2016 IEEE International Symposium on
Safety, Security, and Rescue Robotics (SSRR)
EPFL, Lausanne, Switzerland, October 23-27, 2016

Sound-based Online Localization for an In-pipe Snake Robot

Yoshiaki Bando[1], Hiroki Suhara[2], Motoyasu Tanaka[3], Tetsushi Kamegawa[2],
Katsutoshi Itoyama[1], Kazuyoshi Yoshii[1], Fumitoshi Matsuno[4], Hiroshi G. Okuno[5]

***Abstract*— This paper presents a sound-based online localization method for an in-pipe snake robot with an inertial measurement unit (IMU). In-pipe robots, in particular, snake robots need online localization for autonomous inspection and for remote operator supports. The GPS is denied in a pipeline, and conventional odometry-based localization may deteriorate due to slippage and sudden unintended movements. By putting a microphone on the robot and a loudspeaker at the entrance of the pipeline, their distance can be estimated by measuring the time of flight (ToF) of a reference sound emitted from the loudspeaker. Since the sound propagation path in the pipeline is necessary for estimating the robot location, the proposed sound-based online localization method simultaneously estimates the robot location and the pipeline map by combining the distance obtained by the ToF and orientation estimated by the IMU. The experimental results showed that the error of the distance estimation was less than 7% and the accuracy of the pipeline map was more than 68.0%.**

I. Introduction

While most in-pipe robots for inspecting pipelines in industrial facilities such as power plants and gas or water supply [1]–[5] have been targeted to specific pipes, snake robots are rather general purpose in the sense that they can be deployed to different kinds of pipelines by twisting their high-degree-of-freedom (DOF) bodies [1]–[3] (*e.g.* Fig. 1). Since the maps of pipelines in old facilities may be either lost or obsolete, these robots are often forced to work in unknown pipeline environments.

Online localization of an in-pipe robot is essential for autonomous inspection and for remote operator supports. Conventional localization methods use encoders, visual sensors, and inertial sensors [6]–[11]. For example, visual odometry localizes an in-pipe robot by estimating its movements from time series of video camera images [9]. Such integral-type localization, however, incurs an accumulated error problem and becomes inaccurate when the robot suddenly slips or is unintentionally moved.

Pipeline environments disturb sensor systems that can determine the absolute location outdoors and in most indoor environments. The GPS or magnetometers cannot be used in pipelines that disturb radio waves and magnetic fields [12]; laser range finders only provide local information around the robot in a winding pipeline. Although several methods use the length of the power cable between the robot and the entrance of the pipeline [10], [13], the cable length is unreliable in a large-diameter pipeline because the cable curves or coils.

(a) 2-m in-pipe snake robot (b) Snake robot climbing a vertical pipeline

Fig. 1. In-pipe snake robot used in this paper.

Sound-based distance estimation can measure the shortest length along the pipeline between the robot and the entrance of the pipeline. Putting a microphone on the robot and a loudspeaker at the entrance enables their distance to be estimated by measuring the time of flight (ToF) of a reference sound emitted from the loudspeaker [14], [15]. Although reflection causes sounds that have longer paths than the pipeline, the estimation accuracy can be kept by detecting a sound that propagates through the shortest path. Since the ToF of a sound is affected mainly by the contents and surface of the pipeline, it works in GPS-denied environments. Moreover, the distance from the entrance can be measured even when the pipeline is winding because the sound propagates with diffraction.

This paper presents a sound-based localization method that estimates the 3D robot location and pipeline map. Since the ToF estimation provides only the distance from the entrance, the sound propagation path in the pipeline is necessary for estimating the robot location. The proposed method simultaneously estimates the robot location and pipeline map by combining the distance with the robot orientation estimated from inertial measurement unit (IMU) observations. To achieve this, we formulate a nonlinear state-space model that represents the relationships among the observations, robot location, and pipeline map. The pipeline map is represented as an inner space of the pipeline. The robot location is estimated by using an extended Kalman filter in an online manner [16], and the pipeline map is estimated from the past locus of the robot location after

Demo page: http://sap.ist.i.kyoto-u.ac.jp/members/yoshiaki/demo/ssrr2016

[1]Graduate School of Informatics, Kyoto University, Kyoto, Japan {yoshiaki, itoyama, yoshii}@kuis.kyoto-u.ac.jp
[2]Graduate School of Natural Science and Technology, Okayama University, Okayama, Japan suhara.h.mif@s.okayama-u.ac.jp, kamegawa@okayama-u.ac.jp
[3]Graduate School of Information Science and Engineering, The University of Electro-Communications, Tokyo, Japan mtanaka@uec.ac.jp
[4]Graduate School of Engineering, Kyoto University, Kyoto, Japan matsuno@me.kyoto-u.ac.jp
[5]Graduate Program for Embodiment Informatics, Waseda University, Tokyo, Japan okuno@aoni.waseda.jp

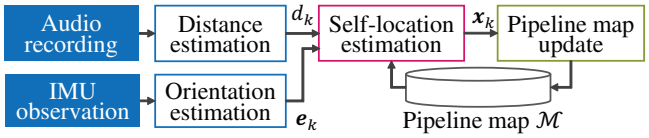
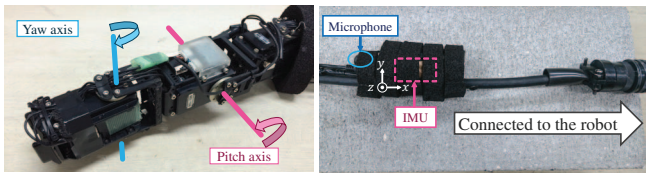
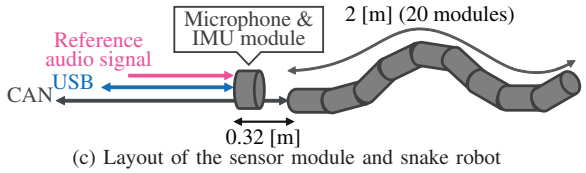


Fig. 2. Overview of the sound-based localization.



(a) Connection of three modules (b) Localization sensor module



(c) Layout of the sensor module and snake robot

Fig. 3. Design of the in-pipe snake robot used in this paper.

each update. The effectiveness of the proposed sound-based localization method was evaluated using a prototype in-pipe snake robot (Fig. 1).

II. Related Work

This section overviews conventional in-pipe localization methods and indoor sound-based localization methods.

A. In-pipe Localization Methods

To cope with GPS-denied pipeline environments, several localization methods have been developed by using visual sensors [9], [17]. Hansen et al. [9], for example, used a fisheye camera to perform simultaneous localization and mapping (SLAM). Since the visual odometry in a pipeline is degraded by limited image textures of the pipeline wall and visual aliasing, this method suppresses odometry errors by introducing a pipeline-shape constraint that assumes the model that a pipeline is composed of straight sections and T-intersections.

To correct accumulated errors without using the pipeline-shape constraints, landmark extraction algorithms are tuned to in-pipe SLAM [7], [11], [18]. When the same landmark is reobserved, the robot is assumed to revisit the same location where the landmark is firstly observed [16]. Krys et al. [7] extracted visual landmarks on the wall of a pipeline to correct accumulated errors of the inertial navigation system. Ma et al. [18] reported an algorithm that detects unique void patterns in or beyond the wall of a pipeline by using ultrasonic echoes. The in-pipe robots used in the above mentioned methods were designed to fit target pipes, and thus slippage or sudden fall were rather rare. Therefore, continuous textures are available. On the other hand, an in-pipe snake robot often encounters slippage and sudden unintended movements such as falling.

B. Indoor Sound-based Localization Methods

Sound-based localization can estimate the distance between a loudspeaker and a microphone immediately without accumulated errors [19]–[21]. The most common features are the time of flight (ToF) and time difference of arrivals (TDoAs) of a reference signal among the microphones. Bando et al. [21] attached loudspeakers, microphones, and IMU on a flexible hose-shaped rescue robot and estimated its 3D shape in a messy damaged environment with TDoAs obtained by recording a special sound from one of the loudspeakers.

TABLE I

Meanings of Mathematical Symbols

Symbol	Meaning
C	Speed of sound in a pipeline (set to 340 m/s in this paper)
k	Measurement index
t	Time
$d_k \in \mathbb{R}$	Distance between the microphone and loudspeaker
$\boldsymbol{e}_k \in \mathbb{R}^3$	Unit vector representing the robot orientation
$\boldsymbol{x}_k \in \mathbb{R}^3$	Location of the robot
$v_k \in \mathbb{R}$	Velocity of the robot
$\boldsymbol{z}_k = [\boldsymbol{x}_k^T, v_k]^T$	Robot state
$\mathcal{M} \subset \mathbb{R}^3$	Pipeline map
$\mathcal{S}(\boldsymbol{p}, r) \subset \mathbb{R}^3$	Sphere ($\mathcal{S}(\boldsymbol{p}, r) = \{\boldsymbol{x} \in \mathbb{R}^3 \mid\mid \boldsymbol{x} - \boldsymbol{p}_i \mid < r_i\}$)

III. Online Sound-based Localization for an In-pipe Snake Robot

The proposed sound-based method estimates the robot location and pipeline map by combining the ToF-based distance estimation and IMU-based orientation estimation as shown in Fig. 2. The robot location is estimated by using an extended Kalman filter [16] and the pipeline map is estimated from the past locus of the robot location in an online manner.

A. Design and Implementation of In-pipe Snake Robot

Fig. 1 shows the in-pipe snake robot used in this paper. The robot body consists of 20 modules of 120 mm in diameter, serially connected with a gap of 95 mm between adjacent modules. Each module has two half joints in yaw and pitch axes, connecting to the previous and next modules, respectively (Fig. 3-(a)). It is covered with foam rubber tape increasing friction with the wall of a pipeline (Fig. 1-(a)). The robot moves forward by making a helical rolling motion [3].

As shown in Figs. 3-(b) and (c), a sensor module with a microphone and IMU is attached to the tail cable of the robot. The microphone on the module is connected to a synchronized stereo A/D converter. The other input of the A/D converter is connected to an audio cable that indicates the onset time of the reference signal emitted from the loudspeaker at the entrance of the pipeline. The IMU sensor has a three-axis gyroscope and a three-axis accelerometer with the axis indicated in Fig. 3-(b).

B. Problem Statement

The sound-based localization problem in this paper is defined as follows:

Input: 1) The distance between the microphone on the robot and loudspeaker at the entrance $d_k \in \mathbb{R}$.
2) The orientation of the snake robot $\boldsymbol{e}_k \in \mathbb{R}^3$.

Output: 1) Location of the robot $\boldsymbol{x}_k \in \mathbb{R}^3$.
2) Pipeline map $\mathcal{M} \subset \mathbb{R}^3$.

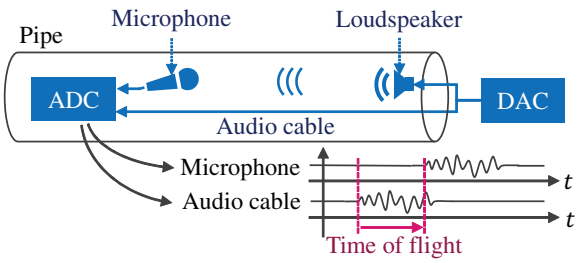


Fig. 4. Overview of Time of Flight estimation.

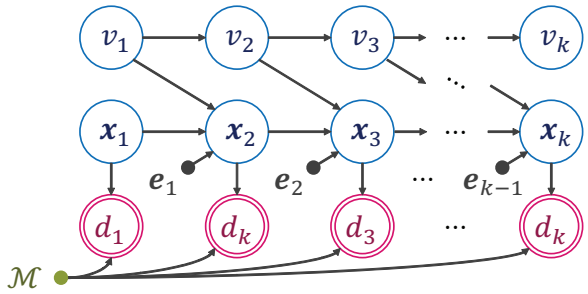


Fig. 5. Graphical model of the online sound-based localization framework.

where k represents the measurement index. We estimate the location of the sensor module as the robot location. Note that we can estimate the head location of a robot instead of the tail location by putting the sensor unit to the head of the robot. The orientation $\boldsymbol{e}_k$ is a unit vector representing the x-axis of the sensor module in the absolute coordinate system. The pipeline map $\mathcal{M}$ represents the inner space of the pipeline. The distance d_t is estimated by measuring the ToF of a reference signal (Sec. III-C), and the orientation $\boldsymbol{e}_k$ is estimated from the IMU measurements (Sec. III-D). In this paper we assume that the radius of pipelines is known in advance and pipelines are straight or connected at right angles for suppressing the accumulative errors of the IMU-based orientation estimation. The other notations are summarized in Table I.

C. ToF-based Distance Estimation

The distance d_k between the robot and the entrance of the pipeline is estimated by measuring the ToF of a reference signal emitted from the loudspeaker. As shown in Fig. 4, the ToF is measured as the onset time difference between the microphone and an audio cable that transfers the audio signal emitted by the loudspeaker.

The onset estimation for the ToF estimation in a pipeline raises following three problems:

1) Environmental noise: The industrial facilities where the robot will be used generate numerous noise sounds.

2) Reverberation and reflection: Long and narrow pipelines cause high reflection and reverberation sounds.

3) Content of the pipeline: The speed of sound depends on the gas or liquid filling the pipeline.

In this paper we assume that the pipeline is filled with gas, and the temperature and pressure are constant in the pipeline. This assumption enables us to estimate the distance between the loudspeaker and microphone by assuming that the speed of sound is the same everywhere in the pipeline:

$$d_k = \left(\tau_k^{\mathrm{mic}} - \tau_k^{\mathrm{ref}}\right) C \tag{1}$$

where τ_k^{mic} and τ_k^{ref} represent the onset times of the reference signal in the microphone and audio signal recording, respectively, and C is the speed of sound in the pipeline.

For the onset estimation robust against the first two problems, we use the onset estimation method proposed in [21]. To tackle the environmental noise problem, a time stretched pulse (TSP) [22] is used as a reference signal emitted by the loudspeaker. Since the auto-correlation of a TSP signal becomes an impulse signal, we can detect the onset time of a TSP signal with high time resolution by taking the cross-correlation between an original TSP signal and its recording. For the reverberation and reflection problem, the onset time of the reference signal is calculated using the generalized cross correlation method with phase transform (GCC-PHAT) which is robust against reverberation [14]. Since the direct sound of a reference signal is observed earlier than reflection sounds, the reflection problem is tackled by extracting the first peak of the GCC-PHAT coefficient.

D. Orientation Estimation from IMU Observations

The current orientation of the robot is estimated by accumulating the angular velocity observed by the gyroscope and correcting the accumulated error with the linear acceleration observed by the accelerometer. These two types of measurements are integrated by a complementary filter proposed by Velanti et al. [23]. This filter is known to have a fast convergence and low calculation cost.

The orientation information is used for detecting the direction of the robot movement. In this paper, x-axis of the sensor module $\boldsymbol{e}_k$, which represents the orientation of the module, is assumed as the movement direction. Since IMU-based orientation estimation has accumulated errors, the raw x-axis vector $\hat{\boldsymbol{e}}_k \in \mathbb{R}^3$ ($|\hat{\boldsymbol{e}}_k| = 1$) estimated by the complementary filter is assigned to one of the axis directions in the absolute coordinate system for suppressing the errors. This assignment is conducted by maximizing the cosine similarity as follows:

$$\boldsymbol{e}_k = \arg\max_{\boldsymbol{e}\in\mathcal{E}} \left(\boldsymbol{e}^T \cdot \hat{\boldsymbol{e}}_k\right) \tag{2}$$

where $\mathcal{E} = \{\boldsymbol{e}^x, -\boldsymbol{e}^x, \boldsymbol{e}^y, -\boldsymbol{e}^y, \boldsymbol{e}^z, -\boldsymbol{e}^z\}$ represents a set of the unit vectors of the candidate axis directions.

E. State-space Model of Robot Location and Pipeline Map

We formulate a state-space model representing relationships among the robot location, pipeline map, and observations (Fig. 5). The latent variables of the proposed state-space model consists of 1) a robot state and 2) a pipeline map. The robot state $\boldsymbol{z}_k \in \mathbb{R}^4$ consists of the location of the robot $\boldsymbol{x}_k \in \mathbb{R}^3$ and its velocity $v_k \in \mathbb{R}$:

$$\boldsymbol{z}_k = [\boldsymbol{x}_k^T, v_k]^T \tag{3}$$

The pipeline map $\mathcal{M} \subset \mathbb{R}^3$ represents the inner space of the pipeline. For the convenience, it is represented as a union of N spheres (Fig. 6):

$$\mathcal{M} = \bigcup_i \mathcal{S}(\boldsymbol{p}_i, r_i) \tag{4}$$

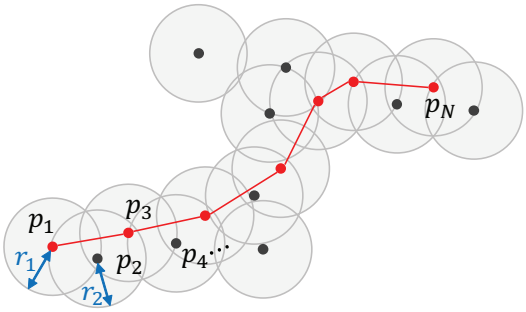


Fig. 6. Pipeline map $\mathcal{M}$ represented by a union of spheres (gray region) and an approximated sound path from $\boldsymbol{p}_1$ to $\boldsymbol{p}_N$ (red line).

where $\mathcal{S}(\boldsymbol{p}_i, r_i) = \{\boldsymbol{x} \in \mathbb{R}^3 \,||\boldsymbol{x} - \boldsymbol{p}_i| < r_i\}$ represents a sphere with a center position $\boldsymbol{p}_i \in \mathbb{R}^3$ and a radius r_i ($i = 1, 2, 3, \ldots, N$).

1) State Update Model: The state update model of the robot state $p(\boldsymbol{z}_{k+1}|\boldsymbol{z}_k)$ is formulated with two individual update models: 1) robot location update model $p(\boldsymbol{x}_{k+1}|\boldsymbol{z}_k)$ and 2) velocity update model $p(v_{k+1}|\boldsymbol{z}_k)$:

$$p(\boldsymbol{z}_{k+1}|\boldsymbol{z}_k) = p(\boldsymbol{x}_{k+1}|\boldsymbol{z}_k)p(v_{k+1}|\boldsymbol{z}_k) \tag{5}$$

The current robot location $\boldsymbol{x}_k$ is updated by the current robot orientation $\boldsymbol{e}_k$ and velocity v_k:

$$p(\boldsymbol{x}_{k+1}|\boldsymbol{z}_k) = \mathcal{N}(\boldsymbol{z}_{k+1} \mid \boldsymbol{x}_k + v_k\boldsymbol{e}_k, \boldsymbol{\Sigma}^{\boldsymbol{x}|\boldsymbol{z}}). \tag{6}$$

where $\boldsymbol{\Sigma}^{\boldsymbol{x}|\boldsymbol{z}} \in \mathbb{R}^{3\times3}$ represents the covariance matrix of the process noise of the robot location. The velocity update model $p(v_{k+1}|\boldsymbol{z}_k)$ is represented as a random walk:

$$p(v_{k+1}|\boldsymbol{z}_k) = \mathcal{N}(v_{k+1} \mid v_k, (\sigma^{v|\boldsymbol{z}})^2) \tag{7}$$

where $\sigma^{v|\boldsymbol{z}} \in \mathbb{R}$ is the standard deviation of the process noise.

2) Measurement Model: Since the orientation measurement $\boldsymbol{e}_k$ is directly formulated in the state update model (Eq. 6), only the measurement model of the estimated distance is formulated. The distance measurement model $p(d_k|\boldsymbol{z}_k, \mathcal{M})$ is formulated by using the length of the reference signal's propagation path between the microphone and the loudspeaker:

$$p(d_k|\boldsymbol{z}_k, \mathcal{M}) = \mathcal{N}(d_k \mid f(\mathcal{M}, \boldsymbol{x}_k), (\sigma^{\tau})^2) \tag{8}$$

where $f(\mathcal{M}, \boldsymbol{x}_k)$ is the length of the propagation path of the reference signal and σ^{τ} is the standard deviation of the measurement noise.

The length of the sound propagation path $f(\mathcal{M}, \boldsymbol{x}_k)$ between the robot (microphone) location $\boldsymbol{x}_k$ and the entrance of the pipeline ($[0, 0, 0]^T$) is defined as the shortest path between them on the pipeline map $\mathcal{M}$. Since it is hard to find the shortest path on the map analytically or numerically, we approximate the path as a polyline connecting the center positions $\boldsymbol{p}_i$ of the spheres $\mathcal{S}(\boldsymbol{p}_i, r_i)$ of the map $\mathcal{M}$ as shown in Fig. 6. $f(\mathcal{M}, \boldsymbol{x}_k)$ is then easily calculated by Dijkstra's algorithm [24]. This algorithm finds the shortest path from the list of the distances between each pair of two center positions $\boldsymbol{p}_i$ of the adjacent spheres.

(a) Bottom part (b) Upper part

Fig. 7. The mockup pipeline used in the evaluation for ToF-based distance estimation.

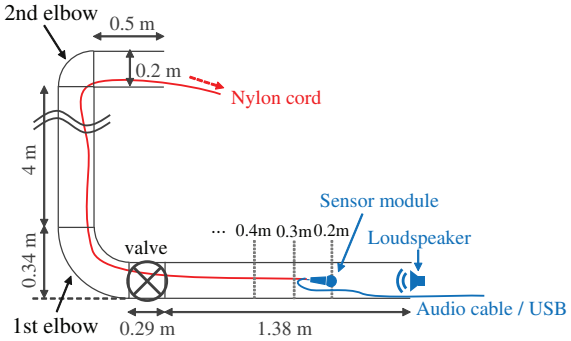


Fig. 8. Configuration of the evaluation for the ToF-based distance estimation.

F. Estimation Algorithm

The current robot state $\boldsymbol{z}_k$ (the robot location $\boldsymbol{x}_k$ and velocity v_k) is estimated from the measurements $d_{1:k}$ and $\boldsymbol{e}_{1:k}$ in an online manner by using an extended Kalman filter (EKF). Since EKF needs the derived function of $\frac{\partial}{\partial \boldsymbol{z}_k} f_d(\mathcal{M}, \boldsymbol{x}_k)$, it is approximated by numerical derivation. The pipeline map $\mathcal{M}$, on the other hand, is estimated after each update of the robot state. Since the robot is in the pipeline, the space around the current location of the robot can be assumed to be in the pipeline. Therefore the pipeline map is updated by adding the space around the robot location $\boldsymbol{x}_k$ as follows:

$$\mathcal{M} \leftarrow \mathcal{M} \cup \mathcal{S}(\boldsymbol{x}_k, r) \tag{9}$$

where $r > 0$ is a parameter representing the radius of a sphere to be added.

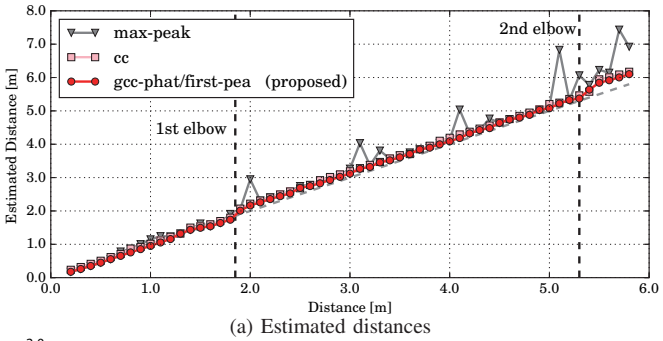
IV. Experimental Evaluation

This section reports experimental results evaluating the ToF-based distance estimation in a 6-m pipe, and the sound-based localization method with a moving in-pipe snake robot.

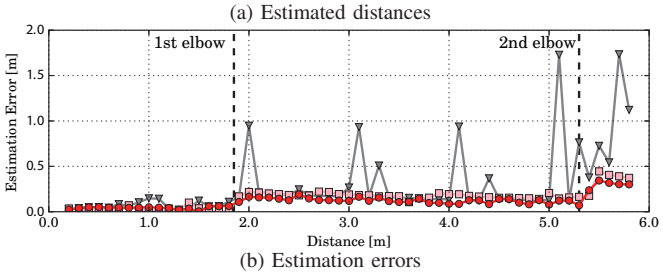
A. Evaluation 1: ToF-based Distance Estimation

The proposed ToF-based distance estimation method was evaluated by varying the distance between a microphone and loudspeaker in a 6-m pipe.

1) Experimental Settings: We evaluated the proposed ToF-based distance estimation method with the recordings in a pipeline as shown in Figs. 7 and 8. It had two elbow sections and was 6 m long and 0.2 m in diameter. A loudspeaker was equipped at the entrance of the pipeline, and a

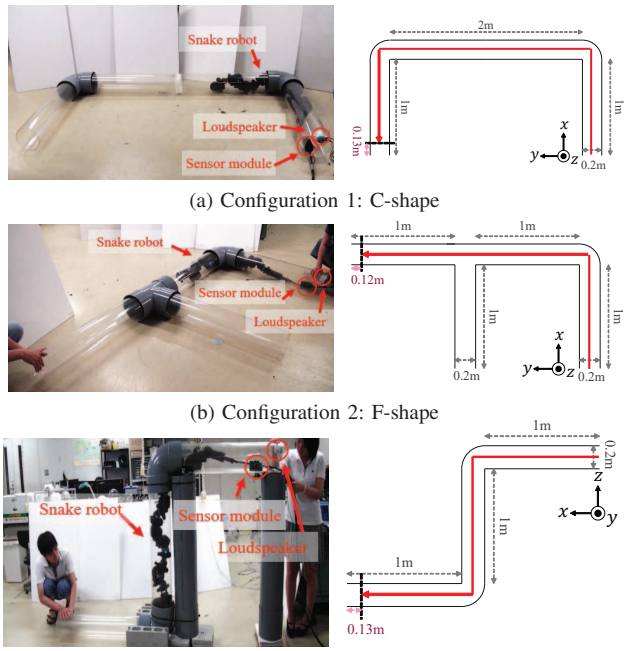


(a) Estimated distances



(b) Estimation errors

Fig. 9. Estimated distances and their errors with the ToF-based distance estimation.



(a) Configuration 1: C-shape

(b) Configuration 2: F-shape

(c) Configuration 3: Z-shape

Fig. 10. Three configurations for the experimental evaluation.The robot moved in the pipeline along the red line.

microphone was suspended by a nylon cord in the pipeline. The distance between the microphone and loudspeaker were increased in 10 cm increments by drawing up the nylon cord. The TSP reference signal used in this evaluation had a length of 16384 samples (1.024 s) at 16 kHz. The reference signals were sampled at 16 kHz and 24 bits. These parameters were determined empirically. The proposed ToF-based estimation method (gcc-phat / first-peak) was compared with the following two baseline methods: 1) using cross correlation instead of GCC-PHAT (cc), and 2) extracting the maximum correlation coefficient instead of extracting the first peak (max-peak) of the GCC-PHAT coefficients.

2) Experimental Results: Fig. 9 shows the estimated distance at each distance setting. When the microphone was placed in front of the first elbow section, the estimation errors of the proposed method (gcc-phat / first-peak) were less than 0.1 m. Moreover, when the microphone was placed beyond the first elbow section, the estimation errors of the baseline methods became significantly greater than those of the proposed method. Although the estimation errors of the proposed method increased when the microphone was beyond the elbow section, the estimation errors were less than 7 % of the distances between the microphone and loudspeaker. This result shows that our distance estimation works precisely even in high-reverberant in-pipe environments.

B. Evaluation 2: Sound-based Localization of a Moving Robot

The proposed sound-based localization method was evaluated with a moving in-pipe snake robot.

1) Experimental Settings: As shown in Fig. 10, the proposed sound-based method was evaluated under the following three configurations:

1) C-shape: Two 1-m pipes and one 2-m pipe, both 0.2 m in diameter, were connected with two elbow pipes to form a horizontal C shape.

TABLE II

PRECISION, RECALL, AND F-MEASURE OF THE ESTIMATED PIPELINE MAPS

Method	Configuration	Precision	Recall	F-measure
tof-imu-pc	C-shape	68.2 %	90.0 %	77.6 %
	F-shape	72.7 %	97.8 %	83.4 %
	Z-shape	57.7 %	83.0 %	68.0 %
tof-imu	C-shape	12.6 %	16.2 %	14.1 %
	F-shape	23.7 %	31.8 %	27.1 %
	Z-shape	18.5 %	27.1 %	22.0 %

2) F-shape: Four 1-m pipes 0.2 m in diameter were connected with one elbow pipe and one T-intersection to form an F shape.

3) Z-shape: Three 1-m pipes 0.2 m in diameter were connected with two elbow pipes to form a vertical Z shape.

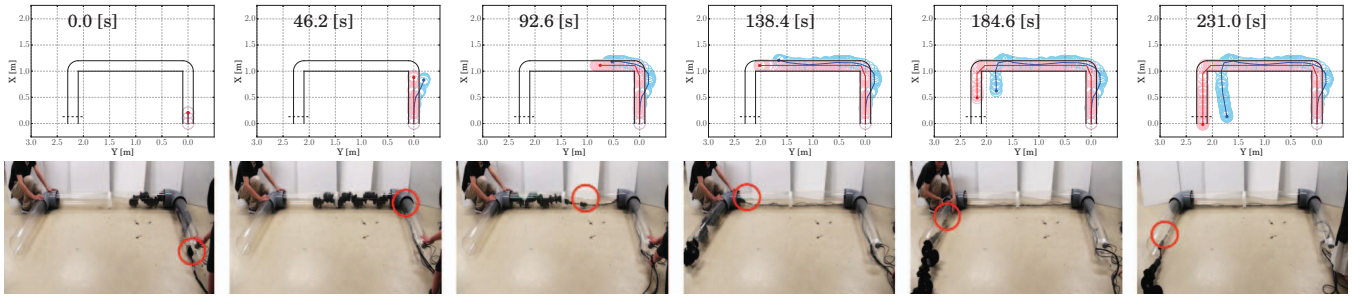
The accelerometer and gyroscope signals were sampled at 200 Hz and 16 bits. We used the same TSP signal we used in Evaluation 1. The initial state $\boldsymbol{z}_0 = [\boldsymbol{x}_0, v_0]$ was set to $\boldsymbol{x}_0 = [0.1, 0, 0]$ (m) and $v_0 = 0$ (m/s). The other parameters were determined experimentally.

In this experiment we evaluated the accuracy of the estimated pipeline map because it was hard to determine the ground truth location of the sensor module accurately. Since the proposed method estimates the pipeline map as the space inside of the pipeline, the pipeline map accuracy is evaluated with the volume ratio of the precision and recall as follows:

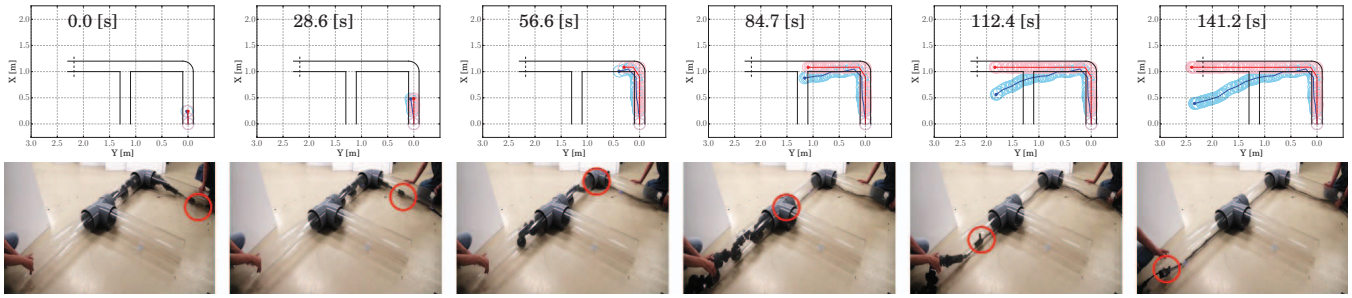
$$\mathrm{Precision}(\bar{\mathcal{M}}, \mathcal{M}) = \frac{\mathrm{V}(\bar{\mathcal{M}} \cap \mathcal{M})}{\mathrm{V}(\bar{\mathcal{M}})} \tag{10}$$

$$\mathrm{Recall}(\bar{\mathcal{M}}, \mathcal{M}) = \frac{\mathrm{V}(\bar{\mathcal{M}} \cap \mathcal{M})}{\mathrm{V}(\mathcal{M})} \tag{11}$$

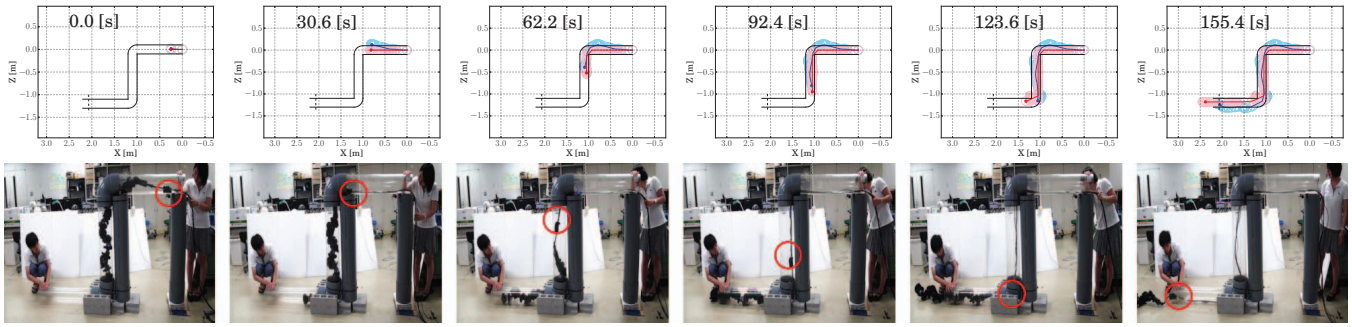
where $\bar{\mathcal{M}}$ and $\mathcal{M}$ are the ground truth and estimated pipeline



(a) Configuration 1: C-shape (red: tof-imu-pc, blue: tof-imu)



(b) Configuration 2: F-shape (red: tof-imu-pc, blue: tof-imu)



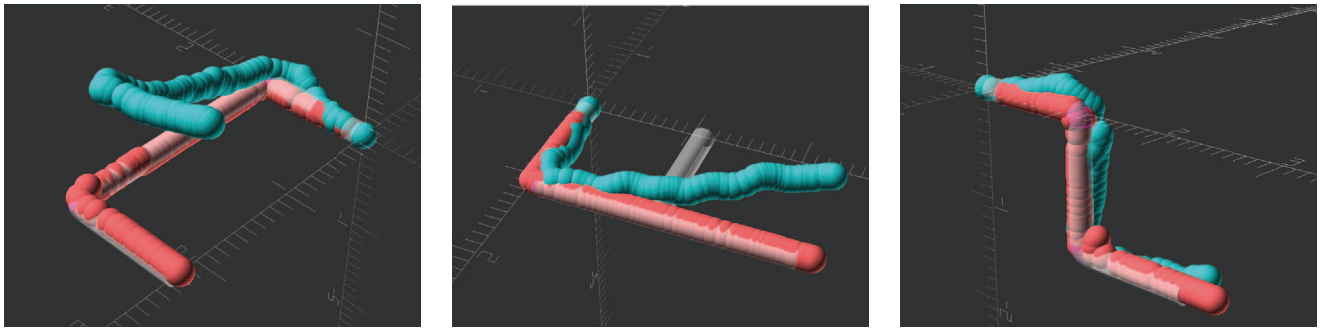
(c) Configuration 3: Z-shape (red: tof-imu-pc, blue: tof-imu)

Fig. 11. The estimated pipeline map and self-location at each measurement, and corresponding picture of the robot and pipeline. The point, line, and circles indicate the estimated location, sound propagation path, and spheres of pipeline map, respectively. Black solid lines indicate the ground truth pipeline map. The sensor module ended up at the dashed black line at each configuration. The red circles in the pictures indicate the location of the sensor module.

maps, respectively, and $\mathrm{V}(\mathcal{A})$ represents the volume of $\mathcal{A}$. Since in the F-shape configuration the robot did not get into the middle 1-m pipe (Fig. 10-(b)), the pipeline map was estimated as L-shape. Therefore, we did not take into the volume of this pipe in this evaluation. To investigate the effectiveness of the perpendicular condition that assumes pipelines to be straight or connected at right angles, we compared the proposed method (tof-imu-pc) with a baseline method (tof-imu) that does not assume the perpendicular condition. That is, tof-imu uses the raw orientation $\hat{e}_k$ estimated by the IMU-based estimation (Sec. III-D).

2) Experimental Results: TABLE II shows the precision, recall and F-measure for the estimated pipeline map at each configuration. In the all three configurations, the precision and recall of the proposed method (tof-imu-pc) have at least 57.7 % and 83.0 %, respectively. Fig. 11 shows the estimated robot location and pipeline map at each time. Although the estimated robot location at the last measurement in each configuration had more than 10 cm of the errors from the actual location, tof-imu-pc correctly estimated the locations of the elbow sections. These results showed the proposed method could robustly estimate the pipeline map even when the pipeline vertically or horizontally curved (C-shape or Z-shape), or the pipeline had a branch (F-shape).

On the other hand, compared to the tof-imu-pc, the precision and recall were significantly degraded when the perpendicular condition was not assumed (tof-imu). Fig. 12 shows the 3D projections of the estimated pipeline maps. The pipeline maps estimated by the tof-imu curved even at the straight sections because the IMU-based orientation estimation has accumulated error problem and ToF information only provides the distance information. One way to improve the proposed sound-based localization method is combining it with visual sensors for estimating curving pipelines. The pipeline shape, such as how the pipeline is curving, can be observed by using the visual odometry or visual-SLAM [9]. Although such a visual-based method also has the accumulated error problem, it will be overcame by integrating sound, IMU, and visual sensors on a unified state-space model as a SLAM framework.



(a) Configuration 1: C-shape (b) Configuration 2: F-shape (c) Configuration 3: Z-shape

Fig. 12. 3D projections of the ground truth pipeline (transparent white) and estimated pipeline maps (red: tof-imu-pc, blue: tof-imu).

V. Conclusion

This paper presented a sound-based online localization method for an in-pipe snake robot. The proposed method simultaneously estimates the robot location and pipeline map in an online manner by combining ToF-based distance estimation and IMU-based orientation estimation. The results of the preliminary experiments showed that our robot moved through the various mockup pipelines and the proposed sound-based method correctly conducted the localization and pipeline mapping. More specifically, the error of the distance estimation was less than 7% in a mockup pipeline and the accuracy of the pipeline map was more than 68.0%.

The next step is to design and implement a wireless sound-visual-based SLAM for an in-pipe snake robot. The reference audio cable that limits the movements of a robot can be replaced with a wireless radio transmitter. The wireless version will be used for more different types of in-pipe robots. Our method will be able to work without the perpendicular assumption by combining visual sensors that can directly observe the conditions of a pipeline. The integration of sound, IMU, and visual sensors will be conducted in a SLAM manner for complementing each other. We will also conduct more detailed evaluations to compare with conventional IMU- or visual-based frameworks, and more practical experiments in simulated pipelines to evaluate the effectiveness in the inspection tasks.

Acknowledgment

This study was partially supported by ImPACT Tough Robotics Challenge and by JSPS KAKENHI No. 24220006 and No. 15J08765.

References

[1] I. N. Ismail *et al.*, "Development of in-pipe inspection robot: A review," in *Proc. of IEEE Conf. on Sustainable Utilization and Development in Engineering and Technology*, 2012, pp. 310–315.

[2] D. Rollinson and H. Choset, "Pipe network locomotion with a snake robot," *J. of Field Robotics*, vol. 33, no. 3, pp. 322–336, 2014.

[3] T. Kamegawa *et al.*, "Realization of cylinder climbing locomotion with helical form by a snake robot with passive wheels," in *Proc. of IEEE Intern'l Conf. on Robotics and Automation*, 2009, pp. 3067–3072.

[4] P. Debenest *et al.*, "PipeTron series - Robots for pipe inspection," in *Proc. of 3rd Intern'l Conf. on Applied Robotics for the Power Industry*, 2014, pp. 1–6.

[5] S.-g. Roh *et al.*, "In-pipe robot based on selective drive mechanism," *Intern'l J. of Control, Automation and Systems*, vol. 7, no. 1, pp. 105–112, 2009.

[6] N. M. Yatim *et al.*, "Automated mapping for underground pipelines: An overview," in *Proc. of 2nd Intern'l Conf. on Electrical, Electronics and System Engineering*, 2015, pp. 77–82.

[7] D. Krys *et al.*, "Development of visual simultaneous localization and mapping (VSLAM) for a pipe inspection robot," in *Proc. of Intern'l Symp. on Computational Intelligence in Robotics and Automation*, 2007, pp. 344–349.

[8] H. Lim *et al.*, "SLAM in indoor pipelines with 15mm diameter," in *Proc. of IEEE Intern'l Conf. on Robotics and Automation*, 2008, pp. 4005–4011.

[9] P. Hansen *et al.*, "Pipe mapping with monocular fisheye imagery," in *Proc. of IEEE/RSJ Intern'l Conf. on Intelligent Robots and Systems*, 2013, pp. 5180–5185.

[10] A. C. Murtra *et al.*, "IMU and Cable Encoder Data Fusion for In-Pipe Mobile Robot Localization," in *Proc. of IEEE Intern'l Conf. on Technologies for Practical Robot Applications*, 2013.

[11] D. Y. Kim *et al.*, "Artificial landmark for vision-based slam of water pipe rehabilitation robot," in *Proc. of 12th IEEE Internat'l Conf. on Ubiquitous Robots and Ambient Intelligence*, 2015, pp. 444–446.

[12] R. R. Murphy, *Disaster Robotics*. MIT Press, 2014.

[13] H. Lim, J. Y. Choi, Y. S. Kwon, E.-J. Jung, and B.-J. Yi, "SLAM in indoor pipelines with 15mm diameter," in *Proc. of IEEE Intern'l Conf. on Robotics and Automation*, 2008, pp. 4005–4011.

[14] C. Zhang *et al.*, "Why does PHAT work well in lownoise, reverberative environments?" in *Proc. of IEEE Intern'l Conf. on Acoustics, Speech, and Signal Processing*, 2008, pp. 2565–2568.

[15] C. Knapp *et al.*, "The generalized correlation method for estimation of time delay," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 24, no. 4, pp. 320–327, 1976.

[16] S. Thrun *et al.*, *Probabilistic robotics*. MIT press, 2005.

[17] P. Hansen *et al.*, "Stereo visual odometry for pipe mapping," in *Proc. of IEEE/RSJ Intern'l Conf. on Intelligent Robots and Systems*, 2011, pp. 4020–4025.

[18] K. Ma *et al.*, "Robot mapping and localisation for feature sparse water pipes using voids as landmarks," in *Towards Autonomous Robotic Systems, LNCS*, vol. 9287, 2015, pp. 161–166.

[19] N. Ono *et al.*, "Blind alignment of asynchronously recorded signals for distributed microphone array," in *Proc. of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2009, pp. 161–164.

[20] H. Miura *et al.*, "SLAM-based online calibration of asynchronous microphone array for robot audition," in *Proc. of IEEE/RSJ Intern'l Conf. on Intelligent Robots and Systems*, 2011, pp. 524–529.

[21] Y. Bando *et al.*, "Microphone-accelerometer based 3D posture estimation for a hose-shaped rescue robot," in *Proc. of IEEE/RSJ Intern'l Conf. on Intelligent Robots and Systems*, 2015, pp. 5580–5586.

[22] Y. Suzuki *et al.*, "An optimum computer-generated pulse signal suitable for the measurement of very long impulse responses," *J. of the Acoustical Society of America*, vol. 97, p. 1119, 1995.

[23] R. G. Valenti *et al.*, "Keeping a good attitude: A quaternion-based orientation filter for imus and margs," *Sensors*, vol. 15, no. 8, pp. 19 302–19 330, 2015.

[24] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische mathematik*, vol. 1, no. 1, pp. 269–271, 1959.