Underwater Simultaneous Localization and Mapping Based on Forward-looking Sonar

Tiedong Zhang[*], Wenjing Zeng and Lei Wan

National Key Laboratory of Technology of Autonomous Underwater Vehicles, Harbin Engineering University, Harbin 150001, China

Abstract: A method of underwater simultaneous localization and mapping (SLAM) based on forward-looking sonar was proposed in this paper. Positions of objects were obtained by the forward-looking sonar, and an improved association method based on an ant colony algorithm was introduced to estimate the positions. In order to improve the precision of the positions, the extended Kalman filter (EKF) was adopted. The presented algorithm was tested in a tank, and the maximum estimation error of SLAM gained was 0.25 m. The tests verify that this method can maintain better association efficiency and reduce navigation error.

Keywords: simultaneous localization and mapping (SLAM); looking forward sonar; extended Kalman filter (EKF)

Article ID: 1671-9433(2011)03-0371-06

1 Introduction

As a new type of underwater instrument, the autonomous underwater vehicle (AUV) is widely used for activities such as ocean resource research and offshore oil exploration (Anderson and Crowell, 2005; Grasmueck *et al.*, 2006). In many applications of an AUV, navigation is a critical issue (Kinsey *et al.*, 2006; Williams *et al.*, 2000). Some means of navigation, such as GPS and radio navigation, can only be used effectively when on the surface (Eustice *et al.*, 2007). Alternatively, the position of the AUV can be calculated based on DVL by the method of dead-reckoning, but the growth of positional errors associated with the method makes its use impractical for long term navigation (Whyte *et al.*, 2006). Although there is highly precise underwater navigation equipment such as inertial navigation, the cost has precluded its widespread use in AUVs until recently. Therefore, a great deal of attention is focused on the method of simultaneous localization and mapping (SLAM) (Mahon and Williams, 2004).

SLAM has been adequately studied in a land environment. Many of the systems use laser scanners, which provide precise 2-D and 3-D depth information (Bailey, 2002). In underwater environments, it is not possible to use lasers or cameras due to lighting conditions. With the development of the technology of imaging sonar, it is possible to obtain the underwater information by sonar images (Folkesson *et al.*, 2007). It is the same as optic images in nature, and the sonar image is also the energy distribution map in the imageing

space. However, because the water mediums in the underwater sound channel and the boundary (seabed, sea level) have complex and changeable characteristics, and an acoustic wave has the ability of penetration, the sonar image has low resolution, strong disturbance, and the obscure boundaries of objects. This not only increases the difficulties of interpreting the information in the images directly, but also frequently leads to more inaccurate measurements.

In this paper, an underwater SLAM system for AUVs using imaging sonar is presented according to terrestrial SLAM theory. First, image processing of single-beam sonar is carried out to obtain information of underwater objects, and extended Kalman filter (EKF) is used to estimate the positions of the vehicle and features. At last, an improved ant colony algorithm is selected to realize data association. The approach is proven to be feasible through the water tank experimental data.

2 Sonar data processing

2.1 Image formation

Single-beam sonar is selected as the image sonar in this paper. It only receives one set of data at a time, namely the echo intensity of objects. The region often comprises a series of such data, so the data should be fused effectively in order to make the acoustic image (Zhang *et al.*, 2008).

At first, considering the relationship between the objects and echo signal, a method of Lagrangian quadratic interpolation is selected to increase sampling numbers.

The data obtained is stored according to the polar coordinate system, and all pixels must be transformed into Cartesian coordinates. In Fig.1, A, B, C, D are adjacency sampled points in a polar coordinate system, and A', B', C', D'

describe the position of sampled points in a Cartesian coordinate system, so no pixel is mapped to point O, which leads to loss of information. To keep this from happening, pixel value P_o of point O is given by the method of the nearest neighborhood. Assume that pixel values of points A, B, C, D are P_A, P_B, P_C, P_D respectively, and the distances away from O are d_A, d_B, d_C, d_D. If

$$d_i = \min[d_A, d_B, d_C, d_D] \tag{1}$$

Thus the pixel value of O is defined as

$$P_o = P_i \tag{2}$$

The image comes into being as seen in Fig.2 according to Eq.(2).
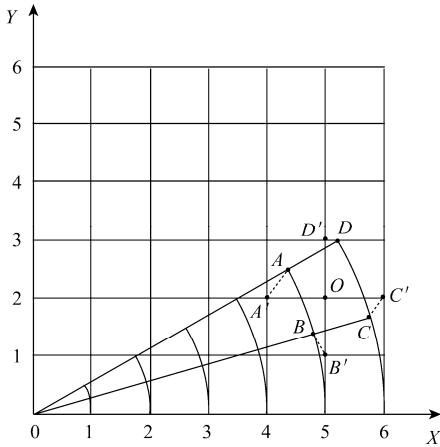

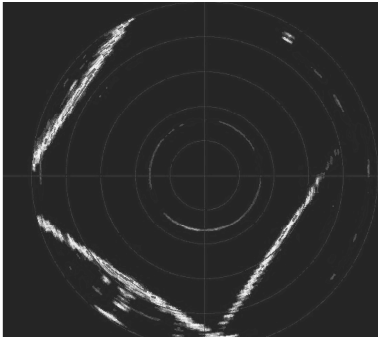**Fig.1 Transformation between polar coordinates and cartesian coordinates**

Fig.2 Looking forward sonar image

2.2 Object extraction and description
Different regions of objects can express different echo intensities in sonar images for the region composition. The method of region growing is adopted to detect the object regions in this paper. If the echo intensity of a certain object is strong, it will be more brightly reflected in the image, so the seed can be selected according to gray value (Zhang, 2004). However, the brightness of objects may be discrepant due to different distances away from sonar, so the seed region is marked off by the threshold value, which is 60% of the gray level of the highest brightness. In the meantime, the growing rule is defined as follows: if the neighbor pixel

value is larger than the threshold value T, it is in the seed region, where value T is 20% of the threshold value.

The image after segmentation should be marked and merged. The marking process is as follows:
(a) If the gray value is 1, the current scanned pixel is marked as the target pixel.
(b) If the current scanned pixel is connected with two or more target pixels, they are seen as the same one and connected together.
(c) If there is only one pixel with a gray value of 1 and the gray values of its neighborhood are all 0, then it should be marked as a new target.

Because of the material of underwater objects, region combination should be selected to reduce region segmentation error. Given that the target region R_k is divided into two individual sub-regions R_k^i and R_k^j, the process based on four connected regions is as follows:
(a) In the same column of the image, measure the shortest distance between regions R_k^i and R_k^j in a horizontal line, and define it as d_1.
(b) In the same line of the image, measure the shortest distance between regions R_k^i and R_k^j in a vertical line, and define it as d_2.
(c) If $d_1 < T_{\text{merge}}$ or $d_2 < T_{\text{merge}}$, merge two regions, and T_{merge} denotes the distance between regions.

The combination order is as follows: R_k^1 is the first subclass of R_k, so it is the basic class and is discriminated with the sequent subclasses; so the new basic class R_k^1 is obtained after combination, then $R_k = \{R_k^1\}$, and target regions $R = \{R_1, R_2, ..., R_k\}$. The extraction results are shown in Fig.3.

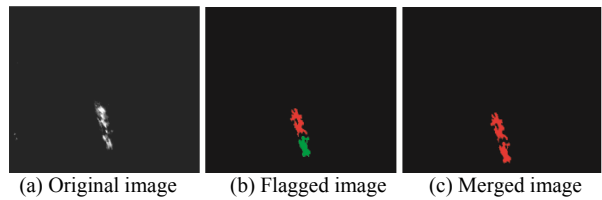(a) Original image   (b) Flagged image   (c) Merged image
Fig.3 Results of flag and merge procedure

3 Extended Kalman filter

3.1 State vector of SLAM system
The state vector of the SLAM system comprises the states of the AUV and map (Ribas, 2005). The state of the AUV is represented by position vectors in Cartesian coordinates and a heading angle, which is defined as

$$\hat{X}_v = [\hat{x}_v, \hat{y}_v, \hat{\varphi}]^T \tag{3}$$

The features gained in Section 2 are still 2-D point features, and these features are used to construct the map. Since the features are thought to be still and their positions don't change, the cross-correlation of these features in the map will be enhanced with the increasing time of observation, which makes the map more precise. The position of the *n*th feature is denoted as $(\hat{x}_n, \hat{y}_n)^T$, and the map can be represented as:

$$\hat{X}_m = [\hat{x}_1, \hat{y}_1, ..., \hat{x}_n, \hat{y}_n]^T \tag{4}$$

So the state vector of the vehicle system is defined as:

$$\hat{X}_a = \begin{bmatrix} \hat{X}_v \\ \hat{X}_m \end{bmatrix} \tag{5}$$

The covariance matrix of the vehicle system is as:

$$P_a = \begin{bmatrix} P_v & P_{vm} \\ P_{vm}^T & P_m \end{bmatrix} \tag{6}$$

3.2 Prediction stage

In Fig.4, the change of situation of the vehicle is represented by $\hat{X}_\delta$, and defined as

$$\hat{X}_\delta = [\hat{x}_\delta, \hat{y}_\delta, \hat{\varphi}_\delta]^T \tag{7}$$
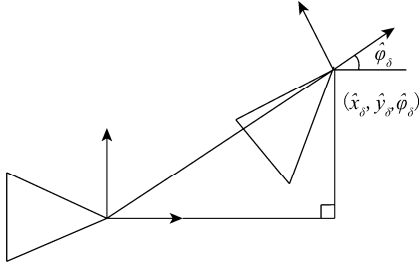

Fig.4 Attitude change of vehicle

So the prediction state of vehicle system is represented as:

$$\hat{X}_a^- = f(\hat{X}_a, \hat{X}_\delta) = \begin{bmatrix} g(\hat{X}_v, \hat{X}_\delta) \\ \hat{X}_m \end{bmatrix} = \begin{bmatrix} \hat{x}_v + \hat{x}_\delta \cos\hat{\varphi}_v - \hat{y}_\delta \sin\hat{\varphi}_v \\ \hat{y}_v + \hat{x}_\delta \sin\hat{\varphi}_v + \hat{y}_\delta \cos\hat{\varphi}_v \\ \hat{\varphi}_v + \hat{\varphi}_\delta \\ \hat{X}_m \end{bmatrix} \tag{8}$$

and the corresponding covariance is:

$$P_a^- = JP_aJ^T + QP_\delta Q^T \tag{9}$$

Where Jacobi matrix J and Q are:

$$J = \frac{\partial f}{\partial X_a}\bigg|_{(\hat{x}_a, \hat{x}_\delta)} = \begin{bmatrix} J_v & 0_{vm} \\ 0_{vm}^T & I_m \end{bmatrix}_{(\hat{x}_a, \hat{x}_\delta)} \tag{10}$$

$$Q = \frac{\partial f}{\partial X_\delta}\bigg|_{(\hat{x}_a, \hat{x}_\delta)} = \begin{bmatrix} Q_v \\ 0_{vm}^T \end{bmatrix}_{(\hat{x}_a, \hat{x}_\delta)} \tag{11}$$

Jacobi matrix J_v and Q_v are:

$$J_v = \frac{\partial g}{\partial X_v}\bigg|_{(\hat{X}_v, \hat{X}_\delta)} = \begin{bmatrix} 1 & 0 & -\hat{x}_v \sin\hat{\varphi}_v - \hat{y}_\delta \cos\hat{\varphi}_v \\ 0 & 1 & -\hat{x}_\delta \cos\hat{\varphi}_v - \hat{y}_\delta \sin\hat{\varphi}_v \\ 0 & 1 & 1 \end{bmatrix} \tag{12}$$

$$Q_v = \frac{\partial g}{\partial X_\delta}\bigg|_{(\hat{x}_v, \hat{x}_\delta)} = \begin{bmatrix} \cos\hat{\varphi}_v & -\sin\hat{\varphi}_v & 0 \\ \sin\hat{\varphi}_v & \cos\hat{\varphi}_v & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{13}$$

3.3 Update stage

Suppose that a feature detected by sonar is already in the map, represented as $(\hat{x}_i, \hat{y}_i)$; then the value of position should be transformed from the global coordinate to the vehicle coordinate (Tardos *et al*, 2002). The transformation is as follows:

$$\hat{Z} = h(\hat{X}_a^-) = \begin{bmatrix} \sqrt{(\hat{x}_i - \hat{x}_v)^2 + (\hat{y}_i - \hat{y}_v)^2} \\ \arctan\left(\dfrac{\hat{y}_i - \hat{y}_v}{\hat{x}_i - \hat{x}_v}\right) - \hat{\varphi}_v \end{bmatrix} \tag{14}$$

The observed value is

$$Z = \begin{bmatrix} r \\ \theta \end{bmatrix} \tag{15}$$

$$R = \begin{bmatrix} \sigma_r^2 & \sigma_{r\theta}^2 \\ \sigma_{r\theta}^2 & \sigma_\theta^2 \end{bmatrix} \tag{16}$$

where (r, θ) denote the distance and the orientation of the feature with respect to the vehicle, R denotes measurement covariance.

If the estimated value $\hat{Z}$ is associated with the observed value Z correctly, the results of SLAM after the update are defined as

$$\hat{X}_a^+ = \hat{X}_a^- + Wv \tag{17}$$

$$P_a^+ = P_a^- - WSW^T \tag{18}$$

where W, S, v are defined as

$$v = Z - h(\hat{X}_a^-) \tag{19}$$

$$S = HP_a^- H^T + R \tag{20}$$

$$W = P_a^- H^T S^{-1} \tag{21}$$

In Eq.(20), H is as follows:

$$H = \begin{bmatrix} -\dfrac{\Delta x}{d} & -\dfrac{\Delta y}{d} & 0 & 0 \cdots 0 & \dfrac{\Delta x}{d} & \dfrac{\Delta y}{d} & 0 \cdots 0 \\ -\dfrac{\Delta y}{d^2} & -\dfrac{\Delta x}{d^2} & -1 & 0 \cdots 0 & -\dfrac{\Delta y}{d^2} & -\dfrac{\Delta x}{d^2} & 0 \cdots 0 \end{bmatrix} \tag{22}$$

where $\Delta x, \Delta y, d$ are defined as

$$\Delta x = \hat{x}_i - \hat{x}_v$$

$$\Delta y = \hat{y}_i - \hat{y}_v$$

$$d = \sqrt{(\hat{x}_i - \hat{x}_v)^2 + (\hat{y}_i - \hat{y}_v)^2}$$

3.4 Extend stage

The state vector and covariance are augmented by the new measurement Z_{new} and its covariance R_{new}, and are as follows:

$$\hat{X}_{\text{aug}} = \begin{bmatrix} \hat{X}_a \\ Z_{\text{new}} \end{bmatrix} \tag{23}$$

$$P_{\text{aug}} = \begin{bmatrix} P_v & P_{vm} & 0 \\ P_{vm}^{\text{T}} & P_m & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{24}$$

The augmentation is initialized by the linear transformation f_i, which is defined as

$$\hat{X}_a^+ = f_i\left(\hat{X}_{\text{arg}}\right) = \begin{bmatrix} \hat{X}_a \\ g_i\left(\hat{X}_v, Z\right) \end{bmatrix} \tag{25}$$

$$P_a^+ = \nabla f_{x_{\text{aug}}} P_{\text{aug}} \nabla f_{x_{\text{aug}}}^{\text{T}} \tag{26}$$

where $\nabla f_{x_{\text{aug}}}$ is the Jacobi matrix and function $g_i(\hat{X}_v, Z)$ is defined as

$$g_i\left(x_v, z\right) = \begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} x_v + r\cos\left(\theta + \varphi_v\right) \\ y_v + r\sin\left(\theta + \varphi_v\right) \end{bmatrix} \tag{27}$$

For the sparse Jacobi matrix, Eq.(26) is turned into

$$P_a^+ = \begin{bmatrix} P_v & P_{vm} & P_v \nabla g_{x_v}^{\text{T}} \\ P_{vm}^{\text{T}} & P_m & P_{vm}^{\text{T}} \nabla g_{x_v}^{\text{T}} \\ \nabla g_{x_v} P_v & \nabla g_{x_v} P_{vm} & \nabla g_{x_v} P_v \nabla g_{x_v}^{\text{T}} + \nabla g_z R \nabla g_z^{\text{T}} \end{bmatrix} \tag{28}$$

4 Data association method

To produce the update, an association hypothesis is needed (Cooper *et al.*, 2005; Gil *et al.*, 2006). For this purpose, the maximum probability association is applied, and an ant colony algorithm is used to improve the association ability. The detailed steps are implemented as follows:

1) Abandon the associations that don't satisfy the individual compatibility rule, which is as follows:

$$M_{ij} = v_{ij}^{\text{T}} S_j^{-1} v_{ij} \tag{29}$$

$$M_{ij} < \gamma^n \tag{30}$$

where M_{ij} denotes the Mahananobis distance between measurement i and feature j, the quantity γ^n is determined from two parameters, i.e. the dimension of the innovation vector n and a free parameter which defines the expected percentage of correct associations that will be excepted.

2) When initializing the pheromone if the association of measurement and feature satisfies the individual compatibility rule, the initial pheromone τ_0 and heuristic information η_{ij} are defined as

$$\tau_{ij} = f_{ij} \tag{31}$$

$$\eta_{ij} = 1 / M_{ij} \tag{32}$$

where f_{ij} denotes the dependence relationship between the estimated value $\hat{Z}$ and the observed value.

3) Randomly place the kth ant on the position of one arbitrarily observed value, and it is defined as the initial position. Select the associated feature based on the probability decision rule as:

$$p_{ij}^k = \begin{cases} \dfrac{\tau_{ij}^{\alpha} \cdot \eta_{ij}^{\beta}}{\sum\limits_{\gamma \notin \text{tab}u_k} \tau_{ir}^{\alpha} \cdot \eta_{ir}^{\beta}} & \text{if } j \notin \text{tab}u_k \\ 0 & \text{otherwise} \end{cases} \tag{33}$$

where α denotes the importance degree of pheromone information, and β denotes the importance degree of heuristics information. The memory $\text{tab}u_k$ is empty at first.

4) Put the features that have been chosen in $\text{tab}u_k$, and the ant has to choose other features for the rest of the measurements until there are no redundant measurements.

5) Clear $\text{tab}u_k$ after the ant has completed all measurements and feature pairs. Pheromone $\Delta\tau_{ij}^k$ will be released on the successfully associated pairs (i, j) as:

$$\Delta\tau_{ij}^k = \begin{cases} \dfrac{Q}{d_k} & \text{if the pair}(i, j) \text{ is on the route of } k\text{th ant} \\ 0 & \text{otherwise} \end{cases} \tag{34}$$

where Q is the reconcilability coefficient, and d_k is the route length of the kth ant.

6) Repeat steps (3)–(5) for another ant to choose its route.

7) After all ants have finished their routes in an iterative, pick out the best solution which has the minimum route length and update the pheromone globally as:

$$\tau_{ij} = (1 - \rho)\tau_{ij} + \sum_{k=1}^{t} \Delta\tau_{ij}^k \tag{35}$$

where ρ denotes the pheromone volatility coefficient, and t is the number of ants.

8) Keep selecting the best solution in the results of all iteratives and select it for the last solution.

5 Experimental results

Experiments were done to verify the method further in the water tank of the National Defense Key Laboratory of Autonomous Underwater Vehicle Technology (See Fig.5); the parameters of the water tank are 50 m(length)×30 m (width)×10 m (height). An imaging sonar was selected, which was mounted on a certain AUV (See Fig.6). The sensor was designed to be applied in underwater tasks such as obstacle avoidance and target tracking. It can perform scans in a 2-D plane by rotating a fan-shaped sonar beam of 3° horizontal beam width and 40° vertical beam width. During the experiment, the sensor was assigned to work within a range of 10 meters. Its scanning rate was set to the mode in which the sonar scan was fastest.

The objects included a reflector, triangle body, and metallic ball (See Fig.7). In experiments, the AUV only showed approximate linear motion due to the scale of the tank, and the position of the AUV was estimated according to the position of objects. The sonar image was processed as seen in Fig.8. The localization of the vehicle is seen in Fig.9 and Tables 1–2.
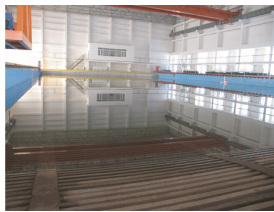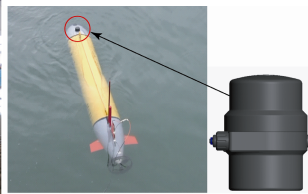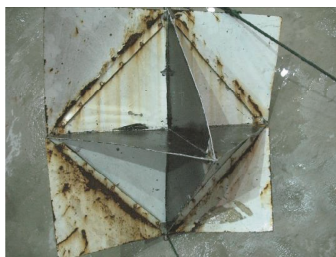

Fig.5 Water tank   **Fig.6 AUV and imaging sonar**

(a) Object

(b) Experiment scene
Fig.7 Object and experiment scene

As seen in Fig.12, under the same noise conditions, the position of the AUV can be gotten by dead-reckoning at first, but the error will increase, and the dead reckoning trajectory has a significant drift. The proposed method of SLAM is not very steady at first, but the error converges gradually with

repeated observation of the same feature. The same trend can be concluded from Table 1 and Table 2. The values on the *x*-axis of the two curves are similar. The maximum estimation error of dead reckoning on the *y*-axis is 0.53 m, and the trend is divergent. Its value will be larger in a longer time period. The maximum estimation error of SLAM is 0.25 m, and the trend is convergent. Its value will be smaller in a longer time period. This shows that the proposed navigation method has better precision than dead-reckoning.
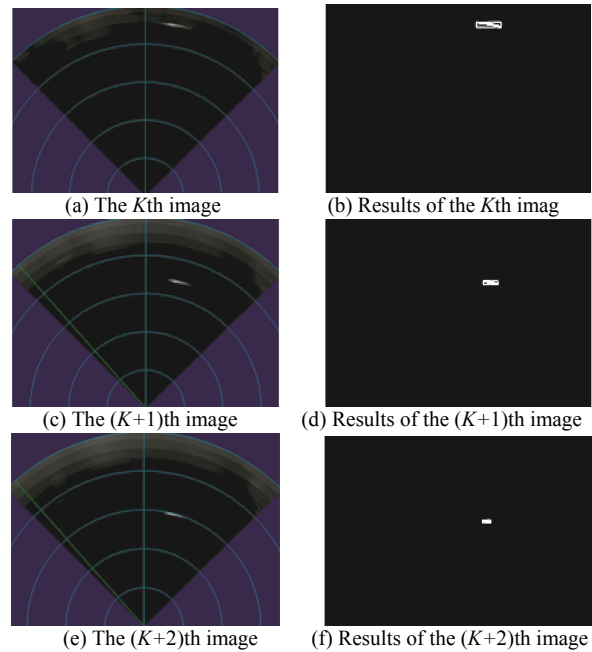

(a) The *K*th image   (b) Results of the *K*th imag

(c) The (*K*+1)th image   (d) Results of the (*K*+1)th image

(e) The (*K*+2)th image   (f) Results of the (*K*+2)th image
Fig.8 The sequential sonar images and processing results
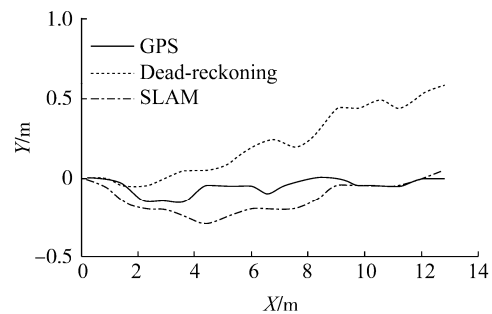

Fig.9 The trajectory comparison of dead-reckoning and SLAM

Table 1 Vehicle position estimation of SLAM

The *i*th image	Estimated result/ (m, rad)	The *i*th image	Estimated result/ (m, rad)
1	(0,0)	2	(0.78, −0.01)
3	(1.5,−0.09)	4	(2.27, −0.14)
5	(2.98, −0.17)	6	(3.71, −0.2)
7	(4.45, −0.2)	8	(5.27, −0.21)
9	(6.03, −0.17)	10	(6.85, −0.22)
11	(7.61, −0.25)	12	(8.35, −0.19)
13	(9.07, −0.09)	14	(9.79, −0.09)
15	(10.53, −0.07)	16	(11.28, −0.05)
17	(12.1, −0.008)	18	(12.77,0.04)

Table 2 Vehicle position estimation of dead-reckoning

| The ith image	Estimated result/ (m, rad)	The ith image	Estimated result/ (m, rad)
1	(0,0)	2	(0.77, −0.001)
3	(1.5, −0.047)	4	(2.26, −0.07)
5	(2.97, −0.04)	6	(3.7, −0.04)
7	(4.43,0.057)	8	(5.23,0.081)
9	(6,0.2)	10	(6.81,0.21)
11	(7.58,0.22)	12	(8.34,0.3)
13	(9.03,0.41)	14	(9.76,0.42)
15	(10.5,0.45)	16	(11.27,0.48)
17	(12.04,0.52)	18	(12.78,0.53)

6 Conclusions

Most of the previous civil work in underwater SLAM has focused on simulation. The main contribution of this work consists of carrying out underwater SLAM using sonar imagery in a water tank environment. A method to process the sonar images was proposed, and an improved association method based on an ant colony algorithm was introduced to improve the association efficiency. The results of the tank test show the superiority of the presented method. In the experiment, the AUV only moves along the line trajectory due to the limitations of the water tank, so more experiments should be carried out in river or sea environments in the future based on the above studies.

References

Anderson B, Crowell J (2005). A cost-sensible new autonomous underwater vehicle for surveys/soundings, search & rescue, and research. *IEEE/MTS OCEANS Conference and Exhibition*, Washington DC, 1228-1233.

Bailey T (2002). *Mobile robot localization and mapping in extensive outdoor environments*. PhD thesis, University of Sydney, Australian Centre for Field Robotics, 45-60.

Cooper AJ (2005). *A comparison of data association techniques for simultaneous localization and mapping*. PhD thesis, Aerospace Engineering and Mechanics, University of Minnesota, 116-128.

Eustice RM, Whitcomb LL, Singh H, Grund M (2007). Experimental results in synchronous-clock one-way-travel-time acoustic navigation for autonomous underwater vehicles. *IEEE International Conference on Robotics and Automation*, Rome, Italy, 4257-4264.

Folkesson J, Leonard J, Leederkerken J, Williams R (2007). Feature tracking for underwater navigation using sonar. *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, San Diego, 3678-3684.

Gil A, Stachniss C, Burgard W (2006). Improving data association in vision-based SLAM. *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Beijing, 2076-2081.

Grasmueck M, Eberli GP, Viggiano DA (2006). Autonomous underwater vehicle (AUV) mapping reveals coral mound distribution, morphology, and oceanography in deep water of the Straits of Florida. *Geophysical Research Letters*, **33**(23), 23616-23622

Kinsey JC, Eustice RM, Whitcomb LL (2006). Underwater vehicle navigation: Recent advances and new challenges. *IFAC Conference on Maneuvering and Control of Marine Craft*, Lisbon, Portugal, 435-445.

Mahon I, Williams S (2004). Slam using natural features in an underwater environment. *IEEE Control, Automation, Robotics and Vision Conference*, Kunming, China, 2076-2081.

Ribas D (2005). *Towards simultaneous localization & mapping for an AUV using an imaging sonar*. PhD thesis, University de Girona, 60-72.

Tardos JD, Newman J, Leonard J (2002). Robust mapping and localization in indoor environments using sonar data. *The Int. Journal of Robotics Research*, **21**(4), 311-330.

Whyte HD, Bailey T (2006). Simultaneous localization and mapping: Part I. *Robotics & Automation Magazine*, **13**(2), 99-108.

Williams SB, Newman P, Dissanayake G, Whyte HD (2000). Autonomous underwater simultaneous localization and map building. *IEEE International Conference on Robotics & Automation*, San Francisco, 1793-1798.

Zhang Tiedong (2004). *The post-processing of the image for the forward looking sonar*. Master thesis, Harbin Engineerng University, 90-103.

Zhang Tiedong, Wan Lei, Xu Yuru, Ma Yue (2008). A preprocess method of the looking forward sonar image. *Acoustics and Electronics Engineering*, **91**(3), 14-18.

Tiedong Zhang was born in 1978. He has a PhD and is an instructor at Harbin Engineering University. His current research interests include AUV technique and underwater image processing.
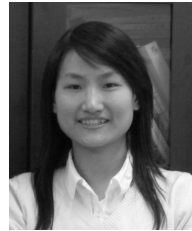
Wenjing Zeng was born in 1985. She is a PhD candidate at Harbin Engineering University. Her current research interests include AUV technique and underwater image processing.

Lei Wan was born in 1965. He is a professor of Harbin Engineering University. His current research interests include AUV technique.