2005 IEEE/RSJ International Conference on Intelligent Robots and Systems

Cooperative relative robot localization with audible acoustic sensing

Yuanqing Lin, Paul Vernaza, Jihun Ham, and Daniel D. Lee
GRASP Laboratory, Department of Electrical and Systems Engineering,
University of Pennsylvania, Philadelphia, PA 19084
{*linyuanq, vernaza, jhham, ddlee*}*@seas.upenn.edu*

***Abstract*— We describe a method for estimating the relative poses of a team of mobile robots using only acoustic sensing. The relative distances and bearing angles of the robots are estimated using the time of arrival of audible sound signals on stereo microphones. The robots emit specially designed sound waveforms that simultaneously enable robot identification and time of arrival estimation. These acoustic observations are then combined with odometry to update a belief state describing the positions and heading angles of all the robots. To efficiently resolve the ambiguity in the heading angle of the observing robot as well as the back-front ambiguity of the observed robot, we employ a Rao-Blackwellised particle filter (RBPF) where the distribution over heading angles is represented by a discrete set of particles, and the uncertainty in the translational positions conditioned on each of these particles is described by a Gaussian. This approach combines the representational accuracy of conventional particle filters with the efficiency of Kalman filter updates in modeling the pose distribution over a number of robots. We demonstrate how the RBPF can quickly resolve uncertainties in the binaural acoustic measurements and yield a globally consistent pose estimate. Simulations as well as an experimental implementation on robots with generic sound hardware illustrate the accuracy and the convergence of the resulting pose estimates.**

I. INTRODUCTION

Work on probabilistic approaches to robot localization have demonstrated the advantages of modeling the uncertainties in sensor readings and odometry for pose estimation [1], [2]. In a typical scenario, a robot needs to determine its own position and orientation by sensing environmental features or landmarks and integrating these measurements with odometry updates [3]. This task can be made more complex if the robot's initial pose is not well-specified (global localization), or if the robot is transported without notification (kidnapped robot). Fortunately, these problems can be overcome by modeling the uncertainty in the robot pose as a probabilistic distribution and properly filtering the noisy sensing and motion measurements over time.

These approaches have been extended to the problem of simultaneously localizing and mapping (SLAM) in an unknown environment. The SLAM problem expands the state space for estimation from that of an unknown robot pose to include the unknown positions of landmarks and other environmental features. In one approach to SLAM, a large Gaussian is used to model the uncertainty over this joint state space and the mean and covariance of the Gaussian are updated using an extended Kalman filter (EKF) [4]. When the system consists of many robots and landmarks, the EKF needs to maintain a very high dimensional covariance matrix. The EKF also cannot properly account for multimodal distributions that may arise, such as when there is a data association problem in the sensor readings. More recent work showed that a Rao-Blackwellized particle filter (RBPF) can be used to represent a more general but factorized distribution over this large state space [5]. The RBPF first represents the path of a robot with a sampling of discrete particles in path space. Then conditioned on the robot path, the uncertainties in the landmark positions are factorized and represented by Gaussians, whose means and covariances are updated with each sensor observation.

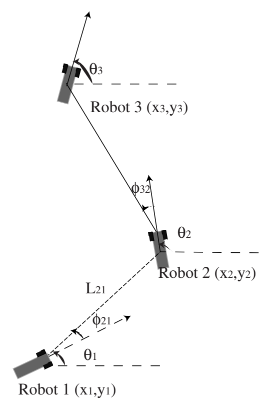

Fig. 1. Typical scenario for cooperative acoustic robot localization. Each mobile robot in the 2-D field is equipped with a speaker and a pair of stereo microphones. The heading angle of robot i relative to the positive x-axis is θ_i, and its position is (x_i, y_i). The distance and bearing angle of robot j relative to robot i are given by L_{ji} and ϕ_{ji}, respectively.

Here, we consider the problem of cooperative localization of a robot team, where the relative poses of all the robots must be simultaneously estimated using only audible acoustic measurements. With only relative observations between the robots, cooperative localization is similar to the SLAM problem in that the estimated pose of a robot is dependent upon the unknown poses of the other robots [6]. We consider the following scenario as depicted in Figure 1. Each robot in the team is equipped with a speaker and a pair of stereo microphones. By means of audible signals, each robot estimates the relative distances (L_{ij}) and relative bearing angles (ϕ_{ij}) of its neighboring robots using time-of-arrival measurements. By integrating these acoustic measurements and odometry

0-7803-8912-3/05/$20.00 ©2005 IEEE

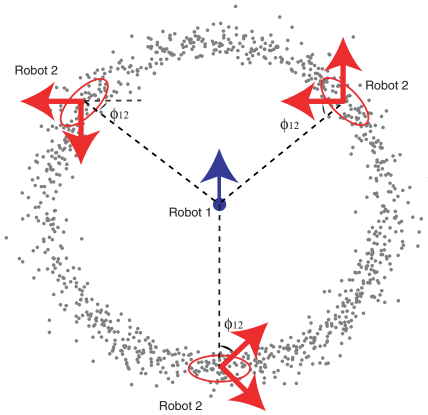

Fig. 2. Example of the ambiguity in pose estimation associated with an acoustic measurement. Robot 2 hears a sound from robot 1, and from the time of flight estimates a relative distance between the two robots L_{12}. From the time difference of arrival on robot 2's microphones, the relative bearing is also consistent with two possible ϕ_{12} due to front-back ambiguity. Thus, from robot 1's perspective, the relative location of robot 2 could be anywhere on a circle with one of two possible heading angles. Additionally, a proper probabilistic model needs to account for noise in these measurements.

information over time, a globally consistent estimate of the relative positions and heading angles of all the robots needs to be determined.

Compared to vision and ranging sensors such as lasers and sonar, robot audition is very underutilized as a sensor. This is unfortunate given that many platforms already incorporate speakers and microphones as standard hardware [7]. A major difficulty in using these acoustic sensors for localization is the uncertainty in the bearing and range measurements arising from time-of-arrival estimates. We show how to improve the accuracy of the time delay estimation by using a recently developed deconvolution algorithm that can estimate the optimal source signal for cross-correlation [8].

One aspect of time-of-arrival acoustic sensing using stereo microphones that is quite different from other modalities is the ambiguity in the relative bearing measurement. Figure 2 illustrates the uncertainty in the relative pose given an acoustic measurement between two robots. The distance measurement constrains the relative position to lie on a circle. For each possible position on the circle, the time difference of arrival measurements yields two possible bearing angles. With noise in these measurements, there will also be some associated uncertainty in the relative pose estimates.

We demonstrate that by using an RBPF, we can properly account for the uncertainties in the heading angles ($\{\theta_i\}_{i=1}^N$) of a team of N robots as well as their positions ($\{x_i, y_i\}_{i=1}^N$). In particular, we factorize the pose probability distribution by modeling the uncertainty over the heading angles as a discrete set of particles, and the conditional distribution over the translational coordinates as a Gaussian. This approach has the advantage of being able to represent the multimodal distributions associated with acoustic measurement ambiguities. The uncertainties in the translational positions are propagated using an update rule similar to a Kalman filter, allowing for fast and efficient implementation. We demonstrate how quickly this approach converges to a globally consistent pose estimate, even in the case when some robots do not have direct observations of some other robots. Once the absolute poses of the robots are determined up to an overall global rotation and translation, the relative poses between any two robots can be easily computed.

The remainder of this paper is organized as follows. In Section II, we describe the signal processing methods used to estimate the relative distance and bearing angles between two robots from acoustic signals. Section III describes the details of the RBPF model that we use to integrate the acoustic measurements with odometry in an online fashion. Finally, in Section IV, we evaluate the performance of the algorithm with simulated data, and then on an experimental implementation using three robots with generic sound hardware.

II. Acoustic Sensing

In this section, we describe the algorithms used to identify the robots and estimate the relative distance and bearing angles from measuring audible sounds. We show how multifrequency tones are detected to identify the robots and describe the algorithms for time of flight and interaural time difference (ITD) estimation used to provide relative distance and bearing measurements.

A. Robot Identification

The robots emit unique sound beacons consisting of a short dual tone multifrequency (DTMF) signal immediately followed by a wideband signal. The DTMF signal is similar to ones used for touch-tone telephone dialing and consists of two sinusoids, the frequencies of which are used to identify the robot emitting the sound. We use a digital signal processing algorithm based upon the Goertzel algorithm [9] that can efficiently recognize and identify the DTMF signal, even in the presence of a large amount of noise. Once a DTMF tone is detected and identified, the sensing robot begins a time delay estimation algorithm or may respond by emitting its own sound beacon. In our implementation, the DTMF signal is successfully identified only if it is detected above a certain energy threshold for four consecutive 8 ms frames. Since the DTMF detection algorithm is much less computationally expensive than the time delay estimation algorithms, it is useful to gate the more computationally demanding algorithms with this initial DTMF detection.

B. Estimating relative distance and bearing angle

In our robot platforms, sounds are acquired on stereo microphones using generic sound hardware with a sampling rate of 16000 Hz. To obtain accurate relative distance measurements, the clocks on the different sensing robots need to be synchronized to within a sample period, corresponding to an uncertainty of approximately 2 cm from the speed of sound. Unfortunately, conventional network time synchronization algorithms such as NTP cannot provide this accuracy due to time delays in the underlying network drivers and

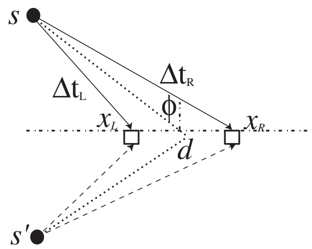

Fig. 3. Interaural time difference ($\Delta t_2 - \Delta t_1$) of arrival for measuring the bearing angle of the sound source.

TCP/IP stack in the robots' operating system. We instead use a pair of sound measurements between two robots to estimate their relative distance as well as their clock differences. Robot i_1 initiates a sound beacon which is heard by itself and robot i_2 at times t_{i_1,i_1} and t_{i_2,i_1}, using their own respective clocks. Robot i_2 responds by emitting its sound beacon which is subsequently heard by itself and robot i_1 at times t_{i_2,i_2} and t_{i_1,i_2}, respectively. The time of flight of the sound between the two robots can then be accurately calculated using $\Delta t_{i_1,i_2} = \Delta t_{i_2,i_1} = [(t_{i_1,i_2} - t_{i_1,i_1}) - (t_{i_2,i_2} - t_{i_2,i_1})]/2$.

The relative bearing of a robot emitting a sound is estimated from ITD of arrival of the wideband acoustic signal on the left and right microphones. As shown in Figure 3, the time difference of arrival $\Delta t = \Delta t_R - \Delta t_L$ gives the relative bearing angle of the sound source through:

$$\phi = \tan^{-1} \frac{c\Delta t}{\sqrt{d^2 - (c\Delta t)^2}} \tag{1}$$

where c is the speed of sound, and d is the microphone separation. Since a sound source located symmetrically behind the microphones, *s'* in Figure 3, gives rise to the same ITD as one in front of the microphones, *s*, there will be a front-back ambiguity in the relative bearing of the robot emitting the sound.

The accuracy of the relative distance and bearing calculations is determined by the accuracy of the time delay estimates. We employ the generalized cross-correlation (GCC) method for determining the time-of-arrival of the wideband signal at the microphones [10]. This method pre-whitens the spectrum of the measured signal before performing the cross-correlation computation:

$$\hat{\tau} = \arg\max_{\tau} \sum_f \frac{X(f)S^*(f)e^{j2\pi f\tau}}{|X(f)||S(f)|}, \tag{2}$$

where $X(f)$ and $S(f)$ are respectively the Fourier transforms of the measured signal and the known source signal, and $*$ denotes complex conjugation. Because wide-band signals are used in the sound beacons, this cross-correlation computation is able to achieve good temporal resolution while being robust to errors in relatively noisy environments.

Unfortunately, we do not know the optimal source signal $S(f)$ to use in the cross-correlation computation, due to uncertainties in the transfer functions of the speaker as well as microphones used to generate and measure the sounds. The wideband signal waveform in the sound beacons are filtered through these transfer functions, and this modified source signal needs to be estimated for optimal detection accuracy. To overcome this difficulty, we employ a newly developed deconvolution algorithm to initially estimate this filtered source signal [8]. The algorithm estimates the filtered sound source signal in addition to the room impulse response by optimizing

$$\min_{h_m(t)\geq 0, s(t)} \sum_{m=1}^{M} \int dt\, |x_m(t) - h_m(t) * s(t)|^2 + \lambda_m(t) h_m(t), \tag{3}$$

where M is the total number of measurements for the estimating task, $s(t)$ is the filtered source signal to be estimated, and $h_m(t)$ is the nonnegative coefficients of the room impulse response associated with the m^{th} measurement $x_m(t)$. Equation 3 is formulated in a Bayesian framework to infer the optimal sparsity regularization parameters $\lambda_m(t)$. Using this estimated source signal in the cross-correlation algorithm results in better detection and time resolution accuracy in the acoustic measurements.

III. RAO-BLACKWELLISED PARTICLE FILTER (RBPF)

In this section, we develop the RBPF framework for estimating the pose states of the robots. In the following, we assume the number of robots is N, and the pose state at time t is given by $\{\mathbf{x}(t), \boldsymbol{\theta}(t)\}$, where $\mathbf{x}(t) = \{\mathbf{x}_i(t)\}_{i=1}^N$ are the translational coordinates and $\boldsymbol{\theta} = \{\theta_i(t)\}_{i=1}^N$ are the heading angles of the robots. $\mathbf{a}(t) = \{a_{di}(t), a_{\theta i}(t)\}_{i=1}^N$ denotes the motion of the robots during the time interval $[t-1, t)$, with a rotation $a_{\theta i}(t)$ followed by a step $a_{di}(t)$. An acoustic measurement at time t is given by $\mathbf{z}(t) = (L_{n(t)}, \phi_{n(t)})$, where $L_{n(t)}$ and $\phi_{n(t)}$ are the relative distance and bearing angles of two robots, with $n(t) = \{i_2, i_1\}$ specifying the indices of the observed and observing robots. The bold variable $\mathbf{t}$ indexes a set of variables from the initial time $t = 1$ to time t.

The RBPF is similar to other Bayesian filtering algorithms in that it iteratively updates the belief state $P(\mathbf{x}(t), \boldsymbol{\theta}(t)|\mathbf{z}(\mathbf{t}), \mathbf{a}(\mathbf{t}), n(\mathbf{t}))$ using motion and acoustic measurements. It accomplishes this by decomposing the distribution over heading angles $\boldsymbol{\theta}(t)$ and translational degrees of freedom $\mathbf{x}(t)$. The heading angles are represented by a discrete set of N_p particles $\boldsymbol{\theta}^j(t)$, each of which is associated with a conditional distribution over Gaussian translational variables:

$$P(\mathbf{x}(t), \boldsymbol{\theta}(t)) = \frac{1}{N_p} \sum_j \delta(\boldsymbol{\theta}(t) - \boldsymbol{\theta}^j(t)) P(\mathbf{x}(t)|\boldsymbol{\theta}^j(t)). \tag{4}$$

Because the motion and measurement models satisfy the Markov property, the belief distribution can be iteratively updated in time:

$$\boldsymbol{\theta}^j(t) \sim P(\boldsymbol{\theta}(t)|\boldsymbol{\theta}^j(t-1), \mathbf{a}(t)) \tag{5}$$

and

$$\begin{aligned} & P(\mathbf{x}(t), \boldsymbol{\theta}^j(t)|\mathbf{z}(\mathbf{t}), \mathbf{a}(\mathbf{t}), n(\mathbf{t})) \\ = \ & \eta \sum_{I(t)} P(\mathbf{z}(t)|\mathbf{x}(t), \boldsymbol{\theta}^j(t), n(t), I(t)) \\ & \times \int d\mathbf{x}(t-1) P(\mathbf{x}(t)|\mathbf{x}(t-1), \boldsymbol{\theta}^j(t-1), \mathbf{a}(t)) \\ & \quad P(\mathbf{x}(t-1), |\boldsymbol{\theta}^j(t-1), \mathbf{z}(\mathbf{t-1}), \mathbf{a}(\mathbf{t-1}), n(\mathbf{t-1})) \end{aligned} \tag{6}$$

where η is a normalization constant, and $I(t) = \pm 1$ is an indicator variable denoting whether the bearing angle from the acoustic measurement corresponds to detecting a robot in the front or back. Equation 5 depicts the sampling of the new particles from the motion model, while Equation 6 describes the joint distribution of a new particle and the translational coordinates as a product of the measurement model and motion model. This leads to a EKF algorithm (as detailed in Section III-A) for updating the mean and covariance of the translational state belief with respect to this particle, as well as the marginal probability for estimating the importance factor of this particle (as detailed in Section III-B):

$$\begin{aligned} & P(\boldsymbol{\theta}^{j,I(t)}(t)|\mathbf{z}(\mathbf{t}), \mathbf{a}(\mathbf{t}), n(\mathbf{t}), I(t)) \\ = \ & \int P(\mathbf{x}(t), \boldsymbol{\theta}^{j,I(t)}(t)|\mathbf{z}(\mathbf{t}), \mathbf{a}(\mathbf{t}), n(\mathbf{t}), I(t)) d\mathbf{x}(t), \end{aligned} \tag{7}$$

where each particle $\boldsymbol{\theta}^j(t)$ is duplicated into two particles $\boldsymbol{\theta}^{j,I(t)}(t)$ with $I(t) = \pm 1$ accounting for the back-front ambiguity.

A. Updating the mean $\boldsymbol{\mu}^{j,I(t)}(t)$ and covariance $\Sigma^{j,I(t)}(t)$ of $P(\mathbf{x}(t)|\boldsymbol{\theta}^{j,I(t)}(t), \mathbf{z}(\mathbf{t}), \mathbf{a}(\mathbf{t}), n(\mathbf{t}))$

For the $[j, I(t)]$ heading angle particle, the motion model gives

$$\mathbf{x}(t) = \mathbf{x}(t-1) + \begin{bmatrix} a_{d1}(t)\cos(\theta_1{}^j(t)) \\ a_{d1}(t)\sin(\theta_1{}^j(t)) \\ \vdots \\ a_{dN}(t)\cos(\theta_N{}^j(t)) \\ a_{dN}(t)\sin(\theta_N{}^j(t)) \end{bmatrix} + \mathbf{u}^j(t) \tag{8}$$

where $\mathbf{u}^j(t)$ denotes the noise uncertainty in the odometry. Then, the mean $\boldsymbol{\mu}^{-j,I(t)}(t)$ and covariance $\Sigma^{-j,I(t)}(t)$ of the translational state in the proposal distribution can be computed from this model.

With a new acoustic measurement of a relative distance L_{i_2,i_1} and bearing angle ϕ_{i_2,i_1}, the distribution over relative coordinates conditional on the heading angle particle, $P(\mathbf{x}_{i_2}^{j,I(t)}(t) - \mathbf{x}_{i_1}^{j,I(t)}(t)|\boldsymbol{\theta}^{j,I(t)}, L_{i_2,i_1}, \phi_{i_2,i_1})$, can be approximated by a Gaussian with mean $\mathbf{z}_{i_2,i_1}^{j,I(t)}(t)$ and covariance $R^{j,I(t)}(t)$, where $n(t) = \{i_1, i_2\}$ specifies the index of the observing robot, i_1, and the observed robot, i_2. With this derived observation, the measurement model is linear:

$$\mathbf{z}_{i_2,i_1}^{j,I(t)}(t) = B\mathbf{x}(t) + \mathbf{v}^{j,I(t)}(t) \tag{9}$$

where the matrix $B = [0, ..., -I, 0..., 0, I, 0, ...0]$ contains the 2×2 identity matrix I, and the column positions of $-I$ and I correspond to the indices i_1 and i_2 respectively. The noise term $\mathbf{v}^{j,I(t)}(t)$ is zero-mean with covariance $R^{j,I(t)}(t)$. This approximation is more accurate than the standard EKF approximation, since the mean position $\boldsymbol{\mu}^{-j,I(t)}(t)$ may not be initially very accurate.

With the motion and measurement model in this form, the mean and covariance of the conditional Gaussians in RBPF can be updated using the standard Kalman equations:

$$\begin{aligned} K^{j,I(t)} &= \Sigma^{-j,I(t)} B^T (B\Sigma^{-j,I(t)} B^T + R^{j,I(t)})^{-1} \\ \boldsymbol{\mu}^{j,I(t)} &= \boldsymbol{\mu}^{-j,I(t)} + K^{j,I(t)} (\mathbf{z}_{i_2,i_1}^{j,I(t)} - B\boldsymbol{\mu}^{-j,I(t)}) \\ \Sigma^{j,I(t)} &= (I - K^{j,I(t)} B)\Sigma^{-j,I(t)} \end{aligned} \tag{10}$$

where the simple dependency on time t is omitted for brevity. This update requires inverting 2×2 matrices, and is thus very efficient computationally.

B. Importance factor of particle $w^{j,I(t)} = P(\boldsymbol{\theta}^{j,I(t)}(t)|\mathbf{z}(\mathbf{t}), \mathbf{a}(\mathbf{t}), n(\mathbf{t}))$

From the derviation in SectionIII-A, we have

$$\begin{aligned} & P(\mathbf{x}(t), \boldsymbol{\theta}^{j,I(t)}(t)|\mathbf{z}(\mathbf{t}), \mathbf{a}(\mathbf{t}), n(\mathbf{t})) \\ = \ & \frac{1}{(2\pi|\Sigma^{-j,I}|)^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}^{-j,I})^T (\Sigma^{-j,I})^{-1} (\mathbf{x}-\boldsymbol{\mu}^{-j,I})} \\ & \times \frac{1}{(2\pi|R^{j,I}|)^{1/2}} e^{-\frac{1}{2}(\mathbf{z}^{j,I} - B\mathbf{x})^T (R^{j,I})^{-1} (\mathbf{z}^{j,I} - B\mathbf{x})} \end{aligned} \tag{11}$$

where the simple dependency on time t is omitted for brevity, the second line comes from updating the state proposal distribution in the motion model, and the third line is the incorporation of the measurement observation. The importance weighting of a particle in the RBPF, $P(\boldsymbol{\theta}^{j,I(t)}(t)|\mathbf{z}(\mathbf{t}), \mathbf{a}(\mathbf{t}), n(\mathbf{t}))$, is computed by integrating Equation 11 over $\mathbf{x}(t)$, yielding:

$$w^{j,I} = \frac{1}{(2\pi|Q|)^{1/2}} e^{-\frac{1}{2}(\mathbf{z}^{j,I} - B\boldsymbol{\mu}^{-j,I})^T Q^{-1} (\mathbf{z}^{j,I} - B\boldsymbol{\mu}^{-j,I})} \tag{12}$$

where

$$Q = B\Sigma^{-j,I} B^T + R. \tag{13}$$

From these importance factors, the resulting particles are resampled to yield N_p particles, from which the recursion can be applied to derive the state belief of the next time step.

To summarize the algorithm, at time step $t-1$, N_p particles represent possible robot headings with equal weights, $\boldsymbol{\theta}^{\boldsymbol{j}}(t-1), j = 1, ..., N_p$, each containing a conditional distribution over translational positions described by mean $\boldsymbol{\mu}^j(t-1)$ and covariance $\Sigma^j(t-1)$. At time step t, the belief state is propagated by first duplicating each particle into two identical particles to account for the front-back ambiguity originating from the acoustic bearing measurement. Then, for each of these particles, the mean and covariance of the translational degrees of freedom, $\boldsymbol{\mu}^{j,I(t)}(t)$ and $\Sigma^{j,I(t)}(t)$ respectively, are updated using the Kalman equations. The importance weights of the $2 \times N_p$ particles are then estimated, and the particles are resampled to yield a new set of N_p particles that will be updated at the next iteration.

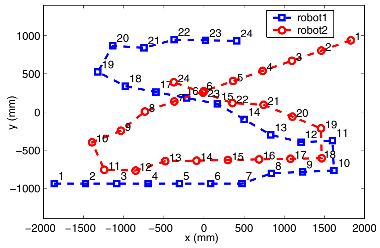

Fig. 4. Pose trajectories of the two simulated robots. The numbers 1–24 indicate the time intervals where the belief states are updated from the odometry and acoustic measurements. The size of the simulated workspace is approximately $4600mm \times 2400mm$.

IV. RESULTS

We present both simulation and experimental results in this section. The simulation shows how well the RBPF performs under controlled noise and measurement models, illustrating the quantitative accuracy of the relative pose estimation between two robots. In our experiments, we evaluate the performance of our acoustic measurement processing and RBPF algorithm on three legged robots in our performance testbed.

A. Simulation Result

For our simulation, Figure 4 displays the simulated trajectories of two robots. With each time step, robot 1 and robot 2 are given simulated acoustic measurements giving relative distance L_{21} (or L_{12}), or relative bearing ϕ_{21} and ϕ_{12}. The measurement noise added to the distance estimate is $0.1L_{21}$ and for bearing angle is 0.1π rad. The motion model for each robot contains noise in the odometry readings $a_{\theta_i}(t)$ and $a_{d_i}(t)$ with added noise of 0.05 rad and $0.1a_{d_i}(t)$, respectively. All measurements are then used to update an RBPF model to compare localization accuracy with the true simulated positions.

As the robots move and observe each other, the RBPF is able to quickly resolve the relative poses of the robots from unknown initial conditions as shown in Figure 5. Initially at time step 1, the relative pose of robot 2 is uniformly distributed across heading angles with a large associated covariance of the translational state. At time step 2, when the robots had their first observations of each other, robots were relatively localized up to a front-back ambiguity originating from the acoustic measurement as explained in II-B. Notably, as the robot walked on and the observations came in, the front-back ambiguity died out and RBPF yielded precise relative pose estimations, as seen in time steps 3-24 in Figure 5. Meanwhile, Figure 6 shows the means of the estimated probability distribution of the translation of robot 2 relative to robot 1, compared to the true values.

To further quantify the localization accuracy, we evaluate the error expectation over the estimated probability distribution, namely,

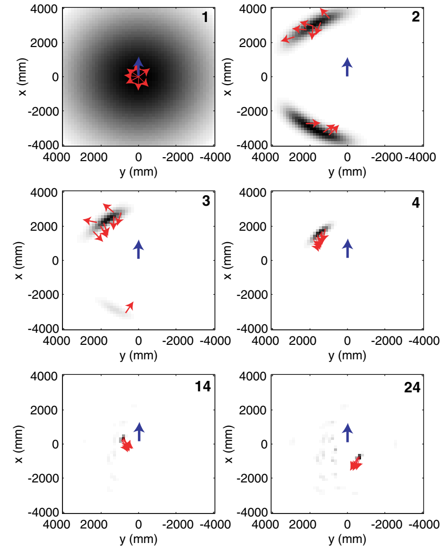

Fig. 5. Updates of the pose of robot 2 relative to robot 1. The shading of the figure indicates uncertainty of the relative translational position of robot 2. The blue arrow indicates the position and orientation of robot 1. The red arrows indicate samples from 100 particles in the RBPF, illustrating the relative heading angles of robot 2. The numerical labels indicate the time step of the simulation.

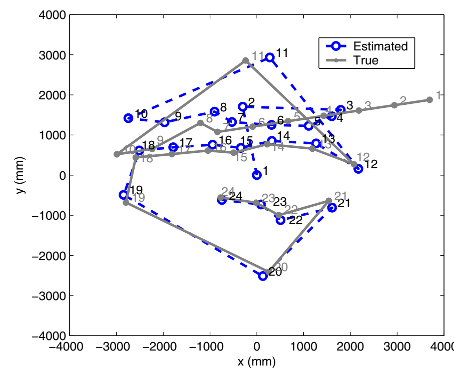

Fig. 6. The localization result of robot 2 relative to robot 1 in the simulation. The solid line with dots is the true relative translation of robot 2, while the dashed line with circles is the estimated relative translation. The 1-24 numbers indicate the time steps.

$$E_{x_{21}} = \sqrt{\sum_{j}^{N_p} \int \|\mathbf{x}_2 - \mathbf{x}_1 - \mathbf{x}_{21}^{j\,*}\|^2 P(\mathbf{x}_1, \mathbf{x}_2 | \theta_1^j, \theta_2^j) d\mathbf{x}_1 d\mathbf{x}_2} \quad (14)$$

and

$$E_{\theta_{21}} = \sum_{j=1}^{N_p} |\theta_{21}^j - \theta_{21}^*| \quad (15)$$

for relative translation and heading angle, respectively, where $\mathbf{x}_{21}^{j\,*}$ is the true translation of robot 2 relative to robot 1 given the j^{th} heading angle particle, and $P(\mathbf{x}_1, \mathbf{x}_2 | \theta_1^{(j)}, \theta_2^{(j)})$ is the estimated joint distribution of the translational states. θ_{21}^j is the estimated relative heading, while θ_{21}^* is the true one. Figure 7 shows the error expectation of the estimated relative

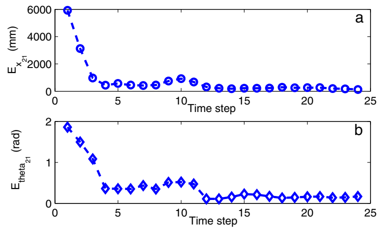

Fig. 7. Error expectation of a) the estimated relative translation, $E_{x_{21}}$ (robot 2 relative to robot 1), and b) the estimated relative heading angle, $E_{\theta_{21}}$.

translation and heading angle at each time step, indicating the convergence of the relative localization via the RBPF framework.

The simulation result qualitatively and quantitatively demonstrated that RBPF is able to resolve the ambiguities intrinsic to acoustic measurements and yield precise relative pose localization. The following examines relative pose localization among *three* real robots using acoustic sensing in the RBPF framework.

B. Experimental results

In the implementation, we had three Sony AIBO robots, indexed by 1, 2, and 3, moving around in a court of 3000mm×4000mm and monitored by a overhead camera, which offered absolute pose information for evaluating the acoustic localization result. At each time step, each robot emitted a sound beacon, and the robot with the following index responded. Note that to illustrate the localization of a robot that might never be observed directly, we only had pairwise measurements between robots 1 and 2 and between 2 and 3. No measurements between robot 3 and 1 were used. Among the measurements, we distinguished valid acoustic measurements from bad ones by checking the specially encoded structure of sound in the acoustic measurement. With the estimated acoustic source using the relevant deconvolution algorithm, the measurement yielded estimates of relative distance and bearing angle with uncertainty of $0.1L_{n_t}$ and 0.3rad, respectively, where L_{n_t} is the relative distance.

Figure 8 shows a cooperative localization result with the estimated translational poses of robot 2 and robot 3 relative to robot 1. The acoustic measurements between robot 1 and 2 at all time steps were valid, while the measurements between robot 2 and 3 were all valid except at time steps 8 and 9. Again, the measurements between robot 3 and 1 were assumed to be unobservable. The localization result indicates that the robots were able to cooperatively localize each other in the case where direct relative measurements were not available, though direct measurements may lead to more precise localization.

V. Discussion

To summarize, we have developed an RBPF framework for cooperative robot localization using relative acoustic measurements. Both simulations and preliminary experiments indicate that the uncertainties and ambiguities in the measurements

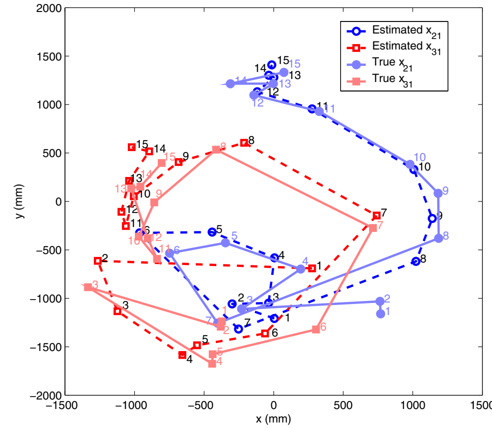

Fig. 8. Cooperative localization result of robot 2 (dark blue dashed line) and robot 3 (dark red dashed line) relative to robot 1, compared with the relative poses estimated from a overhead camera (light solid lines). The numbers indicate the time steps.

can be successfully accounted for using a probabilistic model. Further experiments are currently underway to further characterize the accuracy of our methods using more robots in a more complex environment. Hopefully, these algorithms will enable acoustic sensing, possibly in conjunction with other types of sensors, to be used for other robot teams.

Acknowledgment

The authors would like to thank the support of the U. S. National Science Foundation, the Army Research Office, and the Defense Advanced Research Projects Agency. We also thank Ricky Der for critical reading of the manuscript.

References

[1] J. J. Leonard and H. F. Durrant-Whyte, "Directed sonar sensing for mobile robot navigation," 1992.

[2] S. Thrun, D. Fox, W. Burgard, and F. Dellaert, "Robust Monte Carlo localization for mobile robots," *Artificial Intelligence*, vol. 128, pp. 99–141, 2000.

[3] J. Borenstein, B. Everett, and L. Feng, *Navigating Mobile Robots: Systems and Techniques*. Wellesley, 1996.

[4] J. J. Leonard, I. J. Cox, , and H. F. Durrant-Whyte, "Dynamic map building for an autonomous mobile robot," *Int. J. Robotics Research*, vol. 11, pp. 286–298, 1992.

[5] D. K. B. W. M. Montemerlo, S. Thrun, "Fastslam: A factored solution to the simultaneous localization and mapping problem," in *National Conference on Artificial Intelligence*, 2002.

[6] M. J. M. Andrew Howard and G. S. Sukhatme, "Cooperative relative localization for mobile robot teams: An ego-centric approach," in *Workshop in Multi-Robot Systems: From Swarms to Intelligent Automata*, A. C. Schultz, L. E. Parker, and F. E. Schneider, Eds., 2003, pp. 65–76.

[7] J. Chen, L. Yip, J. Elson, H. Wang, D. Maniezzo, R. Hudson, K. Yao, and D. Estrin, "Coherent acoustic array processing and localization on wireless sensor network," *Proceedings of the IEEE*, vol. 91, pp. 1154–1162, 2003.

[8] Y. Lin and D. D. Lee, "Relevant deconvolution for acoustic source estimation," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2005.

[9] A. V. Oppenheim and R. W. Schafer, *Discrete-time signal processing*. Prentice Hall, 1998.

[10] C. H. Knapp and G. C. Carter., "The generalized correlation method for estimation of time delay," *IEEE Transactions on ASSP*, vol. 24, no. 4, pp. 320–327, 1976.