Analysis of a moving sound source
—Orbit estimation using microphone array—

Tsuyoshi Usagawa, Seizi Nishimura, Masanao Ebata, and Josuke Okda

Department of Electrical Engineering, Kumamoto University,
2-39-1, Kurokami, Kumamoto, 862 Japan
(Received 17 March 1986)

It is not easy to observe the original spectrum information of a moving sound source because of the Doppler effect. Previously, we proposed the method to compensate the Doppler effect to obtain the original spectrum, which is realized by compensating the propagation delay. This method requires the orbit information of target sound source to derive the propagation delay. This paper describes the method to estimate the orbit of moving sound source using sensor array. Proposed method is composed of two parts; the position estimation using a pair of relative propagation delays which are derived from the cross-correlation function, and the orbit estimation by dynamical filtering using the Kalman filter. Simple simulation was carried out to evaluate the performance of the proposed method. The result shows the sufficient accuracy of estimation both of orbit and spectrum.

PACS number: 43.28.Tc, 43.60.Gk

1. INTRODUCTION

When we try to observe the spectral information of a moving sound source, such as a train, an airplane, an automobile and so on, the Doppler effect which is usually observed as the frequency shift, prevents us from observing the original characteristics of signal radiated from the source. Since the Doppler effect is caused by the temporal change of propagation delay, its compensation can be realized by compensating the propagation delay. In order to carry out the compensation, we need information about the propagation delay; i.e. the distance between a moving source and a sensor when the sound velocity is known.

This paper describes the method to estimate the orbit of a moving sound source for the Doppler effect compensation. The proposed method is composed of two parts; the position observation using the cross-correlation technique and the orbit estimation using the adaptive filtering by means of the Kalman filter.

Simple simulation is carried out under the following condition; single point source, which radiates 500 Hz sinusoidal wave, moves on the line which is parallel to the sensor array. The result shows the effectiveness of the proposed method both for the orbit estimation and the spectrum analysis of observed signal.

2. PRINCIPLE OF POSITION OBSERVATION

The geometry of interest is shown in Fig. 1. The sensor array is composed of three microphones M1, M2 and M3, where the distances between sensors are \( L_1 \) and \( L_2 \), respectively. Upon observing the relative delays \( D_{11} \) and \( D_{32} \) which satisfy \( D_{ij} = (R_i - R_j)/c \), we can derive source position as the range \( R_2 \) and the bearing \( \theta_2 \).

That is,

\[
R_2 = \frac{L_1(1-(c\cdot D_{21}/L_1)) + L_2(1-(c\cdot D_{22}/L_2))}{2((c\cdot D_{21}/L_1)-(c\cdot D_{22}/L_2))} \quad (1)
\]

\[
\theta_2 = \cos^{-1}\left\{ \frac{L_1-2R_2\cdot c\cdot D_{22}-(c\cdot D_{22})}{2\cdot R_2\cdot L_2} \right\} \quad (2)
\]

where \( c \) is the sound velocity in an isovelocity medium. As shown in Fig. 1 and Eq. (2), the derived
position has a mirror image in another side of sensor array. However, this ambiguity of source position isn’t a problem to derive the absolute propagation delay, which is required for the Doppler effect compensation.

Relative delay $D_{21}$ and $D_{32}$ can be obtained by means of well-known correlation technique. The relative delay for a moving sound source varies as a function of position; i.e. as a function of time. Consequently the relative delay has to be determined with the adequate length data-segment. If the sound source were fixed, the accuracy of obtained delay can be improved with longer segment. However, in case of moving sound source, the accuracy has obvious limitations due to the movement of the source. The observed delay will be unstable if long segment is used to determine the delay. On the other hand, the maximum observable delay is limited by the segment length. Considering these limits of the segment length, the actual segment length was determined as the minimum value to cover the maximum relative delay.

The iteration process is shown in Fig. 2. First, as shown in Fig. 2 (a), the relative delay at $T/2$ is obtained from the cross-correlation function which is calculated using the first data-segment $\{0 \sim T\}$. That is, the relative delay is determined as the time lag at which the cross-correlation function has a maximum value. As shown in Fig. 2 (b), the second segment is shifted to $(T_n \sim T_n + T)$, which overlaps the first segment by the length $T - T_n$. And the relative delay at $T_n + T/2$ is obtained from the cross-correlation function derived from this segment. The maximum difference between the relative delays of first and second segment is obviously limited, so that the second delay is restricted around the previous delay.

This evidence is utilized to determine the maximum of cross-correlation as discussed in Sect. 4. The iteration of delay observation is carried out up to $\{T_t - T \sim T_t\}$ as shown in Fig. 2 (c).

In addition, considering the case that the time resolution $\Delta$ of the cross-correlation function could not detect a subtle change of the delay, the delay is determined by the linear interpolation. It is realized using the magnitude of the cross-correlation around the maximum point. Namely, as shown in Fig. 3, the interpolated delay $\bar{x}$ is given as follows:

$$\bar{x} = x_i - \frac{y_i-1 - y_{i+1}}{2(y_{i-1} + y_{i+1} - 2y_i)} \Delta$$

where $y_{i-1}$, $y_i$, $y_{i+1}$ are the magnitude of cross-correlation function at time lag $x_{i-1}$, $x_i$, $x_{i+1}$ respectively.
3. DYNAMIC ORBIT ESTIMATION
BY KALMAN FILTER

3.1 Linear Approximation of System

On the assumption that the source moves on the two dimensional plane, the system equation of moving source would be expressed as the sixth order linear equation as follows;

\[ x(k+1) = Fx(k) + w(k) \] (4)

\[ y(k) = Hx(k) + v(k) \] (5)

where

\[ F = \begin{bmatrix} 1 & A & A^2/2 & 0 & 0 & 0 \\ 0 & 1 & A & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & A & 0 \\ 0 & 0 & 0 & 0 & 1 & A \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad H = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \]

\( A \) is the sampling interval \((1/f_{\text{sample}})\), \( x_i \), \( x_s \) and \( x_j \) denote position, velocity and acceleration on \( X \) coordinate, respectively. \( x_4 \), \( x_5 \) and \( x_6 \) denote ones on \( Y \) coordinate.

\( w(k) \) is the random disturbance of state vector.

\( v(k) \) is the observation error of position \( y(k) \).

Though the disturbance of moving source \( w(k) \) is unknown, it is assumed to be the random acceleration when the source moves nearly constant speed. So in this paper, the disturbance of the state vector \( x(k) \) is assumed to be Gaussian random process just for acceleration components. The information of the observation error \( v(k) \) is necessary for accurate and stable estimation of orbit.

3.2 Position Observation Error

The cause of position observation error is classified into three levels; (i) error in observed delays between sensors, (ii) error in bearings, and (iii) error in position. These errors are sequentially connected.

The error of observed delay is due to the statistical characteristics of observed signal and the movement of source. The statistical characteristics of the delay estimation is discussed by Carter,\(^3\) and the Crâmer-Rao lower bound (CRLB) for the variance of estimation errors of delay is given as follows;

\[ \sigma_d^2 > \frac{2}{2T} \int_{-\omega_0}^{\omega_0} \frac{C(\omega)}{1-C(\omega)} \omega^2 d\omega \] (6)

where \( \omega_0 \) is the highest frequency of observed signal, \( T \) is the observation time length, and \( C(\omega) \) is the coherence function between two channel signals. As shown in Eq. (6), the error depends on signal to noise ratio (SNR), signal bandwidth, and observation time. Under the condition of the simulation carried out in this paper, this limitation is negligibly small compared with the resolution of the delay estimation; \( T = (\text{sampling frequency})^{-1} \). So the variance of the error due to the characteristics of observed signal is treated as the resolution of the cross-correlation function.

On the other hand, the error due to the movement of source is significant in the near field because of the rapid change of delay. Since the delay is given as the average over specific correlation time, the difference between the true delay and observed one is approximated as follows;

\[ \sigma_{d_m} = |\bar{d}(t) - d(t)| \]

\[ = \frac{1}{T} \int_{t-T/2}^{t+T/2} d(t)dt - d(t) \] (7)

where \( d(t) \) is the true relative delay at \( t \), and \( \bar{d}(t) \) is the estimated one.

Considering the statistical error and the error due to the movement of source, the error of delay can be treated as

\[ \sigma_d = \max\{\sigma_d, \sigma_{d_m}\} \] (8)

Then the bearing of source from the center of paired sensors is derived from the delay. The variance of observation error of bearing is described as a function of the bearing of source and the variance of delay estimation as follows;

\[ \sigma_{B_i} = \frac{c}{L_i \sin B_i} (i=1, 2) \] (9)

where \( L_i \) is a distance between sensors and \( B_i \) is a bearing of source from the center of paired sensors \( M_i \) and \( M_{i+1} \).

From the geometrical relationship between the bearings and the position, the estimation error of position can be shown as follows;

\[ \sigma_x = \max\{r_i \sin \sigma_{B_i} \sin B_i\} \]

\[ \sigma_y = \max\{r_i \sin \sigma_{B_i} \cos B_i\} \] (10)

where \( r_i \) is the distance between the source and the center of paired sensors. It is given as

\[ r_i = \frac{(L_i + L_i) \sin B_i}{2 \sin (B_i - B_i)} \] (11)
where \( i = 1, 2 \) and \( j \neq i \).

Under the condition that the source moves at 100 m/s (360 km/h) on the line which is parallel to the sensor array, the variance of position observation error, which is derived from Eqs. (10) and (11), is shown in Fig. 4.

4. CONFIGURATION OF THE ORBIT ESTIMATION METHOD

The diagram of the proposed method to estimate orbit using three microphones is shown in Fig. 5. The estimated position \( \hat{y}_{k} \), which is derived from the cross-correlator and position estimator, is the input signal of the Kalman filter, and the observation error, which is calculated using Eq. (10) in the previous section, is introduced into the system.

When the cross-correlation of observed signal has a periodicity, it is very difficult to determine the delay because of the appearance of many local peaks in the cross-correlation function. Also there is the same kind of difficulty when there are two or more sound sources. However as mentioned in Sect. 2, the difference between relative delays of each observation interval is limited. The upper limit of the difference is given as the maximum change of delay, and it is given as \( 2 \cdot T_{s} \cdot V/c \), where \( V \) is the velocity of a moving source and \( T_{s} \) is the interval of the delay observation. So that the observable range of next relative delay, which is derived from the output of the Kalman filter, is utilized to trace the peak of the cross-correlation function as shown in Fig. 6. This treatment has the advantage not only in detecting the delay but also in the calculation time of the cross-correlation function. Because the observing delay is restricted around the predicted delay which is derived from the estimated orbit of source, neither

![Fig. 4 Predicted error in position observation. Velocity of source = 100 m/s [360 km/h].](image)

![Fig. 5 Diagram of the orbit estimator.](image)

![Fig. 6 Block diagram of delay tracking procedure in orbit estimation.](image)
periodicity nor local peak of cross-correlation function affects to the delay determination. Furthermore, we need to calculate only the limited range of the cross-correlation function. So the calculation time of the cross-correlation function is reduced by the factor of $k$, which is expressed by the following equation.

$$k = \frac{\text{the range of delay}}{\text{segment length}}$$

As the signal to noise ratio (SNR) increases in the near field of sensor array, the observed delay in this field is least affected by the surrounding noise. This delay is utilized as the initial value to drive the Kalman filter, and the Kalman filter works both forward (increasing time) and backward (decreasing time).

The actual Kalman filter algorithm used in this paper is Bierman-Thornton's UD filter which has an advantage in stability.\(^4\)

5. SIMULATION

5.1 Simulation Model

Figure 7 shows the geometry of the simulation model. The source moves at 8.33 m/s (30 km/h) on a line which is parallel to the sensor array. The signal radiated by the source is a 500 Hz sinusoidal wave. And uncorrelated background noise is introduced at each sensor.

5.2 Result of Orbit Estimation

Figure 8 shows the result of orbit estimation; the dotted line shows the directly estimated orbit ($\hat{y}_k$) and the solid line shows the filtered orbit ($\hat{y}_k$) using the variance of position observation error discussed in the previous section. The average velocity, the time when the source passed in front of sensor M2, and the distance between the orbit and sensor array are 8.50 m/s, 1.19 s, and 2.95 m respectively, where the true values are 8.33 m/s, 1.20 s and 3.00 m respectively. The estimation errors of these value are 2.04%, 0.83% and 1.67%, respectively. As shown in Fig. 8, the filtered orbit is smoother and stabler than the directly observed orbit, because the filtering algorithm reduced the random error in position observation. According to the simulations under various initial conditions, the initial condition, especially the initial value of velocity, affects the stability of filtering algorithm. In the field experiment, the velocity of source would be predictable by other methods. So this requirement for the initial condition would not be so serious.
5.3 Result of Spectrum Estimation

The results of spectrum analysis are shown in Figs. 9 and 10.

The directly observed auto-spectrum at sensor M₂ is shown in Fig. 9. Also it shows the peak frequency of auto-spectra. The auto-spectra have obvious peaks around 500 Hz, however peak frequency is shifted due to the Doppler effect.

The result of a spectrum analysis with the Doppler effect compensation using the orbit information is shown in Fig. 10. As shown in this figure, the frequency shift of 500 Hz signal is reduced. It means that the compensation of the Doppler effect is effectively carried out using the estimated orbit.

6. CONCLUSION

The method to estimate the orbit of a moving sound source by the cross-correlation technique is described where the sensor array is composed of 3 microphones and the orbit of source is parallel to the sensor array. The proposed method utilizes an adaptive filtering algorithm to reduce the random observation error. The result of the simulation shows the effectiveness of this method and the sufficient accuracy of observed orbit for the Doppler effect compensation.

ACKNOWLEDGMENT

Authors gratefully acknowledge the contribution of Mr. Kouji Matsuno.

REFERENCES