INVESTIGATION ON THE INFLUENCE OF THE INTERAURAL PHASE DIFFERENCE ON THE SOUND LOCALIZATION AND AUDITORY SOURCE WIDTH AT DIFFERENT FREQUENCY BANDS

Peng Wang, Zhibin Lin
Key Laboratory of Modern Acoustics, Institute of Acoustics, Nanjing University, Nanjing, China
email: pwang@smail.nju.edu.cn

Xiaojun Qiu
Centre for Audio, Acoustics and Vibration, Faculty of Engineering and IT, University of Technology Sydney, Sydney, Australia

The phase adjustment method is popular to broaden auditory source width (ASW), however, the influence of the interaural phase difference on the changes of sound image at different frequency bands has attracted less attention in recent years. In this paper, the influence of phase adjustment on the sound localization and the ASW at different frequency bands was investigated, the performance of phase adjustment on the sound localization below 1600 Hz was obtained and the variation of the ASW above 1600 Hz was verified. Experimental results show that the sound image moves to one side as the interaural phase difference shifts from 0° to 90° at frequency bands below 1600 Hz, and the ASW broadens at frequency bands above 1600 Hz as the interaural phase difference shifts from 0° to 90°. The maximum angle offset appears at 200 - 400 Hz and the ASW broadens about 20° at frequency bands above 6400 Hz.

Keywords: interaural phase difference, sound localization, auditory source width

1. Introduction

Phase adjustment is introduced common method in stereo enhancement techniques to broaden the auditory source width (ASW) [1]. However, the influence of phase adjustment on the changes of the sound image in different frequency bands has not been discussed quantitatively.

Early studies have shown that the interaural phase difference (IPD) is an important factor of sound localization when the interaural time difference (ITD) and the interaural level difference (ILD) are constricted to zero [2]. The influence of the phase adjustment on stereo sound reproduction by loudspeakers has been investigated [3]. It has shown that the shift of the phase within 90° will bring about the movement of the sound image to left or right, which will produce a certain effect on widening the sound image. In addition, Xie [4] conducted a preliminary analysis of the effect of phase adjustment on the ASW in different frequency bands using the loudspeakers to playback stereo sound. When the phase-shift is 90°, the sound field has a broadening effect in the frequency bands above 700 Hz.

The effect of the phase adjustment on the sound localization by headphones has been investigated in the different frequency bands. Existing studies have shown that phase shifts can cause a shift in the perception of the sound image at frequencies below 2000 Hz, while at frequencies above 2000 Hz, the influence of phase changes on the sound localization can be neglected [5].
In this paper, the quantitatively researches on the changes of sound localization in the frequency bands below 1600 Hz were carried out. Meanwhile, the effect of the interaural phase difference on the ASW was quantitatively studied in the frequency bands above 1600 Hz. A quantitative experiment method is adopted to investigate the changes of the sound image when the phase changes from 0° to 90° as the frequency varying from 100 Hz to 12800 Hz.

2. Experimental setup and methodology

In order to undertake the experiment in a controlled manner, special stimuli were introduced in this paper. A more detailed discussion of the requirements of the stimuli were described by Mason [6].

To conduct the comparative experiments, two types of stimuli were adopted, which were fixed stimuli and variable stimuli. The fixed stereo signals, \( l \) and \( r \), are sine signals created by Eqs. (1) and (2): where \( f \) is the centre frequency of the single tone and \( \phi \) is the phase-shift in each channel.

\[
l = \sin(2\pi f t + \phi) .
\]

\[
r = \sin(2\pi f t - \phi) .
\]

To avoid the interference of other factors, only the phase between the left and right channel is changed. A range of values of \( \phi \) from 0° to \( \pi/4 \) results in a range of values of phase-shift of 0° to 90° between the left and right channels, meanwhile, the changes of the interaural cross-correlation coefficient (IACC) are inevitable and the values of IACC are varying from 1 to 0 monotonously.

The centre frequency in the left and right channels is set by altering the frequency \( f \) in Eqs. (1) and (2) and the values of \( f \) were used with octave: 100, 200, 400, 800, 1600, 3200, 6400 and 12800 Hz, which covers the frequency range of most music.

\[
s = \sin(2\pi f t + \phi) \ast h .
\]

The variable stimuli \( s \), which are the reference signals, were obtained by convolving the single frequency signal with the HRTF filter coefficients. The reference signals, were generated by Eqs. (3), with \( f = 300 \) Hz and \( \phi = 0° \), which can provide an accurate sense of spatial angle after convolved with the HRTF filter coefficients. The HRTF data of the MIT with a vertical elevation angle of 0° and a horizontal opening angle of 270° to 360° were adopted to generate reference signals of specific spatial angles. The spatial resolution is 5° and it is accurate enough to distinguish the difference in image position between the fixed stimuli and the variable stimuli. For convenience of explanation, the angle of the reference signal of 270 - 360° is defined as 0° of the midline to 90° of the left side.

All the stimuli were generated by MATLAB, which allowed variation of different phase-shift from 0° to 90°. The listening equipment was a pair of AKG K702 headphones.

All the signals were constricted to a sound loudness level corresponding to 71 phons at each ears. Specially, the sound loudness levels of the left and right channels in reference signals are different and the average sound loudness level of the left and right channels was adjusted to 71 phons.

The experiments were conducted using a “method of adjustment” paradigm, where the listeners were asked to adjust a variable stimulus to be the same perceived auditory source width as a given fixed signal [7]. The advantages of this experimental method are relatively effective and highly reproducible [8]. The fixed stimulus was presented firstly and followed by the variable stimulus. The subjects can repeatedly compare two stimuli until they find the closest reference signal. The subjects can only judge perceptual feeling whether the two signals match, without any specific information of the fixed signal and the variable signal, and this can be considered as double blind test.

All experiments were performed in the standard listening room with low background noise and suitable reverberation time. A total of 14 experient listeners were recruited for the listening test, including 11 males and 3 females. The listeners have normal hearing during the regular physical examination and the individual listening results are consistent.
3. **Influence of IPD on the sound localization at frequency bands below 1600 Hz**

There were 20 fixed stimuli in this experiment: five of them were 0° phase-shift signals and the rest were 30° phase-shift signals, 60° phase-shift signals and 90° phase-shift signals, respectively. The centre frequency of the fixed stimuli was set by altering the frequency \( f \) in Eqs. (1) and (2) and the values of \( f \) were used with octave: 100, 200, 400, 800 and 1600 Hz. The phase of left channel is set 0° to 90° ahead of the right channel to ensure the location of the sound image moving from the midline of the head to the left of the head, which matches the angles of the variable stimuli.

During the listening test, only one stimulus was selected to compare with the reference signals at a time. According to the feedback of the listeners, the spatial angle of the reference signal was adjusted until the listeners ensured that the positions of reference signal was the same with the fixed signal. Then the spatial angle of the reference signal was recorded as the position of the fixed signal.

The results of the subjective experiment were shown in Fig. 1, with the centre frequency of each fixed stimulus along x-axis and the angle of the matched variable stimulus on y-axis. In fact, the variable reference signals were synthesized from fixed-angle HRTF data, which brings in little mismatch between the angle of the reference signal and the angle of the real space.

![Figure 1: Means and associated 95% confidence intervals of the angle of variable stimuli that matched the fixed stimuli with the phase-shift from 0° to 90° at different frequency.](image)

The solid line in Fig. 1 represents the 0° phase-shift and the sound localization is still at the midline of the head. The ASW decreases as the centre frequency of the fixed stimuli varying from 100 Hz to 1600 Hz, which is consistent with the study of Mason [9]. The widest ASW appears at 100 Hz and is approximately 41.4° (the twice of the angle of the matched variable stimuli). Meanwhile, the minimal value of the ASW occurs at 1600 Hz and is about 8.6°.

The dotted line in Fig. 1 represents the results of 30° phase-shift. The position of the sound image begins to shift to the left and the angle of the shift is decreasing as the frequency of fixed stimuli varying from 100 Hz to 1600 Hz. The maximum value of the angle appears at 100 Hz and is approximately 31.1°. Meanwhile, the minimal value occurs at 1600 Hz and is about 7.4°.

The dashed line in Fig. 1 represents the results of 60° phase-shift and it has the same change trend as the 30° phase-shift. The maximum value of the angle appears at 200 Hz and is approximately 37.9°. Meanwhile, the minimal value occurs around 1600 Hz and is about 11.8°.

The dash-dot line in Fig. 1 represents the results of 90° phase-shift and it has the same change trend as the 30° phase-shift. The maximum value of the angle appears at 100 Hz and is approximately 43.9°. Meanwhile, the minimal value occurs at 1600 Hz and is about 16.1°.
The means and associated 95% confidence intervals of the results in Fig. 1 shows that the maximum of the movement of sound image occurs at the 90° phase-shift around 100-200 Hz and the minimum appears at 1600 Hz.

It is reasonable to expect that the sound localization moves to one side of the head as the phase-shift varying from 0° to 90°, with the frequency varying from 100 Hz to 1600 Hz. This phenomenon can be explained by the characteristics of the wave-front of the fixed stimuli [3].

It can be interpreted from these results that the angle of the perceived offset achieves the top at low frequency and decreases smoothly to mid-frequency (approximately 1600-2000 Hz) as the phase-shift existing. This results show that sound image shifts can be created by changing the phase between left and right channels at low frequency.

4. Influence of IPD on the ASW at frequency bands above 1600 Hz

There were 16 fixed stimuli in this experiment: four of them were 0° phase-shift signals and the rest were 30° phase-shift signals, 60° phase-shift signals and 90° phase-shift signals, respectively. The centre frequency of the fixed stimuli was set by altering the frequency \( f \) in Eqs. (1) and (2) and the values of \( f \) were used with octave: 1600, 3200, 6400 and 12800 Hz, in which the sound localization don’t change as the phase-shift varying from 0° to 90°.

During the listening test, only one stimulus was selected to compare with the reference signals at a time. According to the feedback of the listeners, the spatial angle of the reference signal was adjusted until the listeners ensured that the position of reference signal was the same with the edge of the sound image of the fixed signal. In this way, the angle of variable stimuli is only the position of the edge of the actual sound image (the leftmost position of the sound image in this experiment), and the ASW is about twice that of this angle.

The results of the subjective experiment were shown in Fig. 2, with the centre frequency of each fixed stimulus along \( x \)-axis and the ASW calculated from the matched variable stimulus along \( y \)-axis.

![Figure 2: Means and associated 95% confidence intervals of the ASW calculated from variable stimuli that matched the fixed stimuli with the phase-shift from 0° to 90° at different frequency.](image)

The solid line in Fig. 2 represents the 0° phase-shift and the ASW of the sound image broadens as the frequency of fixed stimuli rising from 1600 Hz to 12800 Hz. The maximum ASW appears at 12800 Hz and is approximately 18.6°. Meanwhile, the minimal value occurs at 1600 Hz and is about 8.6°.

The dotted line in Fig. 2 represents the 30° phase-shift and the change trend of the ASW is the same as the 0° phase-shift. The maximum ASW appears at 12800 Hz and is approximately 32.1°. Meanwhile, the minimal value occurs at 1600 Hz and is about 14.3°.
The dashed line in Fig. 2 represents the 60° phase-shift and the change trend of the ASW is the same as the 0° phase-shift. The maximum ASW appears at 12800 Hz and is approximately 41.4°. Meanwhile, the minimal value occurs around 1600-3200 Hz and is about 23.6°.

The dash-dot line in Fig. 2 represents the 90° phase-shift and the change trend of the ASW is the same as the 0° phase-shift. The maximum ASW appears at 12800 Hz and is approximately 48.6°. Meanwhile, the minimal value occurs at 1600 Hz and is about 32.1°.

The means and associated 95% confidence intervals of the results in Fig. 2 shows that the ASW increases as the centre frequency of the fixed stimuli increasing from 1600 Hz to 12800 Hz and the maximum occurs at the 90° phase-shift above 6400 Hz.

It can be inferred that the phase-shift has an obvious impact on the ASW at high frequency and the impact is positively correlated with the phase-shift and the rising of frequency. Thus, we can widen the ASW by adjust the phase between left and right channels at high frequency above 1600 Hz.

Figure 3: The increment of the ASW at different phase-shift, along with the centre frequency of fixed stimuli varying from 1600 Hz to 12800 Hz.

The results of the increment of ASW as the phase-shift varying from 30° to 90° were shown in Figure 3, with the centre frequency of each fixed stimulus along x-axis and the increment of ASW along y-axis.

The solid line in Fig. 3 represents the 30° phase-shift. The maximum value of the increment of ASW appears at 12800 Hz with the 30° phase-shift, which is approximately 13.6° and the minimal value of the ASW is about 5.7° around 1600-3200 Hz.

The dotted line in Fig. 3 represents the 60° phase-shift. The maximum value of the increment of ASW appears at 6400 Hz with the 60° phase-shift, which is approximately 24.3°. While the minimal value is about 12.9° at 3200 Hz.

The dashed line in Fig. 3 represents the 90° phase-shift. The maximum value of the increment of ASW appears at 12800 Hz with the 90° phase-shift, which is approximately 33.6°. While the minimal value is about 20.0° at 3200 Hz.

The influence of the frequency itself on the ASW is removed and the influence of the phase-shift on the ASW is exhibited in Fig. 3 as the frequency varying from 1600 Hz to 12800 Hz. It can be inferred that the influence of the phase-shift on the ASW increases slightly above 1600 Hz.

It must be noted that the centre frequency of the fixed signal has greatly affected the ASW. Comparing the results of Experiment 1 and Experiment 2 at 0° phase-shift, it can be interpreted that the low-frequency stimuli were perceived to be the widest, with a decrease in the ASW to mid-frequency (approximately 1600-3200 Hz) and an increase in the ASW above 3200 Hz. This result does not change with the phase adjustment and consists with the conclusion that the perceived ASW is related with the centre frequency of the fixed stimuli [9].
The physiological explanation for the influence of the centre frequency of fixed stimuli on the ASW is as yet unknown. The impact of the phase adjustment on ASW can be explained with the changes of the IACC. As the phase-shift between the left and right channel increases, the IACC of the fixed signal gradually decreases from 1 to 0, which can widen the auditory source width monotonously [10].

5. Conclusions

In this paper, two psychology listening experiments were performed to investigate the relationship between the interaural phase difference and its effects on sound localization and perceived auditory source width. Firstly, the movement of sound localization to one side of heads is quantitative studied as the phase-shift varying from 0° to 90°, with the frequency varying from 100 Hz to 1600 Hz. The maximum of the movement of sound image occurs at the 90° phase-shift and the maximum deviation angle of the sound image is about 43.9° appearing at 100-200 Hz. Secondly, the ASW broadens as the interaural phase difference shifting from 0° to 90° with the frequency rising from 1600 Hz to 12800 Hz. The minimum ASW appears at 1600-3200 Hz and the maximum occurs around 6400-12800 Hz.

As a result, the influence of interaural phase difference on the sound localization and the ASW has been explored in the full frequency range. This study can be used to refine existing perception models of sound localization and auditory sound width. The influence of the loudness level on the ASW will be investigated in the future and a detailed ASW prediction model will be established.

REFERENCES