The 21st International Congress on Sound and Vibration
13-17 July 2014, Beijing/China

BIOLOGICALLY INSPIRED ACOUSTIC LOCALIZATION SYSTEM FOR DIRECTION OF ARRIVAL ESTIMATION

Yaqiong Zhang, Xinlei Zhu, Ming Yang, Na Ta, Zhushi Rao

The State Key Laboratory of Mechanical System and Vibration, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China 200240

E-mail: zsrao@sjtu.edu.cn

The Ormia ochracea, a kind of parasitoid fly, is known to have a mechanical coupling between its ears, which has a special function of enhancing its hearing. Based on this mechanism of coupling, we designed a biologically inspired acoustic localization system, which has three degrees of freedom, for the purpose of estimating the direction of arrival (DOA) in three-dimensional space. Thanks to this kind of coupling, the system shows a high performance on the DOA estimation. In this paper, we first give the vibration differential equations governing the system’s response. Then the response performances under different incident frequencies or incident angles are analyzed. The response differences are magnified because of the coupling, which shows a probability to apply this system to the DOA estimation. By computing the Cramér–Rao bound (CRB), the accuracy of estimating the DOAs of this system is analyzed. Numerical examples are given to compare the CRB for the coupled and the uncoupled systems both for single and multiple incident sound sources, which illustrate the effect of the biologically inspired acoustic localization system on reducing the errors in estimating the DOAs.

1. Introduction

Microphone array, which has time differences of arrival of sound sources between the transducers, is always applied to the sound source localization. Since the positioning accuracy is directly determined by the element spacing of such arrays, large arrays are often required. However, this is costly and may narrow their practical applications. This paper demonstrates a biologically inspired acoustic localization system with very small element spacing, which can estimate the directions of arrival (DOAs) in three-dimensional space, based on the orientation mechanisms of a parasitic fruit fly called the Ormia ochracea.

The fly is well known to have a remarkable ability to detect the direction of an incident sound stimulus. In order to propagate, the female Ormia ochracea must deposit her larvae on a live field cricket, the specific host for the parasite, relying on the cricket’s mating call. However, the interaural distance separating the fly's hearing organs (450 to 520μm) does not match with the wavelength of the calling song (about 7 cm), which is relatively pure in frequency (peak frequency of 4.8 kHz). Unlike most ears, the fly has a mechanical coupling between its ears. And it is believed that the coupling mechanism magnifies the small interaural intensity and time of arrival differences between
two ears and subsequently improves the localization accuracy.\textsuperscript{1-6} Scanning electron micrograph of the tympanal membranes of the fly and the mechanical analogy model are shown as Figure 1.

![Figure 1](image)

**Figure 1.** (Up) Scanning electron micrograph (frontal view) of the tympanal membranes and (Down) mechanical model abstracted from the tympanal membranes. Dots ○ 1 and ○ 2 indicate the position of the tympanic pits and dot ○ 3 shows the pivot point of the intertympanic pits\textsuperscript{3}.

In this paper, we demonstrate a 3D biologically inspired acoustic localization system, which employs the coupling mechanism. First, the mechanical model of the 3D acoustic localization system is illustrated. And the vibration differential equations governing the system’s response are given. Then the response performances under different incident frequencies or incident angles are analyzed. Though comparing the coupled and uncoupled versions of the system, the effect of coupling on intensity and time of arrival differences are demonstrated. A statistical model with stochastic sources and additive noise is then developed to analyze the effect of the system on sound source localization. The asymptotic Cramér–Rao bound (CRB) on estimating the DOAs of sound sources is computed in space. Finally, numerical examples are presented, which show an obviously improvement in the localization performance of this 3D biologically inspired localization system.

## 2. Modeling

The 2D model can describe the direction of arrival in a plane, but the sound spreads in 3D space in real situation, and the sound source may not in the hypothetical plane. So a 3D bionic localization system is designed.\textsuperscript{7-10} In this section, basing on the 3D sound source localization system, the response performance for the coupled and the uncoupled cases for a far-field sound source are compared. Assuming the incoming source is the narrow-band signal, whose center frequency is $\omega$.

A simple mechanical model, composed of three mass-spring-dashpot systems, is shown as Figure 2. The three systems are placed as an equilateral triangle and connected with each other through a coupling spring $k_c$. To simplify, the coupling damping are not taken into consideration. The mass, stiffness and damping parameters of each system are $m$, $k$ and $c$. Establish a coordinate system (see Figure 3). The origin of coordinates represents the pivot point. And the geometric relationship between the sound source (P) and the three masses (1,2,3) is shown in the figure.

$$
\begin{bmatrix}
m & 0 & 0 \\
0 & m & 0 \\
0 & 0 & m
\end{bmatrix}
\begin{bmatrix}
\ddot{x}_1 \\
\ddot{x}_2 \\
\ddot{x}_3
\end{bmatrix}
+ 
\begin{bmatrix}
c & 0 & 0 \\
c & 0 & 0 \\
c & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2 \\
\dot{x}_3
\end{bmatrix}
+ 
\begin{bmatrix}
k+2k_c & k_c & k_c \\
k_c & k+2k_c & k_c \\
k_c & k_c & k+2k_c
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix}
= 
\begin{bmatrix}
f_1 \\
f_2 \\
f_3
\end{bmatrix}
$$

(1)
Where \( f_i(t) = s p_i(t) \) \((i=1,2,3)\), where \( p_i(t) \) is the sound pressure applied at each mass, \( s \) is the surface area of each of them. And \( x_i(t) \) is the unknown displacement response of each mass.

Eq. (1) may be used to determine the transfer functions between each response coordinate, \( x_i(t) \), and the incident pressure at the pivot point, \( p(t) \). Assuming that there is an equal area \( s \) at the pivot point, then the incident force can be written as \( f(t) = sp(t) \). To simplify the notation, let \( H_{f,i}(\omega) (i=1,2,3) \) denotes the transfer function between the force at each mass \( f_i(t) \) and the equivalent force at the pivot point, \( f(t) \). Then \( H_{f,i}(\omega) = e^{j\omega t} \).

Where

\[
\begin{align*}
\tau_1 &= \frac{d}{c} \cos \theta \cos \alpha \\
\tau_2 &= \frac{d}{c} \cos \theta \cos(\alpha + \frac{2\pi}{3}) \\
\tau_3 &= \frac{d}{c} \cos \theta \cos(\alpha - \frac{2\pi}{3})
\end{align*}
\]

\( \tau \)

\[ \text{(2)} \]

**Figure 2.** Mechanical model of the 3D sound source localization system

**Figure 3.** Geometric relationship between sound source and three masses

Where \( d \) is the circumradius of the three masses. By solving Eq. (1), the vector of transfer functions between the responses of the masses and the force at the pivot point may be shown to be

\[
\begin{bmatrix}
H_{x,i}(\omega) \\
H_{y,i}(\omega) \\
H_{z,i}(\omega)
\end{bmatrix} = \begin{bmatrix}
-k \omega^2 + j \omega c + k + 2k_c & k_c & k_c \\
k_c & -k \omega^2 + j \omega c + k + 2k_c & k_c \\
k_c & k_c & -k \omega^2 + j \omega c + k + 2k_c
\end{bmatrix}^{-1} \begin{bmatrix}
e^{j\omega t_1} \\
e^{j\omega t_2} \\
e^{j\omega t_3}
\end{bmatrix}
\]

\[ \text{(3)} \]

To simplify the notation, rewrite the vector of transfer functions as \( H_{m} \). Apparently, it is related to the incident sound source. If there are multiple sources, \( H_{m} \) denotes the vector of transfer functions of the system when the \( m \)th source works alone.

The amplitude responses of the system for the coupled and uncoupled cases are shown in Figure 4. The incident angle of the sound source is \((\theta, \alpha) = (45^\circ, 135^\circ)\). It shows that in a certain frequency range (about 4kHz to 10kHz), the amplitude responses for the coupled system are much different from each other. If the frequency is set at 5kHz, when azimuth angle \( \alpha \) varies from 0 to \( 2\pi \), we can get Figure 5 and Figure 6, which show the variations of phases and amplitudes of \( H_{x,i}, H_{y,i} \) and \( H_{z,i} \). Figure 5 confirms the improvement of the intensity differences among the masses for the coupled system. Similarly, Figure 6 demonstrates how the phase differences are am-
plified, so are the differences in arrival times of the sound source to the three masses. Apparently, the very small differences in intensity and arrival time are increased by the coupling. That is, the mechanical coupling effectively creates a larger element spacing for the system.

![Figure 4](image1.png)

**Figure 4.** Amplitude responses for coupled and uncoupled cases, incident angle \((\theta, \alpha) = (45^\circ, 135^\circ)\)

![Figure 5](image2.png)

**Figure 5.** Amplitudes of \(H_{x_1f}, H_{x_2f}\) and \(H_{x_3f}\) when azimuth angle varies from 0 to \(2\pi\). The frequency is set at 5kHz.

![Figure 6](image3.png)

**Figure 6.** Phases of \(H_{x_1f}, H_{x_2f}\) and \(H_{x_3f}\) when azimuth angle varies from 0 to \(2\pi\). The frequency is set at 5kHz.

### 3. Performance analysis

The mean-squared error for any unbiased estimator has a lower bound, the Cramér–Rao bound (CRB), which defines the ultimate accuracy of any estimation procedure. In this section, a statistical model is presented and the CRB on DOA estimation of sound sources for the localization system is computed.
3.1 Statistical model

The model consists of $M$ multiple stochastic sound sources $s_m(t)$ ($m=1,2,\ldots,M$), which are all narrow-band signals, additive noise $e(t)=[e_1(t), e_2(t), e_3(t)]^T$ and output vector of three measure points $y(t)$. Then the model gives rise to

$$y(t) = Hs(t) + e(t), t = 1,2,\ldots,N$$

(4)

Where $H = [H_1, H_2,\ldots, H_M]$ and $H_m$ is the frequency response vector of the system related to the sound source signal $s_m(t)$ with incident angle $\phi_m = (\alpha_m, \theta_m)$. We can get the frequency response vector $H_m$ through Eq. (3).

It is assumed that $s_m(t)$ and $e(t)$ are both zero mean Gaussian random processes. For simplicity, the following assumptions are made: $e_1(t), e_2(t)$ and $e_3(t)$ are white; and have the same variance $\sigma^2_e$; $s_m(t)$ ($m=1,2,\ldots,M$) and $e(t)=[e_1(t), e_2(t), e_3(t)]^T$ are uncorrelated with each other for all $t$;

$$E[s(t)s(t)'^H] = P\delta_{tt}, \text{ and } E[s(t)s(t')^H] = 0 \quad, \text{with } P \text{ as the } M \times M \text{ unknown source covariance matrix, and for } t, t'=1,2,\ldots,N \text{, } \delta_{tt'} = 1 \text{ when } t=t' \text{ and zero otherwise; } E[e(t)e(t')^H] = \sigma^2_e I \delta_{tt'} \text{ and } E[e(t)e(t')^H] = 0; \text{ and } (\cdot)^H \text{ denotes the conjugate transpose.}

3.2 Cramér–Rao bound

We analyze the system’s localization performance by computing the Cramér–Rao bounds on the azimuth and elevation estimation error. Following the statistical assumptions above, we write the probability density function of the measurements as

$$p(y(t); \psi, \sigma^2_e) = \prod_{i=1}^{N} \frac{1}{\pi^3 |R_y|} \exp[-y^H(t)R_y^{-1}y(t)]$$

(5)

Where $\psi = [\alpha_1, \alpha_2,\ldots, \alpha_M, \theta_1, \theta_2,\ldots, \theta_M]^T$ is the $2M \times 1$ vector of unknown parameters (azimuth and elevation for each source), $R_y = E[y(t)y(t)^H] = HPH^H + \sigma^2_e I$.

Let $\Theta = [\alpha_1, \alpha_2,\ldots, \alpha_M, \theta_1, \theta_2,\ldots, \theta_M, p_1,\ldots, p_M, \sigma^2_e]^T$, the $3M+1 \times 1$ unknown parameter vector. Where $p_m (m=1,2,\ldots,M)$ is the element on the diagonal of matrix $P$. Then, the elements of the Fisher information matrix can be found as follows:

$$[I(\Theta)]_{ij} = Ntr[R_y^{-1} \frac{\partial R_y}{\partial \Theta_i} R_y^{-1} \frac{\partial R_y}{\partial \Theta_j}] \quad i,j=1,2,\ldots,3M+1$$

(6)

Through the concentrated stochastic likelihood function and the CRB estimate of parameters, we can get

$$[I(\Theta)]_{ij} = \frac{2N}{\sigma^2_e} \text{Re}\{tr[UD_i'^H\Pi^1 D_i]\}$$

(7)

Where $U = PH^H R_y^{-1} HP, D_i = \frac{\partial H}{\partial \psi_i} \Pi^1 = I - \Pi, \Pi = HH^H, H^+ = (H^H H)^{-1}H^H$. Hence, the CRB of estimating the DOA $\psi_i (k=1,2,\ldots,2M)$, can be found as:

$$\text{CRB}(\psi_i) = \frac{\sigma^2_e}{2N} \left[\text{Re}\{(D_i'^H\Pi^1 D) \odot U^T}\right]^{-1}$$

(8)
4. Numerical results

The CRBs of estimating the DOAs for the coupled and uncoupled systems are compared. And the coupled system shows a much better localization performance with higher localization accuracy.

In Figure 7, we use the following scenario: single sound source with true value of the azimuth $\alpha$ and elevation $\theta$ as $(\theta, \alpha) = (45', 45')$; and the frequency is 5kHz; $d=0.01 \text{m}$ circumradius. We plot the CRBs on localization error for different signal-to-noise ratio (SNR) values. The SNR is a function of the ratio of each signal power and the noise power. We observe that the CRBs on DOA estimation error are smaller for the coupled system, for both azimuth angle and elevation angle, meaning a decrease in estimation error and an improvement in the localization accuracy.

In Figure 8, the number of sound sources is 2 ($M=2$). The incident angles of the two sources are $(\theta_1, \alpha_1) = (60', 60')$ and $(\theta_2, \alpha_2) = (11.25', 11.25')$; and their frequencies are 5kHz and 5.5kHz, respectively; $d=0.01 \text{m}$. We can also see that the coupled system decrease the minimum bound on the estimation error, demonstrating a better performance of the DOA estimation. While compared with Figure 7, the estimation errors are much larger. It is observed that when the number of sources increases, higher SNR is needed. So does the number of the time samples for the same CRB values. Even so, the localization performance of the system is largely limited to multiple sound sources, which inspiring us to extend the coupled system to a kind of array. In our future work, we will combine the system with array signal processing to solve the problem of multiple sources localization.

![Figure 7. Square-root of the CRB on the DOA estimation vs SNR for single source](image1)

![Figure 8. Square-root of the CRB on the DOA estimation vs SNR for 2 sources](image2)

5. Conclusion

We designed an acoustic localization system inspired by the hearing system of the *O. ochracea*. And the effect of the system was analyzed. By examining the frequency responses of the system, it was shown that, in a certain frequency range (about 4kHz to 10kHz), the coupled system can distinctly increase the intensity and arrival time differences among measure points. That’s to say, it creates a virtual system with a larger element spacing. As a performance analysis, by establishing a statistical model, the mean-squared error bound on estimating the incoming angular of the sound source was asymptotically computed. Numerical results regarding the localization accuracy of the system are given. Through comparing the Cramér–Rao bounds of the coupled and the
uncoupled cases, the improvement in localization performance was confirmed for this acoustic localization system. Therefore, there may be a chance to change the localization performance of microphone array. In the future work, we will extend this coupled system to an array for multiple sources localization and the performance of the system will be improved in the meantime.

REFERENCES