24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

VEHICLE LOCALIZATION AND TRACKING USING ACOUSTIC SENSORS BASED ON TIME DIFFERENCE OF ARRIVAL METHOD

Kai Zhang, Diange Yang, Feng Miao, and Xiaomin Lian

State Key Laboratory of Automotive Safety and Energy, Department of Automotive Engineering, Tsinghua University, Collaborative Innovation Center of Intelligent New Energy Vehicle, Beijing 100084, China
email: ydg@mail.tsinghua.edu.cn

Localizing and tracking objects using sound information has been already applied in many fields. Moving vehicles can radiate many types of noise, which in turn can be employed to detect and track the position of vehicles. TDOA (Time difference of arrival) method is a common sound source localization technique. It only need four sensors (microphones) to realize three-dimensional sound source localization. In this paper, we present a method based on TDOA to localize the sound source generated by moving vehicles, and accomplish the goal of vehicle localization and tracking. This method considers the Doppler effect of moving sound source, and fixes the frequency shift to improve the accuracy of time delay estimation. Besides, this method can timely and efficiently detect the position of vehicles under bad visual conditions, which covers the disadvantage of tracking method based on video data. The performance of the proposed method has been validated by both simulation and experiment.
Keywords: TDOA, sound source localization, vehicle tracking, Doppler effect, kalman filter

1. Introduction

Sound source tracking has been an important area of research. Locating acoustic sources by means of microphone array has been used in various sceneries, including speaker tracking using TDOA method [1][2], mobile robot tracking system [3] , vehicle detection [4][5]with beamforming method. There are three main sound source localization methods in general, high resolution spectral estimation method, beamforming method and TDOA method. The first two are DOA (Direction Of Arrival) estimation method based on spatial search, which need mass computation with large numbers of sensors. Compared with them, TDOA is more suitable for efficient localization for moving sound source. Nevertheless, the traditional method only using a frame of data obtained at the current time to estimate the current source location will lead to degradation of tracking performance in noise and reverberation environment. Therefore a novel method employing the state-space approach based on kalman filtering [6] and particle filtering [7] is proposed by some researchers. However they didn't consider the Doppler effect when the sound source, like vehicles, is moving fast.

This paper proposed a moving sound source tracking method integrated TDOA method and kalman filtering, considering the Doppler effect of moving sound source to improve the tracking performance. GCC (Generalized Cross-Correlation) method is utilized to get the time delay estimation. To avoid the computational demanding solution of a set of non-linear equations for the exact sound source position, Chan algorithm is applied in localization step. And then a kalman filter is employed to smooth the trajectory. In the process of state updates, the elimination of Doppler effect in time domain is working simultaneously. Finally, the feasibility of the proposed method is verified by both simulation and outdoor experiment with a car as the target.

2. Localization method

The TDOA method is a two-step procedure. First, the time delay of signals between different microphone pairs is estimated. In the second step, the time delays are used in combination with the microphone array geometry to localize the sound source.

2.1 Time delay estimation

GCC(Generalized Cross-Correlation) method is the most common [8] and fastest two-channel algorithm for TDOA estimation. Let $x_i(t), x_j(t)$ be the signal of two microphones M_i, M_j, and the GCC function $R_{ij}(\tau)$ is defined as

$$R_{ij}(\tau) = \int_{-\infty}^{+\infty} \psi_{ij}(\omega) X_i(\omega) X_j^*(\omega) e^{i\omega\tau} d\omega \quad (1)$$

Where $X_i(\omega)$ and $X_j(\omega)$ are the DFT of $x_i(t)$ and $x_j(t)$, * denotes the complex conjugate, $\psi_{ij}(\omega)$ is the weighting function.

For real environments the PHAT weighting is the most used function, it's can be written as

$$\psi_{ij}^{PHAT}(\omega) = \frac{1}{\left| X_i(\omega) X_j^*(\omega) \right|} \quad (2)$$

So the relative time delay τ_{ij} is obtained by an estimation of the maximum peak in the GCC function

$$\tau_{ij} = \arg\max_{\tau} \left\{ R_{ij}(t) \right\} \quad (3)$$

2.2 Localization algorithm

To derive the source position from the TDOAs and the microphone array geometry, a set of non-linear equations need be solved. To achieve real-time performance and robustness of the tracking system, the Chan algorithm [9] is used in this paper.

Let the source be at unknown position (x, y, z) and the i^{th} microphone at known location (x_i, y_i, z_i) $(i = 1, 2, \cdots, M)$, then the squared distance between the source and the i^{th} microphone is given by

$$\begin{aligned} r_i^2 &= (x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2 \\ &= K_i - 2x_i x - 2y_i y - 2z_i z + x^2 + y^2 + z^2 \end{aligned} \quad (4)$$

Where $K_i = x_i^2 + y_i^2 + z_i^2$

Let $r_{i,1} = r_i - r_1$, $x_{i,1} = x_i - x_1$, $y_{i,1} = y_i - y_1$, $z_{i,1} = z_i - z_1$,so Eq. (4) can be written as

$$\begin{aligned} &(r_{i,1} + r_1)^2 = K_i - 2x_i x - 2y_i y - 2z_i z + x^2 + y^2 + z^2 \\ &\frac{1}{2}(r_{i,1}^2 - K_i + K_1) = -x_{i,1} x - y_{i,1} y - z_{i,1} z - r_{i,1} r_1 \end{aligned} \quad (5)$$

Where $K_1 = x_1^2 + y_1^2 + z_1^2$, $r_1 = \sqrt{(x_1 - x)^2 + (y_1 - y)^2 + (z_1 - z)^2}$

Let $\boldsymbol{z}_a = [x, y, z, r_1]^T$,then Eq. (5) can be looked as a set of linear equations with unknown vector $\boldsymbol{z}_a$, that is

$$\boldsymbol{h} = \boldsymbol{G}_a \boldsymbol{z}_a \quad (6)$$

Where

$$\boldsymbol{h}=\frac{1}{2}\begin{bmatrix} r_{2,1}^2-K_2+K_1 \\ r_{3,1}^2-K_3+K_1 \\ \vdots \\ r_{M,1}^2-K_M+K_1 \end{bmatrix}, \boldsymbol{G}_a=-\begin{bmatrix} x_{2,1} & y_{2,1} & z_{2,1} & r_{2,1} \\ x_{3,1} & y_{3,1} & z_{3,1} & r_{3,1} \\ \vdots & \vdots & \vdots & \vdots \\ x_{M,1} & y_{M,1} & z_{M,1} & r_{M,1} \end{bmatrix}$$

When $M>4$, a two-step Weighted Least-Squares (WLS) solution is utilized. The first ML estimate of $\boldsymbol{z}_a$ is

$$\boldsymbol{z}_a=(\boldsymbol{G}_a^T Q^{-1}\boldsymbol{G}_a)^{-1}\boldsymbol{G}_a^T Q^{-1}\boldsymbol{h} \tag{7}$$

Where $Q=diag(\sigma_1^2,\sigma_2^2,\cdots,\sigma_{M-1}^2)$, σ is the variance of TDOA estimation.

The solution of $\boldsymbol{z}_a$ assumes that x,y,z and r_1 are independent. Considering that they are related by Eq. (4) at i=1, so an improved estimation can be given with the second WLS.

Suppose the result of Eq. (7) is $\boldsymbol{z}_a^0=[x^0,y^0,z^0,r_1^0]^T$,and let $\boldsymbol{z}_{a,1}=x^0+e_1$, $\boldsymbol{z}_{a,2}=y^0+e_2$, $\boldsymbol{z}_{a,3}=z^0+e_3$, $\boldsymbol{z}_{a,4}=r_1^0+e_4$,where e_1,e_2,e_3,e_4 is the estimation errors of $\boldsymbol{z}_a$. Subtracting the first three components of $\boldsymbol{z}_a$ by x_1, y_1 and z_1 ,and then squaring the elements gives another set of equations

$$\boldsymbol{h}'=\boldsymbol{G}_a'\boldsymbol{z}_a' \tag{8}$$

Where

$$\boldsymbol{h}'=\begin{bmatrix} (z_{a,1}-x_1)^2 \\ (z_{a,2}-y_1)^2 \\ (z_{a,3}-z_1)^2 \\ z_{a,4}^2 \end{bmatrix}, \boldsymbol{G}_a'=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \boldsymbol{z}_a'=\begin{bmatrix} (x-x_1)^2 \\ (y-y_1)^2 \\ (z-z_1)^2 \end{bmatrix} \tag{9}$$

The ML estimation of $\boldsymbol{z}_a'$ is

$$\boldsymbol{z}_a'=(\boldsymbol{G}_a'^T\psi'^{-1}\boldsymbol{G}_a')^{-1}\boldsymbol{G}_a'^T\psi'^{-1}\boldsymbol{h}' \tag{10}$$

Where $\psi'=4B'\operatorname{cov}(\boldsymbol{z}_a)B'$, $B'=diag(x^0-x_1,y^0-y_1,z^0-z_1,r_1^0)$

The final position estimate is then obtained from $\boldsymbol{z}_a'$ as

$$[x,y,z]^T=\pm\sqrt{\boldsymbol{z}_a'}+[x_1,y_1,z_1]^T \tag{11}$$

The selection of sign + or – agrees with the first result $\boldsymbol{z}_a^0$.

3. Tracking method

When the observed position of an object is obtained, the system with the position, velocity and acceleration can be modelled. Hence, the next position is predicted, and the current information can be updated using next observed position. Kalman filter is widely used in tracking the object [10].

3.1 Kalman filter

The motion of the sound source can be modelled with a line time-discrete state space description by means of a state and an observation equation.

Suppose the state equation is defined as

$$X(k)=AX(k-1)+\Gamma W(k) \tag{12}$$

Where $X(k)=[x(k),y(k),z(k),v_x(k),v_y(k),v_z(k)]^T$ is the state vector, $W(k)$ is the process noise and $W(k)\sim N(0,Q(k))$, A is the state transfer matrix, and Γ is the noise input matrix.

Let the observation vector is $Z(k)=[x(k),y(k),z(k)]^T$,so the observation equation is

$$Z(k)=HX(k)+V(k) \tag{13}$$

Where H is the observation matrix, $V(k)$ is the observation noise and $V(k)\sim N(0,R(k))$.

If the target node moves in a line, the matrix A, H, Γ are defined as

$$A=\begin{bmatrix} I_{3\times3} & T\times I_{3\times3} \\ 0 & I_{3\times3} \end{bmatrix}, \Gamma=[0.5T^2, 0.5T^2, 0.5T^2, T, T, T]^T, H=[I_{3\times3}, 0_{3\times3}] \tag{14}$$

Where T is the sample period. In predication step, the time update equations are

$$\hat{X}(k \mid k-1) = A\hat{X}(k-1 \mid k-1) \tag{15}$$

$$P(k \mid k-1) = AP(k-1)A^T + \Gamma Q(k)\Gamma^T \tag{16}$$

Where $P(k)$ denotes the covariance matrix of state estimation. In correction step, the measurement update equations are

$$\hat{X}(k \mid k) = \hat{X}(k \mid k-1) + K(k)[Z(k) - H\hat{X}(k \mid k-1)] \tag{17}$$

$$P(k \mid k) = [I - K(k)H]P(k \mid k-1) \tag{18}$$

Where the kalman gain $K(k)$ is

$$K(k) = P(k-1 \mid k-1)H^T[HP(k-1 \mid k-1)H^T + R(k)]^{-1} \tag{19}$$

3.2 Doppler effect elimination

When the sound source moves at a high speed, the Doppler effect can cause the frequency shift of the signal received by the microphone. Hence, the elimination of Doppler effect can improve the positioning precision.

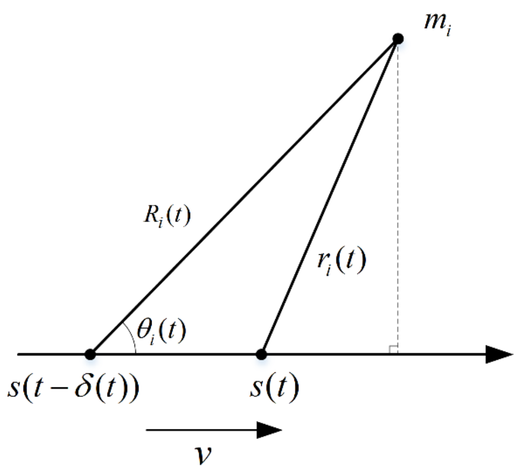

Figure 1: The Doppler effect model

The Morse model of the moving sound source is as Fig. 1. The source moves along the x axis at a uniform speed v, microphone m_i is fixed. At time $t-\delta(t)$, the sound source is at $s(t-\delta(t))$, where $\delta(t)=\frac{R_i(t)}{c}$, c is the sound speed and $R_i(t)$ is the distance between $s(t-\delta(t))$ and microphone m_i. The sound wave received by microphone m_i at time t is generated by sound source locating at $s(t-\delta(t))$ at time $t-\delta(t)$. According to Morse theory of acoustics [11], the sound pressure received by microphone m_i can be expressed as follows:

$$p_i(t)=\frac{q'\left(t-R_i(t)/c\right)}{4\pi R_i(t)\left(1-M\cos\theta_i(t)\right)^2}+\frac{q\left(t-R_i(t)/c\right)\left(\cos\theta_i(t)-M\right)v}{4\pi R_i^2(t)\left(1-M\cos\theta_i(t)\right)^2} \tag{20}$$

With the Doppler effect removed, the signal can be written as:

$$\tilde{p}_i(t) = \frac{R_i(t)\left(1 - M\cos\theta_i(t)\right)^2 p_i(t) + \left(M - \cos\theta_i(t)\right) v \int_0^t p_i(\tau)\,\mathrm{d}\tau}{r_i(t_0)} \cdot o\left(\frac{R_i(t) - r_i(t_0)}{c}\right) \qquad (21)$$

where $o(\xi)$ is defined as $o(\xi)\cdot f(t) = f(t)\cdot o(\xi) = f(t+\xi)$, and $M = \frac{v}{c}$.

In every updating step, the value of Eq. (11) can be used for the elimination of the Doppler effect. Then the result can be regarded as the observation value.

4. Tracking simulation

Simulation is carried out to evaluate the performance of the tracking algorithm. Since the trajectory of vehicle is two dimensional, the condition of the experiment is in the 2D-plane. Six microphones are utilized to get the sound signal, which are located at $(0.3m, 0.3m)$, $(-0.5m, 0.5m)$, $(-0.3m, -0.3m)$ and $(0.5m, -0.5m)$, $(0.4m, 0m)$, $(-0.4m, 0m)$, as Fig. 2 shows. The sound source is white noise. The start point is at $(1m, 0.5m)$, and the source moves at a speed of 5m/s along the y axis. The SNR of the background noise is 20dB, and the sampling frequency is 20 kHz. The sound data were analysed in frames of 20ms to assure quasi-stationary. The average of TDOA estimates in a block out of 12 consecutive frames is calculated. Thus it generates a new position observed estimate every about 0.2s.

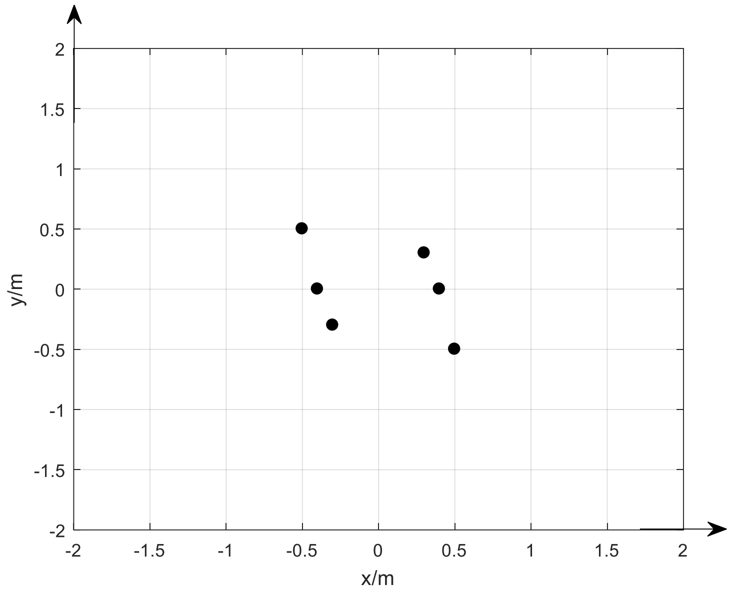

Figure 2: Microphone arrangement

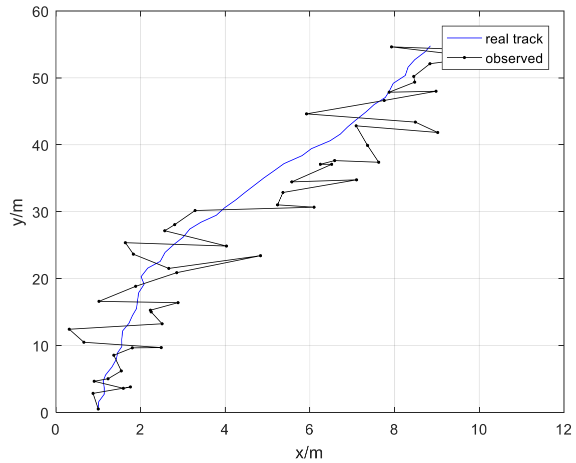

Figure 3: Observational trajectory before filtering

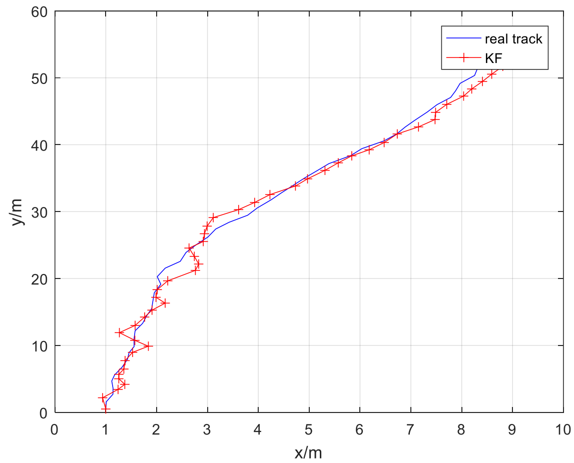

Figure 4: trajectory after filtering

Fig. 3 and Fig. 4 is the tracking result. In Fig. 3 the estimation error is significant only using TDOA algorithm. With the kalman filter, the tracking result is more smooth and reliable in Fig. 4. The RMS of estimation error is presented in Fig. 5. It indicates that the tracking performance gain much with the help of kalman filter, and the error of localization is stable compared with the TDOA method.

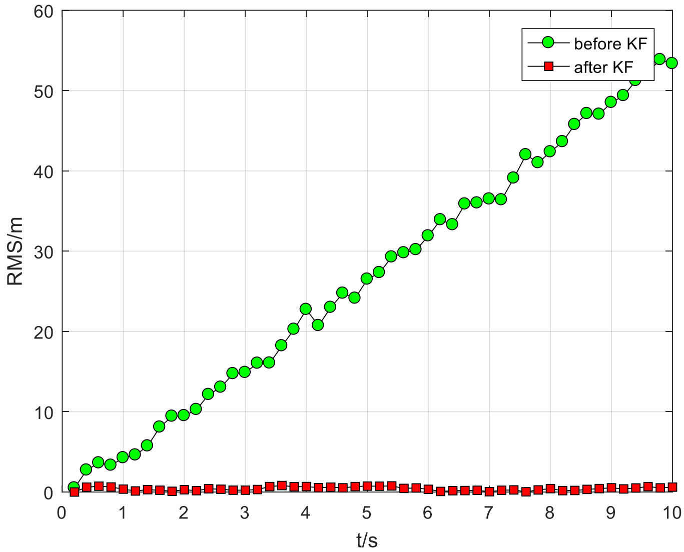

Figure 5: Comparison of position error before and after filtering

5. Experiment

Outdoor experiment was carried out to verify the effectiveness of the proposed algorithm. The setup is shown in Fig. 6.

Figure 6: the experiment setup

The structure of microphone array is the same as Fig. 2. A loudspeaker is fixed at the rear of the car, and the car moves roughly in a straight line at a speed of 20km/h in the front of the microphone array. The sampling frequency is 20 kHz. The loudspeaker generates white noise. The real position of the car is acquired by GPS, which can be regarded as the true value. Fig. 7 indicates the observed values only with TDOA method are unstable, especially when the distance increases. With kalman filter, the estimation errors is evidently less, and the trajectory is more close to the real, as Fig. 8 shows. The RMS of tracking result is given in Fig. 9. The proposed method can performance well in the outdoor environment. Meanwhile the error of position will increase with the distance from the microphone array. So the system need a short range of target in order to ensure the accuracy of tracking.

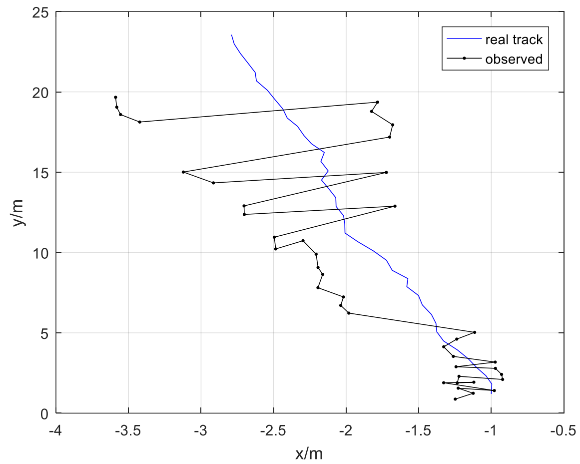

Figure 7: Tracking result only using TDOA method

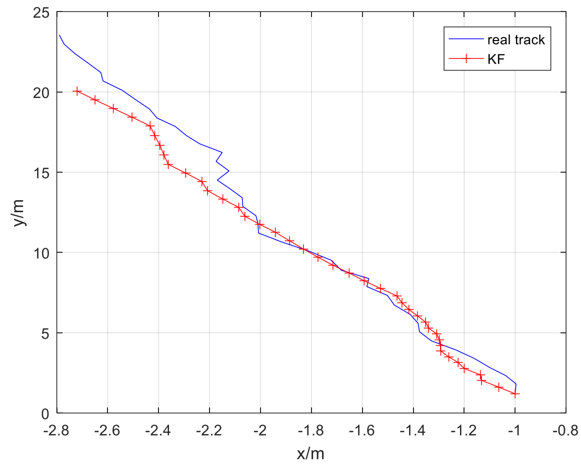

Figure 8: Tracking result with Kalman filter

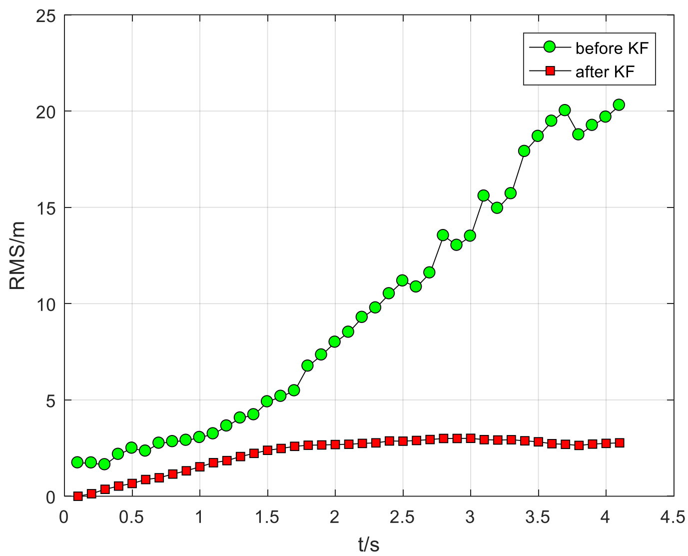

Figure 9: RMS before and after filter

6. Conclusion

This paper proposed a moving sound source tracking method integrated TDOA method and Kalman filtering, taking into account the Doppler effect. With 6 microphones, the system realized car tracking. It provide a novel strategy for vehicle detection. However the current method is only suit for single sound source, further work should be carried out to track multiple sound targets. Furthermore, the performance will degrade rapidly in strong noise environment. So the robustness of the algorithm to noise needs to be further improved.

Acknowledgement

This paper is supported by National Natural Science Foundation of China (51375252).

REFERENCES

1 Bechler D, Grimm M, Kroschel K. Speaker tracking with a microphone array using Kalman filtering. *Advances in Radio Science*, **1**(B. 3), 113-117, (2003).

2 Liang Z, Ma X, Dai X. Robust tracking of moving sound source using multiple model Kalman filter. *Applied acoustics*, **69**(12), 1350-1355, (2008).

3 Valin J M, Michaud F, Rouat J. Robust localization and tracking of simultaneous moving sound sources using beamforming and particle filtering. *Robotics and Autonomous Systems*, **55**(3), 216-228, (2007).

4 Chellappa R, Qian G, Zheng Q. Vehicle detection and tracking using acoustic and video sensors. *Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on.*, 3, IEEE, (2004).

5 Shon S, Kim E, Yoon J, et al. Sudden noise source localization system for intelligent automobile application with acoustic sensors, *Consumer Electronics (ICCE), 2012 IEEE International Conference on. IEEE*, 233-234, (2012).

6 Bechler D, Schlosser M S, Kroschel K. System for robust 3D speaker tracking using microphone array measurements, *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*. 3, IEEE, (2004).

7 Liang Z, Ma X, Dai X. Robust tracking of moving sound source using scaled unscented particle filter. *Applied Acoustics*, **69**(8), 673-680, (2008).

8 Knapp C, Carter G. The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics*, *Speech, and Signal Processing*, **24**(4), 320-327, (1976).

9 Chan Y T, Ho K C. A simple and efficient estimator for hyperbolic location. *IEEE Transactions on signal processing*, **42**(8), 1905-1915, (1994).

10 Kalman R E. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, **82**(1), 35-45, (1960).

11 Morse PM, Ingard KU. *Theoretical Acoustics*, Yang X, Lv R, Dai G, trans. Beijing: Science Press, (1986).