Research on Recognition Method of UUV Vision Based on PSO-BP Network

Wei Zhang, Detao Meng, Zhicheng Liang, Yi Guo, Jiajia Zhou
College of Automation
Harbin Engineering University
Harbin, Heilongjiang Province, China
Email: toere123@126.com

Yunfeng Han
College of underwater acoustic engineering
Harbin Engineering University
Harbin, Heilongjiang Province, China

***Abstract*—During unmanned underwater vehicle (UUV) recovery process, the vision sensor needs to extract and recognize all light sources from guided images. Firstly, Aiming at restoring underwater image without exact depth map, a segmented-linear-mapping approach based on scattering model is put forward. Restoring results indicate that the approach is capable of highlighting subtle details and improving image quality effectively. Then the light source edge and its invariant moments are extracted by Canny edge detection algorithm combined with improved Snake model. At last, a kind of Back Propagation network based on particle swarm optimization algorithm (PSO-BP) for UUV recognition is designed, which demonstrates higher recognition rate than traditional BP network.**

***Keywords*—unmanned underwater vehicle (UUV); scattering model; restoring underwater image; PSO-BP; recognition**

I. INTRODUCTION

With the development of marine science and technology, UUVs are widely used in different marine tasks and have been important tools for exploring and exploiting the underwater resources [1-2]. Various ways are put forward for UUV close range navigation in UUV recovery process, such as sound navigation. Meanwhile, vision sensor has much higher positioning accuracy in short distance and can supply sufficient visual information of environment [3]. It can be helpful in ocean inspection, measurement, and task execution, which is applied for UUV close range navigation. UUV needs to identify the specified targets from the image backgrounds when it carries out docking task. The accuracy and speed of target recognition have close relationships with the ways of underwater image restoration, image edge extraction and target recognition. This study is about the whole process from restoring underwater image to recognizing the guided light source.

Currently the model-based restoration concerns mainly about atmosphere images and underwater image restoration is still underdeveloped. In literature [4], polarized imaging technology with underwater white light lamp was studied. It concluded that both linear and circular polarizers could improve image quality to an extent related to its rotatory direction. Schechner and Karpel searched a polarized-imaging based restoration algorithm [5], which effectively extends camera's underwater visible range. Queiroz-Neto introduced the degradation model in haze into underwater stereo vision and as a result, improved the accuracy of scene reconstruction [6]. However, an applicable restoring approach is still in expectation. At the same time, the accuracy of underwater target recognition will affect the UUV's operational capacity, scholars have done a lot of theoretical and applied researches for target recognition. K. T Kim extracted rotation invariant and scale-invariant feature vectors and used Bayesian classification to recognize five different targets [7]. K. Iqbal calculated the similarity distance of reference image and test image by moment features to achieve the image classification [8].

The paper is organized as follows. In section II, Scattering model based underwater image restoration and a segmented-linear-mapping approach are described. Section III develops GVF-Snake model to extract edge of the heart-shaped light source. Underwater target recognition based on PSO-BP algorithm is presented in section IV. Finally, Recognition results of the heart-shaped light source are given in section V.

II. UNDERWATER IMAGE RESTORATION

The image restoration focuses on the physical course of light propagation in medium, builds light propagation model (degradation model) for it and removes the image noises through model inversion. It is especially effective in bad weather or underwater condition, although it is more complex.

As light propagates through water from scene point to CCD element, its key characteristics are changed mostly due to scattering and absorption by water particles. It assumes that the intensity of each pixel has two components: attenuated scene intensity and water light radiance [6]. The intensity received by camera sensors can be expressed as.

$$\begin{aligned} E(d,\lambda) &= E_s(d,\lambda) + E_n(d,\lambda) \\ &= E_0(\lambda)e^{-\alpha(\lambda)d} + E_\infty(\lambda)(1-e^{-\beta(\lambda)d}) \end{aligned} \quad (1)$$

which is called underwater light scattering model, where d is the propagation distance and λ is wavelength. $E_s(d,\lambda)$ and $E_n(d,\lambda)$ are respectively the attenuated scene intensity and the water light radiance. $E_s(d,\lambda)$ has an exponential relation

978-0-933957-43-5 ©2015 MTS

with $E_0(\lambda)$, the radiance that exits the scene's surface, $\alpha(\lambda)$ is called optical density measuring the strength of absorption and scattering in the medium. $E_\infty(\lambda)$ is the water light radiance in infinite distance and $\beta(\lambda)$ is a property of the medium called water light coefficient. For monocular camera in several meters' range, the influence of wavelength λ can be neglected. Furthermore, α and β are both coefficients reflecting the property of the medium, hence, Eq. 1 can be simplified as

$$E = E_0 e^{-\beta d} + E_\infty(1 - e^{-\beta d}) \tag{2}$$

where βd is called scaled depth (or distance).

A. Calibration of model parameters

Before the scattering model (Eq. 2) is used for image restoration, the coefficients β and E_∞ must be calibrated. β can be considered as a constant in the same water body. E_∞ can also be considered as a constant under indoor illuminations. A plane calibration board with white background and black-dots foreground is used, the board's surface is always parallel to the camera lens at each position, which justifies a sole value of d for one image frame. The experiment was executed step by step as follows:

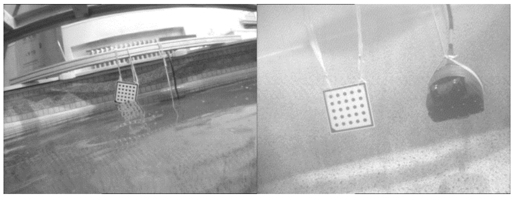
Step 1: Capture an image of calibration board in air at a moderate distance. As Fig. 1(a) captured at 1.8 meters. Compute the average gray values respectively for the black dots area and the background area, represented by E_b and E_w;

Step 2: Immerse the camera and the calibration board into the pool and capture n frames at different distances as degraded images. Fig. 1(b) is frame captured at 0.97 m;

Step 3: Compute respectively the average gray values for the black dots area and the black background area for all the n frames, represented by E_{bi}' and E_{wi}', $i = 1, 2, \cdots, n$. The $2n$ equations involving two unknowns, E_∞ and β, can be obtained:

$$\begin{aligned} E_{bi}' &= E_b e^{-\beta d_i} + E_\infty(1 - e^{-\beta d_i}) \\ E_{wi}' &= E_w e^{-\beta d_i} + E_\infty(1 - e^{-\beta d_i}) \end{aligned} \quad i = 1, 2, \cdots, n \tag{3}$$

Step 4: Solutions for the equation set are $\beta_i d_i$, $E_{\infty i}$, $i = 1, 2, \cdots, n$. It can be represented by n points: $(d_i, \beta_i d_i)$, $i = 1, 2, \cdots, n$, plotted in Fig. 2. Linear fitting is adopted to estimate β: $\beta d = 0.64d - 0.11$, hence $\beta = 0.64$. E_∞ is estimated by averaging: $E_\infty = 115$.



(a) (b)

Fig. 1. Images used in model calibration: (a) is captured in air as the unattenuated image; (b) is captured in pool

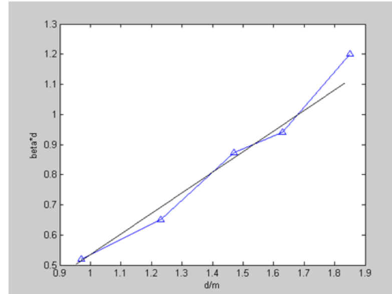


Fig. 2. $\beta d - d$ map: β is the coefficient reflecting property of water medium, d is the scene's distance and βd is scaled depth

B. Segmented-linear-mapping Restoration

Theoretically, once the model has been calibrated, it can be used to restore images by model inversion. The inverse function of the model is demonstrated in Eq. 4. The process of image restoration can be considered as mapping from degraded image to true image. First, decide a benchmark for d according to prior information. Then leave alone the physical meaning of d, instead, consider d as an array of coefficients of the segmented linear mapping with certain restrictions, according to which the array of d are decided.

$$E_0 = e^{\beta d} E + (1 - e^{\beta d}) E_\infty \tag{4}$$

The process of restoration can be considered as a model based mapping from the measured pixel value E to its true value E_0. All possible values for the degraded image pixel form the definition domain of the mapping, denoted by E. The set of possible values for the restored pixel constitutes the codomain of the mapping, denoted by E_0. Both E and E_0 are subsets of U_0, i.e. $E \subseteq U_0$, $E_0 \subseteq U_0$, U_0 = [0, 255]. For different values of d mean different mapping slopes, a segmented linear mapping approach was proposed to approximate the mapping model from the degraded pixels to the true pixels.

The proposed segmented-linear-mapping approach was employed on Fig. 1(b), whose histogram was calculated firstly, as shown in the upper part of Fig. 3. The bottom part of Fig. 3 is the segmented-linear-mapping fold line, whose functional representation is in Eq. 5. The highest frequency appears at grayscale 139, and there are three other local maximums at 16, 157 and 191. The steepest mapping segment was located spanning over the grayscale 139. Furthermore, each local maximum and its neighborhood fall into the same mapping segment. To refine the mapping model, three more segments (d_3, d_4 and d_5) were inserted among the maximums. Figure 4 is the restored image. It can be seen that the restored image gets the highest recovery and the overall image quality was unquestionable improved.

$$E_0=\begin{cases}0 & E\in[0,25]\\ E-25 & E\in[26,84]\\ (E-115)e^{0.6\times1.0}+115 & E\in[85,132]\\ (E-115)e^{0.6\times1.2}+115 & E\in[133,148]\\ (E-115)e^{0.6\times1.15}+115 & E\in[149,165]\\ (E-115)e^{0.6\times1.13}+115 & E\in[166,177]\\ 255 & E\in[178,255]\end{cases} \tag{5}$$

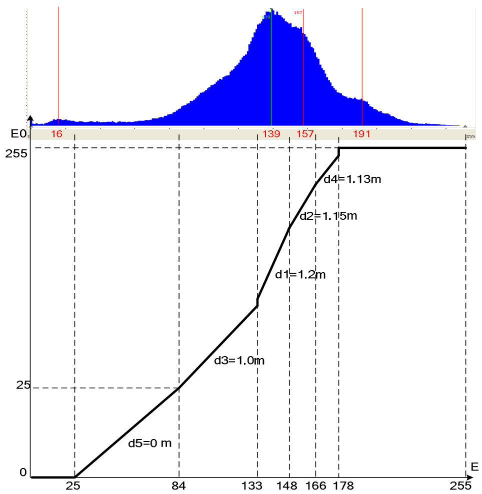


Fig. 3. Segmented-linear-mapping diagram

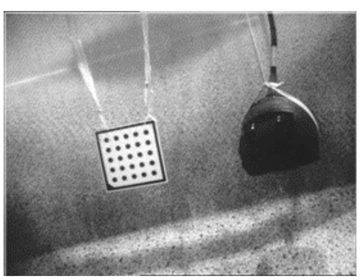


Fig. 4. Segmented-linear-mapping restoration result

III. THE LIGHT SOURCE EDGE EXTRACTION

In order to identify the guided light source accurately, we need to extract the precise contours of the light source. In this paper, Canny edge detection algorithm is applied to the restored image segmentation and edge extraction. Then the improved Snake model is adopted to extract the light source edge further.

Snake model is the mobile contour under the actions of internal force and external constraining force [9]. Snake curve can be bended and elongated, it converges to the energy minimum autonomously under the actions of internal energy, image energy and external constraint energy. The energy function can be described as:

$$E=\int_0^1[E_{\text{int}}(V(s))+E_{ext}(V(s))+E_{con}(V(s))]ds \tag{6}$$

The process of approaching the contour is to minimize the energy function. Place n control points on the Snake curve, with variational method and the Euler-Lagrange equation the iterative equations can be expressed as follows.

$$\begin{aligned}X^{t+\delta_t}&=M^{-1}\left(X^t-\delta_t\frac{\partial E_{ext}}{\partial X^t}\right)\\ Y^{t+\delta_t}&=M^{-1}\left(Y^t-\delta_t\frac{\partial E_{ext}}{\partial Y^t}\right)\end{aligned} \tag{7}$$

The traditional Snake model exists defects which can't converge to the edge of the recessed position, while the depression on heart-shaped light source is the most important feature. However, GVF (Gradient Vector Flow) Model is gradient vector field [10], the minimum energy function is shown as following

$$\varepsilon=\iint[\mu(u_x^2+u_y^2+v_x^2+v_y^2)+\left|\nabla f(x,y)\right|^2\left|w-\nabla f(x,y)\right|^2dxdy \tag{8}$$

Gradient Vector Field is got by solving partial differential equations before the Snake curve coordinates iterative computation. $\left(\frac{\partial E_{ext}}{\partial X^t},\frac{\partial E_{ext}}{\partial Y^t}\right)$ is replaced in the iterative calculation.

This paper adopts GVF-Snake model to extract edge of the light source, and greedy algorithm to optimize the iterative process of Snake model [11].

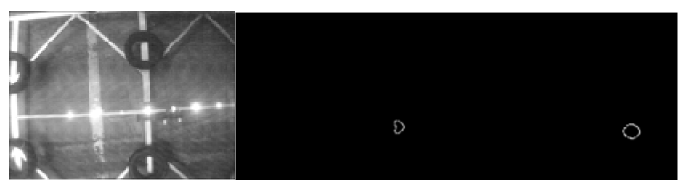


Fig. 5. Edge extractions of heart-shaped light source and round light source

Moment invariant is invariable under shifting, scaling and rotation. This paper takes moment invariants of edge extractions as recognition features. The moment invariants of the round and heart-shaped contours are shown in Fig. 6. It can be seen that the moment invariants between round light source edge and heart-shaped light source edge are different. So they can be identified easily.

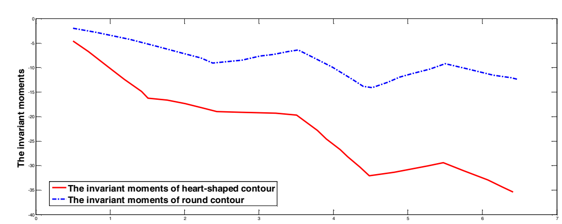


Fig. 6. The moment invariants of round contour and heart-shaped contour

IV. Recognition Method Based on PSO-BP Network

A. The learning process of BP network

The back propagation (BP) algorithm includes forward propagation and error back propagation. Forward propagation calculates input and output of each node and the network output error based on the input data, weights, thresholds and excitation function. Back propagation aims at minimizing the error by adjusting weights and thresholds of the respective layers in reverse when errors do not meet the expectations.

Input vector of BP network is $P_k=(x_1,x_2,\cdots,x_n)$, hidden layer output vector is $B_k=(y_1,y_2,\cdots,y_m)$, output vector is $O_k=(o_1,o_2,\cdots,o_p)$, the weights of input layer and hidden layer is w_{ij}, the weights of output layer and hidden layer is v_{jp}, expected output is t_p, threshold of hidden layer nodes is θ_j, threshold of output layer nodes is γ_p. Activation function with S-type has been used in this experiment. BP network learning and training process is as follows:

Step 1: Initialize the weights and thresholds of BP network and set a random number in (-1, 1) for each one.

Step 2: Input the learning samples (P_k,T_k), where $T_k=(t_1,t_2,\cdots,t_p)$ means expected output sample, it is usually a 0 and 1 code combination in target recognition.

Step 3: The output value of hidden layer nodes can be calculated as

$$y_j=f(\sum_i w_{ij}x_i-\theta_j) \tag{9}$$

Step 4: The output value of output layer nodes is calculated as

$$o_p=f(\sum_j v_{jp}y_j-\gamma_p) \tag{10}$$

Step 5: The generalization error d_s^k and the global error e of output layer nodes can be calculated according to $T_k=(t_1,t_2,\cdots,t_p)$ and O_p.

$$d_s^k=(t_s^k-o_s)o_s(1-o_s) \qquad s=1,2,\cdots p \tag{11}$$

$$e=\sum_{s=1}^{p}d_s^k \tag{12}$$

Step 6: The generalization error of hidden layer can be calculated as:

$$e_j^k=\left[\sum_{s=1}^{p}d_s\times v_{js}\right]y_j(1-y_j) \tag{13}$$

Step 7: The weights and thresholds of output layer are adjusted based on step 5.

$$v_{js}(N+1)=v_{js}(N)+\eta\cdot d_s^k\cdot y_j \tag{14}$$

$$\gamma_s(N+1)=\gamma_s(N)+\eta\cdot d_s^k \tag{15}$$

where $0<\eta<1$, $s=1,2,\cdots,p$, $j=1,2,\cdots,m$

Step 8: The weights and thresholds of hidden layer can be adjusted as follows.

$$\begin{aligned} w_{ij}(N+1)&=w_{ij}(N)+\eta e_j^k x_i^k \\ \theta_j(N+1)&=\theta_j(N)+\eta e_j^k \end{aligned} \tag{16}$$

where $i=1,2,\cdots n,\quad j=1,2,\cdots m,\quad 0<\beta<1$.

Step 9: Select the next set of learning samples, go to step 3 until all the k samples are trained.

Step 10: Continue training the samples and go to step 3 until the global error is less than the preset value.

B. The Principle of PSO-BP Network

Particle swarm optimization algorithm (PSO) takes optimal solution as a particle of the search space. The quality of each particle is determined by the fitness function of optimization problem, flight direction and distance are determined by the velocity of each particle. It is assumed that a particle swarm has M particles and the search space is D, property settings of particle I at moment t can be described as follows:

Position: $x_i^t=(x_{i1}^t,x_{i2}^t,\cdots,x_{id}^t)^T$, $x_{id}^t\in\left[L_d,U_d\right]$, L_d and U_d are the upper limit and lower limit of the search space respectively.

Velocity: $v_i^t=(v_{i1},v_{i2},\cdots,v_{id})^T$, $v_{id}^t\in\left[v_{\min,d},v_{\max,d}\right]$, $v_{\min}$ and $v_{\max}$ are the minimum and maximum speeds of the particle respectively.

Personal best position: $p_i^t=(p_{i1}^t,p_{i2}^t,\cdots,p_{iD}^t)^T$.

Global best position: $p_g^t=(p_{g1}^t,p_{g2}^t,\cdots,p_{gD}^t)^T$, $1\le d\le D$, $1\le i\le M$.

The position of the particle at moment t+1 can be obtained as follows [12]:

$$\begin{aligned} v_{id}^{t+1}&=wv_{id}^t+c_1r_1(p_{id}^t-x_{id}^t)+c_2r_2(p_{gd}^t-x_{gd}^t), \\ x_{id}^{t+1}&=x_{id}^t+v_{id}^{t+1} \end{aligned}$$

Where, r_1 and r_2 are random numbers well distributed in (0, 1), c_1 and c_2 are learning factors, w is inertia weight whose value decides its current speed inheritance. It is beneficial for balancing the searching ability and developing ability to choose proper inertia weight value.

The gradient descent method is used in weight adjustment process of BP network algorithm, so it needs a long training time because of the differential, derivative of target function. Particle swarm optimization algorithm is applied to weight and threshold modification instead of traditional gradient descent method, it can improve the performance of BP network, avoid generating local minimum, shorten training time and strengthen the generalization ability of network.

The method of optimized BP network with particle swarm optimization algorithm (PSO) is as follows: Particle replaces the weight and threshold value of BP network and all particles form a particle swarm. The mean square error of network is taken as fitness function. During the network training process, optimal position and speed are adjusted according to particle swarm optimization algorithm until global minimum mean square error is found. Finally, the weight and threshold value are obtained and the training process ends. The algorithm steps are given as follows:

Step 1: The nodes of the input layer, hidden layer and output layer of BP network are initialized.

Step 2: The size, iterations, inertia factor size, maximum speed, acceleration factor of particle swarm are initialized.

Step 3: Samples are inputted.

Step 4: Fitness value of each particle is calculated to update particle extreme value and global extreme value.

Step 5: Position and speed of particles are updated.

Step 6: Determine whether the algorithm reaches desired accuracy and the maximum number of iterations. Turn to Step 3 if it doesn't reach the conditions. Output global optimal position as weight and threshold value of BP network in the end.

V. EXPERIMENTAL RESULTS

Experiments used a three-layer BP network structure and took the extracted moment invariants as input data of the network. Select 7 nodes of the input layer, 9 nodes of the hidden layer, 1 node of the output layer, as is shown in Fig. 7. The underwater camera captured 100 guided light source images in different distances and different orientations. These images were restored with the method described in Section Ⅱ and the heart-shaped light contours were extracted. 50 of them were taken as training samples and others as testing samples. Simultaneously 30 circular light source contours were extracted, half of them were used for training interference and others for test interference.

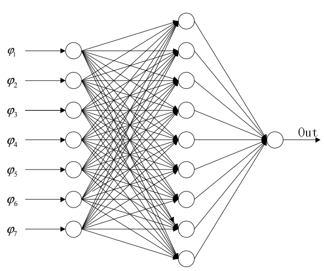
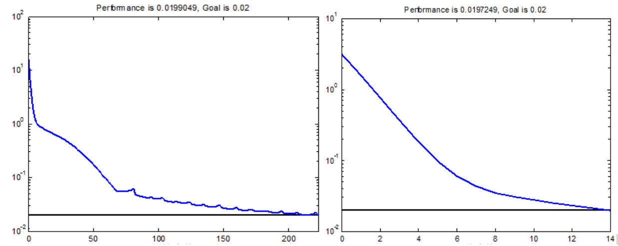


Fig. 7. The structure of recognition network

The heart-shaped light source recognition training is carried out by the traditional BP network and PSO-BP algorithm. The allowable error is 0.02, 50 heart-shaped contour images and 15 round contour images are selected, seven moment invariants are extracted as the input of the training network. The expected network output of the heart-shaped contour is 1 and the round contour is 0, the initial weights is a random number in (-1, 1). The training effect of BP network is best when the learning rate is 0.85. The population size of PSO-BP network is $D=(9\times7+9)+(1\times9+1)=82$, the maximum speed is 21, the inertia factor is 0.8, the acceleration factor is 2. The training results are shown in Fig. 8.



(a) (b)

Fig. 8. (a) Training error curve of BP network (b) Training error curve of PSO-BP network

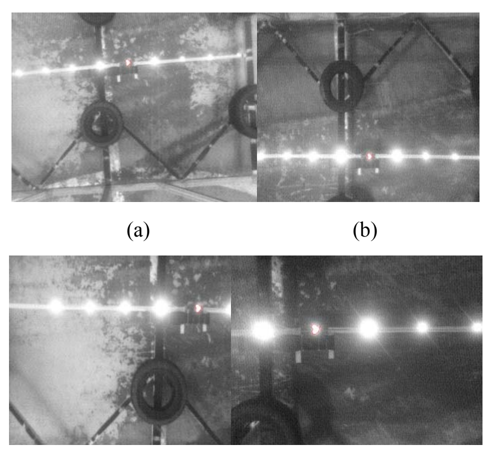
It can be seen that the two recognition networks can converge to the maximum permissible error. However, the PSO-BP network can improve the convergence speed and reduce the training steps. The trained weights and threshold value were stored in the computer for test.

Select 4 guided images for experiments, restore them in a segmented-linear-mapping approach, extract the precise contour curve with the improved Snake model and moment invariant features, take PSO-BP network as recognition method and mark the edge by the red line. Identification data are shown in Table 1.

TABLE I. IDENTIFICATION DATA

Number	Distance	Location	Time
(a)	3.0m	(154,353)	84.68ms
(b)	2.5m	(415,379)	58.43ms
(c)	1.9m	(135,623)	60.32ms
(d)	1.4m	(215,236)	34.87ms

PSO-BP network can recognize the heart-shaped light source from these images successfully, the results are shown in Fig. 9. The average time is 59.5ms, which can meet the real-time requirement of the target recognition.



(a) (b)

(c) (d)

Fig. 9. Recognition results of the heart-shaped light source

VI. CONCLUSION

The objective of this paper is to recognize the heart-shaped light source during UUV docking process. First of all, the scattering model is used to restore the underwater image. Secondly, Canny edge detection algorithm and improved Snake model are applied for extracting the precise guided light source edge. Then the moment invariants of the round and heart-shaped contours are taken as input data of neural network. Finally, because of the shortcomings of BP network, the PSO-BP network is designed for recognizing heart-shaped light source. The results show that it can recognize the heart-shaped light source accurately, which meant that the proposed PSO-BP network would be a reliable target recognition method for UUV docking.

ACKNOWLEDGMENT

This work is supported jointly by National Natural Science Foundation (NNSF) of China under Grant 51309067/E091002 and the Fundamental Research Funds for the Central Universities under Grant HEUCFX41402.

REFERENCES

[1] E. Bovio, D. Cecchi, and F. Baralli, "Autonomous underwater vehicles for scientific and naval operations," Annual Reviews in Control, vol. 30, pp. 117-130, 2006.

[2] M. Chyba, "Autonomous underwater vehicles," Ocean Engineering, vol. 36, no. 1, 2009.

[3] J. Horgan, D. Toal, "Review of machine vision applications in unmanned underwater vehicles," Control, Automation, Robotics and Vision, 2006. ICARCV'06. 9th International Conference on. IEEE, 2006, pp. 1-6.

[4] C. Nian-wen, L. Wen-qing, and Z. Yu-jun, "Quantitative study of improvements of the imaging contrast and imaging range by the polarization technique," Acta Physica Sinica, vol.49, pp. 61-66, 2000.

[5] Y. Y Schechner, N. Karpel, "Clear underwater vision," Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 536-543.

[6] J. P Queiroz-Neto, R. Carceroni, W. Barros, and M. Campos, "Underwater stereo," Computer Graphics and Image Processing, 2004. Proceedings. 17th Brazilian Symposium on. IEEE, 2004 pp. 170-177.

[7] K. T Kim, D. K Seo, and H. T Kim, "Efficient radar target recognition using the MUSIC algorithm and invariant features," Antennas and Propagation, IEEE Transactions on vol. 50, no. 3, pp. 325-337, 2002.

[8] K. Iqbal, M. O Odetayo, A. James, "Content-based image retrieval approach for biometric security using colour, texture and shape features controlled by fuzzy heuristics," Journal of Computer and System Sciences, vol. 78, pp. 1258-1277, 2012.

[9] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," International journal of computer vision, vol. 1, pp. 321-331, 1988.

[10] C. Xu, J. L Prince, "Snakes, shapes, and gradient vector flow," Image Processing, IEEE Transactions on vol. 7, pp. 359-369, 1998.

[11] D. J Williams, M. Shah, "A fast algorithm for active contours and curvature estimation," CVGIP: Image understanding, vol. 55, pp. 14-26, 1992.

[12] M. Clerc, J. Kennedy, "The particle swarm-explosion, stability, and convergence in a multidimensional complex space," Evolutionary Computation, IEEE Transactions on, vol. 6, pp. 58-73, 2002.