ISSN 2229-5518

Fish Counting from Underwater Video Sequences by Using Color and Texture

Suman Sharma, Aman Shakya, Sanjeeb Prasad Panday

Abstract— Fish population estimation and classification of fish species have been an integral part of marine science. These tasks are important for the assessment of fish abundance, distribution and diversity in marine environments. Underwater video measurement systems are used widely for counting and measuring fish in aquaculture, fisheries and conservation management. This paper is presented the techniques used for the detection, identification, and counting of fish in underwater video image sequences, including consideration of the changing body shape, color and texture of fish. It presented simple method of counting fish as blob counting, which automatically using image processing techniques. The detection algorithm, canny edge detection algorithm is used. Coral-blackening process is used to distinct fish and the background. A video sequence was taken from Sesnarayan Pond and its every frame is processed singly and independently. Finally each Zernike moment is calculated for each blob with the template of three different types of fish, the blobs is counted the number of different template of fish for the blobs whose amplitude moment is near about zero and the average number of type fish over the frame is recorded. Finally template is matched with the blob by using color and texture of three different types of template.

Index Terms— *blob counting, template matching, underwater video, and Zernike moment.*

1 Introduction

THE study of underwater species is a fascinating topic to marine biologists and environmental experts. Video survey is a popular approach to study marine life. To determine the size and distribution, or to study the behavior of species, the researchers need to locate them in images or videos, but this can be very time consuming when processing a large volume of data. Automated image and video segmentation helps to speedup this tedious work [1]. The purpose of this paper is to identify type of fish in the video or images from given sample, by analyzing and exploring from various visuals information.

The monitoring of fish for stock assessment in aquaculture, commercial fisheries and in the assessment of the effectiveness of biodiversity management strategies such as Marine Protected Areas and closed area management is essential for the economic and environmental management of fish populations. Video based techniques for fishery independent and non-destructive sampling are widely accepted. The advantages of using video for counting the numbers of fish, measuring their lengths and defining the sample area have been well demonstrated. In order to study the effects that climate change and pollution has on the environment, long-term monitoring of the environment is necessary. One of the most important natural environments on earth is coral reefs, however monitoring the fish population and biodiversity is still a challenging task. Data collection in this kind of environment is labor intensive, requiring divers to count the fish species in a certain area. In recent years, video recording has become much cheaper which makes underwater cameras a good alternative for data collection. Furthermore, automatic video processing and pattern recognition is able to process this kind of data [2], [3], [4].

However, the time lag and cost of processing video imagery decreases the cost effectiveness and uptake of this technology. Current research aims to minimize or completely eliminate the involvement of the human operator in the process of recognition and length measurement of fish recorded by underwater video surveys. Advances in automated techniques will substantially decrease the cost of processing and make the technology more accessible to a broad spectrum of end users [4], [5].

2 General Procedure

2.1 Histogram Analysis

Intensity transformation functions based on information extracted from image intensity histograms play a central role in image processing, in areas such as enhancement, compression, segmentation, and description. The focus of this section is on obtaining, plotting, and using histograms for image enhancement. The histogram of a digital image with L total possible intensity levels in the range [0, G] is defined as the discrete function.

$$h(r_k) = n_k \quad \text{……………………………………..……………. (1)}$$

Where, r_k is the k^{th} intensity level in interval [0, G] and n is the number of pixels in the image whose intensity level is r_k. The value of G is 255 for given images. Sometimes it is necessary to work with normalized histograms, obtained simply by dividing all elements of n_k by the total number of pixels in the image, which we denote by n.

$$p(r_k) = \frac{n_k}{n} \quad \text{……………………………………….…………………. (2)}$$

Where, for integer images, k = 0, 1, 2…L-1. From basic probability, we recognize $p(r_k)$ as an estimate of the probability of occurrence of intensity level r [6],[7].

Fig 1 is original frame capture from video taken at **Sesna-**

ISSN 2229-5518

rayan pond and corresponding its RGB color histogram. A color histogram is a representation of the distribution of colors in an image. The color histogram can be built for any kind of color space, although the term is more often used for three dimensional spaces like RGB or HSV. Color histograms are flexible constructs that can be built from images in various color spaces, whether RGB, chromaticity or any other color space of any dimension. A histogram of an image is produced first by discretization of the colors in the image into a number of bins, and counting the number of image pixels in each bin [8]. The histogram analysis shows that the video taken also depends on intensity of light as there is a different intensity level at different time so it can go for further processing.

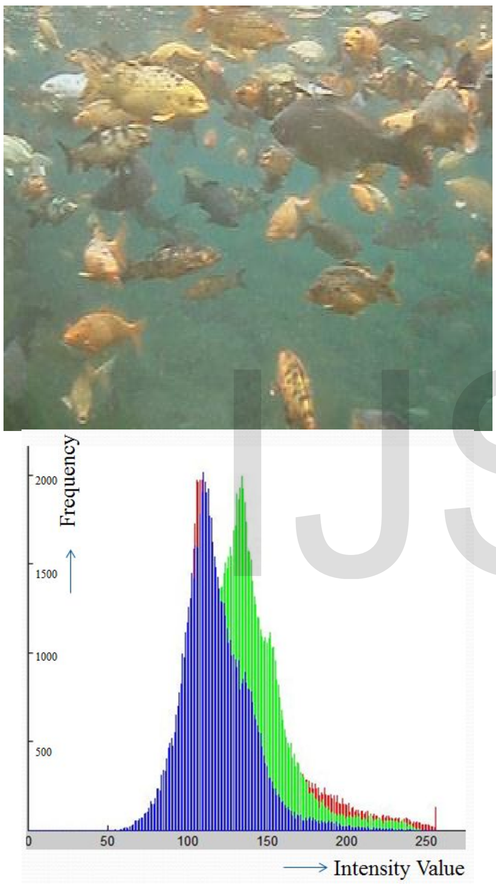


Fig 1 Original frame for 1st video sequence and its Histogram

2.2 Canny Edge Detection and Coral Blackening

The purpose of edge detection in general is to significantly reduce the amount of data in an image, while preserving the structural properties to be used for further image processing. Several algorithms exists, and this worksheet focuses on a particular one developed by John F. Canny (JFC) in 1986. Even though it is quite old, it has become one of the standard edge detection methods and it is still used in research. The aim of JFC was to develop an algorithm that is optimal with regards to the following criteria. **Detection:** The probability of detecting real edge points should be maximized while the probability of falsely detecting non-edge points should be minimized. This corresponds to maximizing the signal to noise ratio. **Localization:** The detected edges should be as close as possible to the real edges. **Number of responses:** One real edge should not result in more than one detected edge (one can argue that this is implicitly included in the first requirement) [1], [9].

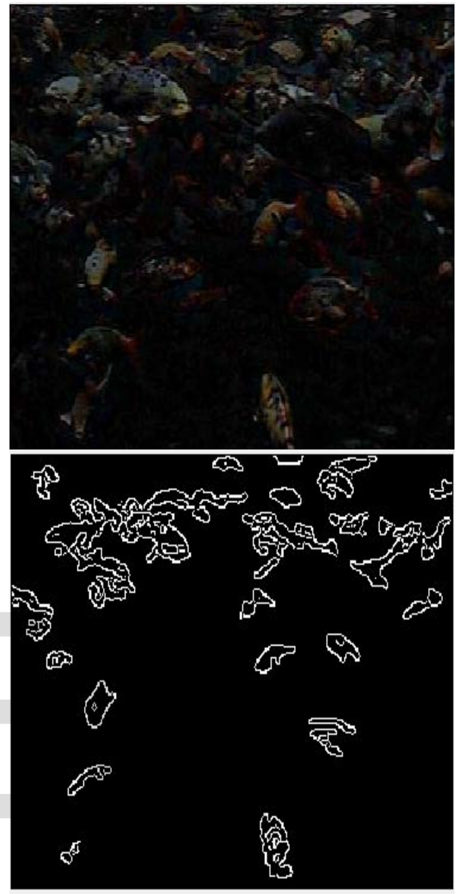


Fig 2 Coral Blackening and Canny Edge Detection

In order to determine blackened parts, the entire image is divided into different pixel blocks and histograms are computed for each block. Since the background against which the fish are clearly visible is predominantly water color. Coral blackening and Canny edge detection is used for corresponding frame shown in fig 2. Due to unclear water and unnecessary moving particles, it is difficult to implement background subtraction method by taking maximum number of frames, so coral blackening is used. As canny high pass filter is used for edge detection in any orientation and even in more noisy condition, it uses probability for finding error rate, localization and response..

2.3 Zernike Moments

Zernike moments are the mappings of an image onto a set of complex Zernike polynomials. Since Zernike polynomials are orthogonal to each other, Zernike moments can represent the properties of an image with no redundancy or overlap of information between the moments. Zernike moments are significantly dependent on the scaling and translation of the object in an ROI. Nevertheless, their magnitudes are independent of rotation angle of the object. Hence, we can utilize them to describe shape characteristics of the objects. For instance, we took the advantage of Zernike moments to extract the shape information of benign and malignant breast masses [1].

The set of orthogonal Zernike moments are known to be superior compared to other image moments due to their nice rotational, translational and scale invariant properties. Here choosing Zernike moments for this system because of these important properties match the requirements for fish species identification.

$$Z_{nm} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 f(\rho, \theta) V_{nm}^{*} \, \rho \, d\rho \, d\theta \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Where ρ = (x2 + y2)1/2 is the length of the vector from the origin to the pixel (x, y) and θ = arctan(y/x) is the angle that the vector makes with the axis. The order n and repetition m are integers that satisfy

$$n \geq 0,\ n - |m| = (even) \text{ and } |m| \leq n. \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

The complex-valued 2-D Zernike basis functions (which are defined within a unit circle) are formed by function.

$$V_{nm}(\rho, \theta) = R_{nm} \exp(jm\theta), \quad |\rho| \leq 1 \ldots\ldots\ldots\ldots\ldots\ldots\ldots (5)$$

Where j = $\sqrt{-1}$ and the real valued Zernike 1-D radial polynomial is given by:

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} C(n, m, s) \rho^{n-2s} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

Where, $C(n, m, s) = (-1)^s \frac{(n-s)!}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n-|m|}{2}-s\right)!}$

The radial polynomials satisfy the orthogonal properties for the same repetition

$$\int_0^{2\pi} \int_0^1 R_{nm}(\rho, \theta) R_{n',m}(\rho, \theta) \rho d\rho d\theta = \begin{cases} \frac{1}{2(n+1)} & if\ n = n' \\ 0 & otherwise \end{cases} \ldots (7)$$

The Zernike basis functions are orthogonal which implies that there is no redundancy of information among the Zernike moments with different orders and repetitions. Thus, each moment is unique and independent of each other [6].

Algorithm of Proposed System

Step 1: Raw input video
Step 2: Convert video to frame.
Step 3: Apply the identified thresholds and create a binary image, remove any regions less than threshold, crop the original image to the rectangle containing the remaining candidate regions.
Step 4: Apply coral blackening for a selected image sample.
Step 5: Apply canny edge detection for the detection of fish.
Step 6: Image areas classified as non-background fish candidates are then subject to morphological filters, erosion and closing, dilation and opening, and a median filter;
Step 7: The count of fish in the frame is then determined by a blob counting. Apply dilation to remove noise, expand and merge adjacent regions, then apply erosion to restore the external boundaries of the regions;
Step 8: Count the total blob on a selected frame.
Step 9: Calculate Zernike moment form a template fish of type one with the selected frame.
Step 10: If the Zernike moment is found as less than one then the count the blob is type one fish and go back to step 9.
Step 11: Display the total fish count on a selected frame.

2.4 Implementation and Experiment

Sample water and non-water color histogram templates are obtained from the selected frame of the video sequences. It is then used to generate the respective mean values for the two types of image templates. When all of the benthic portions of the image have been blackened out, the Canny Detector is used to detect fish contours. In order for fish counting to be accomplished, blobs have to be formed from the fish outlines. Since blobs are computed per frame, difficulties arise when two fish outlines overlap thereby resulting in a single blob count. This brings about inaccurate counts, so that some procedural adjustments are undertaken. Fish templates from known species of fish captured in the video are obtained. The Zernike moments for the template is computed and stored as previously discussed. The image blobs are colored according to its identified species.

The video was taken from the Sesnarayan pond located at Farping. The selected underwater video was the least viewpoint changes. This video also contains more fish thereby making population estimation and identification more feasible. Any fish that is against the benthic background will be blackened by the algorithms thus excluding it from the fish count. Average of data is counted from a ten conjugative frames. Average data is counted for a three different sequence of video. Each fish was manually counted. This was then compared with the program's count. The counts are shown in fig 3, fig 4, and fig 5 in which machine versus human count for each frame. In fig 3, samples are taken in time interval of 2 minutes. As the human count is the exact data, machine count does not varies with a large gap. The machine count is 5 to 7 counts less than that of human count. Observing in fig 5, with time interval of 5 minutes, the difference between human and machine count is even lesser than that of data shown in fig 3 and 4. But in fig 5, the outcome is bit different than initial statistics. Fig 6 indicates average count of fish in different prospective taken reference as in fig 3, fig 4, and fig 5. Here machine count and human count is different slightly for overall cases. In each time human count is greater than machine count. Here the accuracy for machine count versus human count is ranging from 70 to 84 percentages. The count of fish in not matched because of the reflection, movement of fish, and it may be occlusions. For counting different types of fish in a single frame, it is necessary to classify the types of fish. This paper work is based on the technology to classify the type of fish through matlab software. Such as fish gray, binarization, image enhancement, contour extraction, and characteristic parameters for dimension reduction. Three different types were analyzed to determine the techniques which could be used effectively for automatic species identification. Feature vectors were extracted from each of the template and use to classify species using the feature vectors with each blob. Whereas a previous analysis of this image dataset required to make able to achieve a greater level of accuracy without manual specification of a region of interest at all. The features that extracted from the images are various experiments that using different classifiers and datasets that were performed.

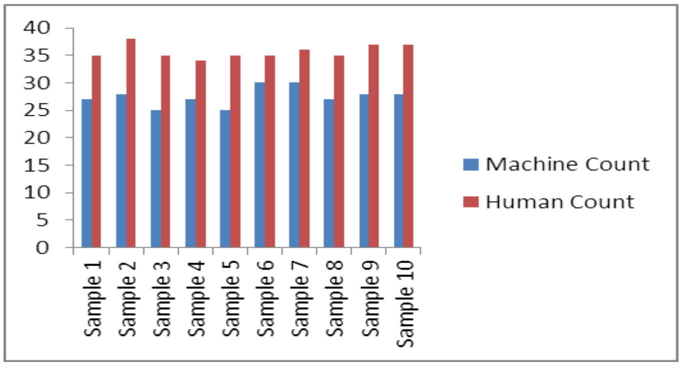


Fig 3 Machine versus Human count for 1st sequence at 00:02:00 (Hr:Min:Sec)

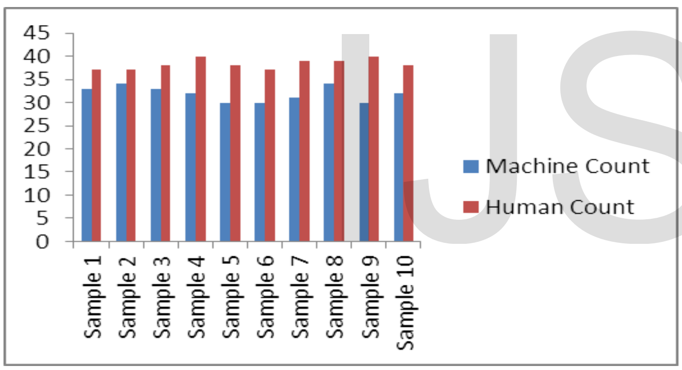


Fig 4 Machine versus Human count for 2nd sequence at 00:05:00(Hr:Min: Sec)

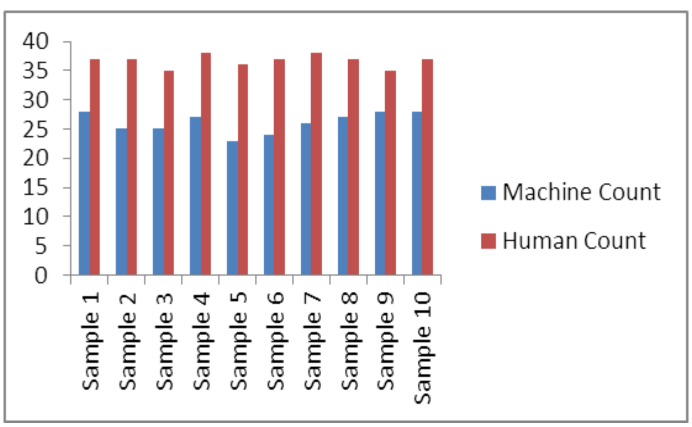


Fig 5 Machine versus Human count for 3rd sequence at 00:13:00(Hr:Min:Sec)

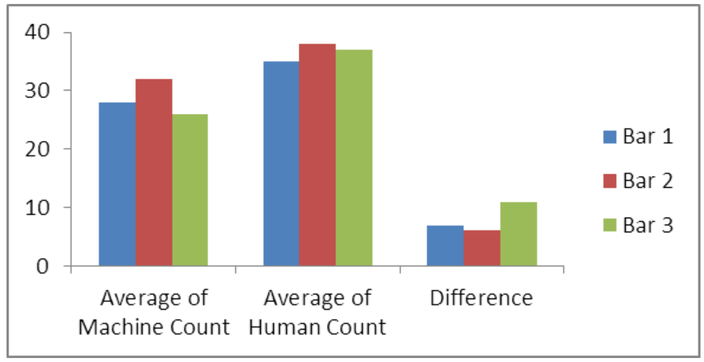
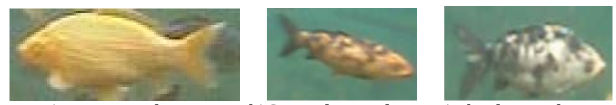


Fig 6 Bar diagram showing average data of different video sequences

Species Identification



a)First Template b)Second Template c)Third Template

Fig 7 Three different types of template

2.5 Template matching by using Zernike Moment

Template matching based on Zernike moments is a very popular scheme because moment descriptors are invariant to translation, rotation and scaling. On one hand, they simply define the region properties that can easily be used for feature classification and object recognition. On the other hand, moment descriptors do require huge computation cost because of many integration and multiplication operations. Zernike moments acquire a phase shift on rotation and their magnitudes remain constant. Thus the magnitudes of an original image and its rotated image are identical. Thus they are invariant to rotation [10], [11]. Fig 7 shows three different templates, to verify the matching performance several examples were tested. The templates and the test images are rotated and vary as in real time. Fig 8 shows the result of template matching with first testing image. The blobs shown in figure are approximately matched template. The total count of the blobs is 16. Similarly fig 9 and fig 10 depicts the result of template matching with second and third testing image respectively. The blobs shown are approximately matched template with corresponding template. The total count of the blobs is 7 and 0 when matching with second and third template respectively.

2.6 Template matching by using Color and Texture

Since the image color information is dominated by the water's hue, color is not much useful to differentiate these fish types. Shape also provides little discernment, so in our work we focused on color and texture-based classification. The problem of classifying objects can be solved by using color, texture and shape. We used texture and color features to find a small number of images in the database, and identify regions in the candidate images which share similar texture and color. To speed up the processing, the texture and color features are directly extracted from the Discrete Cosine Transform (DCT) compressed domain. The color and texture features are direct-

ISSN 2229-5518

ly extracted from the compressed images and sample of the result can be seen in different figures. In fig 11, the result of template matching with first testing image is shown by using color and texture. The selected 10 blobs are approximately matched with given template observing fig 8 and fig 11, the better result is obtained by using color and texture in which count is 6 less than that of Zernike moment. The result of template matching with second testing image is shown in fig 12; selected 8 blobs are approximately matched with given template. Observing fig 9 and fig 12, the better result is obtained by using color and texture in terms of selected type of fish although the number of blob count is similar.

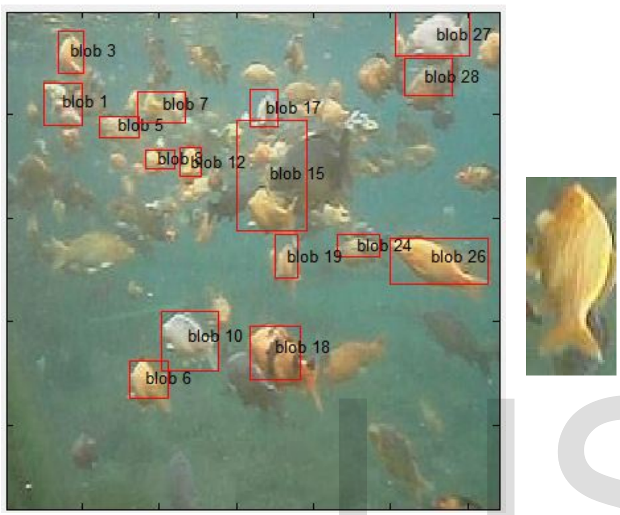


Fig 8 Template matching with 1st testing image

Fig 9 Template matching with 2nd testing image

Fig 10 Template matching with 3rd testing image

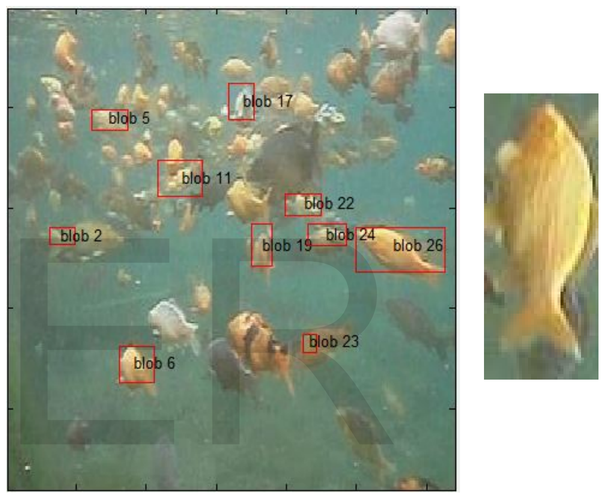


Fig 11 Template matching with 1st testing image by using color and texture

Fig 12 Template matching with 2nd testing image by using color and texture

ISSN 2229-5518

Fig 13 Template matching with 3rd testing image by using color and texture

Similarly the result of template matching with third testing image is shown in fig 13. Here 2 selected blobs are approximately matched with corresponding template. Comparing with fig 10, the result is more accurate by using color and texture in terms of selected type of fish as well as the number of blob count.

Conclusion

The goal of this paper is to explore methods to segment fish components, investigate the possible methods of counting fish components and develop classifiers that can discriminate different component categories. Automated detection of underwater species from a large volume of underwater video is useful in marine studies. However, the task can be difficult because of the complexity and variability within the benthic environments. Some key factors include the water depth and visibility, lighting condition, time of day, and certainly the image resolution.

References

[1] Fabic, J. N., et al. "Fish population estimation and species classification from underwater video sequences using Blob Counting and Shape Analysis" Underwater Technology Symposium (UT), 2013 IEEE International. IEEE, 2013.

[2] Spampinato, Concetto, et al. "Automatic fish classification for underwater species behavior understanding." Proceedings of the first ACM international workshop on Analysis and retrieval of tracked events and motion in imagery streams. ACM, 2010.

[3] Shortis, Mark R., et al. "A review of techniques for the identification and measurement of fish in underwater stereo-video image sequences." SPIE Optical Metrology 2013. International Society for Optics and Photonics, 2013.

[4] Boom, Bastiaan J., et al. "Long-term underwater camera surveillance for monitoring and analysis of fish populations." VAIB12 (2012).

[5] D.R. Edgington, D.E. Cline, D. Davis, I. Kerkez, and J. Mariette, Detecting, Tracking and Classifying Animals in Underwater Video OCEANS 2006, Moss Landing, CA, Sept. 2006.

[6] Toh, Y. H., T. M. Ng, and B. K. Liew. "Automated fish counting using image processing." Computational Intelligence and Software Engineering,2009. CiSE 2009. International Conference on. IEEE, 2009.

[7] Yu, Ning, et al. "Underwater range-gated laser imaging system design with video enhancement processing." Instrumentation and Measurement, Sensor Network and Automation (IMSNA), 2013 2nd International Symposium on. IEEE, 2013.

[8] Mahmood, Arif, and Sohaib Khan. "Correlation-coefficient-based fast template matching through partial elimination." Image Processing, IEEE Transactions on21.4 (2012): 2099-2108.

[9] Cootes, Timothy F., et al. "Active shape models-their training and application."Computer vision and image understanding 61.1 (1995): 38-59.

[10] Schuster, Stefan, and Silke Amtsfeld. "Template-matching describes visual pattern-recognition tasks in the weakly electric fish Gnathonemus petersii."Journal of experimental biology 205.4 (2002): 549-557.

[11] Sclaroff, Stan, and Lifeng Liu. "Deformable shape detection and description via model-based region grouping." Pattern Analysis and Machine Intelligence, IEEE Transactions on 23.5 (2001): 475-489.

[12] Clement, Pierre M., et al. Fishery applications of optical technologies. Eds. James Churnside, Michael Jech, and Eirik Tenningen. International Council for the Exploration of the Sea, 2012.

[13] Boom, Bastiaan J., et al. "A research tool for long-term and continuous analysis of fish assemblage in coral-reefs using underwater camera footage." Ecological Informatics (2013).

[14] Mokhtarian, Farzin, and Riku Suomela. "Robust image corner detection through curvature scale space." Pattern Analysis and Machine Intelligence, IEEE Transactions on 20.12 (1998): 1376-1381.

[15] Lee, D. J., et al. "An automated fish species classification and migration monitoring system." Industrial Electronics Society, 2003. IECON'03. The 29th Annual Conference of the IEEE. Vol. 2. IEEE, 2003.

Suman Sharma received the Bachelor's degree in Electronics and Communication Engineering from Tribhuvan University, Nepal in 2009, and Master's degree in Information and Communication Engineering from Tribhuvan University, Nepal in 2014. Currently he is working as a Senior lecturer at Kathmandu Engineering College, Department of Electronics and Communication Engineering. He is currently the year coordinator for Electronics and Communication Engineering. His research interest includes the wireless communication, image processing, adaptive signal processing, and neural network application on wireless communication.

Dr. Aman Shakya is an Assistant Professor at the Department of Electronics and Computer Engineering, Institute of Engineering, Tribhuvan University, Pulchowk campus, Nepal. He received his Doctor of Philosophy from the National Institute of Informatics, Department of Informatics of The Graduate University for Advanced Studies (SOKENDAI), Tokyo, Japan in 2009. He received his Masters of Engineering in Information and Communication Technologies from the Asian Institute of Technology, Pathumthani, Thailand in 2006. He did Bachelors in

ISSN 2229-5518

Computer Engineering from the Institute of Engineering, Tribhuvan University, Pulchowk campus, Lalitpur, Nepal in 2003. His research interests include the Semantic Web, social semantics, Web 2.0 and structured information sharing. He has been teaching courses on Semantic Web, XML, web technologies, enterprise application development and theory of computation. He has also worked as an IT consultant in several projects.

Dr. Sanjeeb Prasad Panday was born in Kathmandu, Nepal. He received the Bachelor's Degree in Electrical Engineering from the University of Engineering and Technology, Lahore, Pakistan in 2001 and Master's Degree in Information and Communication Engineering from Tribhuvan University, Nepal in 2006. He completed his Doctoral Degree in Information Systems Engineering at the graduate school of engineering, Osaka Sangyo University, Japan. He has been working as an assistant professor in the Department of Electronics and Computer Engineering at Institute of Engineering, Central Campus Pulchowk, Tribhuvan University, Pulchowk, Lalitpur, Nepal since 2002 A.D. He is currently the program coordinator for M.Sc. in Computer Systems and Knowledge Engineering. He is also the Deputy Director of Information and Communication Technology Centre (ICTC). His research interests include Disaster Communications, RF and Microwave Techniques, Image Processing, Digital Holography and Optimization Algorithms. He has been working as an academic advisor to Multinational Corporation Interoperability (MCIP) since their capstone event Pacific Endeavor 2012. He has been teaching Electromagnetics in Department of Electronics and Computer Engineering, Central Campus Pulchowk, Institute of Engineering, Tribhuvan University since 2002 and has nearly 13 years of experience in teaching this subject.