Intel Serv Robotics (2014) 7:175–184
DOI 10.1007/s11370-014-0153-y

ORIGINAL RESEARCH PAPER

Artificial landmark-based underwater localization for AUVs using weighted template matching

Donghoon Kim · Donghwa Lee · Hyun Myung · Hyun-Taek Choi

Received: 22 April 2013 / Accepted: 1 April 2014 / Published online: 22 April 2014

Abstract This paper deals with vision-based localization techniques in structured underwater environments. For underwater robots, accurate localization is necessary to perform complex missions successfully, but few sensors are available for accurate localization in the underwater environment. Among the available sensors, cameras are very useful for performing short-range tasks despite harsh underwater conditions including low visibility, noise, and large areas of featureless scene. To mitigate these problems, we design artificial landmarks to be utilized with a camera for localization, and propose a novel vision-based object detection technique and apply it to the Monte Carlo localization (MCL) algorithm, a map-based localization technique. In the image processing step, a novel correlation coefficient using a weighted sum, multiple-template-based object selection, and color-based image segmentation methods are proposed to improve the conventional approach. In the localization step, to apply the landmark detection results to MCL, dead-reckoning information and landmark detection results are used for prediction and update phases, respectively. The performance of the proposed technique is evaluated by experiments with an underwater robot platform and the results are discussed.

D. Kim · D. Lee · H. Myung (✉)
Urban Robotics Lab., Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon 305-701, Republic of Korea
e-mail: hmyung@kaist.ac.kr

H.-T. Choi
Ocean System Engineering Research Division, Korea Research Institute of Ships and Ocean Engineering (KRISO), 32 1312 Beon-gil, Yuseong-daero, Yuseong-gu, Daejeon305-343, Republic of Korea

Keywords Vision processing · Object detection · Segmentation · Localization · Autonomous underwater vehicle

1 Introduction

Underwater robots including remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs) are one of the challenging research fields in ocean engineering related to exploration and investigation. ROVs can perform complicated work by remote operation with a tether, and have been developed for underwater construction and rescue purposes [1–8]. AUVs, in contrast, operate without remote operation, and thus advanced autonomous navigation techniques are needed. Autonomous navigation is thus a very important feature for AUVs. In relation to this, the present study concentrates on vision-based underwater localization.

Underwater environments are three dimensional and dynamic. Due to this medium, the available sensors for localization are limited. Underwater localization for AUVs is hence very challenging and an important research field. A variety of underwater applications require an underwater robot to be able to successfully perform autonomous navigation. Examples include investigation of underwater structures, path following to a docking place, and area coverage for monitoring purposes [9–13].

The visual data from cameras are useful for underwater navigation. Although vision sensors have only relatively short range compared to other underwater sensors such as acoustic sensors and their performance may be limited by the featureless conditions in underwater environments, visual data play an important role, especially for object detection at a close distance. Vision-based object detection in underwater environments is challenging because of the rough conditions

in these environments including poor visibility due to poor light conditions with depth, turbidity, and undesired effects such as aquatic snow.

Many researchers have studied this challenging task using artificial landmarks, and various artificial landmarks have been designed for underwater localization. Some studies on vision-based guidance for docking and homing of AUVs have meanwhile been conducted by adopting computer vision techniques. Yu et al. [14] proposed a color extraction method for the localization of their AUV by detecting artificial underwater landmarks. Park et al. [15] developed an active landmark composed of an array of LED lights. Their AUV is guided to the docking place by detecting the landmark using a single camera. Dudek et al. [16] introduced reef detection techniques using a color correction filter and their underwater robot platform. Sattar and Dudek [17] demonstrated the performance of underwater object detection using an ensemble tracker. Negre et al. [18] proposed an underwater target named a self-similar landmark for estimation of range and bearing. Maire et al. [19] proposed a docking system with Haar rectangular features to detect the encoded poles in the docking location. These studies proposed particular landmarks for underwater localization. But the landmarks have to be placed in front of the AUV to accurately estimate localization information because of their two-dimensional design. The sides or the area behind the landmark are blind spots. Lee et al. [20] demonstrated the effectiveness of a color correction method for landmark detection and tracking for underwater localization using their AUV. As an extension of this work, Kim et al. [21] proposed an enhanced template-based landmark detection technique for detection of simple-shaped three-dimensional landmarks using a single camera. These two studies used three-dimensional landmarks to demonstrate their landmark detection technique, but the AUV had to move along a line oriented toward the landmarks at a fixed distance from them because the technique is susceptible to scale change of the landmarks. Other researchers have used manmade objects for visual inspection using AUVs. Balasuriya et al. [22] developed an autonomous underwater cable tracking system with a single camera and an acoustic sensor. Hover et al. [23] used bolts, holes, and weld lines on the surface of a submerged hull as visual features for vision-based ship inspection. For these inspection missions, a camera as well as sensors for distance measurement such as an acoustic sensor and an imaging sonar is needed.

This paper proposes a novel vision-based underwater localization technique for underwater robots in a structured underwater environment. For localization purposes, simple-shaped three-dimensional landmarks are designed. The landmarks are detected based on a template-based approach [24]. However, the conventional template matching is not robust to changes in the illumination condition and affine transform, and thus a weighted correlation coefficient-based template matching with a multiple-template-based object selection scheme is proposed. In addition, to enhance the detection performance when the sizes of landmarks in images are varying, a color-based image segmentation technique is applied to the proposed technique. The proposed landmark detection algorithm is applied to a map-based underwater localization method based on a Monte Carlo localization (MCL) method [25]. MCL is widely used for robot localization problems by representing the probabilistic density of a vehicle's current pose with a number of particles. At the first stage, the current pose is estimated with velocity and attitude sensors of the vehicle. Subsequently, in the second stage, the object detection algorithm updates the vehicle's position based on range and bearing information of detected landmarks. Experiments involving landmark detection and MCL localization are performed with an AUV and artificial landmarks in a water tank. A video clip is taken during the localization experiment. Results of the proposed landmark detection algorithm are compared to those of the conventional method. The vehicle's pose is estimated on a two-dimensional plane while moving along a rectangular path. Results are compared to odometry-only data and the ground truth obtained from an indoor GPS system.

The remainder of this paper is organized as follows. Section 2 introduces the design of artificial landmarks, a novel vision-based landmark detection technique, and a map-based localization technique in detail. Section 3 describes details of the experimental setup and results for landmark detection and map-based localization. Section 4 summarizes and concludes this paper and presents directions for future work.

2 Vision-based landmark detection and localization

2.1 Design of artificial landmarks

The ultimate goal of underwater robots is to explore unstructured ocean environments with complete autonomy. Although some studies have been conducted on localization and mapping using vision data of natural features in underwater environments [26–28], it is still difficult to perform missions in the field, especially in a deep-sea environment, because these environments are characterized by very harsh conditions including poor light conditions with depth, decreasing visibility with turbidity, and noise in images due to aquatic snow. A structured environment using artificial landmarks can be helpful to mitigate this problem. Artificial landmarks are designed for localization while an AUV performs a mission. The main challenge is that the landmarks in water cannot be seen clearly due to the aforementioned conditions. Thus, four simple-shaped three-dimensional landmarks have been designed, as shown in Fig. 1 [20]. They have been used

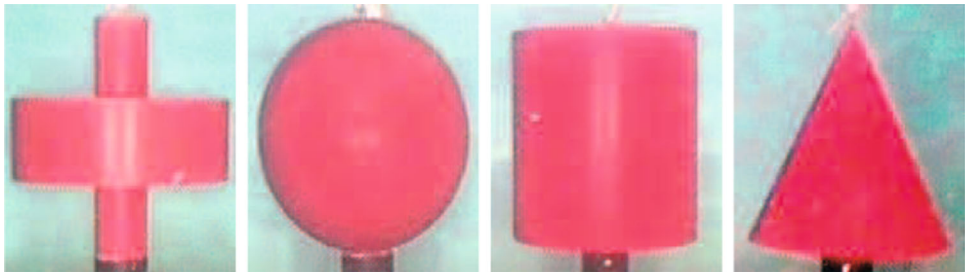

Fig. 1 Designed artificial underwater landmarks: *cross*, *sphere*, *cylinder*, and *cone*

to estimate the visual odometry and the location in structured underwater environments.

The artificial landmarks are designed with geometric symmetry, and thereby their appearances are almost unchanged with horizontal movement of the robot [20]. There are three reasons why the artificial landmarks are designed. First, because of the symmetric three-dimensional shapes, the robot can observe the landmarks in any direction toward the landmarks. Second, the shapes seen in the camera images such as cross, triangle, rectangle, and circle are simple enough to be detected through cameras in conditions such as high turbidity and low visibility. Third, red color has been selected as the color of the landmarks because it is easily seen in the underwater environment due to high contrast with dominant blue color of water. The landmarks have been designed with the dimensions of 0.4 m height and maximum 0.4 m diameter.

2.2 Landmark detection using template matching

The harsh visual conditions in underwater environments impede image processing. The attenuation and scattering of light due to turbidity of the medium blur the edges in images obtained from cameras. Aquatic snow underwater creates high-frequency noise in images. Moreover, the illumination conditions projected on surfaces of landmarks change greatly in deep-sea environments because often only one light source is installed on an underwater robot. Thus, preprocessing methods are employed to minimize the effects of these conditions and to maximize detection performance. The preprocessing includes image undistortion and noise reduction. An image from a camera has distortion due to the characteristics of the camera and the medium. Thus, the distorted images must be corrected using camera parameters through the camera calibration process [29]. In addition, noise reduction processes including Gaussian smoothing, a mean-shift filter [30], and a bilateral filter [31] are applied to remove the high- frequency noise. In particular, the bilateral filter is shown to remove noise while preserving edges of objects [31].

Landmark detection As the landmarks are used for underwater localization, computer vision techniques for object detection and recognition are needed. Various techniques, including SIFT [32], SURF [33], GLOH [34], and HOG [35], are available to analyze the correspondences between the feature points detected in two different images. An object can be recognized by measuring the similarities of the feature descriptors and their spatial distributions. However, these approaches are not suitable for designed landmarks [20]. The generated feature descriptors are not sufficient to judge the differences because the landmarks are designed as simple and textureless objects. Hence, considering the characteristics of the designed landmarks, template-based image matching techniques are used to solve the landmark detection problem. The template matching technique can be used to localize the position of the objects of interest in images based on a probabilistic approach. A statistical comparison of the template and the camera images generates a correlation coefficient. The values in the correlation coefficient denote probabilities of successful image matching. The positions with higher values than a threshold value can be considered as matched objects. This technique works well in small affine transformation, but it performs poorly for large affine transforms and for large variation in illumination conditions. In the following, a few techniques are introduced to enhance the robustness to illumination condition changes and scale changes.

Weighted correlation coefficient-based template matching The template and camera images are usually compared in *gray level* images for matching in a template-based approach. This approach is thus sensitive to changes in environmental conditions such as illumination and visibility, which is severe in underwater environments. To compensate the weaknesses of the template-based approach and to enhance its detection performance, two techniques are introduced: (i) an adaptive threshold-based approach and (ii) a weighted correlation coefficient. The detailed processes of these techniques are illustrated in Fig. 2. The *first and second columns* of the figure show the procedures of the conventional template matching and the template matching with adaptive thresholded images, respectively. The *bottom row* illustrates correlation coefficient, and a landmark is detected with the weighted correlation coefficient, as shown in the *third column* of the figure. The brightness in the correlation coefficient indicates the correlation coefficient value.

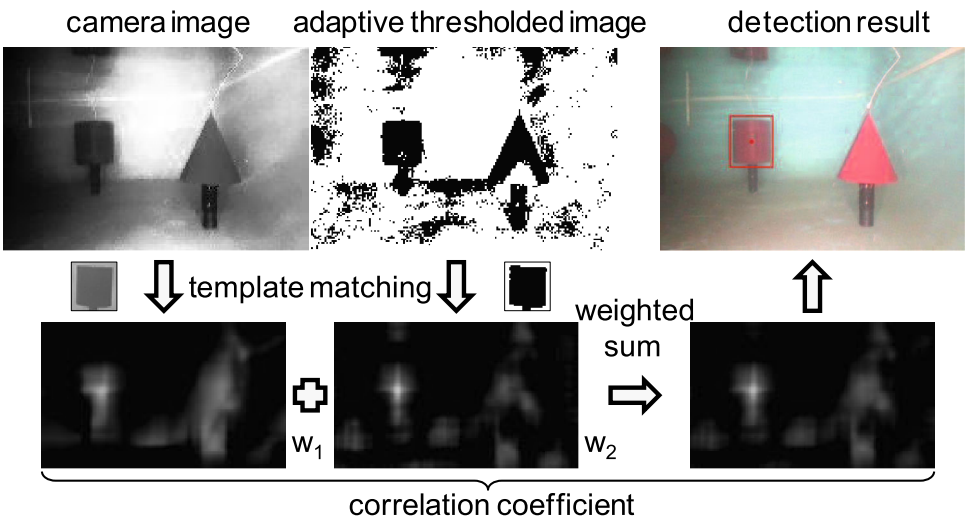

Fig. 2 Process of weighted correlation coefficient-based template matching. The *arrows* indicate the direction of the process. The *left two columns* show the conventional template matching and the template matching with adaptive thresholded images, respectively. The *bottom row* illustrates correlation coefficient. The weighted correlation coefficient on the *bottom right* is obtained by summing the left two correlation coefficient with weights , respectively. The brightness of the correlation coefficient denotes the correlation coefficient value. The result of landmark detection is denoted in the *top right with a box*

The adaptive threshold-based approach generates a correlation coefficient using the adaptive thresholded template and camera images. The adaptive threshold method is an approach to binarize an image with local threshold values. The local threshold values are calculated for all position of an image by statistically examining the intensity values of the local neighborhood, e.g., average, median, Gaussian weighted sum, etc. In this paper, the Gaussian weighted sum is used to calculate the local threshold values. The adaptive threshold method is used in uneven lighting conditions to segment foreground objects from its background and to binarize the image, which provides a rough shape of the object. Compared with the first correlation coefficient in Fig. 2, the second correlation coefficient provides the prominent correlation coefficient values at point where the template is matched. The third correlation coefficient shows a novel correlation coefficient that is the weighted sum of correlation coefficient of the conventional template matching and the template matching using adaptive thresholded images with the weights w_1 and w_2, respectively. Compared to the conventional correlation coefficient, this correlation coefficient is prominent at the locations of landmarks.

Multiple-template-based object selection method A selection method based on multiple templates is proposed to enhance the detection rate and to reduce the number of threshold values for the template matching. This is because selecting the position with the highest correlation coefficient may not be effective for template-based landmark detection. In other words, local maxima of correlation coefficient values greater than the threshold value might appear at the landmarks' positions, and the highest value might not be an object

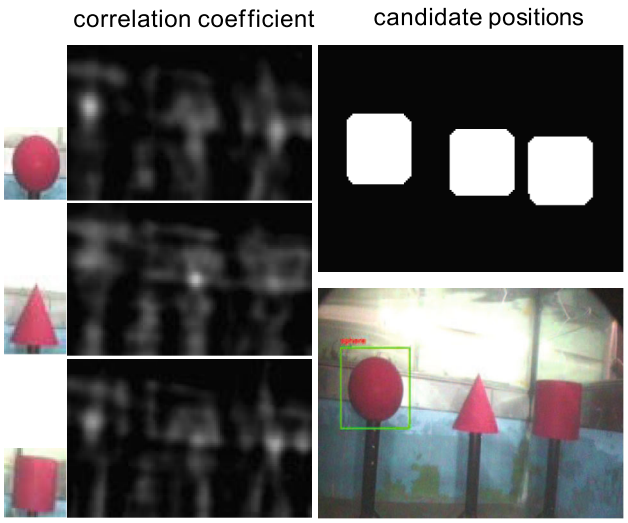

Fig. 3 Multiple-template-based object selection method. The *left two columns* indicate template images and correlation coefficient, respectively. The *top-right image* illustrates candidate positions with higher correlation coefficient values than a threshold value. The *bottom-right image* shows the detected landmarks with boxes

of interest due to illumination changes. Thus, a threshold value for a landmark might not ensure high landmark detection performance, and adjusting it for diverse landmarks is very difficult.

To solve this problem, a selection method is proposed by comparing correlation coefficient for all templates. Before applying this method, the correlation coefficient for all templates is obtained. The positions with higher correlation coefficient values than a proper threshold value, i.e. 0.5, in the correlation coefficient are extracted and merged into the white

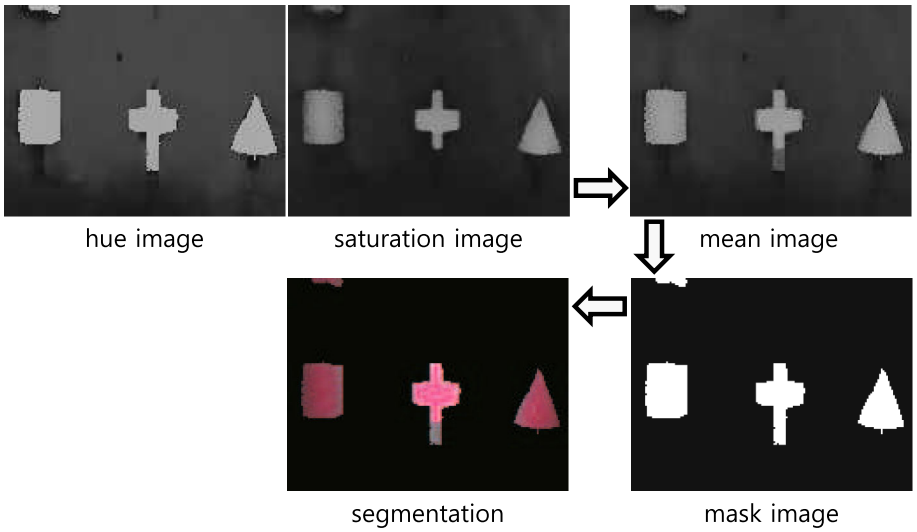

Fig. 4 Process of color-based image segmentation. The *arrows* indicate the directions of the process. Extracted hue and saturation images are used to generate a mask image by averaging two images and by applying the adaptive threshold method. Image segmentation is done using the mask image

regions, as shown in the top right of Fig. 3. Each white region indicates a candidate position of a landmark. For each white region, the template with the highest correlation coefficient value among all template candidates is selected as a matched object.

Image segmentation The detection performance of the template matching algorithm depends on the size of the template image. This dependency could restrict the movements of an AUV to ensure adequate landmark detection performance. Thus, a color-based image segmentation method is proposed, as shown in Fig. 4. With the light source installed on an AUV, the color of landmarks is observed with high hue and saturation values when the red color of the landmarks is dominant. The characteristics of the color are utilized for the proposed method. Hue and saturation images are extracted from camera images, and the mean image of the two images is converted to a mask image by thresholding through an adaptive threshold method. The regions of the landmarks are segmented from camera images using the generated mask image and resized to the size of the template images. The resized regions are matched with template images using the proposed template-based approach described earlier. When using this method, there is no need to scan all positions in the camera images for template matching, which results in a reduction of the computational load for template matching.

2.3 Map-based localization

For underwater localization with the vision-based landmark detection method suggested before, the MCL method is applied to our work [25]. This method employs particles $S_k = \{s_k^i\}$ $(i = 1 \ldots N$ where N is the number of the particles) at the current time k to represent the posterior density $p(\mathbf{x}_k|Z^k)$ of the current pose $\mathbf{x}_k = [x, y, \theta]$ given all measurements $Z^k = \{\mathbf{z}_k^j\}$ where $j = 1 \ldots k$. The procedure for estimating localization consists of two phases: prediction and update.

In the prediction phase, each particle s_{k-1}^i from the previous time step moves to $\bar{s}_k^i$ by applying a motion model $p(\mathbf{x}_k|s_{k-1}^i, \mathbf{u}_{k-1})$ where $\mathbf{u}_{k-1}$ is dead-reckoning data. In our work, an underwater robot motion is estimated using a Doppler velocity log (DVL) and an inertial measurement unit (IMU) on a two-dimensional plane. It is assumed that random noises of the translation and rotation values follow normal distributions with standard deviations σ_{trans} and σ_{rot}, respectively.

In the update phase, the particles from the prediction phase are weighted by $w_k^i = p(\mathbf{x}_k|\bar{s}_k^i)$ using a measurement model. The landmark detection method with underwater vision is used for obtaining the measurement model, and hence range and bearing values to the landmarks are estimated. In particular, the range value is obtained by referring to the segmented size from the landmark detection algorithm. It is also assumed that the noises of these values follow normal distributions with standard deviations σ_{range} and σ_{bear}, respectively. Finally, s_k^i is obtained by resampling from the weighted sample $\bar{S}_k = \{\bar{s}_k^i, w_k^i\}$.

3 Experimental results

3.1 Experimental setup

Experiments for demonstrating the performance of the proposed technique were conducted in an indoor test water tank at Korea Institute of Ocean Science and Technology (KIOST) in Daejeon, Korea. The indoor water tank has dimensions of $7 \times 4.5 \times 1.2$ m (length $\times$ width $\times$ depth) and four designed landmarks are installed inside it as shown in Fig. 5. For the

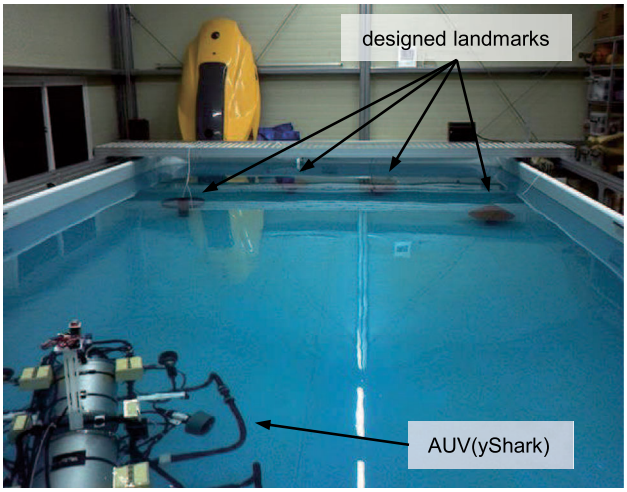

Fig. 5 Indoor water tank and an AUV for underwater experiments. Designed landmarks are set up inside the water tank

experiments, we used an AUV named yShark described in [20] developed by KIOST. This AUV can maneuver in five degrees of freedom (DOF) and has the following sensors: front and downward cameras (Bowtech Divecam-550C), a DVL, an IMU, and an indoor GPS. The indoor GPS system used is USAT A105 made by KoreaLPS [36]. This system measures the time-of-flight of ultrasound waves (40 kHz) with four transmitters and one receiver. Repeatability of this product is 0.1–0.01 m, which varies according to air conditions. All sensor data are saved with a synchronized timestamp in embedded computers. Four artificial landmarks are positioned as shown in Fig. 7 so that at least two landmarks are always visible by the front-facing camera, and their absolute positions are measured using the indoor GPS for map-based localization prior to the experiments. Two of the landmarks placed at close distance from yShark are set at a distance of at least 1 m up to 4 m, and the others are set at a distance of at least 2 m up to 5 m.

3.2 Landmark detection

Experiments for detection of the artificial landmarks have been performed. The front-facing camera and two white LED lights installed in front of yShark without any external light sources have been used to simulate deep-sea environmental conditions. A video clip taken with a resolution of 720×480 pixels at 15 fps, has been used for the experiments. A set of experiments have been performed using the video clip to examine the performance of the proposed landmark detection technique described in Sect. 2.2. The conventional template matching has been implemented for Experiment 1. The proposed methods also have been implemented without and with the image segmentation method for Experiments 2 and 3, respectively. The weight values w_1 and w_2 for the proposed methods are heuristically set to 0.6 and 0.4, because the small variation of the weight values makes little difference in the results. The results of Experiment 1 are compared with the results of others to verify the performance of the proposed methods. The results of Experiments 2 and 3 are compared to examine the effectiveness of the image segmentation method when sizes of landmarks are varying.

Table 1 Detection results of Experiment 1

Metrics	Cross	Cone	Sphere	Cylinder	Total
TP	58	120	29	122	329
FN	193	137	243	138	711
FP	0	151	97	0	248
TN	359	202	241	350	1,152
TPR	0.231	0.467	0.107	0.469	0.316
FPR	0.000	0.428	0.287	0.000	0.177

Table 2 Detection results of Experiment 2

Metrics	Cross	Cone	Sphere	Cylinder	Total
TP	56	176	118	226	576
FN	195	81	154	34	464
FP	0	36	0	106	142
TN	359	317	338	244	1,258
TPR	0.223	0.685	0.434	0.869	0.554
FPR	0.000	0.102	0.000	0.303	0.101

Table 3 Detection results of Experiment 3

Metrics	Cross	Cone	Sphere	Cylinder	Total
TP	88	175	256	139	658
FN	163	82	16	121	382
FP	0	13	22	2	37
TN	359	340	316	348	1,363
TPR	0.351	0.681	0.941	0.535	0.633
FPR	0.000	0.037	0.065	0.006	0.026

To evaluate the performance, true positive rate (TPR) and false positive rate (FPR) have been used as the performance evaluation metrics [37]. TPR, or sensitivity, represents the ratio of true positives (TP) to actual trues ($\text{TPR} = \text{TP}/(\text{TP} + \text{FN})$). FN stands for false negative, and FPR is calculated as $\text{FP}/(\text{FP} + \text{TN})$, where FP and TN denote false positive and true negative, respectively. TPR is the proportion of correct detection of a landmark of interest that yields positive test outcomes. FPR is the rate of absent detection events that also yield positive test outcomes. Larger TPR and smaller FPR are desirable.

For the analysis, 610 frames of the video clip were used. Landmark detection results of the experiments are shown in Tables 1, 2 and 3. As shown in three tables, the detection

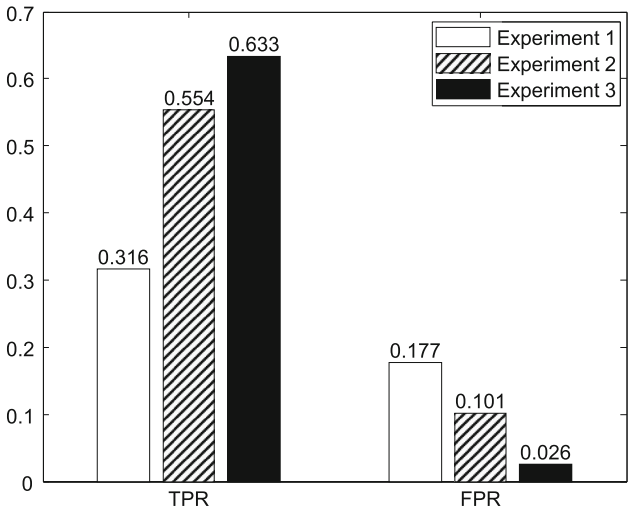

Fig. 6 The TPR and FPR of detection results of three experiments. The *bars* denote values of the TPR and FPR

with conventional template matching (Experiment 1) shows the worst results with TPR of 0.316 and FPR of 0.177. This is obvious since the sizes of landmarks and illumination condition on the landmarks are varying. In Experiment 2, with the proposed detection methods without the image segmentation method, TPR is 0.554 and FPR is 0.101. Compared to Experiment 1, the performance is better in both metrics. When the image segmentation method is applied in Experiment 3, TPR is increased to 0.633 and FPR is decreased to 0.026. The graph in Fig. 6 shows improvement of the detection performance using the proposed method. In all experiments, the TPR of the cross landmark is low compared with the results of other landmarks. The reason is that the cross landmark in a large portion of camera images was taken at excessively close distance, causing its red color to be seen as white color due to the white LED light, and the size of the landmark is enlarged in the images such that it is bigger than the field of view of the camera. In addition, FPR for the sphere is relatively high. When the distance between the camera and the landmark is large in underwater environments, the landmarks in an image become blurred and darkened due to attenuation and scattering effects of the medium. This trend occurred at distance of more than 3.5 m, and consequently the cylinder landmark was frequently detected incorrectly as a sphere. We can see the effectiveness of the proposed template-based approach with the images taken while the sizes of objects are varying, by comparing the results from these three experiments.

The results of this study show that the proposed methods enhance the robustness to the changes in environmental conditions such as illumination and visibility. This is because a novel correlation coefficient is calculated by weighted sum of the two correlation coefficients and an adaptive thresholded image provides a rough shape of the landmark regardless of the illumination changes and noise. In addition to that, since the correlation coefficient values for all template images are compared to select the correct landmarks among the candidate objects, the detection performance is also enhanced. The images segmented by the proposed method are used as the images to be matched with the template images, which makes the proposed methods robustly detect the landmarks when their sizes vary.

3.3 Map-based localization

Map-based localization of the underwater robot is conducted in a water tank. Four landmarks that have different shapes are positioned as shown in Fig. 7. The trajectory of the robot is

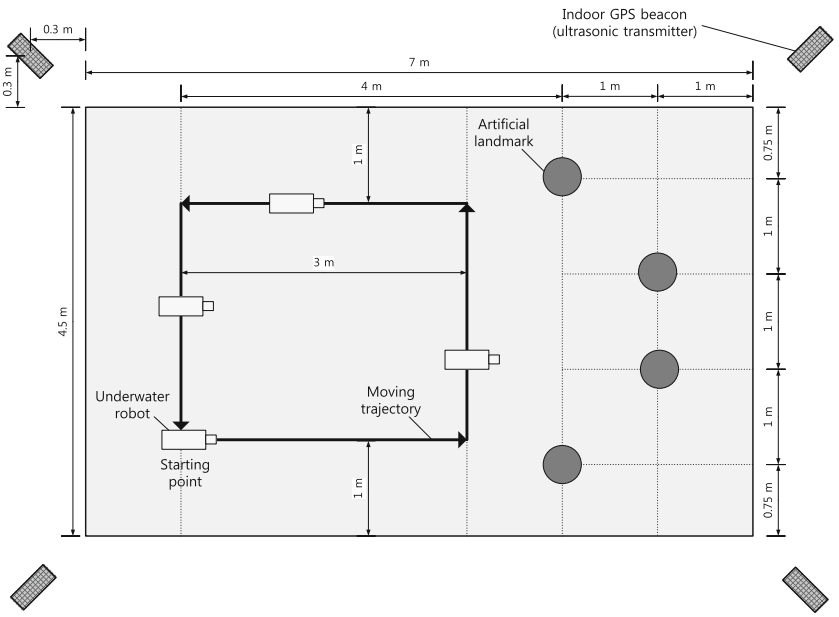

Fig. 7 Trajectory of an underwater robot equipped with a vision sensor and positions of artificial landmarks for an MCL experiment in a water tank. The robot moves on two-dimensional plane and always sees forward

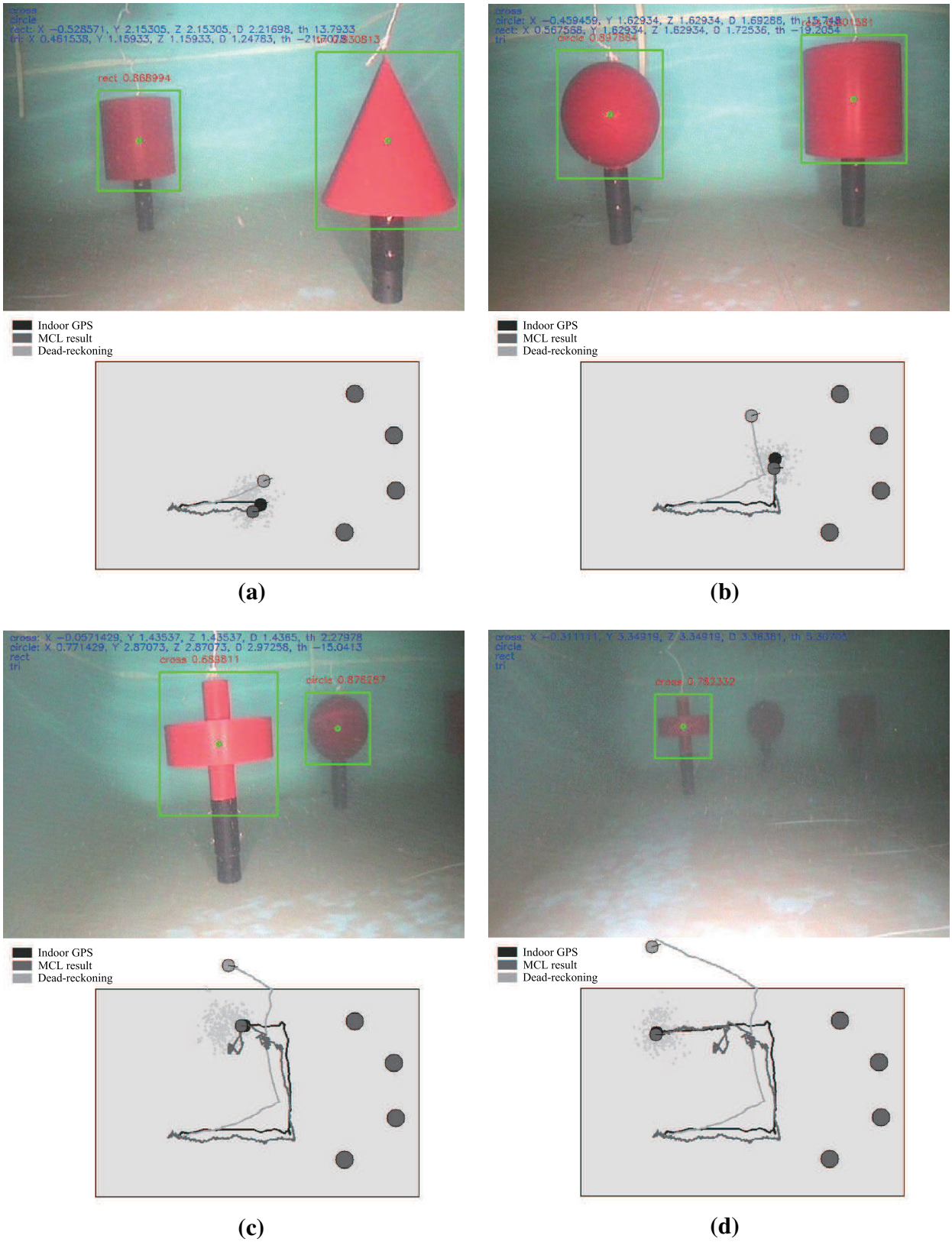

Fig. 8 *Top of each figure* shows landmark detection result. MCL results are shown on *bottom* which are compared to indoor GPS and dead-reckoning data. **a** 10 s, **b** 20 s, **c** 30 s and **d** 40 s

also indicated in Fig. 7. The robot moves along a rectangular path and sees forward to detect artificial landmarks. An indoor GPS receiver based on ultrasonic signal is equipped on top of the robot so that the receiver can receive ultrasonic signals, and four transmitters are installed on top of the corners in air as shown in Fig. 7. While the body of the robot moves underwater, the receiver is outside the water and hence ground truth data are obtained in air with the indoor GPS system.

In this experiment, the MCL technique is applied for probabilistic localization with 200 particles. In the prediction phase, the DVL and IMU sensors of the robot are used for the motion model. For the prediction of each particle, the standard deviations σ_{trans} and σ_{rot} are set to 0.1 m and 1°, respectively. In the update phase, the results of the landmark detection are applied to the measurement model. To give importance weight to each particle from the range and bearing estimation of the landmarks, the standard deviations σ_{range} and σ_{bear} are set to 0.5 m and 5°, respectively. The results of the MCL experiment are shown in Fig. 8. The top of each figure depicts the landmark detection result. Although the landmarks' size, illumination condition, and the amount of noise are different, most landmarks are correctly detected except the two landmarks placed at a distance longer than 3.5 m shown in Fig. 8d. The bearing is obtained from a landmark position on a two-dimensional image plane and the distance is calculated by comparing the size of landmarks with that of the matched template. The MCL result is shown at the bottom along with the trajectories of indoor GPS and dead-reckoning data. The trajectory of dead reckoning gradually goes outside the water tank. However, the MCL result follows the trajectory of the indoor GPS data. The root mean square error (RMSE) of Euclidean distance between the MCL result and the indoor GPS data is 0.28 m. In Fig. 8c and d, relatively large errors are shown in the upper right hand of the MCL trajectory because the Cross landmark is sometimes erroneously recognized as a rectangle at close distance due to a limited field of view.

4 Conclusion

This paper proposed a robust vision-based landmark detection technique and applied it to a map-based underwater localization for AUV. Four simple-shaped three-dimensional artificial landmarks were designed for vision-based underwater localization, and to detect them, the template-based object detection technique was employed. The performance of the conventional template matching is not robust when illumination conditions and sizes of the landmarks vary. We hence proposed an enhanced template-based object detection technique combining weighted correlation coefficient-based template matching, multiple-template-based object selection, and color-based image segmentation techniques. Map-based underwater localization algorithm is applied to our work using the proposed object detection method. The localization framework is based on the MCL algorithm. The robot motion is predicted using DVL and IMU sensors. In the update phase, the range and the bearing information of the landmarks are given by the object detection algorithm. The experiments performed in the indoor water tank using yShark AUV demonstrated that the proposed template-based landmark detection technique enhanced the detection performance, when the illumination conditions and size of the landmarks are varying. The experiment involving the localization algorithm was performed in the same environment, and the results show that the RMSE to the ground truth data is within tens of centimeters.

For improvement of the proposed approach, we will concentrate on (i) improving the image segmentation and template-based landmark detection techniques for landmarks at greater distances than current possible ranges and (ii) employing other forms of landmarks such as natural objects and textured landmarks to enrich the features in underwater environments.

Acknowledgments This research was supported by Korea Institute of Ocean Science and Technology (KIOST) (Project title: "Development of technologies for an underwater robot based on artificial intelligence for highly sophisticated missions") and by grant No. 10043928 from the Industrial Source Technology Development Programs of the MOTIE (Ministry Of Trade, Industry and Energy), Korea. The students are supported by Korea Ministry of Land, Transport and Maritime Affairs (MLTM) as U-City Master and Doctor Course Grant Program.

References

1. Caccia M (2006) Laser-triangulation optical-correlation sensor for ROV slow motion estimation. IEEE J Ocean Eng 31(3):711–727
2. Caccia M (2007) Vision-based ROV horizontal motion control: near-seafloor experimental results. Control Eng Pract 15(6):703–714
3. Ferreira F, Veruggio G, Caccia M, Bruzzone G (2012) Real-time optical SLAM-based mosaicking for unmanned underwater vehicles. Intell Serv Robot 5(1):55–71
4. Leabourne KN, Rock SM, Fleischer SD, Burton R (1997) Station keeping of an ROV using vision technology. In: Proceedings on MTS/IEEE OCEANS'97, vol 1, pp 634–640
5. Negahdaripour S, Firoozfam P (2006) An ROV stereovision system for ship-hull inspection. IEEE J Ocean Eng 31(3):551–564
6. Nomoto M, Hattori M (1986) A deep ROV "DOLPHIN 3K": design and performance analysis. IEEE J Ocean Eng 11(3):373–391
7. Whitcomb L, Yoerger D, Singh H, Howland J (1999) Advances in underwater robot vehicles for deep ocean exploration: navigation, control, and survey operations. Navigation, control and survery operations. In: 9th International symposium on robotics research, pp 346–353
8. Yoerger D, Newman J, Slotine JJ (1986) Supervisory control system for the Jason ROV. IEEE J Ocean Eng 11(3):392–400
9. Hollinger GA, Englot B, Hover FS, Mitra U, Sukhatme GS (2013) Active planning for underwater inspection and the benefit of adaptivity. Int J Robot Res 32(1):3–18

10. Jun BH, Park JY, Lee FY, Lee PM, Lee CM, Kim K, Lim YK, Oh JH (2009) Development of the AUV isimiand a free running test in an ocean engineering basin. Ocean Eng 36(1):2–14
11. Kim A, Eustice R (2009) Pose-graph visual SLAM with geometric model selection for autonomous underwater ship hull inspection. In: Procedings on IEEE/RSJ international conference on intelligent robotics and systems, pp 1559–1565
12. Marani G, Choi S (2010) Underwater target localization. IEEE Robot Autom Mag 17(1):64–70
13. Webster SE, Eustice RM, Singh H, Whitcomb LL (2012) Advances in single-beacon one-way-travel-time acoustic navigation for underwater vehicles. Int J Robot Res 31(8):935–950
14. Yu SC, Ura T, Fujii T, Kondo H (2001) Navigation of autonomous underwater vehicles based on artificial underwater landmarks. In: Procedings on MTS/IEEE OCEANS 2001, pp 409–416
15. Park JY, Jun BH, Lee PM, Oh J (2009) Experiments on vision guided docking of an autonomous underwater vehicle using one camera. Ocean Eng 36(1):48–61
16. Dudek G, Jenkin M, Prahacs C, Hogue A, Sattar J, Giguere P, German A, Liu H, Saunderson S, Ripsman A et al (2005) A visually guided swimming robot. In: Procedings on IEEE/RSJ international conference on intelligent robotics and systems, pp 3604–3609
17. Sattar J, Dudek G (2009) Robust servo-control for underwater robots using banks of visual filters. In: Procedings on IEEE international conference on robotics and automation, pp 3583–3588
18. Negre A, Pradalier C, Dunbabin M (2008) Robust vision-based underwater homing using self-similar landmarks. J Field Robot 25(6–7):360–377
19. Maire FD, Prasser D, Dunbabin M, Dawson M (2009) A vision based target detection system for docking of an autonomous underwater vehicle. In: Proceedings of the Australasian conference on robotics and automation. Australian Robotics and Automation Association, Sydney
20. Lee D, Kim G, Kim D, Myung H, Choi HT (2012) Vision-based object detection and tracking for autonomous navigation of underwater robots. Ocean Eng 48:59–68
21. Kim D, Lee D, Myung H, Choi HT (2012) Object detection and tracking for autonomous underwater robots using weighted template matching. In: Procedings on MTS/IEEE OCEANS 2012. IEEE, pp 1–5
22. Balasuriya B, Takai M, Lam W, Ura T, Kuroda Y (1997) Vision based autonomous underwater vehicle navigation: underwater cable tracking. In: Procedings on MTS/IEEE OCEANS'97, vol 2, pp 1418–1424
23. Hover FS, Eustice RM, Kim A, Englot B, Johannsson H, Kaess M, Leonard JJ (2012) Advanced perception, navigation and planning for autonomous in-water ship hull inspection. Int J Robot Res 31(12):1445–1464
24. Brunelli R (2009) Template matching techniques in computer vision: theory and practice. Wiley, New York
25. Dellaert F, Fox D, Burgard W, Thrun S (1999) Monte Carlo localization for mobile robots. In: Procedings on IEEE international conference on robotics and automation, vol 2, pp 1322–1328
26. Aulinas J, Carreras M, Llado X, Salvi J, Garcia R, Prados R, Petillot YR (2011) Feature extraction for underwater visual SLAM. In: Procedings on MTS/IEEE OCEANS 2011, pp 1–7
27. Salvi J, Petillot Y, Thomas S, Aulinas J (2008) Visual slam for underwater vehicles using video velocity log and natural landmarks. In: Procedings on MTS/IEEE OCEANS 2008, pp 1–6
28. Singh H, Can A, Eustice R, Lerner S, McPhee N, Pizarro O, Roman C (2004) Seabed AUV offers new platform for high-resolution imaging. Eos Trans Am Geophys Union 85(31):289–296
29. Zhang Z (2000) A flexible new technique for camera calibration. IEEE Trans Pattern Anal 22(11):1330–1334
30. Comaniciu D, Meer P (2002) Mean shift: a robust approach toward feature space analysis. IEEE Trans Pattern Anal 24(5):603–619
31. Tomasi C, Manduchi R (1998) Bilateral filtering for gray and color images. In: Procedings on 6th international conference on computer vision, pp 839–846
32. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60(2):91–110
33. Bay H, Tuytelaars T, Van Gool L (2006) Surf: speeded up robust features. In: Proceedings on European conference on computer vision, pp 404–417
34. Mikolajczyk K, Schmid C (2005) A performance evaluation of local descriptors. IEEE Trans Pattern Anal 27(10):1615–1630
35. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: IEEE computer society conference on computer vision and pattern recognition, vol 1, pp 886–893
36. KoreaLPS. http://korealps.co.kr. Accessed 12 Apr 2014
37. Fawcett T (2006) An introduction to ROC analysis. Pattern Recogn Lett 27(8):861–874