Characterizing a Moving Source in Wireless Sensor Networks from the View of Maximum Likelihood

Viet-Hung Dang
Department of Computer Engineering, KyungHee University, Gyeonggi-do, 446-701, Korea.
Email: dangviethung@oslab.khu.ac.kr

Young-Koo Lee
Department of Computer Engineering, KyungHee University, Gyeonggi-do, 446-701, Korea.
Email: yklee@oslab.khu.ac.kr

Sungyoung Lee
Department of Computer Engineering, KyungHee University, Gyeonggi-do, 446-701, Korea.
Email: sylee@oslab.khu.ac.kr

***Abstract*—While trying to solve the problem of tracking unwelcome objects in wireless sensor networks (WSNs) based on acoustic signals, we notice that the sources' frequency components (f-components) and velocities play very important roles in improving the positioning results and for higher applications including source recognition. The information of a source like position, velocity, direction of moving and f-components are twisted in the data recorded at different sensors. This paper aims to derive all these characteristics based on the Doppler effect information. Since the outputs of Discrete Fourier Transformation of the collected data suffer from frequency leakage and Gaussian noise, we consider them as distribution functions on frequency domain. This is the novel idea to treat these functions as the likelihood distributions and calculate the hidden parameters during the attempt to maximize the join likelihood function. The method is described and then illustrated with experiments in the regard of estimation accuracy.**

I. Introduction

In principle, the speed of a moving source can be calculated based on Doppler effect if the emitted acoustic signal is recorded, and the main frequency of the source and the direction of moving are known [1]. For the tracking applications in Wireless Sensor Networks (WSNs) where sensors are deployed randomly for tracking object, the frequency of the source and the direction of moving are not known. Fortunately, the combination of data from different sensors can help figure out most characteristics of the source like source position, velocity, direction of moving and the original f-components of the source. These features are very important in many applications. First of all, they can be used to improve the tracking accuracy because when the speed of the moving source and the position are estimated, predictions can be carried out to reduce the uncertainty of the next step of positioning and velocity estimating. Secondly, while performing the position estimation based on sound or acoustic signals, the system usually faces the error when two or more sources meet together and make the system confuse that there is only one source. Knowing the position and the direction of the source will eliminate this error, improve tracking accuracy and enhance the system's robustness. Thirdly, when all the features of a source are indicated, the applications of source classification or source recognition are easily implemented.

Related positioning methods so far are mainly based on the nodes in which the arrays of microphones are used for calculating the directions of arrivals (DoAs) [2][3][4][5][6][7][8]. DoA is the relative angle from the source to the array of sensors. Several tracking approaches estimate the velocity of the source and then use the techniques of Particle Filter, hidden Markov model or Kalman Filter [9] to smooth up the traced sources [7][10]. However, the velocity estimation is usually computed based on two successive estimated positions which suffer high errors. In this paper, we try to indicate each f-component and its velocity based on the sensor network in which each node has only one isotropic sensor like in [11]. Isotropic sensors help reduce communication cost which is the most critical cost aimed to compress by any wireless network designer. We do not attempt to do the tracking to smooth the moving path or to narrow down the likely area that covers the source in the time sequence. Instead, we provide a frame work for estimating the original frequencies, the velocity and the position of the source based on the data collected in one frame where the Doppler effect is the key knowledge for solving the problem. By regarding the data on Fourier domain as likelihood function and then maximizing the join likelihood, we can obtain all the hidden features of the source in term of position, velocity, direction of moving and its f-components.

II. Problem formulation

Consider a source moving in the deployed area where its signal $s(t)$ is recorded by N location-known sensors. The received signal at each sensor is denoted as $x_i(t)$, where $i = 1, .., N$. $x_i(t)$ is modeled similarly to [3],

$$x_i(t) = a_i s_j(t - \tau_i(t)),\ i = 1, .., N. \quad (1)$$

$a_i \in \mathbf{R}$ is the amplitude gain of the source signal measured at sensor i and $\tau_i(t)$ is the signal propagation time. If the source is fixed, the delays $\tau_i(t)$ are constants. However, if it moves with the velocity of v_s, then the time delays change at different time points,

$$\tau_i(t) = \frac{d_i(t)}{v_c}. \quad (2)$$

$d_i(t)$ in equation (2) is the immediate distance from sensor i to source s and v_c is the velocity of sound propagation, $v_c = 343m/s$ in air. Source s's movement makes $\tau_i(t)$ increase or decrease over time, resulting in a stretched or compressed image of the source signal on the time domain

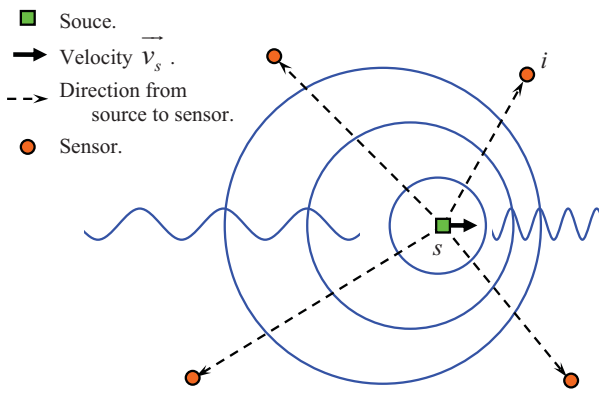

Fig. 1. Illustration of Doppler effect when the source moves. Different angles of observation therefore record different frequency shifts.

at receiver i (see Fig.1). As a result, different shifts result into different f-components at the receivers. This phenomenon is called Doppler effect [1] and is described via the formula

$$f_{im} = \left(\frac{v_c}{v_c + v_s \cos(\theta_i(t))}\right) f_m, \tag{3}$$

where f_m is some f-component of source s, f_{im} is the shifted version of f_m received at sensor i, and $\theta_i(t)$ is the immediate angle between $\overrightarrow{v_s}$ and $\overrightarrow{is}$. A source is usually composed of many f-components and the components are shifted with different amounts on the frequency domain, building up the twisted information in the sensed data at different sensors. Our goal is to extract the frequency information from some f-component and calculate the hidden characteristics of the source. Therefore, indicating a group of recorded f-components that are generated from an original f-component is necessary. The higher f-component is shifted with bigger amount of frequency range given that the source speed is fixed. That makes the clustering task difficult even on the frequency domain. Fortunately, when considering frequency domain under logarithmal scale, we see that the distance on log-frequency domain between every pair of dominant f-components of the source is the same as that from sensor. Indeed, consider two arbitrary f-components f_m and f_n of the source, the distance on log-frequency domain is

$$\Delta f_{source}(dB) = \log_{10}(f_m) - \log_{10}(f_n) = \log_{10}(\frac{f_m}{f_n}). \tag{4}$$

Meanwhile, at sensor i, this distance is

$$\begin{aligned} &\Delta f_{sensor}(dB) = \log_{10}(f_{im}) - \log_{10}(f_{in}) = \\ &= \log_{10}\left(\frac{v_c}{v_c + v_s \cos(\theta_i(t))}\right) + \log_{10}(f_m) - \\ &- \left(\frac{v_c}{v_c + v_s \cos(\theta_i(t))}\right) - \log_{10}(f_n) = \\ &= \log_{10}(\frac{f_m}{f_n}). \end{aligned} \tag{5}$$

Evidently, comparing equations (4) and (5) shows that the shifts of different f-components observed at sensors are fixed on the log scale of frequency domain. Therefore, each dominant f-component will be indexed again on the log scale so that grouping f-components based on maximum correlation technique is performed to find the matching. Then, f-components that originate from some specific source's f-component are clustered into one group.

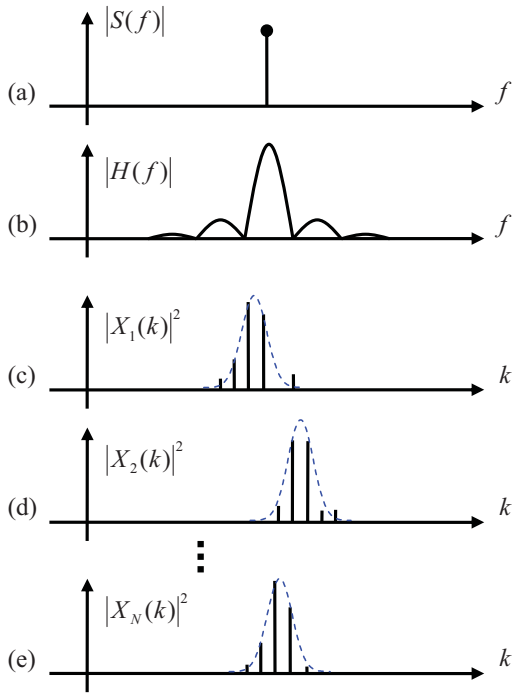

Fig. 2. Illustration of frequency leakage, frequency shifts, and the uncertainty of indicating the exact recorded frequency of a specific original f-component. The results on Fourier domain can be considered as PMFs constrained by the Gaussian distributions.

Now given that all the dominant f-components have been marked and clustered, the method of extracting the source characteristics is described as following.

III. Proposed Method

Owing to the frequency leakage [12] and the background noise in the recorded data, the Discrete Fourier Transform (DFT) result on the frequency segment of an f-component presents the uncertainty law. That means the higher accuracy for indicating an f-component is required, the more samples must be needed in a frame due to the frequency leakage. The frequency of an f-component is just known to be around the dominant bins but not exactly known(see Fig.2). Moreover, due to the Parseval's theorem, the energy of an f-component is conserved, which means at a fixed distance from the sensor to the source position, the DFT frequency intervals of an f-component can be different in term of sharp, but it holds

$$\sum_k |X_{im}(k)|^2 = \text{const}, \tag{6}$$

where $X_{im}(k)$ are the bins resulted from the f-component m of the source at sensor i.

Fig.2a demonstrates an f-component on frequency domain, while Fig.2b is for the leakage frequency from that f-component due to the limited length of window frame on the time domain. Fig.2c, Fig2d and Fig.2e are the illustrations for the DFT results of an f-component at different sensors. The received frequencies are the shifted versions of the original f-component and the DFT results suffer from the frequency leakage. However, due to the Parseval's theorem, since the energy of an exact frequency can drop to other bins and the energy is conserved for a fixed distance from sensor to source, we can consider the square of the absolute DFT results as the discrete versions of probability functions, or probability mass functions (PMFs). It is similar to the situation when we try to measure a parameter with a set of quantified values, the measured values fall around the parameter's mean, and the sum of their probabilities is equal to 1, or conserved. Because of the fact that the power of the f-component is only affected by the distance, the change in the function by the distance merely plays the role of a normalization constant. Meanwhile, the shapes of these PMFs depend on the speed and moving direction of the source. We will apply optimization technique to find these parameters during the attempt of maximizing the join likelihood of these PMFs. Each of the PMFs is considered as a Gaussian distribution with its individual mean $\overline{f}_{im}$ and the common deviation σ_m. The σ_m is the same because the time frame length is the same at sensors, causing the same width of the lopes on frequency domain as can be seen from the illustrations of Fig.2c, Fig.2d and Fig.2e.

For some f-component f_m, let $\mathbf{p}$ be the immediate coordinates of the source and $\mathbf{v}$ be its immediate velocity vector, $\mathbf{v} = (v_x, v_y)$. A shifted f-component f_{im} of a given f_m at a sensor depends only on that sensor's location, not on other shifted f-components recorded at other sensors. Therefore the shifted f-components at different sensors are independent of each other. Then the joint likelihood function of the shifted f-component, denoted with $p(f_{1m}, f_{2m}, ..., f_{Nm}|f_m, \mathbf{p}, \mathbf{v})$, is expressed as

$$p(f_{1m}, f_{2m}, ..., f_{Nm}|f_k, \mathbf{p}, \mathbf{v}) = \prod_{i=1}^{N} p(f_{im}|f_k, \mathbf{p}, \mathbf{v}). \quad (7)$$

In equation (7), f_{im} is calculated as in equation (3) and $p(f_{im}|f_m, \mathbf{p}, \mathbf{v}) \sim N(f_{im}|\overline{f}_{im}, \sigma_m^2)$, where

$$\overline{f}_{im} = \frac{\sum\limits_{k} \left(|X_{im}(k)|^2 f(k) \right)}{\sum\limits_{k} |X_{im}(k)|^2}. \quad (8)$$

Again, $X_{im}(k)$ are the bins resulted from the f-component m of the source at sensor i as in (6). Because the $\log$ function is monotone, maximizing likelihood function is equivalent to maximizing the log-likelihood function. In other words, we will try to maximize the below function,

$$\ln\left(p(f_{1m}, f_{2m}, ..., f_{Nm}|f_m, \mathbf{p}, \mathbf{v})\right) = \sum_{i=1}^{N} \ln\left(p(f_{im}|f_m, \mathbf{p}, \mathbf{v})\right). \quad (9)$$

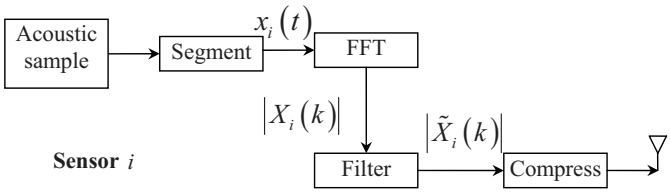

Fig. 3. Work flow of a sensor for reducing communication cost.

Meanwhile,

$$p(f_{im}|f_k, \mathbf{p}, \mathbf{v}) = \\ = \frac{1}{Z_m} \exp\left\{ -\frac{1}{2\sigma_m^2} \left(\frac{v_c}{v_c + |\mathbf{v}| \cos\left(\theta_i(t)\right)} f_m - \overline{f}_{im} \right)^2 \right\} \quad (10)$$

where Z_m is the normalization constant. Note that $|\mathbf{v}| = v_s$ and $\cos\left(\theta_i(t)\right) = \frac{(\mathbf{p}_i - \mathbf{p})^T \mathbf{v}}{|\mathbf{p}_i - \mathbf{p}||\mathbf{v}|}$ where $\mathbf{p}_i$ is the location of sensor i, equation (9) is rewritten

$$\ln\left(p(f_{1m}, f_{2m}, ..., f_{Nm}|f_m, \mathbf{p}, \mathbf{v})\right) = \\ = -N \ln(Z_m) - \frac{1}{2\sigma_m^2} \sum_{i=1}^{N} \left(\frac{v_c}{v_c + |\mathbf{v}| \cos\left(\theta_i(t)\right)} f_m - \overline{f}_{im} \right)^2 \\ = -N \ln(Z_m) - \frac{1}{2\sigma_m^2} \sum_{i=1}^{N} \left(\frac{v_c}{v_c + \frac{(\mathbf{p}_i - \mathbf{p})^T \mathbf{v}}{|\mathbf{p}_i - \mathbf{p}|}} f_m - \overline{f}_{im} \right)^2. \quad (11)$$

The normalization constant Z_m, sensor number N and deviation σ_m are independent of f_m, $\mathbf{p}$ and $\mathbf{v}$. Therefore maximizing likelihood function with respect to f_m, $\mathbf{p}$ and $\mathbf{v}$ is equivalent to maximizing the last term in equation (11) with respect to f_m, $\mathbf{p}$ and $\mathbf{v}$. Applying an iteration optimization technique that updates alternately these parameters, we will obtain the source's f-component f_m, position $\mathbf{p}$, speed v_s and direction of moving. It is remarked that in order for the method to work, we have to consider a frame work in which a central computer is needed because the data must be collected together for calculation. Even so, the communication cost should be low because sound data is usually large for one frame due to high sampling frequency. Therefore, in this paper, we utilize the same frame work as in [11] (see Fig.3). The flow work in Fig.3 gives the sensor the ability of performing Fast Fourier Transformation (FFT), eliminating background Gaussian noise and compressing the data before sending to the central base computer. As a result, the transmitted data includes only few values of dominant f-components on the frequency domain. Moreover, only half of the f-component number is needed owing to the symmetric property of the FFT image (recorded data is real). The frame work helps reduce the communication load and makes the method applicable into WSNs. At the base, dominant f-components are put to logarithmal scale and clustered into groups each of which contains all f-components resulted from a specific original f-component. Then the f-components are again considered on frequency domain where

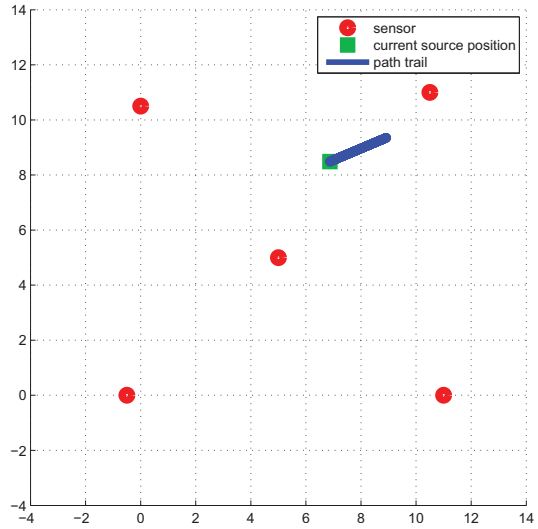

Fig. 4. Deployed area for the simulation set, [0m,12m]x[0m,12m].

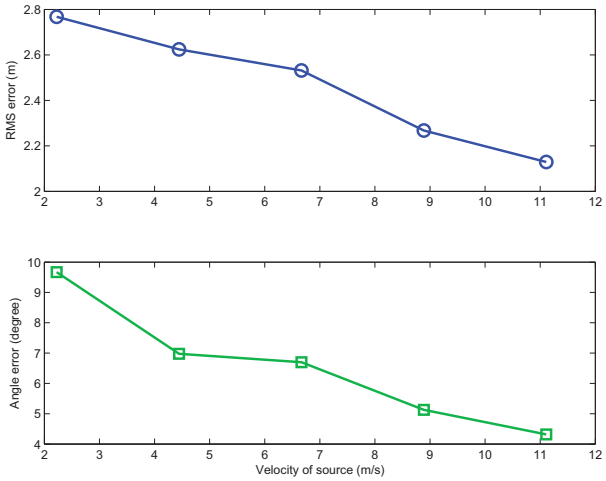

Fig. 5. RMS error for the source position estimation and error of moving direction with respect to the velocity change when $SNR = 3.05$.

the maximum likelihood view is applied for finding the best set of hidden parameters during the attempt of maximizing the function defined in equation (9).

IV. Experiments and Discussion

Experiments in this section aim to find out the hidden features of the source within only one frame time. That means we do not focus on keeping characterizing the source in sequence of frames and continuously using the results to refine the parameter estimations. We parametrically generate a source, which mimics the sound of a vehicle, so that Doppler effect can be generated as described in equation (3). The examples of recorded data with Doppler effect are available on the site [13] in which the mechanism of generating are the same as in this paper. Data is collected in the time of one frame where the source position $\mathbf{p}$ and the direction of moving are generated randomly. The position of the source is somewhere within the monitored area which is deployed with four sensors around the corners and one in the middle (see Fig.4). The source signal is composed of three dominant f-components, $f_1 = 10.18Hz$, $f_2 = 34.43Hz$ and $f_3 = 48.70Hz$. The sampling frequency at the sensors is $F_s = 16,384Hz$ and the length of time frame is $T_f = 0.2s$. For evaluating the result with respect to noise level, we choose the parameter signal to noise ratio $SNR = \frac{P_{mean}}{P_{noise}}$. P_{mean} is the average power of the recorded signals at sensors when noise level is zero and P_{noise} is the power of Gaussian noise at one sensor. Note that in these simulations, we consider the noise of the setting around the sensors to be Gaussian and the noise levels to be the same at all sensors.

In this paper, two sets of simulation are conducted, showing the influences of the speed of the source and noise level on the estimated parameters. Fig.5 and Fig.6 illustrates the relations of the velocity and other hidden source characteristics. The velocity is varied in the increment of $8km/h$ (or $2.22m/s$) up to $40km/h$ (or $11.1m/s$). Since the source is composed of three major f-components, the estimations of position, velocity, direction of moving and frequency error are the average values computed from these major f-components. Specially, for a reasonable evaluation of frequency error, we consider the relative error in percent because the higher frequency has the higher Doppler shift. Each point in the figure is obtained as the mean result of 500 trials.

For each trial, the data is recorded in one frame time $T_f = 0.2s$ at the five sensors. Each sensor performs FFT, Gaussian elimination, compression and then sends the data to the base (see Fig.3). At the base, dominant f-components are indexed in the log-scale. The base marks a set of shifted f-components which belong to some original f-component by maximizing the correlation. After that, each set of shifted f-components is used to infer all the hidden features of the source as described in equation (11). Negative gradient method is used for the optimization task. The initial position is the mean position of the sensors, the initial velocity is zero, and the initial frequency is the mean value of the f-components in the group.

The result in Fig.5, which is obtained when the $SNR = 3.05$, shows that the position of the source is indicated more accurately if the source speed increases. The root mean square error (RMS) decreases from 2.8m to 2.1m. It is reasonable because when the source speed is higher, the path trail it leaves is longer during the monitoring time of one frame. Mean while, the accuracy of the moving direction is high, which is illustrated as angle error in the second sub-figure of Fig.5. The angle error is the difference between the actual moving direction and its estimated result. The maximum angle error is less than 10 degrees. When the source moves faster, the angle error reduces significantly, even to around 4 degrees at the speed of $40km/h$. It can be seen that the velocity estimation suffers from the increment of the source speed. However, the error is likely to be stable when the source speed is high. When the source velocity is $8.88m/s$ and $11.1m/s$, the error are about only $1.7m/s$. In addition, the relative error of frequency estimation is approximately 5% (see Fig.6). Evidently, the

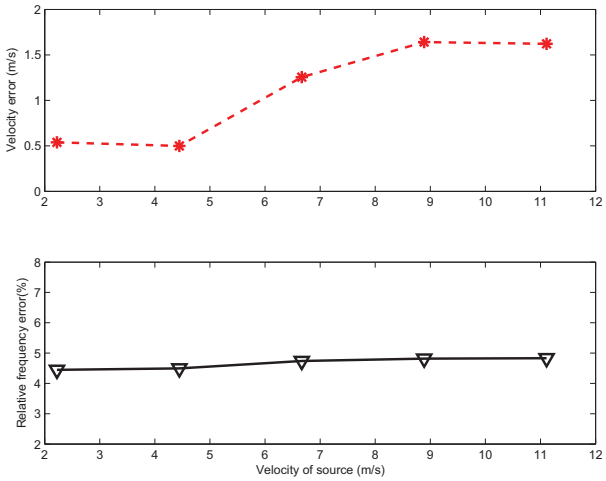

Fig. 6. Error of velocity estimation and error of frequency estimation with respect to the velocity change when $SNR = 3.05$. The error of frequency estimation is the relative error comparing to the original frequency (%).

TABLE I
CHARACTERISTIC RESULTS OF THE SOURCE WITH RESPECT TO THE CHANGE OF NOISE LEVEL.

SNR	**43.44**	**5.02**	**2.35**	**1.78**	**0.66**
RMS (m)	2.10	2.08	2.07	2.08	2.25
angle error (degree)	3.76	4.61	5.07	5.54	6.92
velocity error (m/s)	1.59	1.62	1.47	1.62	1.67
frequency error (%)	4.87	4.85	4.84	4.80	4.80

system works robustly with different source speeds.

Meanwhile, TABLE.I displays the estimated results of all the source characteristics under the influences of noise level when the source speed is $40km/h$. Each value is obtained as the average result of 500 trials like in the first set of simulation. The deviation of the Gaussian noise is set to increase linearly, so the SNR value decreases rapidly from 43.44 to 0.66. Under these changes, except for the angle error, other estimations are at high accuracy and are independent of the noise level. The RMS error is around $2.1m$, the mean velocity error is within the range from $1.45m/s$ to $1.67m/s$, and the frequency error is less than 5% even with SNR of 0.66. It is because the system works based mainly on the Doppler shifts when the source moves in the monitored area. It tries to maximize the join likelihood function of the shifted f-components on Fourier domain. The Fourier transformation can help indicate the frequency of shifted f-component with little influence of the noise, thus during that attempt, the hidden parameters are revealed with high accuracy. In other words, the results are quite independent of the noise level and are dependent of the source velocity. However, the dependence is not significant since the accuracy is very high either when the source speed are low as can be seen in Fig.5 and Fig.6, or when the noise level is very high as can be seen in TABLE.I. Obviously, the system is robust and reliable.

V. CONCLUSIONS

This paper introduces a novel method for characterizing an acoustic source in which, after stages of FFT and Gaussian filtering, we consider the energy on Fourier domain of each shifted f-component as a likelihood function. The optimization is used to maximize the join likelihood function and the hidden features of the source are obtained in terms of position, velocity, moving direction and dominant original frequencies. The system can be implemented as a WSN application since the communication cost is very low for each lengthen frame of collected data with the isotropic microphone sensors. The experiments show that the system provides high accuracy results for all kinds of estimations. A continuous inference mechanism for these hidden features can be used as a combination to improve their own accuracies because the prediction can be carried out to narrow the uncertainty of the next estimations. Moreover, the method is also useful for source classification or recognition when all of its characteristics are revealed.

ACKNOWLEDGMENT

This work was supported by a grant from the Kyung Hee University in 2010, (KHU-10201372).

REFERENCES

[1] R. Serway and R. Beichner, *Physics for Scientists and Engineers*, 5th ed. Brooks/Cole Publishing Company, 1999.
[2] Y. Sasaki, S. Kagami, and H. Mizoguchi, "Multiple sound source mapping for a mobile robot by self-motion triangulation," in *International Conference on Intelligent Robots and Systems*. IEEE, Oct 2006, pp. 380–385.
[3] M. Stanacevic, "Micropower gradient flow acoustic localizer," *IEEE Transations on Circuits and systems*, vol. 52, pp. 2148–2157, 2005.
[4] S. T.Birchfield and D. K. Gilmor, "A unifying framework for acoustic localization," in *IEEE International Conference*. IEEE Computer Society, 2001.
[5] A. Mahajana and M. Walworth, "3-d position sensing using the differences in the time-of-flights from a wave source to various receivers," *IEEE Transaction on Robotics and Automation*, vol. 17, pp. 91–94, 2001.
[6] J.-M. Valin, F. Michaud, J. Rouat, and D. Letourneau, "Robust sound source localization using a microphone array on a mobile robot," in *In Proc. ICASSP98*, 2003, pp. 1228–1233.
[7] F. Antonacci, D. Riva, D. Saiu, A. Sarti, M. Tagliasacchi, and S. Tubaro, "Tracking multiple acoustic sources using particle filtering," in *European Signal Processing Conference, EUSIPCO-2006*, 2006.
[8] K. Kaouri, "Left-right ambiguity resolution of a towed array sonar," in *Ph.D dissertation*. MSc in Mathematical Modelling and Scientific Computing, 2000.
[9] C. Rohig and M. Muller, "Localization of sensor nodes in a wireless sensor network using the nanoloc trx transceiver." 69th IEEE Vehicular Technology Conference, 2009.
[10] N. Roman and D. Wang, "Binaural tracking of multiple moving sources," *IEEE Transactions on Audio, Speech, and Processing*, vol. 16, pp. 728–139, 2008.
[11] V.-H. Dang, T. Le-Tien, Y.-K. Lee, and S. Lee, "Acoustic multiple object positioning system," in *International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks, PE-WASUN-2010*. ACM, 2010.
[12] R. A. Haddad and T. W. Parsons, *Digital Signal Processing Theory, Applications and Hardware*. Computer Science Press, 1991.
[13] V.-H. Dang, "Generated examples of multi-object sound data," in *http://uclab.khu.ac.kr/ext/dvhung/*.