2025 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)

Evaluation of auditory and tactile perception for augmented sound-image enhancement using pre-virtual-leading hypersonic signals

Ryota Imanaka*, Yuting Geng*, Masato Nakayama* and Takanobu Nishiura*
* Ritsumeikan University, Osaka, Japan
E-mail: {is0545rs@ed, geng@fc, mnaka@fc, nishiura@is}.ritsumei.ac.jp

***Abstract*—Recently, stereo audio technologies have been proposed to control sound-image localization based on binaural effects. Precedence effect, which occurs when sounds arrive from different directions with a slight time difference, allows listeners to perceive a sound-image in the direction of the first-arriving sound. Moreover, hypersonic effect suggests that inaudible high-frequency sounds influence the perception of audible sounds. Considering the potential of ultrasound in controlling sound-image localization, we have previously proposed a sound-image augmentation method with ultrasonic signals, named pre-virtual-leading hypersonic signals. In this method, a sound-image is constructed using stereo technologies, while these signals enhance directional perception. These signals are designed based on the tactile perception of ultrasounds. This method has been demonstrated to be effective in limited conditions, though the sensory perception mechanism has not yet been fully clarified. This paper focuses on whether the signals are perceived auditorily or tactilely. We conducted evaluation experiments and demonstrated the effectiveness of the proposed method.**

I. INTRODUCTION

In recent decades, there has been growing interest in technologies for reproducing highly realistic sound fields that can represent sound to listeners with a high sense of presence [1]. Highly realistic sound reproduction technologies can be broadly classified into two categories: those based on psychoacoustic models, such as 5.1 ch. surround systems [2] and 22.2 ch. surround systems [3]; and those based on physical acoustic models, such as wavefront synthesis [4] and ambisonics [5]. Among the psychoacoustic models, various methods for reproducing the direction of sound have been proposed [6], [7], [8]. These methods are based on extensions of stereo sound-image control and have been proposed to control sound-image using binaural effects.

Some of these methods utilize the precedence effect, which is one of the binaural effects. The precedence effect is a phenomenon in which a sound-image emphasized toward the first-arriving direct sound when similar sounds arrive from different directions with a slight time difference [9], [10]. This allows sound-images to be localized in directions otherwise difficult with conventional methods, making it useful for augmenting sound-image perception in stereo audio. Moreover, sound-image control methods using the precedence effect have been proposed [11], [12]. In these methods, audible sound is delayed appropriately to augment localization of sound-images. These methods have been applied to evacuation guidance. In these methods, the sound-image can be presented only in the direction of the loudspeaker.

Moreover, it has been reported that emitted ultrasounds can influence auditory perception, and this phenomenon is known as hypersonic effect [13]. Regarding the hypersonic effect, two variations of this effect have been reported: a hypersonic negative effect, where the effect is suppressed, and a hypersonic positive effect, where the effect is enhanced. Reports have indicated that the occurrence of these effects is related to the ultrasound frequency band. Furthermore, it has been reported that when an ultrasound is amplitude-modulated with a low-frequency signal, the ultrasound can be perceived on the skin surface [14]. This phenomenon is known as vibrotactile perception, and it is possible to enhance ultrasound perception. The vibrotactile perception is the most sensitive at approximately 250 Hz [15]. Therefore, using the envelope can enhance ultrasound perception.

Considering the potential of ultrasound in controlling sound-image localization, we have previously proposed a sound-image augmentation method with ultrasonic signals, named pre-virtual-leading hypersonic signals[16], [17]. In this method, a sound-image is constructed with stereo technology, while these signals enhance directional perception. These signals are designed based on the tactile perception of ultrasounds. This method has been demonstrated to be effective in limited conditions, though the sensory perception mechanism has not yet been fully clarified. Moreover, it has not been investigated whether these hypersonic signals are perceived auditorily or tactilely. In this paper, we conducted two evaluation experiments. First, we conducted evaluation experiments to confirm whether the pre-virtual-leading hypersonic signals are perceived auditorily or tactilely and to clarify the sensory perception mechanism. Second, we conducted evaluation experiments on augmented sound-image enhancement to evaluate the effectiveness of the proposed method.

II. SOUND-IMAGE PERCEPTION CUES

In psychoacoustic models, sound-images are presented by reproducing the sensations at the listening point caused by binaural effects such as interaural time difference (ITD) and interaural level difference (ILD) [18]. These effects involve

979-8-3315-7206-8/25/$31.00 ©2025 IEEE

differences in arrival time and sound pressure at each ear, enabling perception of sound direction and spatial impression, especially in the horizontal plane. In addition, the precedence effect is one of the binaural effect about the time difference between each ear, and the difference exceeds a physically calculable threshold. The precedence effect can lead the listener to localize the sound-image in a direction that is difficult to perceive with other binaural effects. Therefore, it can be utilized to augment sound-image perception. Moreover, it has been reported that ultrasounds change the sensation of listeners. This phenomenon is known as the hypersonic effect [13], which occurrence depends on the frequency band of ultrasound. Furthermore, humans can perceive ultrasounds other than by hearing. For example, tactile feedback methods have shown that applying low-frequency envelopes to ultrasounds enables perception through skin.

A. Horizontal Sound-image Perception Cues

In the horizontal direction, sound direction is mainly perceived based on ITD and ILD. The ITD is the arrival time difference in sound emitted from a certain position due to the distance between the ears[19]. In sound-image construction using stereo technology, the direction and position of a sound-image perceived by the listener can be controlled by adding a time difference to the signals emitted from each loudspeaker in accordance with the sound-image construction direction.

The interaural time difference τ_{ITD} can be obtained by

$$\tau_{\mathrm{ITD}} = \frac{d\sin\theta}{v}, \tag{1}$$

where θ denotes the sound-image construction direction, d denotes interaural distance, and v denotes sound speed and the difference value is about 0.5 ms at the maximum. Up to this value, as the difference between the signals emitted from each channel increases, the sound-image shifts from the center between the two loudspeakers toward the loudspeaker that emits the preceding sound. When the difference between the sound emitted from each channel exceeds this maximum value, the sound-image localizes to the loudspeaker position due to the precedence effect [20].

The ILD is the energy difference between the sounds reaching each ear. This difference is caused by the distance between the ears when sound is emitted from a certain position. It is also one of the cues to perceive the horizontal direction and position of sound along with the ITD.

The precedence effect is a perceptual phenomenon related to the localization of the sound direction and position. When similar sounds are emitted from different directions with a slight time difference that exceeds a physically calculable threshold, the sound-image localizes toward the direction of the first direct sound [9], [10]. This effect can lead the sound-image to a direction that is difficult to perceive with other binaural effects. Therefore, it can be utilized to augment sound perception. The precedence effect typically occurs with time differences between 2 and 100 ms [10], which is much larger than the maximum ITD. Therefore, this effect is a psychological perception that does not correspond with physical phenomena. When the time difference is large, listeners perceive the sound-images as separating into two sounds. When the time difference is larger than about 0.5 ms and smaller than about 2 ms, listeners perceive the sound-image in the middle of the two sound sources.

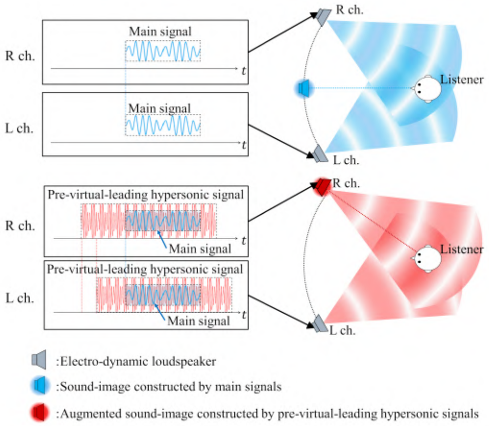


Fig. 1. Concept of the pre-virtual-leading hypersonic signals.

B. Discussion of Ultrasound Influence on Sound-image Perception Cues

It has been reported that emitted ultrasounds change the listener's sensation of sound, and this phenomenon is known as the hypersonic effect [13]. Regarding the hypersonic effect, both a hypersonic negative effect, in which the effect is suppressed, and a hypersonic positive effect, in which the effect is enhanced, and the reports have indicated that the occurrence of these effects is related to the frequency band of the ultrasound. The hypersonic positive effect is strongest around 80,000 Hz. In addition, it has been reported that humans can perceive ultrasounds other than by hearing. For example, tactile feedback methods have been proposed in which amplitude-modulated ultrasonic signals emitted from loudspeakers exert on the skin. Since the frequency response of the vibrotactile sensation peaks around 250 Hz, ultrasounds with low-frequency envelopes are utilized in this study. Consequently, the precedence effect caused by these signals is expected to be perceived easily.

III. PROPOSED METHOD FOR AUGMENTED SOUND-IMAGE ENHANCEMENT USING PRE-VIRTUAL-LEADING HYPERSONIC SIGNALS

Figure 1 shows the concept of the pre-virtual-leading hypersonic signals. In this paper, we focus on the precedence effect, and propose an augmented sound-image enhancement using pre-virtual-leading hypersonic signals. First, a delay is added to main signals so that they are emitted after the pre-virtual-leading hypersonic signals. Next, we design the pre-virtual-leading hypersonic signals. In this method, we use ultrasonic signals for the leading signals because ultrasounds are inaudible, and thus do not interfere with the main signal content. However, ultrasound signals are difficult to perceive.

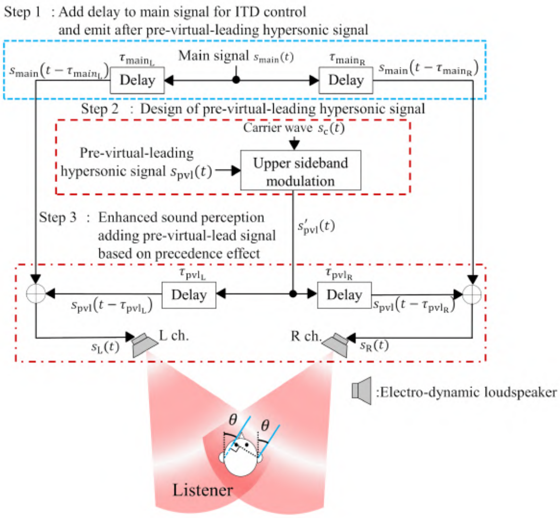


Fig. 2. Overview of the proposed method.

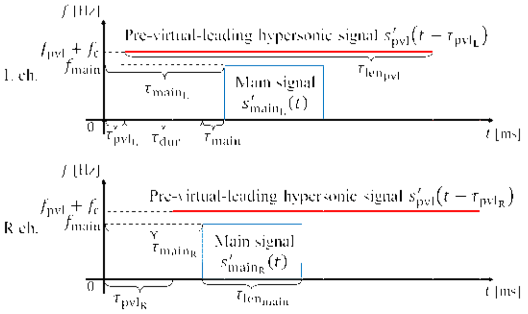


Fig. 3. Concept of the proposed method in time-frequency domain.

Human vibrotactile sensitivity is highest at around 250 Hz, we use ultrasound signals generated by upper sideband (USB) amplitude modulation [21] of a low-frequency signal around 250 Hz as pre-virtual-leading hypersonic signals. In this way, these signals in the ultrasonic band can be designed while maintaining the low-frequency envelope. As a result, these signals stimulate the listener's vibrotactile sensations, and enhance the perception of the hypersonic signals. In addition, a phenomenon known as the hypersonic positive effect, in which the effect of the hypersonic effect is enhanced, has been reported. This effect is strongest around 80,000 Hz [13]. A delay is also added to the designed pre-virtual-leading hypersonic signals. These signals provide the precedence effect and are added to the delayed main signals. In this way, augmented sound perception using pre-virtual-leading hypersonic signals is achieved.

Figure 2 shows the overview of the proposed method. Here $s_{\mathrm{main}}(t)$ is the main signal, $\tau_{\mathrm{main}_\mathrm{L}}$ and $\tau_{\mathrm{main}_\mathrm{R}}$ are the delay times added to the main signals emitted from each channel, $s_{\mathrm{pvl}}(t)$ is the pre-virtual-leading hypersonic signal, $s_\mathrm{c}(t)$ is the carrier wave, $s'_{\mathrm{pvl}}(t)$ is the USB modulated pre-virtual-leading hypersonic signal, $\tau_{\mathrm{pvl}_\mathrm{L}}$ and $\tau_{\mathrm{pvl}_\mathrm{R}}$ are the delay time added to pre-virtual-leading hypersonic signals emitted from each channel, and $s_\mathrm{L}(t)$ and $s_\mathrm{R}(t)$ are the signals emitted from each channel. Figure 3 shows the concept of the signals in the proposed method, where $s'_{\mathrm{main}_\mathrm{L}}(t)$ and $s'_{\mathrm{main}_\mathrm{R}}(t)$ are the main signals emitted from each channel after adding delay, f_{pvl} is the envelope frequency of the pre-virtual-leading hypersonic signal, f_c is the carrier frequency of the USB amplitude modulation, f_{main} is the highest frequency of main signal, $\tau_{\mathrm{len}_{\mathrm{pvl}}}$ is the length of the pre-virtual-leading hypersonic signal, and $\tau_{\mathrm{len}_{\mathrm{main}}}$ is the length of the main signal.

First, in Step 1 in Fig. 2, delay is added to the main signals so that they are emitted after the pre-virtual-leading hypersonic signals, which modifies the sound-image localization of the main signals based on the ITD. $\tau_{\mathrm{main}_\mathrm{L}}$ and $\tau_{\mathrm{main}_\mathrm{R}}$ are calculated based on the time panning method as shown in Eq.(1). Equation(2) shows the time difference between L ch. and R ch. main signals, and Eqs. (3) and (4) show the signals to which delay is added.

$$\begin{aligned}\tau_{\mathrm{main}} &= \tau_{\mathrm{main}_\mathrm{L}} - \tau_{\mathrm{main}_\mathrm{R}},\\ &= \frac{d\sin\theta}{v}, &(2)\\ s'_{\mathrm{main}_\mathrm{L}}(t) &= s_{\mathrm{main}}(t-\tau_{\mathrm{main}_\mathrm{L}}), &(3)\\ s'_{\mathrm{main}_\mathrm{R}}(t) &= s_{\mathrm{main}}(t-\tau_{\mathrm{main}_\mathrm{R}}). &(4)\end{aligned}$$

Next, in Step 2, we design the pre-virtual-leading hypersonic signals. These signal is an ultrasonic signal, hence we add a low-frequency envelope to it and listeners can perceive the signal easily. In this study, sine wave signals that induce the hypersonic effect are used as the carrier signals $s_\mathrm{c}(t)$, and sine wave signals that induce vibrotactile sensations are used as pre-virtual-leading signals $s_{\mathrm{pvl}}(t)$. These signals are shown in Eqs. (5) and (6).

$$\begin{aligned}s_\mathrm{c}(t) &= A_\mathrm{c}\cos(2\pi f_\mathrm{c} t), &(5)\\ s_{\mathrm{pvl}}(t) &= A_{\mathrm{pvl}}\cos(2\pi f_{\mathrm{pvl}} t). &(6)\end{aligned}$$

In this way, a pre-virtual-leading hypersonic signal can be designed with a low-frequency envelope while keeping it in the ultrasonic band. Eqs. (7) and (8) show the equations for the USB modulated wave and its modulation factor, respectively.

$$\begin{aligned}s'_{\mathrm{pvl}}(t) &= (1+ms_{\mathrm{pvl}}(t))A_\mathrm{c}\cos(2\pi f_\mathrm{c} t)\\ &\quad -m\hat{s}_{\mathrm{pvl}}(t)A_\mathrm{c}\sin(2\pi f_\mathrm{c} t), &(7)\\ m &= \frac{A_{\mathrm{pvl}}}{A_\mathrm{c}}, &(8)\end{aligned}$$

where A_C is the maximum amplitude of the carrier wave, m is the modulation factor ($0 < m \leq 1$), A_{pvl} is the maximum amplitude of the pre-virtual-leading hypersonic signal, and $\hat{s}_{\mathrm{pvl}}(t)$ is the pre-virtual-leading hypersonic signal with a phase shift of $\pi/2$. In addition, from Eqs.(5), (6), (7), and (8), the USB modulated pre-virtual-leading hypersonic signal can be expanded by using the trigonometric addition formulas as in the following.

$$\begin{aligned}s'_{\mathrm{pvl}}(t) &= A_\mathrm{c}\cos(2\pi f_\mathrm{c} t) + mA_\mathrm{c}s_{\mathrm{pvl}}(t)\cos(2\pi f_\mathrm{c} t)\\ &\quad -mA_\mathrm{c}\hat{s}_{\mathrm{pvl}}(t)\sin(2\pi f_\mathrm{c} t),\\ &= A_\mathrm{c}\cos(2\pi f_\mathrm{c} t)\\ &\quad +A_{\mathrm{pvl}}\cos(2\pi f_{\mathrm{pvl}} t)\cos(2\pi f_\mathrm{c} t)\\ &\quad -A_{\mathrm{pvl}}\sin(2\pi f_{\mathrm{pvl}} t)\sin(2\pi f_\mathrm{c} t),\\ &= A_\mathrm{c}\cos(2\pi f_\mathrm{c} t)\\ &\quad +A_{\mathrm{pvl}}^2\cos(2\pi(f_\mathrm{c}+f_{\mathrm{pvl}})t). &(9)\end{aligned}$$

In this case, if the frequency of the pre-virtual-leading hypersonic signal f_{c} is audible or below f_{main}, the signal affects listening to the contents of the main signals. Therefore, the constraint conditions of the pre-virtual-leading hypersonic signals are shown in Eqs. (10) and (11).

$$\begin{aligned} f_{\mathrm{c}} &> f_{\mathrm{main}}, & (10) \\ f_{\mathrm{c}} &\geq 20000\ \mathrm{Hz}. & (11) \end{aligned}$$

Finally, in Step 3, we add time delays to the designed pre-virtual-leading hypersonic signals. These signals provide the precedence effect, and are added to the delayed main signals. In this way, augmented sound perception using pre-virtual-leading hypersonic signals is achieved. $s'_{\mathrm{pvl}}(t)$ modulated by USB in Step 2 is delayed on each channel and added to the main signal $s'(t)$ generated in Step 1. Equations (12) and (13) show the equations for the signals emitted from each channel.

$$\begin{aligned} s_{\mathrm{L}}(t) &= s'_{\mathrm{main_L}}(t) + s'_{\mathrm{pvl}}(t - \tau_{\mathrm{pvl_L}}), & (12) \\ s_{\mathrm{R}}(t) &= s'_{\mathrm{main_R}}(t) + s'_{\mathrm{pvl}}(t - \tau_{\mathrm{pvl_R}}). & (13) \end{aligned}$$

As a result, the precedence effect occurs in the pre-virtual-leading hypersonic signals, and an augmented sound-image enhancement using pre-virtual-leading hypersonic signals is achieved.

IV. EVALUATION EXPERIMENT ON SENSORY PERCEPTION OF PRE-VIRTUAL-LEADING HYPERSONIC SIGNALS

A. Experimental Conditions on Sensory Perception of Pre-virtual-leading Hypersonic Signals

To confirm whether the pre-virtual-leading hypersonic signals are perceived auditorily or tactilely, we conducted evaluation experiments. Subjects wore earmuff to block auditory perception and a combination of a face mask, face shield and gloves to block tactile perception in the corresponding body regions. This equipment and conditions are showed in Tables I and II, respectively. In this experiment, only the hypersonic signals were emitted, and they were emitted simultaneously from both channels. We used three low-frequencies and a steady envelopes: f_{pvl} = (100 Hz, 250 Hz, 600 Hz). Also, we used four carrier frequencies: f_{c} = (20,000 Hz, 40,000 Hz, 80,000 Hz, 100,000 Hz). In each condition, we used the modulation factor as $m = 1.00$. In addition, we conducted the experiment under five sensory deprivation conditions involving auditory and tactile perception as shown in Table II. The All free condition allowed auditory and tactile perception, the Auditory only condition allowed for auditory perception, the Tact.(face) only condition allowed for tactile perception on the face, the Tact.(hand) only condition allowed for tactile perception on the hand, and the All block condition deprived both auditory and tactile perception. Thus, a total of $4 \times 4 \times 5 = 80$ patterns were evaluated in this experiment. The equipment and the subject arrangement in the experiments are shown in Fig. 4. Subjects were asked whether the hypersonic signals can be perceived in each condition. The experimental environment is shown in Table III. At the listening position, the sound pressure level was 100 dB for the hypersonic signal with f_{c} = 20,000 Hz and a steady envelope, and 45 dB for the main signal. The number of subjects was 10 (1 female and 9 males), aged 21 to 24 years.

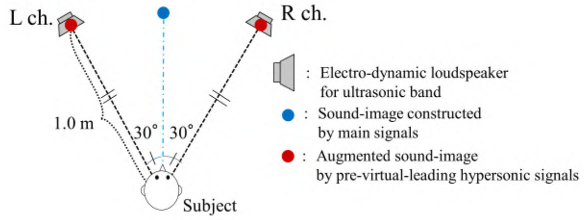


Fig. 4. Experimental arrangement on sensory perception of pre-virtual-leading hypersonic signals.

TABLE I
EXPERIMENTAL EQUIPMENT ON SENSORY PERCEPTION OF PRE-VIRTUAL-LEADING HYPERSONIC SIGNALS.

Electro-dynamic loudspeaker	Technics, SB-G90M2
Power amplifier	Technics, SU-G700M2
Audio interface	RME, ADI-2 Pro FS
Raincoat	Kajimeiku, 1222
Earmuff	PELTOR, HTM79A
Face mask	osakimedical, 51420
Face shield	Yamasho, YPS-001
Gloves	Atsutashizai, No.606

TABLE II
EXPERIMENTAL EVALUATION PATTERNS ON SENSORY PERCEPTION OF PRE-VIRTUAL-LEADING HYPERSONIC SIGNALS.

Pattern	Description
All free	No raincoat, no earmuff, no face mask and shield, no gloves
Auditory only	Wear raincoat, face mask and shield, and gloves, no earmuff
Tact.(face) only	Wear raincoat and earmuff, no face mask and shield, no gloves
Tact.(hand) only	Wear raincoat, earmuff, face mask and shield, no gloves
All blocked	Wear raincoat, earmuff, face mask and shield, and gloves

TABLE III
EXPERIMENTAL CONDITIONS ON SENSORY PERCEPTION OF PRE-VIRTUAL-LEADING HYPERSONIC SIGNALS.

Reverberation time	T_{20} = 0.3 s
Ambient noise level	L_{A} = 24.6 dB
Sampling frequency	384,000 Hz
Quantization	24 bits

B. Experimental Results on Sensory Perception of Pre-virtual-leading Hypersonic Signals

Figure 5 shows the results of the experiment whether the signals can be perceived in auditory or tactile. The horizontal axis is the frequency of the pre-virtual-leading hypersonic signals, and the vertical axis is the correct answer rate. Error bars indicate the range of standard deviations added positively or negatively to the percentage correct for each condition. The results show that the pre-virtual-leading hypersonic signals can

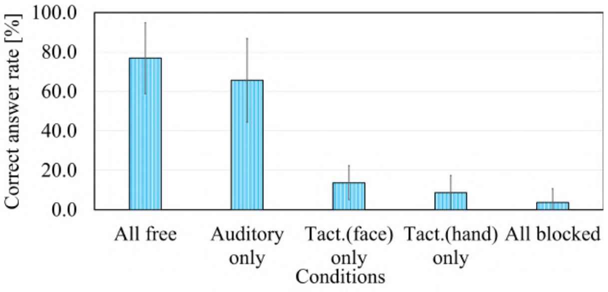


Fig. 5. Correct answer rate for each condition of sensory deprivation.

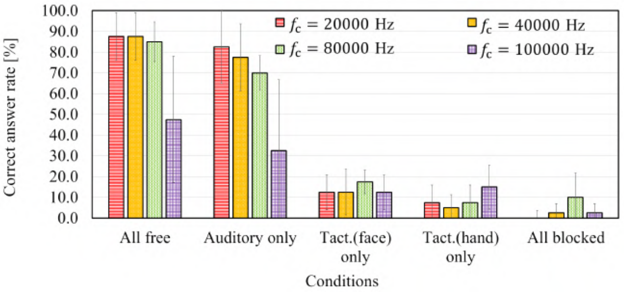


Fig. 6. Correct answer rate for carrier frequencies in each condition of sensory deprivation.

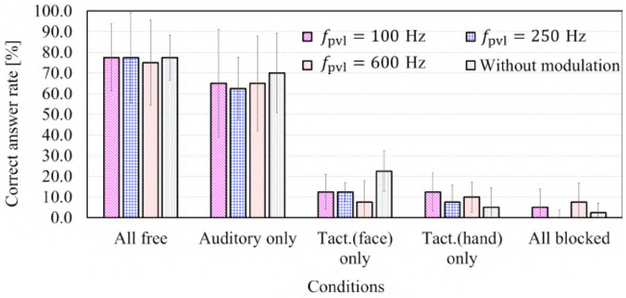


Fig. 7. Correct answer rate for envelope frequencies in each condition of sensory deprivation.

be perceived by auditory. Also, Fig. 6 shows the results of the experiment for carrier frequencies in each condition, these results reflect the auditory perception of the hypersonic signals. These results shows that conditions of f_c = (20,000 Hz, 40,000 Hz, 80,000 Hz) are perceived sufficiently, but the condition of f_c = 100,000 Hz is not perceived well. This may be due to individual differences in ultrasonic perception or limitations of the equipment used. In addition, Fig. 7 shows the results of the experiment for envelope frequencies in each condition, which reflect the tactile perception of the hypersonic signals. The results show no clear trend or notable difference among the f_{pvl} in each condition.

V. EVALUATION EXPERIMENT FOR AUGMENTED SOUND-IMAGE ENHANCEMENT

A. Experimental Conditions for Augmented Sound-image Enhancement

To evaluate the effectiveness of the proposed method, we conducted evaluation experiments for augmented sound-image enhancement. In our experiments, we used three low-frequency envelopes: $f_{\mathrm{pvl}} =$ (100 Hz, 250 Hz, 600 Hz), and four carrier frequencies: $f_c =$ (20,000 Hz, 40,000 Hz, 80,000 Hz, 100,000 Hz), so we obtained 12 conditions in total. In each condition, we used the modulation factor as $m = 1.00$. The delay times of the pre-virtual-leading hypersonic signals in each channel $\tau_{\mathrm{pvl_L}}$, $\tau_{\mathrm{pvl_R}}$ are $(\tau_{\mathrm{pvl_L}}, \tau_{\mathrm{pvl_R}}) =$ (2 ms, 0 ms), (0 ms, 2 ms), (0 ms, 0 ms), and either the left or right pre-virtual-leading hypersonic signal preceded the other or the left and right pre-virtual-leading hypersonic signals were emitted simultaneously. This delay time is equivalent to the direction of each loudspeaker in these experimental conditions. The delay times of main signals $\tau_{\mathrm{main_L}}$ and $\tau_{\mathrm{main_R}}$ were set in $\tau_{\mathrm{main_L}} = \tau_{\mathrm{main_R}}$. This makes the sound-image localized in front of the subject. The time difference between the preceding pre-virtual-leading hypersonic signal and main signals was 1,000 ms. The length of the pre-virtual-leading hypersonic signals $\tau_{\mathrm{len_{pvl}}}$ was set to 5,000 ms, and the length of the main signal $\tau_{\mathrm{len_{main}}}$ was set to 20 ms. For the main signal, we used white noise with a bandwidth of 0 to 16,000 Hz. In the experiments, we had subjects sit in the position shown in Fig. 4. The subjects were asked to answer whether the perceived sound-image was “shifted to the left,” “shifted to the right,” or “stationary” compared to when the signal was emitted without the pre-virtual-leading hypersonic signal. We define the correct answer as being when the direction of the sound-image perceived by the subject matched the direction of the preceding pre-virtual-leading hypersonic signal, and we calculated the correct answer rate. In the case when the left and right pre-virtual-leading hypersonic signals were emitted simultaneously, we specify that the correct answer is “stationary”.

TABLE IV
EXPERIMENTAL CONDITIONS FOR AUGMENTED SOUND-IMAGE ENHANCEMENT.

Reverberation time	T_{20} = 0.4 s
Ambient noise level	L_A = 32.4 dB
Sampling frequency	384,000 Hz
Quantization	24 bits

TABLE V
EXPERIMENTAL EQUIPMENT FOR AUGMENTED SOUND-IMAGE ENHANCEMENT.

Electro-dynamic loudspeaker	Technics, SB-G90M2
Power amplifier	Technics, SU-G700M2
Audio interface	RME, ADI-2 Pro FS

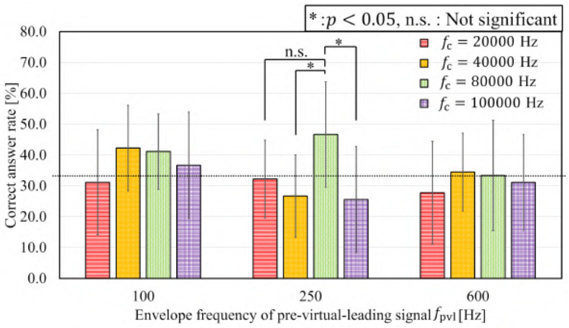
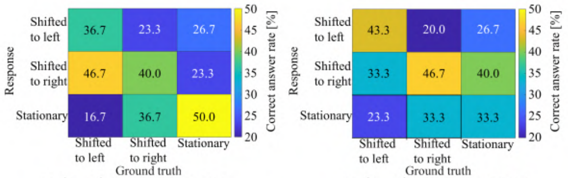


Fig. 8. Correct answer rate for combinations (f_{pvl}, f_c).

Experimental environment and equipment used are listed in Tables IV and V. At the listening position, the sound pressure level was 100 dB for the hypersonic signal with f_c = 20,000 Hz and a steady envelope, and 45 dB for the main signal. The experimental arrangement were same as in the experiment on sensory perception, and it is shown in Fig. 4.

B. Experimental Results for Augmented Sound-image Enhancement

Figure 8 shows the correct answer rate for each condition. The horizontal axis is the frequency of the pre-virtual-leading hypersonic signals, and the vertical axis is the correct answer rate. Error bars indicate the range of standard deviations added positively or negatively to the percentage correct for each condition. The dashed line indicates a reference level of 33.3%, which is the expected value for the correct answer rate when answering randomly. From Fig. 8, $(f_{\mathrm{pvl}}, f_c) =$ (100 Hz, 40,000

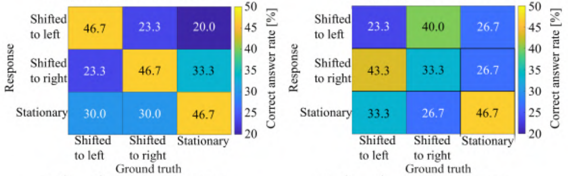


Response \ Ground truth	Shifted to left	Shifted to right	Stationary
Shifted to left	36.7	23.3	26.7
Shifted to right	46.7	40.0	23.3
Stationary	16.7	36.7	50.0

(a) $(f_{\mathrm{pvl}}, f_{\mathrm{c}}) = (100\ \mathrm{Hz}, 40{,}000\ \mathrm{Hz})$.

Response \ Ground truth	Shifted to left	Shifted to right	Stationary
Shifted to left	43.3	20.0	26.7
Shifted to right	33.3	46.7	40.0
Stationary	23.3	33.3	33.3

(b) $(f_{\mathrm{pvl}}, f_{\mathrm{c}}) = (100\ \mathrm{Hz}, 80{,}000\ \mathrm{Hz})$.



Response \ Ground truth	Shifted to left	Shifted to right	Stationary
Shifted to left	46.7	23.3	20.0
Shifted to right	23.3	46.7	33.3
Stationary	30.0	30.0	46.7

(c) $(f_{\mathrm{pvl}}, f_{\mathrm{c}}) = (250\ \mathrm{Hz}, 80{,}000\ \mathrm{Hz})$.

Response \ Ground truth	Shifted to left	Shifted to right	Stationary
Shifted to left	23.3	40.0	26.7
Shifted to right	43.3	33.3	26.7
Stationary	33.3	26.7	46.7

(d) $(f_{\mathrm{pvl}}, f_{\mathrm{c}}) = (600\ \mathrm{Hz}, 40{,}000\ \mathrm{Hz})$.

Fig. 9. Correct answer rate for combinations $(f_{\mathrm{pvl}}, f_{\mathrm{c}})$.

Hz), (100 Hz, 80,000 Hz), (250 Hz, 80,000 Hz), and (600 Hz, 40,000 Hz), the correct answer rate obtained exceeded the expected value (33.3%). Figure 9 shows detailed answer rates for combinations $(f_{\mathrm{pvl}}, f_{\mathrm{c}})$. Especially, from Fig. 9(c), the percentages of responses that matched the correct answers were higher than the percentage of the other options chosen. In addition, Tukey's multiple comparison test was conducted on the correct answer rates between $(f_{\mathrm{pvl}}, f_{\mathrm{c}}) =$ (250 Hz, 80,000 Hz) and $(f_{\mathrm{pvl}}, f_{\mathrm{c}}) =$ (250 Hz, 20,000 Hz), (250 Hz, 40,000 Hz), (250 Hz, 100,000 Hz). In this case, there are significant differences at the significance level $\alpha = 0.05$ in the conditions of $(f_{\mathrm{pvl}}, f_{\mathrm{c}}) =$ (250 Hz, 40,000 Hz) and (250 Hz, 80,000 Hz). Therefore, we found the proposed method to be effective when $(f_{\mathrm{pvl}}, f_{\mathrm{c}}) =$ (250 Hz, 80,000 Hz).

From the results of this experiment for augmented sound-image enhancement, the pre-virtual-leading hypersonic signals were perceived most strongly at $(f_{\mathrm{pvl}}, f_{\mathrm{c}}) =$ (250 Hz, 80,000 Hz). In addition, from the experiment on sensory perception, hypersonic signals are mainly perceived auditorily. Also, The frequency band in which the hypersonic positive effect occurs the most strongly is around 80,000 Hz [13]. Therefore, we concluded that the proposed method has high correlation with the hypersonic positive effect.

VI. CONCLUSIONS

In this paper, we proposed a method of augmented sound perception by adding pre-virtual-leading hypersonic signals to the main signal presented in front of the listener. We examined whether these signals are perceived auditorily or tactilely, and the carrier and envelope of the pre-virtual-leading hypersonic signals. Evaluation experiments revealed that the signals were mainly perceived auditory, and we demonstrated a strong correlation between the hypersonic positive effect and our method. In the future, we plan to conduct experiments to investigate the time difference between the pre-virtual-leading hypersonic signal and the main signal. Moreover, we also aim to increase the number of participants considering their diversity, in order to obtain accurate conclusions.

ACKNOWLEDGMENTS

This work was partly supported by Ritsumeikan University R-GIRO, ARC and RARA, and JSPS KAKENHI Grant Numbers JP23K28115, JP24K20803 and JP25H01158.

REFERENCES

[1] C. T. Jin, "A tutorial on immersive three-dimensional sound technologies," Acoustical Science and Technology, vol.41, no. 1, pp. 16-27, 2020.

[2] ITU-R Recommendation BS. 775-4, "Multichannel stereophonic sound system with and without accompanying picture," International Telecommunication Union, 2022.

[3] K. Hamasaki, T. Nishiguchi, R. Okumura, Y. Nakayama, and A. Ando, "A 22.2 multichannel sound system for ultrahigh-definition TV (UHDTV)," SMPTE Motion Imaging Journal, vol. 117, no. 3, pp. 40-49, 2008.

[4] J. Ahrens, "Analytic Methods of Sound Field Synthesis," Springer Berlin Heidelberg, 2012.

[5] R. K. Furness, "Ambisonics-An Overview," Journal of the Audio Engineering Society, vol.8, no. 24, pp.181-190, 1990.

[6] V. Plukki, "Virtual sound source positioning using vector base amplitude panning," Journal of the Audio Engineering Society, vol. 45, no. 6, pp. 456-466, 1997.

[7] J. Blauert, "Spatial Hearing: The Psychophysics of Human Sound Localization," The MIT Press, 1996.

[8] D. M. Leakey, "Some Measurements on the Effects of Interchannel Intensity and Time Differences in Two Channel Sound Systems," The Journal of the Acoustical Society of America, vol. 31, no. 7, pp. 977-986, 2005.

[9] R. Y. Litovsky, H. S. Colburn, W. A. Yost, and S. J. Guzman, "The precedence effect," The Journal of the Acoustical Society of America, vol. 106, no. 4, pp. 1633-1654, 1999.

[10] H. Wallach, E. B. Newman, and M. R. Rosenzweig, "The Precedence Effect in Sound Localization," The American Journal of Psychology, vol. 62, no. 3, pp. 315-336, 1949.

[11] H. Itou, S. Yazawa, Y. Watanabe, H. Chiba, T. Kawase, T. Kako, M. Isogai, K. Noguchi, A. Nakayama, "A Study on the Law of the First Wavefront for the Realization of Cyber-Real Wavefront-Matching Acoustic Processing Sound SyReal," Proceedings of Autumn Meeting on Acoustical Society of Japan, pp. 405-406, 2024(In Japanese).

[12] T. Miyoshi, "Evacuation Guidance Assistance System Using Emitting Sound," Noise Control, IntechOpen, 2022.

[13] T. Oohashi, E. Nishina, M. Honda, Y. Yonekura, Y. Fuwamoto, N. Kawai, T. Maekawa, S. Nakamura, H. Fukuyama, and H. Shibasaki. "Inaudible high-frequency sounds affect brain activity: hypersonic effect," Journal of Neurophysiology, vol.83, no. 6, pp. 3548-3558, 2000.

[14] E.R. Kandel, J.H. Schwartz, and T. Jessell, "Principles of Neural Science," McGraw-Hill Companies, 2000.

[15] A. Gescheider, S. J. Bolanowski, and K. R. Hardick, "The frequency selectivity of information-processing channels in the tactile sensory system," Somatosensory & Motor Research, vol. 18, no. 3, pp.191-201, 2001.

[16] R. Imanaka, Y. Geng, M. Nakayama, and T. Nishiura, "Augmented sound-image perception using pre-virtual-leading ultrasounds based on precedence effect," 2024 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Paper ID: 112, 2024.

[17] R. Imanaka, Y. Geng, M. Nakayama, and T. Nishiura, "Augmented sound-image perception using pre-virtual-leading hypersonic signals with bass frequency envelopes," 188th Meeting of the Acoustical Society of America joint with 25th International Congress on Acoustics, Paper ID: 3pSP4, 2025.

[18] A. Roginska and P. Geluso, "Immersive Sound The Art and Science of Binaural and Multi-Channel Audio," Routledge, 2017.

[19] A. Brughera, L. Dunai, and W. M. Hartmann, "Human interaural time difference thresholds for sine tones: The high-frequency limit," The Journal of the Acoustical Society of America, vol. 133, no. 5, pp. 2839-2855, 2013.

[20] S. Okuro and Y. Kajikawa, "Reproduction of varied sound image localization for real source in stereo audio system," Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference(APSIPA ASC), Paper ID: 70, 2012.

[21] Y. Wang, X. Li, L. Xu, and L. Xu, "SSB Modulation of the Ultrasonic Carrier for a Parametric Loudspeaker," International Conference on Electronic Computer Technology 2009, pp. 669-673, 2009.