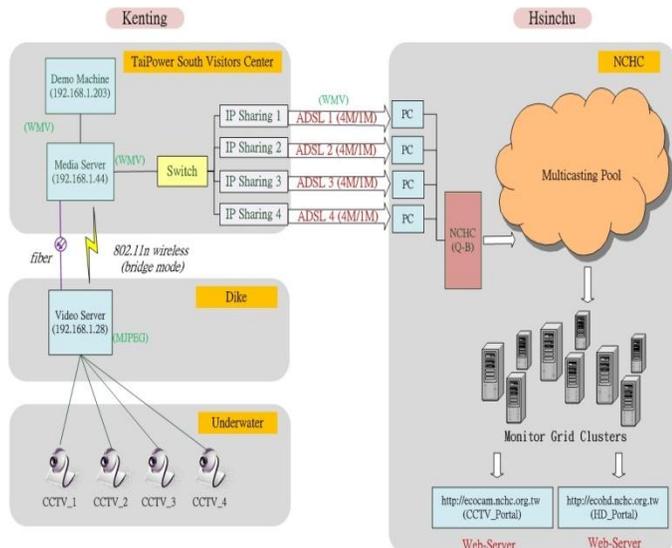


Figure 7. The distributed architecture of the real-time underwater video stream system in Kenting.

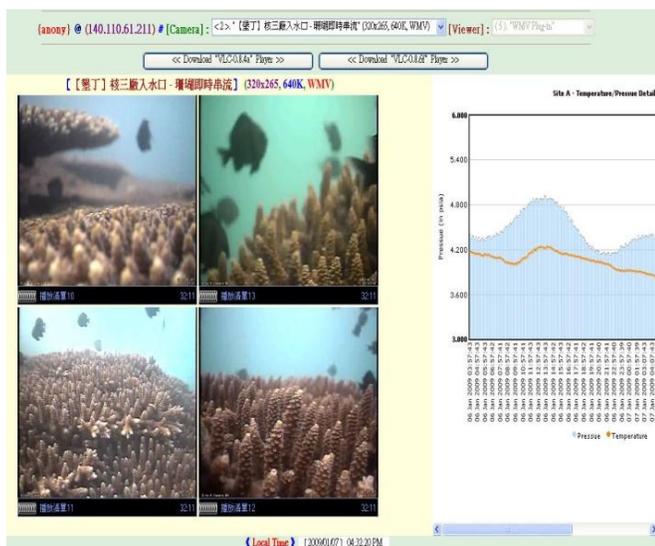


Figure 8. The real-time underwater video streams with CCTV cameras.

### B. Underwater Fish Observation with High-Definition Cameras

In order to support marine ecologists to get more detail data, the high-definition case is adopted. Two high-definition cameras with the resolution of 1280x1080 pixels are located on two different sites inside a fairly large lagoon in Taiwan. We set up a waterproof case to protect the high-definition camera, 1394 repeater and optical fiber. The challenge of this case is network bandwidth and to decode a high-definition video stream in real-time. The video compression method is implemented to compress the raw data from 20mb to 4mb and 1mb per second. Our experimental result shows that the proposed distributed real-time video stream system is successful for high-definition camera with only about 3~5 seconds latency. Figure 9 illustrates the distributed architecture of the high-definition real-time underwater video stream system. Figure 10 shows the high-definition real-time underwater video stream.

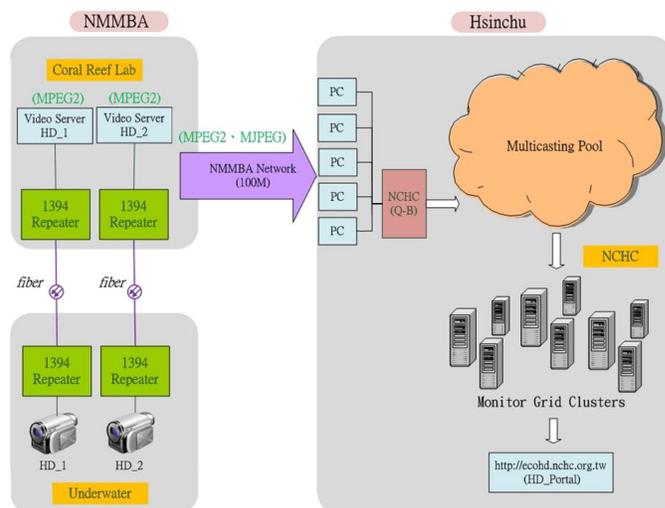


Figure 9. The distributed architecture of the high-definition real-time underwater video stream system.

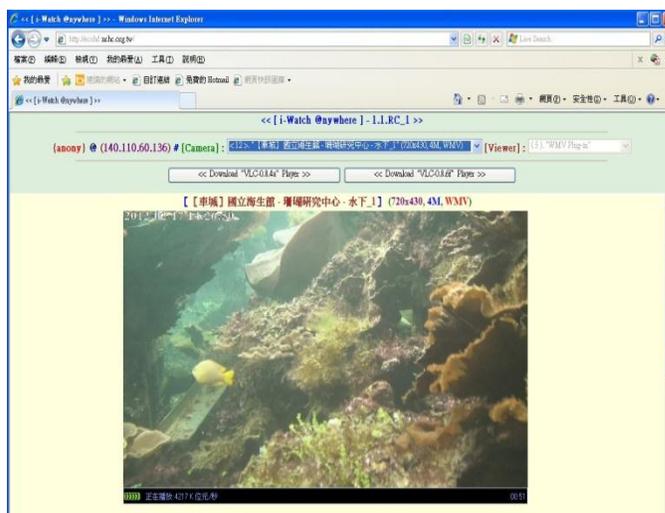


Figure 10. Figure 10. The real-time underwater video stream with high-definition cameras.

### C. Fish Category Database Construction

After implementing our proposed method, we enabled to obtain multiple species of fish images with varied angles, sizes, shapes, and illumination to construct a fish category database. The fish category database that we constructed is composed of 1,000 fish images of 180 rows and 130 columns with JPEG file format. Totally, there are 25 different species of fish. Each one contributed 40 images. The 5 images of 6 species of fish are illustrated in Figure 11. The total 40 images of 2<sup>th</sup> species of fish are illustrated in Figure 12. The fish category database can further for fish recognition and verification.

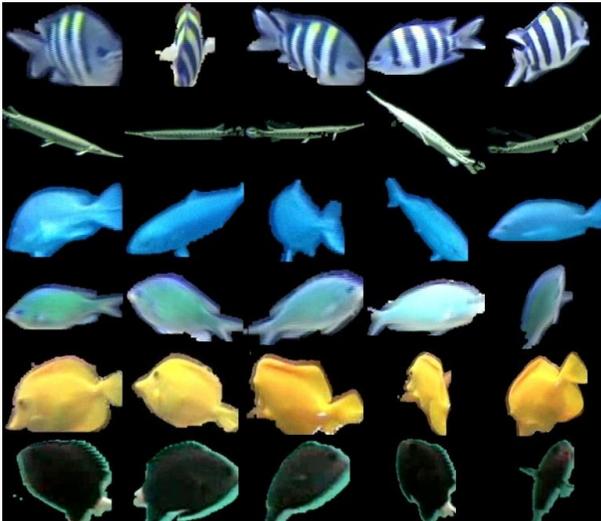


Figure 11. The 5 images of 6 species of fish.

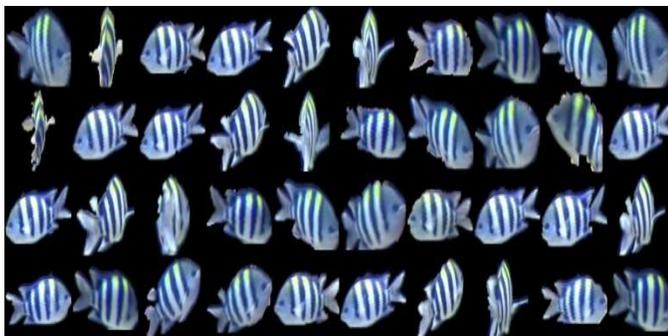


Figure 12. Examples of total 40 images of 2<sup>th</sup> species of fish.

### V. CONCLUSIONS

In this paper, a distributed architecture for real-time underwater video stream system was developed. The system had been demonstrated in Kenting and NMMBA, Taiwan, for long-term fish observation in the real world. Four CCTV cameras and two high-definition cameras were set up to test our developed video stream system. The distributed servers are located on Kenting and NMMBA, and video streams are transmitted back to NCHC's multicasting pool for real-time observation. Experimental results showed that the proposed distributed video stream system is robust, adaptive, and powerful.

In this paper, a bounding-surrounding boxes method had been proposed to reduce the stored video data capacity. It

efficiently discriminated moving fish as the foreground objects and swaying water plants as the background objects. Then, it enabled to remove the irrelevant information (without fish) and only save the data containing fish. It reduced the massive amount of the video data greatly. After that, we implemented fish tracking to acquire multiple species of fish images with varied angles, sizes, shapes, to construct a fish category database.

### ACKNOWLEDGMENT

This research was funded from Taiwan Power Company. We thank the Third Nuclear Power Plant of Taipower for logistical support. We also thank the National Museum of Marine Biology & Aquarium and Ecogrid team at National Center for High-Performance Computing for video stream data supply.

### REFERENCES

- [1] H. M. Chou, Y. H. Shiau, S. W. Lo, S. I. Lin, F. P. Lin, C. C. Kuo, and C. L. Lai, "A Real-Time Ecological Observation Video Streaming System based on Grid Architecture," In HPC Asia, Taiwan, 2009.
- [2] Y. H. Shiau, J. S. Cheng, S. I. Lin, Y. H. Chen, K. T. Tseng, H. M. Chou, and S. W. Lo, "A Distributed Architecture for Real-Time High-Resolution Video Streaming," International Conference on Parallel & Distributed Processing Techniques & Applications, pp. 345-349, 2009.
- [3] Y. H. Shiau, Y. H. Chen, K. T. Tseng, J. S. Cheng, S. I. Lin, S. W. Lo, and H. M. Chou, "A Real-Time High-Resolution Underwater Ecological Observation Streaming System," International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, Japan, pp. 517-521, 2010.
- [4] E. Strandell, S. Tilak, H. M. Chou, Y. T. Wang, F. P. Lin, P. Arzberger, T. Fountain, T. Y. Fan, R. Q. Jan, and K. T. Shao, "Data management at Kenting's underwater ecological observatory," The third International Conference on Intelligent Sensors, Sensor Networks and Information Processing, Melbourne, Australia, pp. 715-720, 2007.
- [5] D. R. Edgington, I. Kerkez, D. E. Cline, J. Mariette, M. Ranzato, and P. Perona, "Detecting, tracking and classifying animals in underwater video," IEEE International Conference on Computer Vision and Pattern Recognition, pp. 634-638, 2007.
- [6] H. Nguyen, P. Duhamel, J. Brouet, and D. Rouffet, "Robust vlc sequence decoding exploiting additional video stream properties with reduced complexity," In IEEE International Conference on Multimedia and Expo, Taiwan, pp. 375-378, 2004.
- [7] C. Traiperm, and S. Kittitumkun, "High-performance MPEG-4 multipoint conference unit," In Networks and Communication Systems, Thailand, 2005.
- [8] S. I. Lin, F. P. Lin, C. Chang, S. W. Lo, Y. C. Mai, P. W. Chen, and Y. Shiau, "Development of grid-based tiled display wall for networked visualization," 9<sup>th</sup> International Workshop on Cellular Neural Networks and Their Applications, pp. 315-318, 2005.
- [9] D. S. Lee, J. J. Hull, and B. Erol, "A Bayesian framework for gaussian mixture background modeling," IEEE International Conference on Image Processing, vol.3, pp. 973-976, 2003.
- [10] M. Piccardi, "Background subtraction techniques: a review," IEEE International Conference on System, Man, and Cybernetics, pp. 3099-3104, 2004.
- [11] V. Mahadevan, and N. Vasconcelos, "Background Subtraction in Highly Dynamic Scenes," IEEE Computer Society Conf. Computer Vision and Pattern Recognition, pp. 1-6, 2008.
- [12] C. Stauffer, and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," IEEE Computer Society Conf. Computer Vision and Pattern Recognition, vol. 2, pp. 246-252, 1999.
- [13] S. Y. Yang and C. T. Hsu, "Background Modeling from GMM Likelihood Combined with Spatial and Color Coherency," IEEE International Conference on Image Processing, pp. 2801-2804, 2006.