

Reliable pose estimation of underwater dock using single camera: a scene invariant approach

Shatadal Ghosh¹ · Ranjit Ray² · Siva Ram Krishna Vadali² · Sankar Nath Shome² · Sambhunath Nandy²

Received: 29 July 2014 / Revised: 4 July 2015 / Accepted: 3 November 2015 / Published online: 28 November 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract It is well known that docking of Autonomous Underwater Vehicle (AUV) provides scope to perform long duration deep-sea exploration. A large amount of literature is available on vision-based docking which exploit mechanical design, colored markers to estimate the pose of a docking station. In this work, we propose a method to estimate the relative pose of a circular-shaped docking station (arranged with LED lights on periphery) up to five degrees of freedom (5-DOF, neglecting roll effect). Generally, extraction of light markers from underwater images is based on fixed/adaptive choice of threshold, followed by mass moment-based computation of individual markers as well as center of the dock. Novelty of our work is the proposed highly effective scene invariant histogram-based adaptive thresholding scheme (HATS) which reliably extracts positions of light sources seen in active marker images. As the perspective projection of a circle features a family of ellipses, we then fit an appropriate ellipse for the markers and subsequently use the ellipse parameters to estimate the pose of a circular docking station with the help of a well-known method in

Safae-Rad et al. (IEEE Trans Robot Autom 8(5):624–640, 1992). We analyze the effectiveness of HATS as well as proposed approach through simulations and experimentation. We also compare performance of targeted curvature-based pose estimation with a non-iterative efficient perspective-n-point (EPnP) method. The paper ends with a few interesting remarks on vantages with ellipse fitting for markers and utility of proposed method in case of non-detection of all the light markers.

Keywords AUV · Perspective projection · Pose estimation · Underwater docking · Vision guidance

1 Introduction

There has been a steady increase in need for long duration-cum-autonomous deep-sea operations in survey applications supporting offshore oil and gas facilities emplacement, maintenance of intercontinental communication cables, etc. [1]. A major constraint for deep-sea operations with an AUV is limited on-board energy and data storage capacity, which necessitates frequent launching and recovery of AUV for basic operations like energy re-fuelling and data uploading. It is for this reason, the concept of underwater docking has evolved, which helps in battery recharging, dynamic path planning, mission reloading, data retrieval, etc., while AUV is on a long duration subsea mission. Figure 1 depicts a typical dock–AUV arrangement.

Needless to highlight, realization of docking is highly challenging as one needs to design develop AUV controllers for navigation in dynamic underwater ambience. It is observed that researchers have exploited optical [2], acoustic [3] and electromagnetic [4,5] means to design docking systems. Though all three methods are recommended in

✉ Ranjit Ray
ranjitray@cmeri.res.in

Shatadal Ghosh
shatadal88@gmail.com

Siva Ram Krishna Vadali
srk_vadali@cmeri.res.in

Sankar Nath Shome
snshome@cmeri.res.in

Sambhunath Nandy
snandy@cmeri.res.in

¹ Academy of Scientific and Innovative Research (AcSIR), CSIR-CMERI, Durgapur, India

² Robotics and Automation Division, CSIR-CMERI, Durgapur, India

literature a comparative chart in Deltheil et al. [6] summarizes that optical means of docking have lower vulnerability to external disturbances and also possess good directional accuracy. It is for this reason, in the present work, we emphasize on vision-based autonomous docking of an AUV.

Generally, vision-based docking of an AUV is a complex task which requires 3D pose estimation of dock entrance in its vicinity. Several variants of vision-based docking systems have been proposed, including—color detection of markers [2], usage of self-similar landmarks [7], 3D shape identification [8], etc. Before we explore existing docking systems, we first broadly classify them into two categories, namely, active and passive imaging-based systems; where we refer to passive imaging systems as those which require external light source to illuminate the scene to be captured, and active ones as those which capture the scene to be an orderly arrangement of lights. To illustrate, Frederic et al. [9] use passive markers (i.e., black and white patterned sticks as landmarks) for docking an AUV. Assuming the dock to be vertically aligned with AUV, the authors dynamically measure the dis-

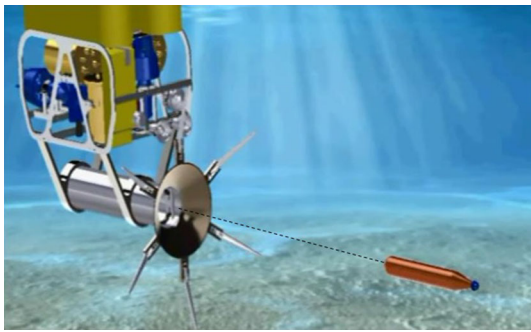


Fig. 1 An artist's view of AUV entering into a floating dock

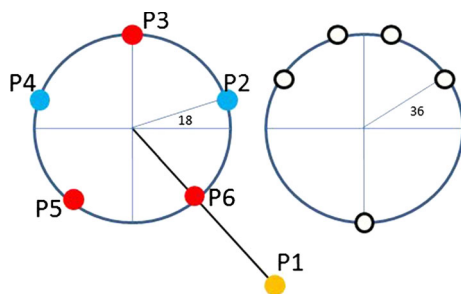


Fig. 2 Placement of beacons by Hong et al. [2] (*left*) and Park et al. [13] (*right*)

Fig. 3 Effect of non-uniform illumination on choice of threshold—original image (*left*), effect of choosing lower (*middle*) and higher (*right*) thresholds



tance between camera and dock with the help of an acoustic sensor. Negre et al. [7] propose to dock underwater vehicles using self-similar landmarks. However, designing and realizing active landmarks proposed by them are an extremely challenging task. Further placing such landmarks in front of a dock is very difficult in a mechanical design perspective. Maki et al. [10] proposed a docking method based on color detection of markers for hovering type AUV's. Their method requires 3D placement of colored markers whose mechanical arrangement once again is difficult. Moreover, it is shown in [11] that using active markers in underwater context has several vantages over passive ones.

Coming to active marker-based docking schemas, Hong et al. [2] propose to use colored lights, five on the periphery of dock and one away from the dock (Fig. 2, left). Since perspective projection of a circle looks like an ellipse, they propose to fit an ellipse for the identified markers and subsequently estimate 6-DOF pose of the dock. In their work, the basic assumption was that the diameter of a circular dock (D_C) would be the same as the length of major axis of an ellipse (MA_E) when seen in perspective projection, which may not always true as difference between D_C and MA_E increases with increasing relative angle between camera and dock [12]. Besides reliable detection of light markers with colored beacons in underwater ambience gets difficult due to spectral absorption.

Park et al. [13] used a similar geometric arrangement (5 colored lights for dock Fig. 2, right) to estimate 3D pose of the dock with respect to AUV. To estimate positions of light markers in spatial domain, they recommend to choose a random (but higher gray) intensity for binarization. However, as seen in Fig. 3, arbitrary choice of threshold can lead to unreliable detection of markers due to non-uniform illumination and scattering. Another problem with their approach is that their algorithm fails if all the lights are not detected. Besides, it is not lucid as to how the camera-dock range is measured through pixel count comparison.

1.1 Problem statement

Though good amount of literature is available, the concept of vision-based docking would prove successful only when near-accurate control parameters acquired from vision data are fed to the AUV controller. However, as mentioned earlier,

realizing a vision-based docking system is complex task and necessitates to address the following important issues:

- *Feature of docking station* Choice of a feature for entrance of the dock as seen by an AUV plays an important role in vision-based docking. Though different features are exploited by researchers, we believe that choosing a regular shape for the dock definitely aids in 3D pose estimation. For example, perspective projection of a circle features a family of ellipses, a fact we exploit in our work.
- *Threshold selection* Traditionally, 3D pose of a docking station with respect to the AUV is obtained from a binarized image. As we demonstrate in this work, choice of suitable threshold for binarization especially in presence of physical effects of the underwater channel definitely plays an important role.
- *Identification of dock center* For progressive maneuvering of AUV towards the dock, accurate computation of the center of light markers in the image plane is essential, especially during final stage docking. It is observed that to dock, it is common to use mass moment-based centroid of identified positions of markers. However, this may impose the stringent condition of accurate detection of all the light markers, failing which can lead to incorrect computation of the center, and eventual collision of AUV with the dock.
- *Pose estimation* It is known that the success of the conceptual docking relies in accurate and fast estimation of the docks' pose. However, this proves extremely challenging in a practical scenario as:
 - (a) in general, a docking station is in relative motion with respect to AUV, for fixed as well floating docks,
 - (b) it is highly likely that an AUV drifts from its targeted path (possibly) due to sudden currents resulting in need for repeated pose estimation.

It is for this reason, generally, accurate cum non-iterative solutions are sought for real-time 3D pose estimation. A review of available non-iterative 3D pose estimation methods may broadly be classified to be based on—point/line features (DLT [14], EPnP [15], LHM [16], RPnP [17]) and curvature [2, 18]. It is also noted that the most accurate point feature-based non-iterative solution is EPnP [15]. On the other hand, curvature-based methods as those developed in [18] are also found to be quite suitable for non-iterative 3D pose estimation. However, considering the 3D pose estimation requirement of AUVs docking into a circular cross section (as in the present context), we note that:

- Typically, arrangement of lights on the dock is arbitrary and obtaining correspondence between 3D light coordinates and 2D image points is not a simple task,

thus necessitating auxiliary methods like RANSAC be employed. Moreover, accuracy of RANSAC method depends on the number of iterations which in turn increases the computational time, making it unsuitable for real-time implementation. On the other hand, reportedly, curvature-based methods are particularly suitable for 3D pose estimation of circular sections. To illustrate, exploiting the fact that perspective projection of a circular section in an arbitrary orientation is always an exact ellipse and a circle has the property of high image-location accuracy, according to [18]—“the complete boundary or an arc of a projected circular feature (i.e., ellipse) can be used for 3-D pose estimation without knowing the exact point correspondence”.

- It is also observed from literature that existing methods fail to accurately estimate the pose of the dock when a few light markers remain undetected in the image. For instance, EPnP is unstable when number of detected light markers are less than five. However, elliptical curvature-based methods are capable of estimating the pose of circular sections even with four points (e.g., 4 light markers in the present context).

In view of the lacuna in available literature, in the present work, we make an attempt to estimate the pose of a circular dock relative to AUV upto 5-DOF via curvature-based pose estimation methods. Before we conclude the section, major contributions of the present work may be summarized as follows:

- Propose a novel adaptive thresholding scheme for feature extraction from active marker images. Also with the help of a point spread function (PSF) model, validate the proposed method through simulation and experimental analysis.
- Analyze and report vantages of employing ellipse fit-based methods in estimating pose of a circular dock.
- Estimate near-accurate relative pose of a circular dock with respect to AUV upto 5-DOF, even when some of the light markers are not detected.

The rest of the paper is organized as follows: Sect. 2 presents the proposed 5-DOF pose estimation method for vision-based docking an AUV. Section 3 presents the experimental results for the proposed pose estimation method. Section 4 concludes the paper with a few remarks on utility of proposed method.

2 Proposed pose estimation methodology

The present section describes the proposed 5-DOF relative pose estimation methodology for vision-based autonomous

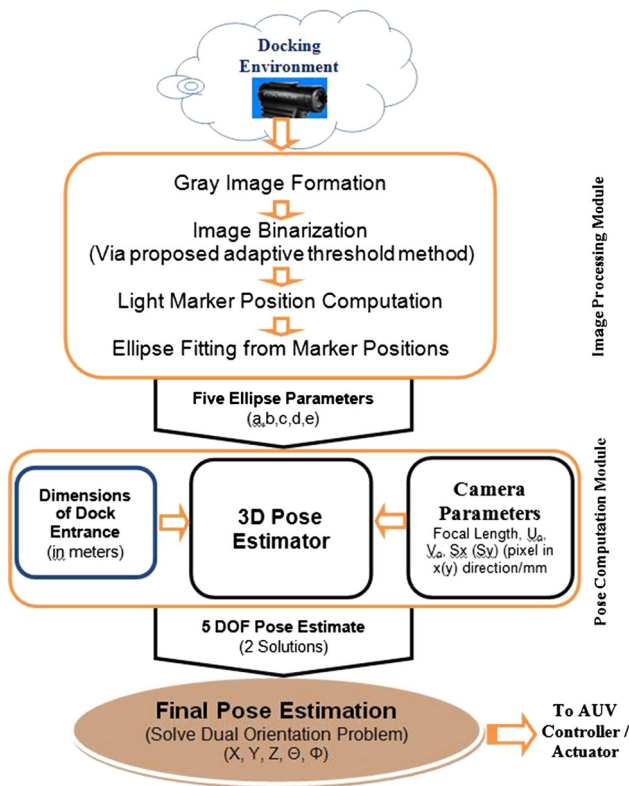


Fig. 4 Flow chart of proposed method to 3D pose estimation

docking of an AUV. Four assumptions made in devising the proposed method are:

1. AUV is homed up to vicinity point (10–20 m from dock [19]) and is ready for final stage docking.
2. Shape of the dock is assumed circular, which effectively eliminates the need to correct roll effects. This also implies that 5 DOF pose estimation is sufficient.
3. Underwater dock being tracked by AUV is, in principle, a circular arrangement of lights and hence images captured by camera are active marker images.
4. Effect of external light is negligible on imaging assuming docking is performed at larger depths.

Figure 4 depicts processing sequences of the proposed method which may be categorized into two phases, namely,

- (a) *Image processing phase* This phase includes detection of peripheral light markers through binarization of captured image using proposed HATS, identification of final position of each of the light marker via mass moment method and elliptical curve fitting for markers identified in binarized image.
- (b) *Pose estimation phase* In this phase, estimated ellipse parameters are fused with dimensions of dock and camera parameters to estimate the pose of dock for progres-

sive maneuvering of AUV towards, and finally into the dock. The following subsections describe each of these sequence of operations in detail.

2.1 Detection of light markers from captured image

As seen in Fig. 4, the first step in our pose estimation methodology is to binarize the captured image. However, instead of conventional fixed thresholding, we propose an image histogram-based adaptive thresholding method for the purpose. The utility of the proposed HATS would be apparent from the intuitive discussion in Sect. 2.1.1. Also, subsequently, we model the explained phenomenon and support our results through simulations. We also aid our results by verifying with a versatile model as well as with experimental images.

2.1.1 Proposed histogram-based adaptive thresholding method: an intuitive reasoning

To explain the basis of proposed adaptive histogram-based choice of threshold, first we recollect the fact that image of an object is the convolution of original image signal, say $f(x, y)$, with the system point spread function, $PSF(x, y)$, which, by definition is the effect of a point source propagating towards the camera through a scattering medium [20]. We also recall that the $PSF(x, y)$ of a medium is a combination of direct and glow field components which may be expressed as:

$$I(x, y) = f(x, y) * PSF(x, y) \\ = f(x, y) * [PSF_d(x, y) + PSF_g(x, y)] \quad (1)$$

Intuitively, since glow component is due to scattering and absorption effects, one may accept that direct-path component carries larger power as compared to scattered ones, i.e., strength of $PSF_d(x, y) \gg PSF_g(x, y)$. The point we have here is, it then follows that the direct component carries maximum intensity light to the camera, thereby creating a maximum intensity pixel in the image. Further, this behavior of direct component may also be ascertained from Volume Scattering Function (VSF) plot [light distribution with scattering angle (θ)] for ocean water. Further, according to [20], the behavior of direct-path stands valid up to a distance 75 m and is hence applicable in the present context, as typical dock–AUV distance ranges from 10 to 20 m.

Based on above discussion, it was felt that the histogram of an active marker image could contain a peak corresponding to the direct-path component. It is for this reason, we recommend to identify and use the gray value corresponding to the peak in the histogram, particularly on the higher intensity side, for binarization.

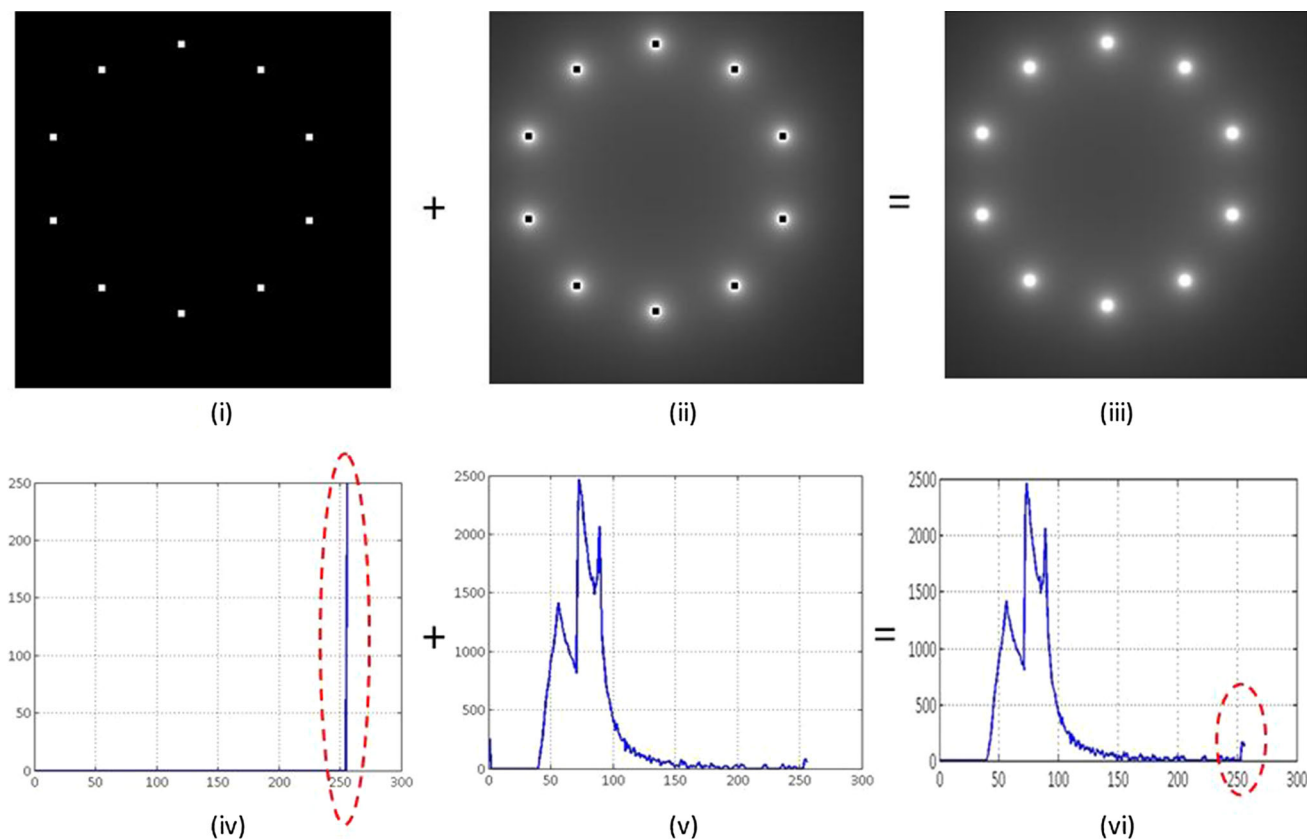


Fig. 5 Simulation of image formation using PSF in [22]: (i) direct and (ii) glow field components, (iii) image formed by summing (i) and (ii); corresponding histograms are shown in (iv) to (vi)

2.1.2 PSF-based modeling and simulation

In the following, we analyze our recommendation in Sect. 2.1.1 with the help of a PSF model and through simulations. First, we note that the phenomenon of light scattering from a surface is, generally, approximated to follow “Lambertian” distribution. Thus, an equivalent PSF can be formulated by computing the radiant distribution of an omnidirectional point source at the entrance of an imaging device, which is typically due to scattering and absorption effects. In underwater context, forward scattering is one of the main reasons for spreading of a point source, and depending upon the characteristics of surrounding medium can be modeled as a low-pass filter that generates symmetric/asymmetric image cone. Hou et al. [21] provide a versatile PSF model for a given angle (θ) , at a range (r) as:

$$PSF(\theta) = K(\theta_0) \frac{br e^{-\tau}}{2\pi\theta^m} \tag{2}$$

where $m = 1/(w_0 - 2\tau\theta_0)$, w_0 is scattering albedo, b is scattering coefficient, τ is optical length and $K(\cdot)$ is a functional constant dependent on mean scattering angle (θ_0) .

On the other hand, in our previous work in [22], we presented a *modified* version of (2) considering the fact that scattering of light (after a traveling r meters) varies according to a *Poisson* distribution. A near-accurate PSF model was for this reason provided as:

$$PSF(\theta) = K(\theta_0) \frac{poisson_rand(\lambda)e^{-\tau}}{2\pi\theta^m} \tag{3}$$

where $poisson_rand(\lambda)$ is a random number generated according to poisson distribution and $\lambda (=br)$ is *poisson* parameter.

Figure 5 simulates formation of an image using an *assumed* direct component and the *modified* PSF in (3). The direct and glow field components used for simulating the image are shown in Fig. 5, (i) and (ii), respectively, whereas their sum effect (forming the image) is shown in Fig. 5, (iii). Also shown directly below these figures are their corresponding histograms shown in Fig. 5, (iv)–(vi). Clearly, all three histograms indicate presence of a vivid peak close to higher gray value (253 in this case), which indeed proves our observation.

We repeat above simulation exercise with a widely used PSF model proposed by Voss et al. [20] especially for ocean

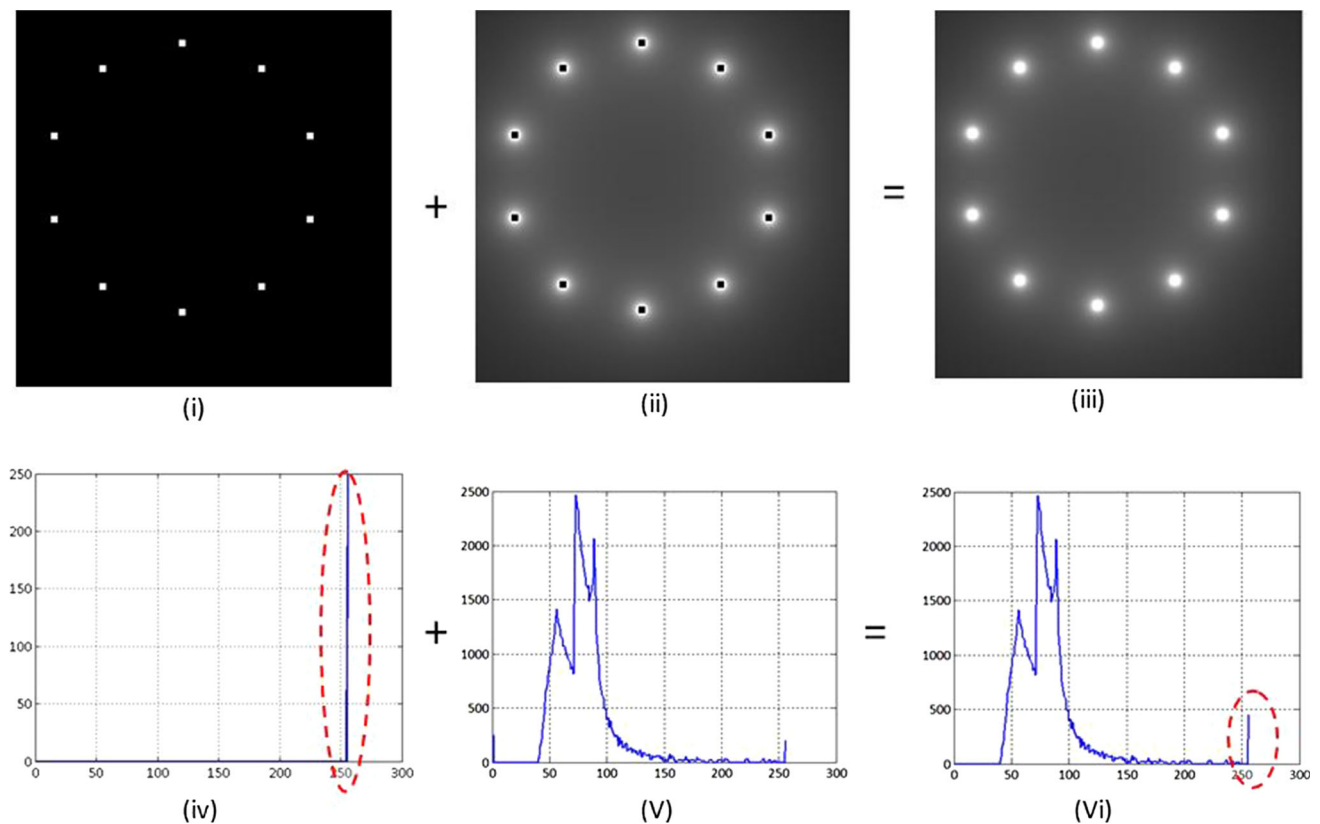


Fig. 6 Simulation of image formation using Voss's PSF in [20]: (i) direct and (ii) glow field components, (iii) image formed by summing (i) and (ii); corresponding histograms are shown in (iv) to (vi)

waters. They provide a modulation transfer function (MTF) whose zeroth order Hankel transform produces an equivalent PSF. Direct, glow field components of PSF, the image formed and their corresponding histograms are shown in Fig. 6, (i)–(iii) and (iv)–(vi), respectively. Clearly, Fig. 6 also validates our observation that a vivid peak is present in the histogram at higher intensities.

A point to be mentioned here is that the two discussed PSFs were formulated based on the assumption that light source is symmetric in nature. In general, light sources are asymmetric when viewed from an angle and such an asymmetric mask can be generated from a symmetric mask by rotating the PSF. Figure 7 shows simulated images using an asymmetric mask generated from PSF in (3). From the corresponding image intensity distributions shown in Fig. 7 (iv)–(vi), our claim once again stands justified. To validate our intuitive reasoning, we present histograms of experimental images obtained with the setup explained in Sect. 3. Figure 8 indeed advocates our recommendation to adaptively choose the threshold from histogram for binarization.

However, in practical scenario depending on the view angle, the rightmost intensity values in histograms of individual light markers (in an image) are not necessarily the same. This in turn reflects as multiple local peaks on higher intensity

side of the histogram of the entire image, thus necessitating extraction of individual light markers for finding the corresponding threshold value from their respective histograms. In other words, a “hard” choice of threshold (i.e., rightmost peak from the histogram of whole image) may result in detection of fewer distinct groups of light markers (for each of the individual light markers) as opposed to detection of all the light markers. Further, it is also observed that the variation in histograms of individual light markers is not significant as well. It is for this reason, in this work, considering the histogram of the whole image we recommend to adaptively find the most suitable threshold by iteratively progressing towards a lower intensity value (corresponding to the next peak in the histogram) on a voting basis which ensures detection of maximum groups of light markers.

2.2 Elliptical conicoid fitting for markers identified from binarized image

To estimate pose of the dock, as per our approach, we fit an appropriate conic for the markers identified from the thresholded image in Sect. 2.1. As noted from Sect. 1, in the context of vision-based docking, researchers have exploited the fact that perspective projection of a circle would feature an ellipse.

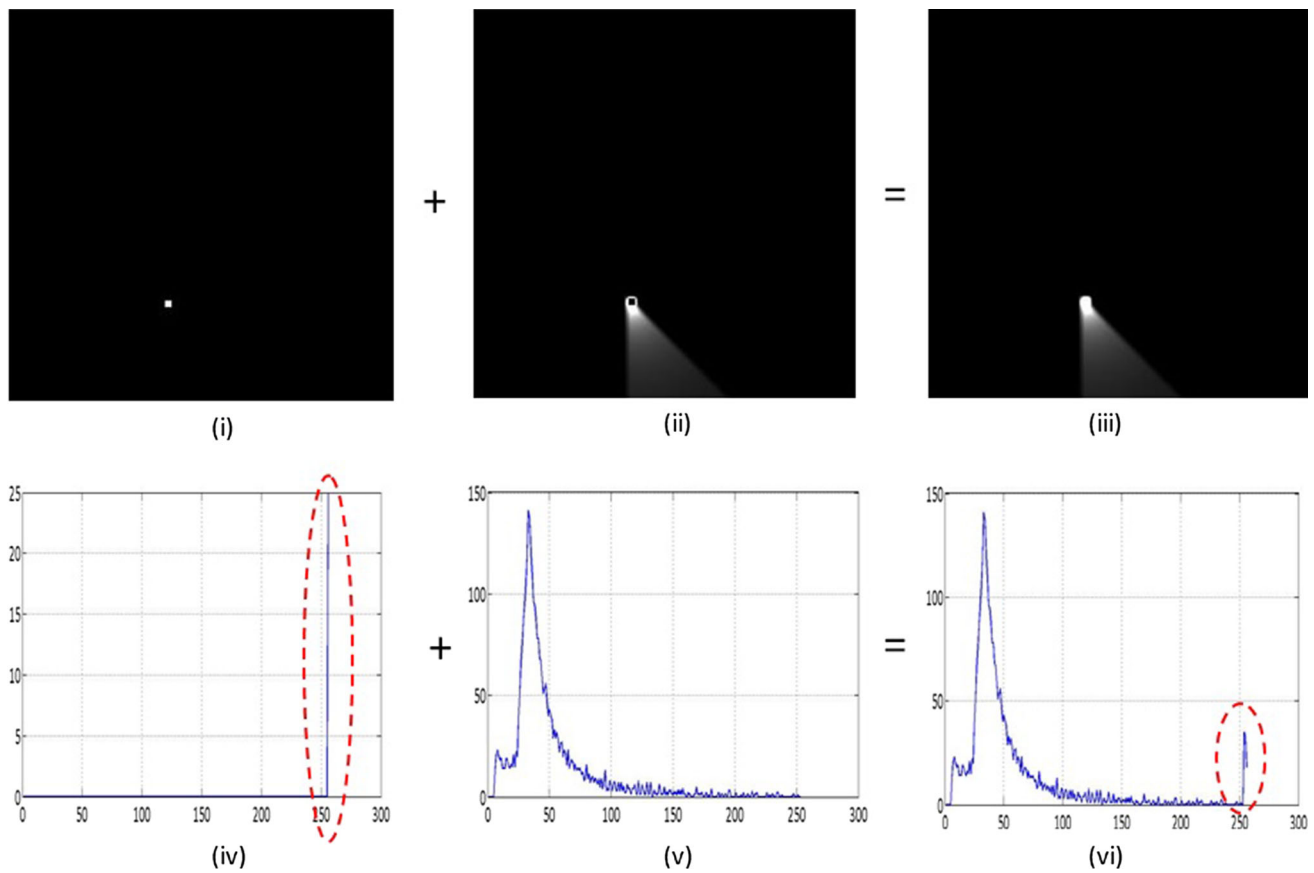


Fig. 7 Simulation of image formation by rotating PSF in [22]: (i) direct and (ii) glow field components, (iii) image formed by summing (i) and (ii); corresponding histograms are shown in (iv) to (vi)

However, very little work is observed on fitting an ellipse, particularly in the context of docking.

In general, the problem of elliptical curve fitting on a set of data requires choosing numerical methods which are supposed to be both optimal as well as stable. Several methods are available for fitting an ellipse, which in principle, may be classified into two categories, namely Clustering and Optimization techniques [23,24]. It is observed that clustering methods are robust to outlier data, but consume large memory and call for higher computational time. On the other hand, literature [23] suggests that least squares-based optimization methods are highly accurate, but suffer from non-convergence problems. It is for this reason, in the present work, we adopt a direct least squares-based optimization method proposed by Halir et al. [24] (which reportedly is highly numerically stable) for elliptical curve fitting on the positions of light markers. As derived in [23], the objective function to near-accurately fit a ellipse over a set of N points is,

$$\min_a \left\{ f^2(\mathbf{a}, \mathbf{x}) | \mathbf{a}^T \mathbf{C} \mathbf{a} = 1 \right\} \tag{4}$$

where $f(\mathbf{a}, \mathbf{x}) = \mathbf{D} \cdot \mathbf{a} = 0$ is the equation of an ellipse in the most general form. In (4), $\mathbf{x} = (x_i, y_i)$, (where (x_i, y_i) is center point of i th light marker obtained after conversion from intrinsic “pixel” coordinate to “axes” coordinate using calibrated camera parameters); $\mathbf{D} = (x_i^2, x_i y_i, y_i^2, x_i, y_i, 1)$, where $i = 1, 2, \dots, N$ and N is no. of light markers; $\mathbf{a} = (a', b', c', d', e', 1)$ and \mathbf{C} is a (6×6) constraint matrix filled with zeros, except $\mathbf{C}(1, 3) = \mathbf{C}(3, 1) = 2$ and $\mathbf{C}(2, 2) = -1$. Finally, parameters of ellipse, \mathbf{a} , are solved using the algorithm proposed by [24], which is subsequently used for estimating the entrance pose of the circular dock under consideration.

2.3 3D Pose estimation of circular dock

Having elliptic fit the perspective projection for the circular dock in the 2D image, in this section, we use the optimized ellipse parameters in \mathbf{a} from (4) to estimate the 3D pose of the docking station (with respect to AUV) as seen in Fig. 9.

Let γ be the focal length of camera. As shown in Fig. 9, taking ellipse as a base at $z = \gamma$ in the imagery plane and

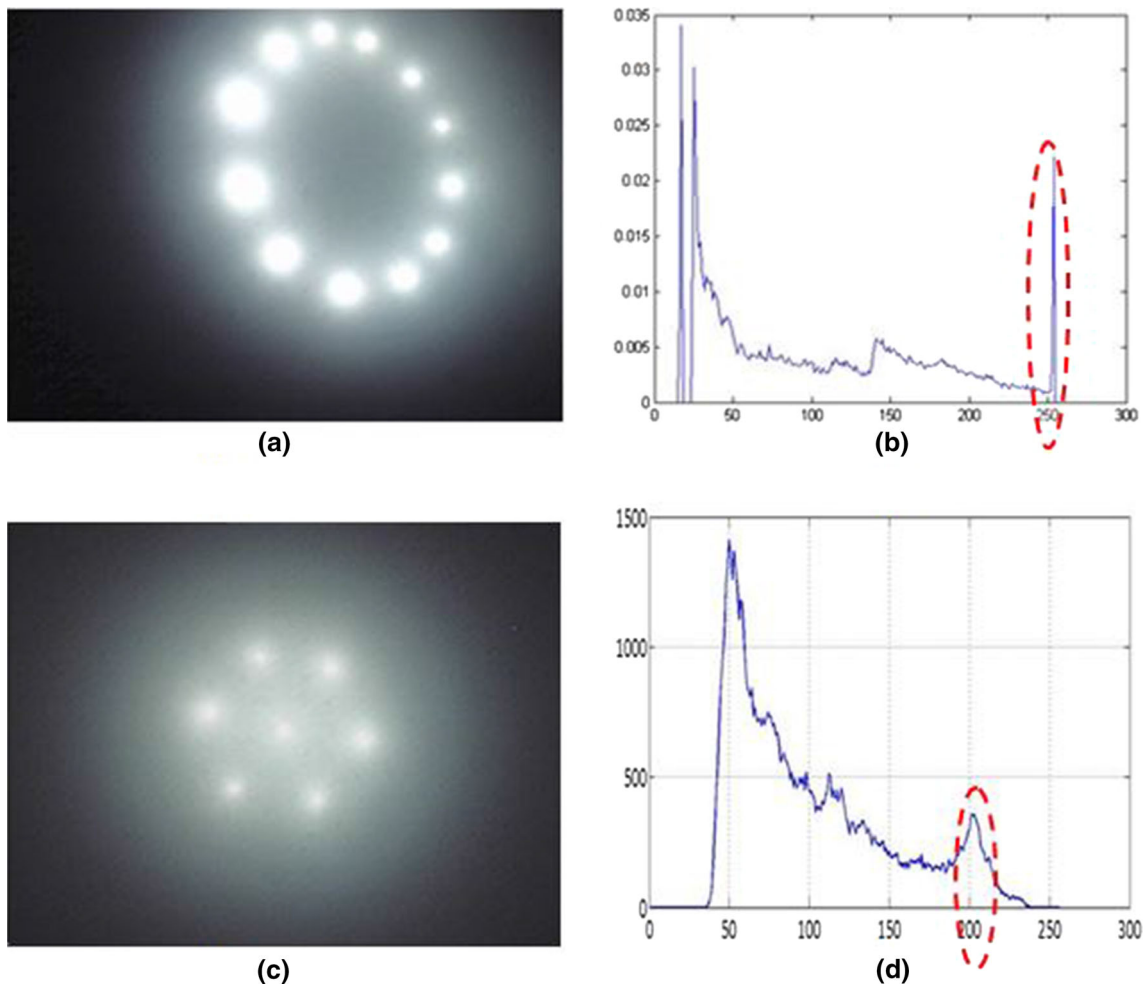


Fig. 8 a, c Experimental images; b, d corresponding histograms

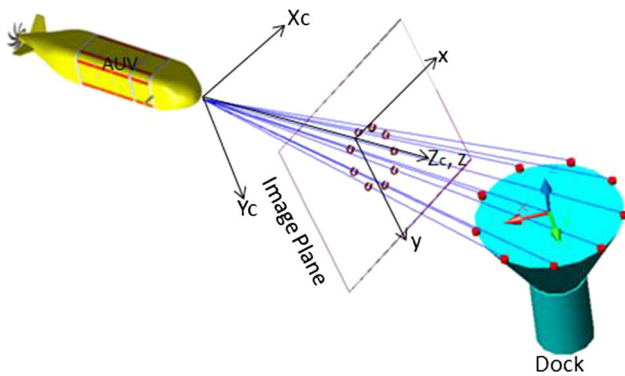


Fig. 9 A coordinate system perspective of a vision-based docking scenario

the vertex at $(0, 0, -\gamma)$, the equation of a conicoid so formed can be expressed in matrix form as:

$$[x \ y \ z] \cdot Q \cdot [x \ y \ z]^T + P \cdot [x \ y \ z]^T + \gamma^2 = 0 \quad (5)$$

where $Q = \begin{bmatrix} a'\gamma^2 & b'\gamma^2 & d'\gamma \\ b'\gamma^2 & c'\gamma^2 & e'\gamma \\ d'\gamma & e'\gamma & 1 \end{bmatrix}$ and $P = \begin{bmatrix} d'\gamma^2 & 0 & 0 \\ 0 & e'\gamma^2 & 0 \\ 0 & 0 & \gamma \end{bmatrix}$.

It may be noted that all parallel planar ‘sections’ $lx + my + nz = 0$ would feature a similar type of a conic. Hence, relative orientation of the conic (ellipse in the context) can be calculated by solving for (l, m, n) , subjected to the condition that the intersection of the conicoid with the following surface is a circle, i.e., $l^2 + m^2 + n^2 = 1$. However, in general such an attempt to estimate 3D pose from an image produces two highly nonlinear equations, solving of which necessitates adoption of suitable numerical methods.

In the present context, we solve the pose estimation problem by adopting the method proposed by Safaee-Rad et al. [18] which produce a single solution for the ‘position’ estimate and two solutions for the ‘orientation’ estimate. In other words, they provide two sets of solutions. Figure 10 depicts the necessary coordinate transformations required for estimating the 3D pose from an image, for detailed transformation procedure, we refer the reader to [18]. Also, the

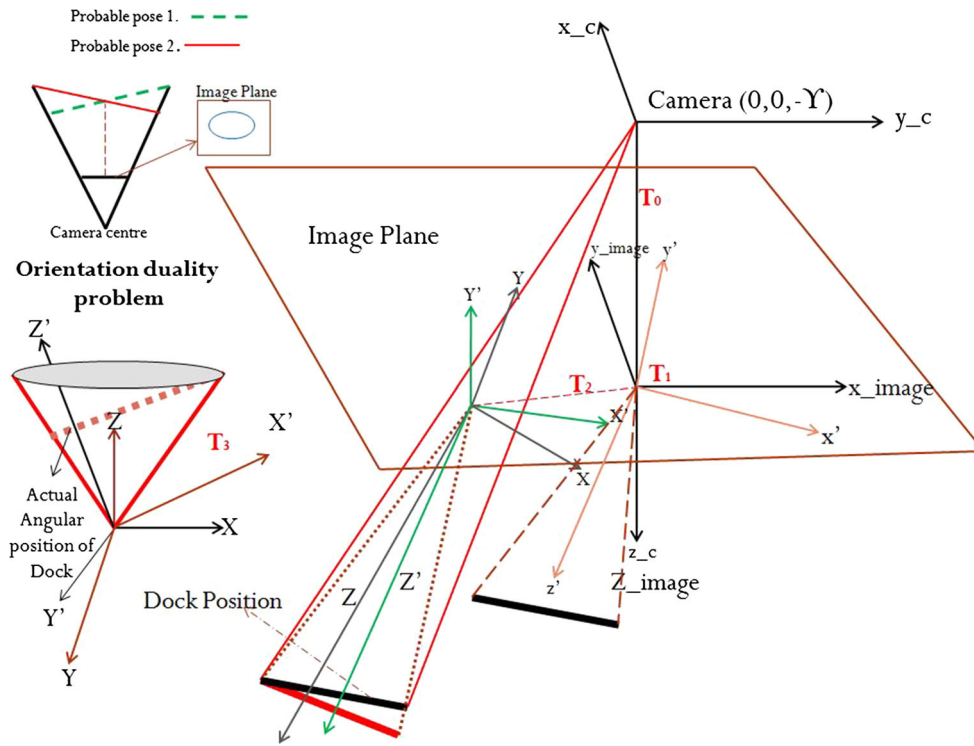


Fig. 10 Coordinate transformations for pose estimation using [18] and two probable solutions for orientation of dock (top left)

top-left portion depicts two probable solutions for orientation of the docking station and to the extent of our knowledge, no work is available to solve this dual orientation problem at least in the context of vision-based docking.

We now move to final pose estimation phase, where to identify the correct solution for orientation, we propose to use the eccentricity parameter (e) of the ellipse in the frame. We exploit the fact that the eccentricity of an ellipse monotonically decreases with decreasing view angle between the Dock and AUV. For this, we compute eccentricity of ellipse from progressively captured images and move in the direction such that eccentricity $\rightarrow 0$ and hence ellipse tends to a circle; thus allowing AUV to align with the dock and gradually enter into it. A pseudo-code for the proposed stepwise methodology to find the correct orientation solution is given in Table 1.

3 Experimental results and discussion

In this section, we present the analysis of observations from experimentation for 5-DOF pose estimation of a circular dock. In the following, we first present the experimental setup and the procedure of experimentation in brief. Section 3.1 discusses the experimental results and Sect. 3.2 analyzes other vantages of proposed approach.

A. Setup for Experimentation For experimentation, we use a Kongsberg underwater camera (Model No. OE14-110), a

Table 1 Pseudo-code to solve orientation duality problem

Step1:	Randomly choose one of the two orientation solutions ($j : 1, 2$) obtained from first frame ($i = 1$) to estimate initial pose of AUV. Let chosen solution be $S_1^1 \mapsto S_j^i$
Step2:	Compute eccentricity (e_1) of the ellipse from the first frame
Step3:	Give a nominal motion control input to AUV and acquire second frame. Once again estimate the solutions and compute the eccentricity (e_2) of the fitted ellipse from second frame
Step4:	Compare e_1 and e_2 Considering pitch (θ_x) as comparison parameter, if $e_1 < e_2$ Choose $\min(S_{2 \theta_x}^1 - S_{2 \theta_x}^2 , S_{2 \theta_x}^1 - S_{1 \theta_x}^2)$ else Choose $\min(S_{1 \theta_x}^1 - S_{2 \theta_x}^2 , S_{1 \theta_x}^1 - S_{1 \theta_x}^2)$ end as the solution for orientation of dock

Next frame onwards the correct orientation solution is identified by comparing with previous solution and minimum deviated solution is accepted

Digital Video System (DVS), a ring of 10 LEDs (dia. 30 cm) and necessary measuring instruments. A schematic of electrical connections with equipments and instruments is shown in Fig. 11. It needs to be noted that the same setup is used

Fig. 11 Schematic electrical connection and coordinate frame of experimental setup

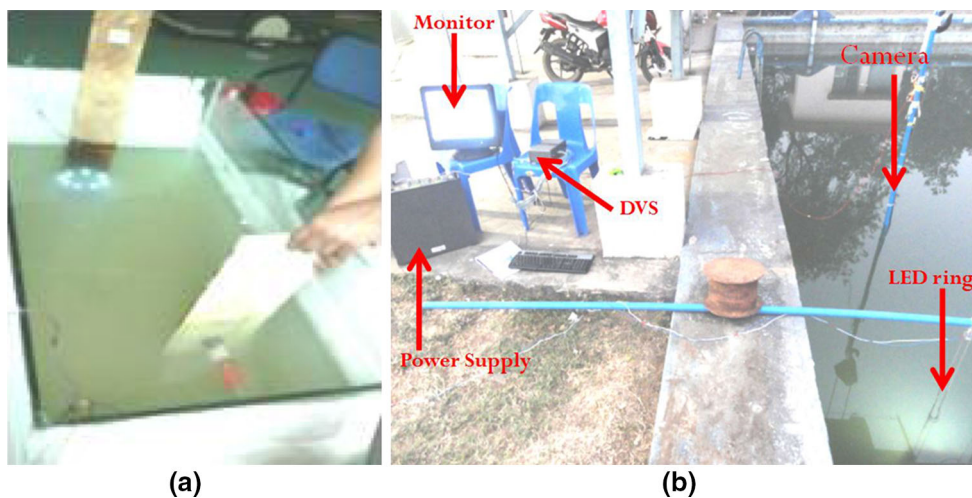
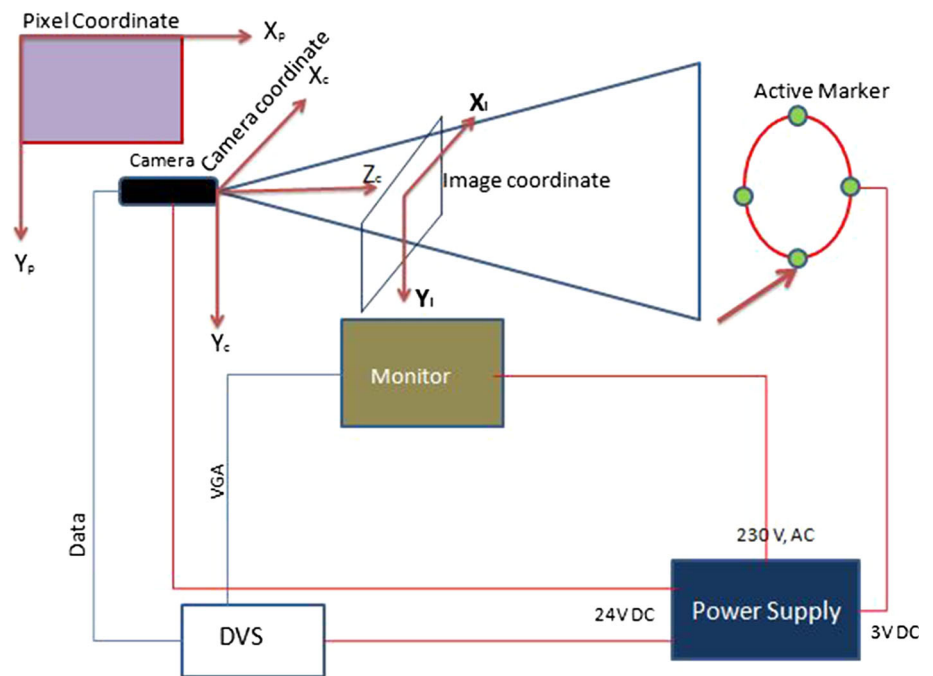


Fig. 12 Experimental setup **a** in-house tank, **b** shallow basin facility at CSIR-CMERI Durgapur

for in-house water tank as well as shallow basin experiments, both seen in Fig. 12a, b respectively.

B. Procedure for Experimentation The LED ring is placed on the wall of the tank and was allowed to move in horizontal and vertical directions as well as in angular fashion inside the tank, thus bringing 5-DOF into consideration. Also, the camera is fixed in different poses with respect to the ring to capture images. To ensure the accuracy of pose estimation, relative pose between camera and the ring is measured (using tapes/large angles) and a series of images were taken from different distances and angles. Due to constriction of space during in-house trials, each time camera's pose gets altered with respect to ring, care is taken to see the ring remained inside camera's field of view (FOV).

3.1 On the estimated pose and GUI development

One hurdle for our experiments was lack of infrastructure and hence the experiments were carried only for shorter ranges for both tank and basin. However, we make a point to mention that the proposed approach renders similar performance with larger infrastructure as well.

In our experiments, assuming that the camera is positioned at $(0, 0, 0)$, the relative 5-DOF pose of the ring is computed as (X, Y, Z, θ, ϕ) . Tables 2 and 3 tabulate experimental results for estimated pose of the dock, for images captured with the setup from in-house tank and shallow basin, respectively. First column of both the tables depicts images acquired in different poses. Columns 2–6 represent 5-DOF pose para-

Table 2 Basin experiments: comparison of estimated and measured pose (in brackets)







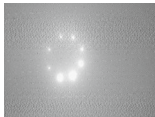
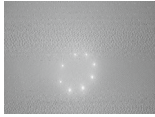

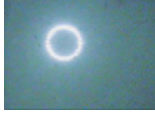
Image	X (cm)	Y (cm)	Z (cm)	θ_x (deg)	θ_y (deg)
	8.07 (8)	-15.85 (-15.5)	145.32 (145)	8 (8)	16.2 (16)
	7.96 (8)	-13.58 (-13.5)	130.81 (131)	8 (8)	18.3 (18)
	7.01 (7)	-11.59 (-11.5)	113.15 (113)	7.3 (7)	20.9 (21)
	-12.79 (-12.5)	11.52 (11.5)	157.36 (157)	29.5 (30)	29.5 (30)
	6.75 (7)	4.92 (5)	96.5 (96)	16.2 (16)	23.2 (23)

Table 3 In-house tank experiments: comparison of estimated and measured pose (in brackets)

Image	X (cm)	Y (cm)	Z (cm)	θ_x (deg)	θ_y (deg)
	-1.72 (-2)	-0.22 (0)	70.76 (70)	33.3 (32)	0 (0)
	0.03 (0)	4.17 (4)	79.4 (80)	28.2 (30)	-6 (-5)
	-5.12 (-5)	0.28 (0)	78.96 (80)	-30.8 (-30)	-18.4 (-19)
	-1.92 (-2)	1.04 (1)	79.45 (80)	34.6 (35)	27.1 (26)
	3.33 (3)	6.27 (6)	97.01 (95)	7.06 (7)	-16.2 (-15)

meters (where Z is horizontal distance between ring and the camera) obtained from experimental images using proposed algorithm. Actual values of the parameters (measured a priori) are shown just below the estimates in brackets, and it is seen that the estimated pose closely matches actual pose. It is to be noted that for in-house experiments, the maximum range of imaging is constrained by length of the tank and by near-field radiation pattern for that of basin.

To estimate the 5-DOF pose for a given image, for the sake of convenience, a MATLAB-based graphical user interface (GUI) is developed, a screenshot of which is shown in Fig. 13. The GUI takes camera parameters, dimensions of dock and an image, all as inputs, executes the sequence of operations (refer Fig. 3) and provides one solution for position, and two solutions for orientation. Taking these two sets of solutions and eccentricity of the ellipse in the frame as hypothesis,

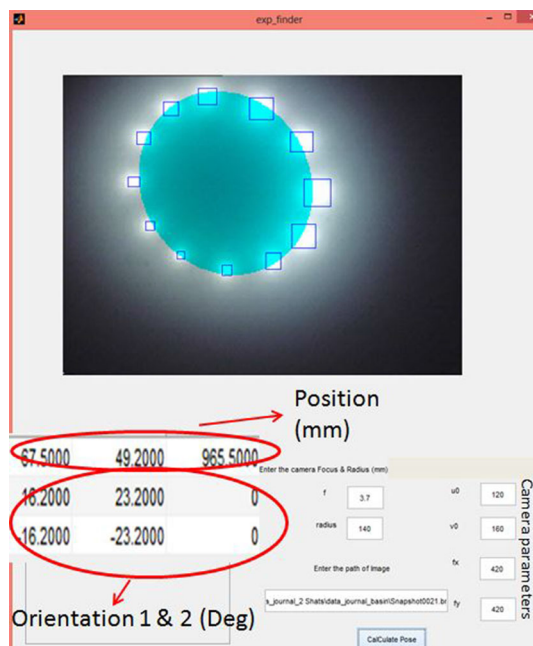


Fig. 13 GUI for image simulation and pose estimation

using procedure in Table 1 the AUV controller gradually estimates the pose, and ultimately tracks the entrance of dock.

3.2 Comparison of curvature-based 3D pose estimation with PnP method

As described in Sect. 1.1, non-iterative pose estimators are more suitable for the present docking application and hence we compare the pose estimated from curvature-based method with EPnP [15] method. Since EPnP requires point correspondences, a new colored LED marker configuration has been setup wherein the RED and BLUE markers are placed in quadrature as shown in Table 4 which also compares pose estimated via [15, 18] with the measured (true) 5D pose for two cases. It is evident from the table that the pose estimated by adopted curvature based as well as EPnP is almost close to the measured pose, the primary reason for which may be attributed to successful detection of all the light markers (in this case 12 lights).

One worth mentioning aspect with regards to colored light marker configuration is the effect of spectral absorption on threshold evaluation. It is observed that colored configuration of markers makes near exact evaluation of the threshold via HATS extremely difficult, which in turn influences the detection of correct positions of light marker. It is also observed that this problem is prominent especially when the camera is close to light markers (refer 2nd row figure in Table 4), which is due to varying intensity levels for different colored LEDs.

3.3 Other advantages of proposed approach

Before we present additional vantages of proposed approach, we note that two commonly prevailing problems in vision-based docking are, ‘spreading of features’ and ‘non-detection of all the light markers’, from the image. In this subsection, we demonstrate reliability of proposed method in these respects as compared to available methods.

Table 4 Comparison of estimated and measured pose detecting 12 light markers: HATS followed by—[18] (top), [15] (middle) and measured (in brackets)

Image	X (cm)	Y (cm)	Z (cm)	θ_x (deg)	θ_y (deg)
	-24.6	2.5	139.9	25.2	24.4
	-24.3	1.96	135.1	25.3	22.23
	(-24)	(2.5)	(137)	(25)	(24)
	-10.2	-0.65	87.8	30.4	32.9
	-9.6	-1.4	82.2	29.1	31.6
	(-10)	(-1)	(85.0)	(30)	(32)

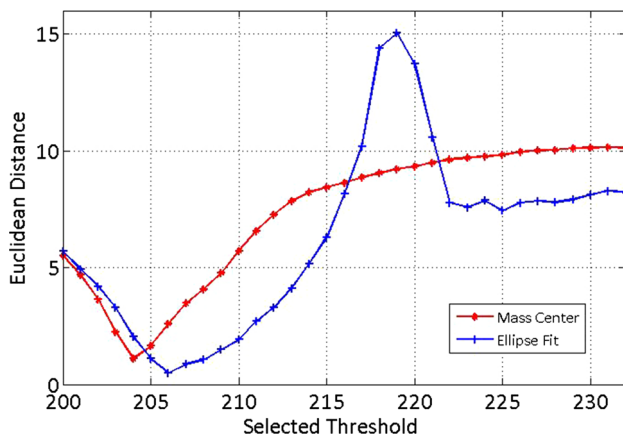


Fig. 14 Comparison of Euclidean distance between actual center and estimated center (from detected markers) with threshold

3.3.1 Feature spreading case

As mentioned earlier, if the incidence angle of light is non-zero with respect to the camera normal, the resultant effect on the image formed is known as “non-uniform spreading”. Such spreading effect renders extraction of oval-shaped feature points, (instead of circular ones). Further, this leads to improper detection of center of light and ends up in erroneous computation of the center of dock. Intuitively, the error in computing dock center would be minimal if and only if light markers are identified accurately which indicates that choice of threshold plays a vital role.

Since the objective is to find the center of dock for each frame, it is felt that the error of computed position of the dock center with that of the actual one would provide more insight on proposed HATS as well as the entire approach. To explore this aspect, we took up an exercise to evaluate the error of computing center of the ring on one of the non-uniform spread images (in Fig. 8c).

Figure 14 shows the effect of choosing an arbitrary threshold on detection of center of dock. It may be seen that in case of choosing only mass moment (red) for both individual marker detection and center of the dock, the Euclidean distance between actual center and the estimated center is minimum for a threshold of 203. However, in case of choosing mass moment method for individual markers, followed by ellipse fitting for computing the centroid of ellipse and hence the dock center, the Euclidean distance is minimum when the threshold is close to 206. Clearly, ellipse fitting procedure closely identifies the center of dock (i.e., produces relatively smaller error), when compared to solely mass moment method. Further, Fig. 8d also indicates that HATS in fact extracts 206 as the suitable threshold for the image.

In the context of feature spreading, as a further analysis, we compare performance of HATS with four versatile feature extraction methods, namely, (1) Otsu [25], (2) Sezan [26], (3)

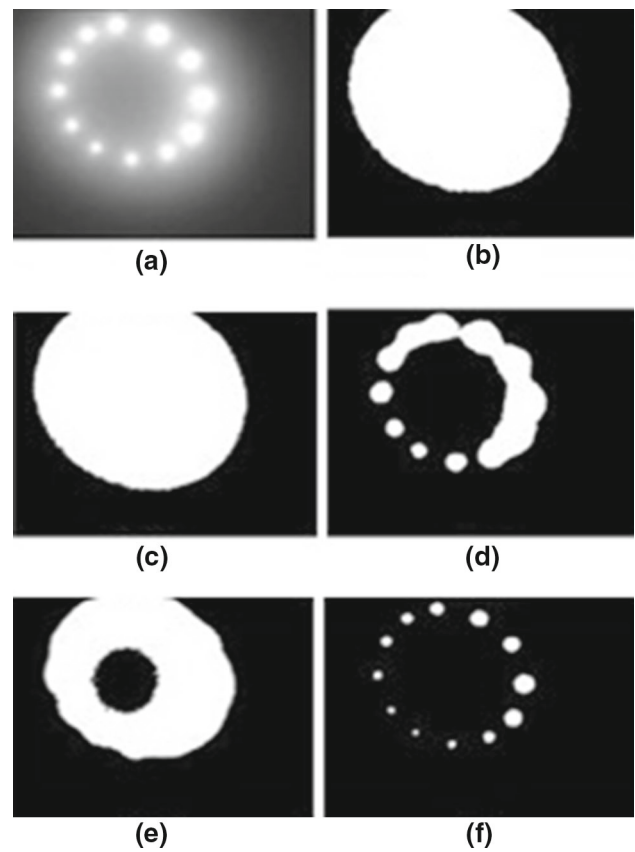


Fig. 15 a Original active image. Feature extraction by thresholding using b Otsu [25], c Sezan [26], d K-I [27], e Max-entropy [28], f proposed adaptive threshold

K-I [27] and (4) Max-entropy [28]. When these four versatile methods and HATS are applied on the LED ring image (in Fig. 15a), it is observed that HATS-based thresholding proves to be able to neatly extract all the feature points (see Fig. 15f). This once again experimentally validates the proposed HATS method.

We, therefore, conclude that HATS indeed is quite accurate in choosing a suitable threshold from active marker images and ellipse fit on the detected markers accurately extracts the center of the dock.

3.3.2 Missing feature/marker case

Coming to the problem of missing a few of the light markers, we note that such situation can arise in following cases:

- when dock is partially in cameras’ FOV and
- turbid environments.

We make a point to mention that in both these cases, the proposed method is highly effective. To address this issue, we first binarize image in Fig. 16a with a threshold (=237)

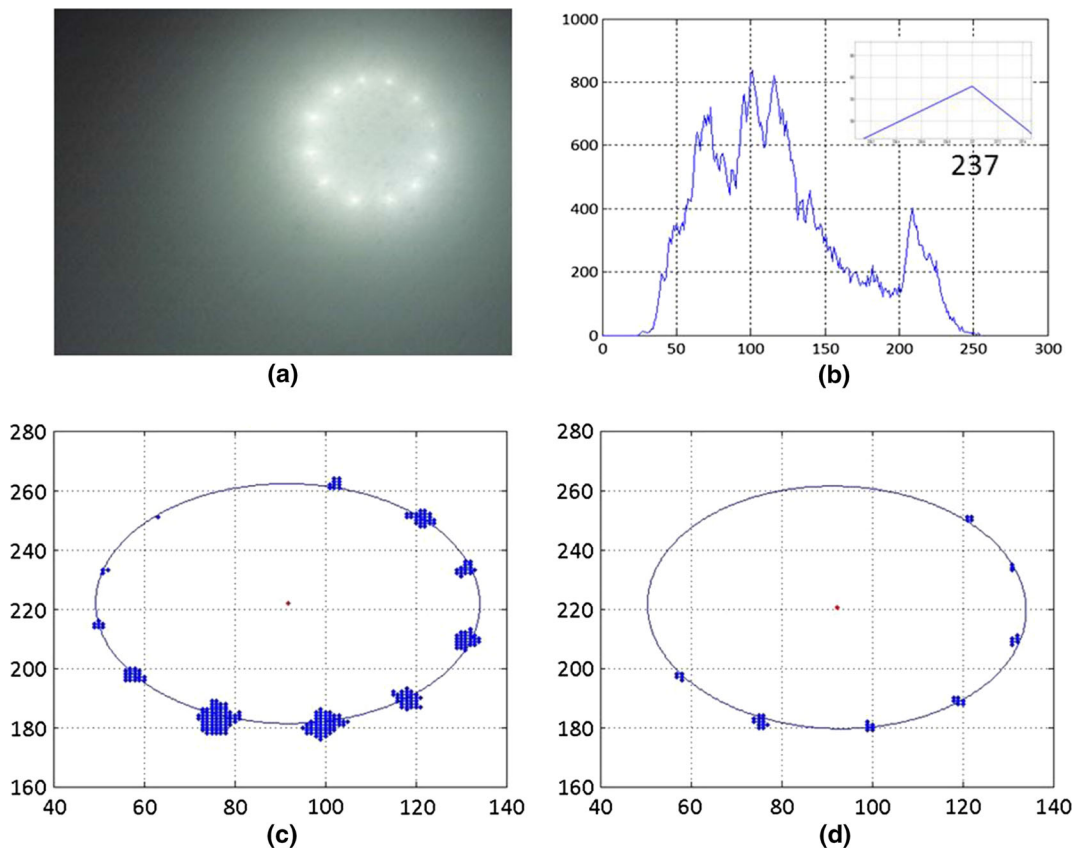


Fig. 16 Usefulness of ellipse fit procedure with HATS

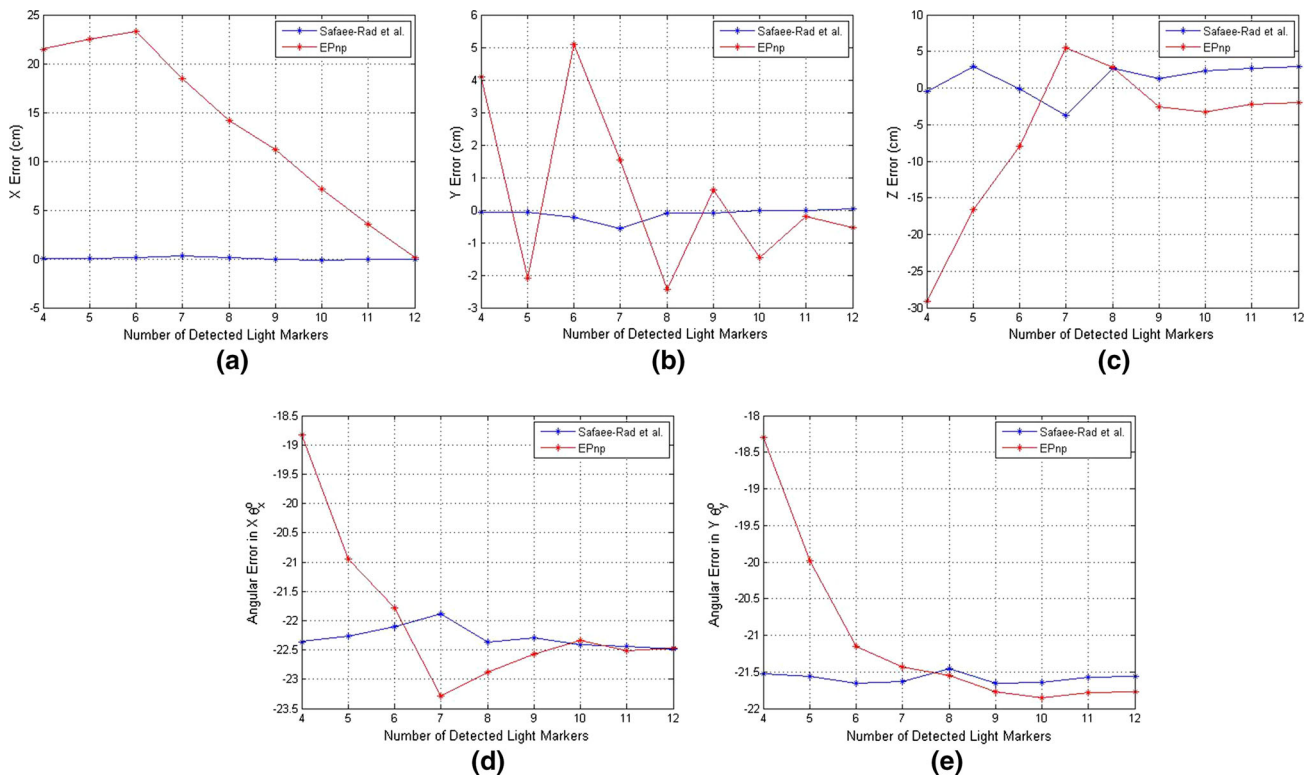


Fig. 17 Comparison of error in 5D pose computed via curvature-based method proposed by [18] and EPnP [15] methods

obtained via HATS from its histogram in Fig. 16b and observe that almost all the markers are detected (in Fig. 16c). We then fit an ellipse on the detected markers, whose estimated parameters are $\mathbf{a} = \{89.56, -198.9, 751.4, 14.3, 16.45\}$. We repeat the above exercise on the above image by choosing an arbitrarily higher threshold ($=245$) (fewer markers seen in Fig. 16d) and fit an appropriate ellipse only to observe that the parameters of the ellipse are almost the same as obtained earlier.

Further, when fewer 2D points so obtained are used for EPnP-based pose estimation, it is observed that the accuracy of the estimated pose severely degrades. On the other hand, the pose estimated by the curvature-based method relies on the fitted ellipse, and hence, the performance of the curvature-based pose estimation remains almost intact even with fewer light markers. As seen in Fig. 17, dock pose estimated by the adopted curvature-based method in [21] is highly stable even with fewer (≥ 4) detected light markers. This in turn implies that the pose to be estimated would once again be exactly the same, thus validating reliability of our method for autonomous docking of an AUV in applications employing circular entrance docks even when all the lights marker are not detected.

4 Conclusions

A method is proposed to reliable relative 5-DOF pose estimation of a circular-shaped docking station. In the process of devising the pose estimation methodology, it is observed that typically success of the vision-based docking largely depends on reliable detection of light markers in the captured image. The other issue with active marker images is the resultant non-uniform spreading. To cater to these issues, a novel method (HATS) has been developed, which, besides being scene invariant, can also reliably extract positions of light sources in the image. Another worth mentioning aspect of the present work is, though ellipse fitting for identified markers is not a newer concept, the proposed curvature-based approach to fit an ellipse upon the identified markers renders reliable estimation of pose of dock in comparison with PnP method. The point we have here is that available point-based methods fail if fewer number of markers are detected whereas our proposed method works well, as validated, at least for circle-shaped docks. Above all, simulation and experimental analysis of the entire method is indeed highly accurate, which successfully validates the proposed method.

Acknowledgments The research was carried out in the “UnWaR” project (Project No. ESC-0113) funded by CSIR, India. The authors wish to express sincere thanks to all Robotics and Automation Group members for their help and support.

References

- Bingham, D., Drake, T., Hill, A., Lott, R.: The Application of Autonomous Underwater Vehicle (AUV) Technology in the Oil Industry—Vision and Experiences. In: Proceedings of FIG XXII International Congress, pp. 1–13. Washington, D.C. USA (2002)
- Hong, Y.H., Kim, J.Y., Lee, P.M., Jeon, B.H., Oh, K.H., Oh, J.H.: Development of the homing and docking algorithm for AUV. In: Proceedings of ISOPE 2003, pp. 205–212. Honolulu, USA (2003)
- Allen, B., Austin, T., Forrester, N., Goldsborough, R., Kukulya, A., Packard, G., Purcell, M., Stokey, R.: Autonomous docking demonstrations with enhanced REMUS technology. In: MTS/IEEE Proceedings of OCEANS, pp. 1–6. Boston, Massachusetts (2006)
- Zhang, Y.W., Yang, L.P., Zhu, Y.W., Ren, X.H., Huang, H.: Self-docking capability and control strategy of electromagnetic docking technology. *J. Actaastro.* **69**(1112), 10731081 (2011)
- Feezor, M.D., Blankinship, P.R., Bellingham, J.G., Sorrell, F.Y.: Autonomous underwater vehicle homing/docking via electromagnetic guidance. In: Proceedings of OCEANS, MTS/IEEE, pp. 1137–1142 (1997)
- Deltheil, C., Didier, L., Hospital, E., Brutzman, D.P.: Simulating an optical guidance system for the recovery of an unmanned underwater vehicle. *J. Ocean. Eng.* **25**(4), 568–574 (2000)
- Negre, A., Pradalier, C., Dunbabin, M.: Robust vision-based underwater target identification and homing using self-similar landmarks. In: Laugier, C., Siegwart, R. (eds.) *Field and Service robotics*, STAR 42, pp. 51–60 (2008)
- Wirtz, M., Hildebrandt, M., Gaudig, C.: Design and test of a robust docking system for hovering AUVs. In: IEEE Proceedings of OCEANS, pp. 1–6. Hampton Roads, VA (2012)
- Maire, F.D., Prasser, D., Dunbabin, M., Dawson, M.: A vision based target detection system for docking of an autonomous underwater vehicle. In: Proceedings of Australasian Robotics and Automation. Australian Robotics and automation Association, University of Sydney, Sydney (2009)
- Maki, T., Mizushima, H., Ura, T., Sakamaki, T., Yanagisawa, M.: AUV navigation around jacket structures I: relative localization based on multi-sensor fusion. *J. Mar. Sci. Technol.* **17**, 330–339 (2012)
- Ghosh S.: Noises in underwater active and passive marker—a comparative study. In: Proceedings of IEEE International Conference Communication and Signal Processing (2014)
- Davies, E.R.: *Computer and Machine Vision: Theory, Algorithms, Practicalities*, pp. 460–461. Academic Press, Waltham (2012)
- Park, J.Y., Jun, B.H., Lee, P.M., Oh, J.: Experiments on vision guided docking of an autonomous underwater vehicle using one camera. *Ocean Eng.* **36**(1), 48–61 (2009)
- Abdel-Aziz Y.I., Karara H.M.: Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates in Close-Range Photogrammetry. In: Proceedings of ASP/UI Symp Close-Range Photogrammetry, pp. 1–18. Washington, D.C (1971)
- Lepetit, V., Moreno-Noguer, F., Fua, P.: EPnP: an accurate o(n) solution to the PnP problem. *Int. J. Comput. Vis.* **81**(2), 155–166 (2009)
- Lu, C.P., Hager, G.D., Mjolsness, E.: Fast and globally convergent pose estimation from video images. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(6), 610–622 (2000)
- Shiqi, L., Xu, C., Xie, M.: A robust o(n) solution to the perspective-n-point problem. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(7), 1444–1450 (2012)
- Safaei-Rad, R., Tchoukanov, I., Smith, K.C., Benhabib, B.: Three-dimensional location estimation of circular features for machine vision. *IEEE Trans. Robot. Autom.* **8**(5), 624–640 (1992)

19. Singh, H., Bellingham, J.G., Hover, F., Lerner, S., Moran, B.A., Heydt, K., Yoerger, D.: Docking for an autonomous ocean sampling network. *IEEE J. Ocean. Eng.* **26**(4), 498–514 (2001)
20. McLean, J.W., Voss, K.J.: Point spread function in ocean water: comparison between theory and experiment. *Appl. Opt.* **30**(15), 2027–2030 (1991)
21. Hou, W., Weidemann, A.D., Gray, D.J., Fournier, G.R.: Imagery-derived modulation transfer function and its applications for underwater imaging. In: Tescher, A.G. (eds.) *Proceedings of SPIE, Applications of Digital Image, Processing XXX*, vol. 6696 (2007)
22. Ghosh, S., Ray, R., Vadali, S.R.K., Shome, S.N.: Light-Particle interaction in underwater: a modified PSF. In: *Proceedings of IEEE International Communication and Signal Processing, India* (2014)
23. Fitzgibbon, A., Pilu, M., Fisher, R.B.: Direct least square fitting of ellipses. *IEEE Trans. Pattern Anal. Mach. Intell.* **21**(5), 476–480 (1999)
24. Halir, R., Flusser, J.: Numerically stable direct least squares fitting of ellipses. In: *Proceedings of International Conference Central Europe on Computer Graphics and Visualization*, pp. 125–132 (1998)
25. Otsu, N.: A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern. SMC* **9**(1), 62–66 (1979)
26. Sezcan, M.I.: A peak detection algorithm and its application to histogram-based image data reduction. *Comput. Vis. Graph. Image Process.* **49**(1), 36–51 (1990)
27. Kittler, J., Illingworth, J.: On threshold selection using clustering criteria. *IEEE Trans. Syst. Man Cybern. SMC* **15**(5), 652–655 (1985)
28. Kapur, J.N., Sahoo, P.K., Wong, A.K.C.: A new method for gray-level picture thresholding using the entropy of the histogram. *Comput. Vis. Graph. Image Process.* **29**, 273–285 (1985)



Shatadal Ghosh received the B.Tech in Electronics and Communication from West Bengal University of Technology, Kolkata, India and M.Tech in Mechatronics from Academy of Scientific and Innovative Research (AcSIR), New Delhi, India, in 2011 and 2014 respectively. From September 2012 to February 2015, he was a Trainee Scientist in Council of Scientific and Industrial Research, India. Currently he is with National Council of Science Museums,

India as Curator. His main areas of research interests are machine vision, under water image processing, directional audio and nonlinear acoustics.



Ranjit Ray obtained his PhD in Applied Mechanics & Aerospace Engineering from Bengal Engineering and Science University (Now IEST), Shibpur, India in 2010. He is working with Robotics & Automation Group of CSIR-CMERI, Durgapur, India as Principal Scientist. He is also a faculty of AcSIR, CSIR, India. His research interests include Mobile robot path planning, navigation, localization and machine vision.

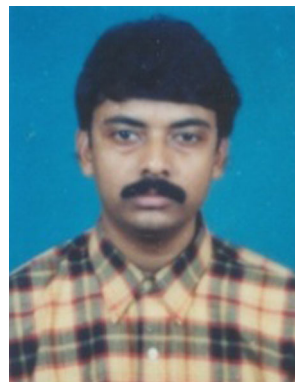


a Senior Scientist at CSIR-CMERI, Durgapur, India in field of underwater robotics and his research interests include signal processing, wireless communication and machine vision for underwater applications.



Sankar Nath Shome a 1980 Mechanical Engineering graduate from Bengal Engineering College (Now IEST), Shibpur, India received the M.Tech degree from IIT, Kanpur, India and the PhD degree in Mechanical Engineering from National Institute of Technology, Durgapur, India. He is working as a chief scientist at CSIR-CMERI, Durgapur, India. His areas of interest include robotics, mechanisms and autonomous underwater vehicle technologies.

Presently he is involved in the design and development of hybrid modular AUV with cooperative modules.



Sambhunath Nandy received B.E. (1994) and M.E. (1996) in Mechanical Engineering from University of Calcutta & Bengal Engineering College (Now IEST), Shibpur, India respectively. He obtained PhD (2014) from Indian Institute of Technology (IIT), Kharagpur in the area of robust control. He is working in the field of Robotics at CSIR-CMERI, Durgapur, India as a Principal Scientist. His research interests are mainly on navigation, guidance & robust control

of non-holonomic dynamic systems and underwater vehicles.