Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

Effects of pitch of sounds on vertical localization of the sounds with frequencies changing over three octaves

Eiichi Miyasaka

Tokyo City University, 3-3-1,Ushikubo-Nishi, Tsuzuki-ku, Yokohama, JAPAN

miyasaka@tcu.ac.jp

PACS: 43.66.Qp

ABSTRACT

Effects of pitch of sounds on vertical localization were investigated using the sounds consisting of 22 tone-bursts ascending/descending along the whole-tone scale between C4 (262Hz) and C7(2093Hz). The kinds of the sounds used were pure tones, one-third octave band noises and piano-tones. The sounds were presented through a fixed loudspeaker (SP-A) set up just in front of listeners without any additional loudspekers but only with numbered cards set perpendicularly (case-1) or with seven dummy loudspeakers attached to the numbered cards (case-2). The results show that most observers perceived the locations of the sound images moved upward from a loudspeaker around SP-A for the ascending sounds or downward for the descending sounds in both cases, although the sounds were reproduced through the fixed loudspeaker(SP-A). These results imply that pitch of the sound used in this experiment will be effective on vertical localization of the sound images.

INTRODUCTION

When a sound is presented monophonically through a loudspeaker set up just in front of a listener, the sound image may be localized at a fixed position at or around the loudspeaker even if the sound is ascending or descending in the funadmental frequency. It will be same in sterophonic reproduction with a couple of loudspeakers.

Agatsuma and Miyasaka[1] investigated vertical sound localization when sound sources of one-third octave band noises were moved upward or downward with the moving speeds of 10 to 300 cm/s along the eight loudspeakers set up perpendicularly just in front of observers. The center frequencies of the noises were fixed at 1, 4 or 8 KHz. The results show that most observers perceived that the sound images moved upward or downward when the sounds were moved upward or downward, respectively. Some observers, however, perceived opposite movement. Nearly a half of the observers perceived sound images moved approximately same distance as the sound source at the speed of 10 to 30 cm/s. Not a few observers perceived the sound images moved higher than the maximum positions of moved sounds. However, the perceived moving distance was shortend at the speed of 120cm/s.

This paper presents an experiment in order to clear whether the location of the sound image will move or not for sound sources with the fundamental frequerncies varying from low to high when they are presented through a loudspeaker set just in front.

EXPRIMENT-1

Experimental set up

Fig.1 shows an experimental set up in which the sounds were presented through only the loudspekers numberd 2 with other dummy loudspeakers numbered 1,3,4,5,6,7,8. An observer was positioned 4.5m apart from the array consisting of eight ALR/JORDAN Entry/S loudspeakers set up perpendicularly with 0.3m apart to each other. The loudspeaker numbered 1 was set 0.8m above the floor in a middle-large room (floor range: 315.6m2 , height : 5.2m) with a small reverberation time. Each loudspeaker was named after 1 to 8 (for example, the first loudspeaker at the bottom is named as the loudspeaker-1) and a plate with its number was put at the side of each loudspeaker. Every stimulus was presented only through the loudspeaker-2, and was calibrated to 60dBA at the observer's position.

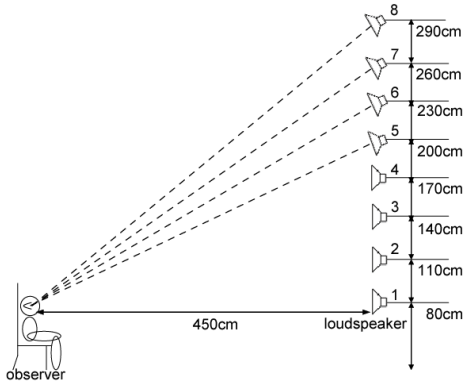

Figure 1.- Diagram of experimental setup. The loudspeakers were vertically set up just in front of an observer, and the numbered cards from 1 to 8 were put aside of the loudspeakers, respectively.

In most experiments for vertical localization with multi-loudspeakers in the median plane [2][3], it will be common that the loudspeakers are set on a vertical circle orbit so as to keep a constant distance between an observer and each loudspeaker. In this experiment, however, the loudspeakers were set perpendicularly as Ferguson and Cabrera [4] performed in their experiments, and the head was not fixed with a head-band in order that the experimental condition could be closer to a real listening condition.

Stimuli

The sound sources used in the experiment were pure tones, one-third octave band noises and piano tones. Each stimulus consists of 22 tone-bursts with the duration of 250ms ascending along the whole-tone scale over three octaves from C4 (262Hz) to C7 (2093Hz) at the speed of 1.83 s/oct. or descending from C7 to C4 at the same speed. Figure 2 shows the sound sequences ascending and descending.

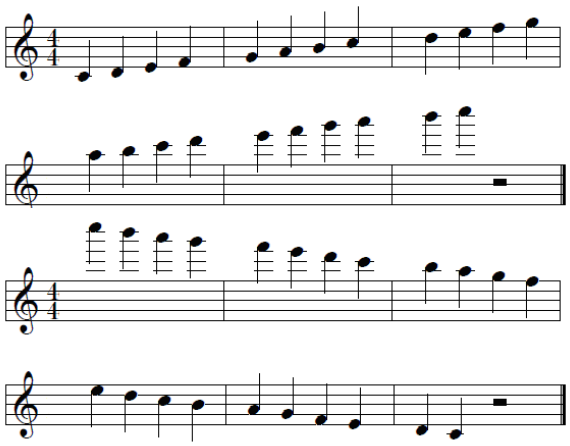

Figure 2.- Sound sequences used in the experiments.Upper: ascending sequences, Lower: descending sequencies.

Procedure

Seventeen normal-hearing students of between 19 to 23 years old participated in the experiment as the observers. Each observer was required to identify one of eight loudspeakers from which the first segment of a stimulus was heard and to identify one from which the last segment of the stimulus was heard. If he/she perceived that the sound images did not move, they assigned the same number of the loudspeaker for the first segment as for the last. The fact that every stimulus was presented only through the fixed loudspeaker-2 was not known to any observer. Each observer was asked to continuously keep his/her head pointed toward a fixed loudspeaker so as to detect all loudspeakers without head-movement.

Before presentation of each stimulus, each observer was instructed by an experimenter one of the following two messages. (1) The stimulus presented just after has been already performed some sort of artificial processing contibuting to localization. (2) The stimulus presented just after has not been performed any artificial processing. The aim of the instructions is to investigate whether the instructions may affect perception of localization or not. The stimuli were presented randomly with fifth repetitions of each stimulus. Total number of the presented stimuli to each observer was 60 (= 3 (types of the stimuli) x 2 (ascending/descending) x 2 (instructions) x 5 (repetitions)).

Results

Two observers perceived movement of the sound images opposite to the rest 15 observers. Fig.3 shows the results of the latter 15 observers for the stimuli ascending in pitch. The ordinate indicates the averaged distances between the center of the loudspeakers through which the first and the last segments of the stimulus were heard and that of the loudspeaker-1, respectively. The empty boxes in the figure show the perceived heights for the first segments, while the hatched boxes show those for the last segments. The abscissa indicates the types of the stimuli with instructions. “Ins-A” indicates the instruction saying that the stimulus presented just after has already performed some sort of artificial processing contibuting to localization, while “Ins-B” indicates the instruction saying that the stimulus presented just after is not performed any artificial processing. The vertical bars on the boxes indicate the standard deviations.

For the ascending stimuli, the heights of the loudspeakers through which the last segments were heard are 30 to 60cm higher than those of the loudspeakers through which the first segments were heard for all stimuli irrespective of the instructions. Most observers perceive that the first segment of every stimulus is radiated from a loudspeaker positioned higher than the loudspeaker-2 through which the real stimulus is radiated. These results imply the following items;

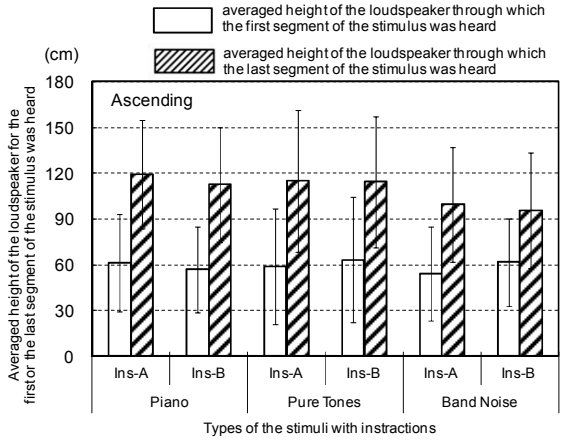

Figure 3.- The averaged distances between the center of the loudspeaker through which the first and the last segment of the stimuli ascending in pitch was heard and that of the loudspeaker-1, respectively. The empty boxes indicate the results for the first segment, while the hatched boxes indicate them for the last segment. Ins-A: Instruction given to each observer just before each stimulus, saying that the stimulus presented just after has been already performed some sort of artificial processing contributing to localization. Ins-B: Instruction given to each observer just before each stimulus, saying that the stimulus presented just after has not been performed any artificial processing. The vertical bars indicate the standard deviations.

(1) The sound images localize at a higher loudspeaker than the loudspeaker-2 radiating the real sounds.

(2) The sound images move upward with increase in pitch.

Fig.4 shows the results of the experiment for the stimuli descending along the whole-tone scale over three octaves from C7 (2093Hz) to C4 (262Hz) at the speed of 1.83 s/oct. The empty boxes in the figure indicate the perceived heights of the sound images of the first segment of each stimulus, while the hatched boxes show those of the last segments. For the descending stimuli, the perceived heights of the sound images of the last segments are approximately 30cm lower

than those of the first segments for all stimuli irrespective of the instructions.

Compared with the results shown in Fig.3, the perceived heights of the sound images of the first segments (note name: C7) for piano-tones and pure-tones are approximately 30cm lower than those of the last segments (note name: C7) of the ascending stimuli. The following items are also indicated.

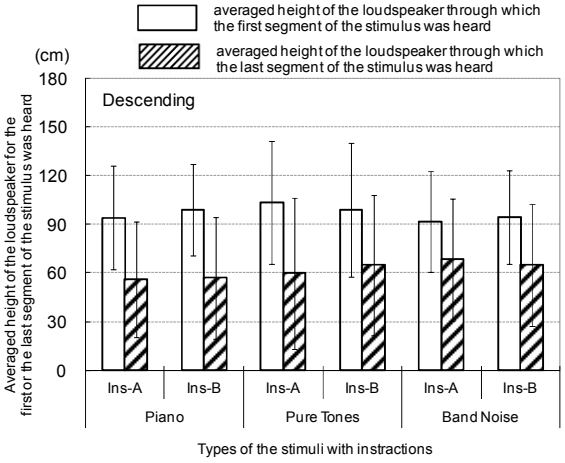

Figure 4.- The averaged height of the loudspeaker through which the first or the last segment of the stimuli descending in pitch was heard. The empty boxes show the results for the first segment, while the hatched boxes show them for the last segment. For Ins-A and Ins-B and the vertical bars, see the caption in Figure 3.

(1) The sound images localize at a loudspeaker higher than the loudspeaker-2 radiating the real sounds.

(2) The sound images move downward for the stimuli with decrease in pitch.

Both in the ascending stimuli and the descending, the instructions (Ins-A, Ins-B) affected nothing on the results contrary to the author's expectations.

EXPRIMENT-2

Experimental set up and Procedure

In this experiment, only the loudspeaker-2 in Fig.1 was left and the other loudspeakers were removed except the eight numbered cards from one to eight. An observer could look at the only one loudspeaker just in front and the eight numbered cards vertically set up with 0.3m apart to each other. The same sound sources in the experiment-1 were used as the stimuli. The observers in the experiment were seventeen normal-hearing students who participated also in the experiment-1. The observers were required to identify the locations of the sound images produced by the first segment and the last segment of a stimulus. Before presentation of each stimulus, each observer was instructed by an experimenter one of the same two messages as used in the experiment-1. The stimuli were presented randomly with fifth repetitions of each stimulus. Total number of the presented stimuli to each observer was 60 (= 3 (types of the stimuli) x 2 (ascending/descending) x 2 (instructions) x 5 (repetitions)).

Results

The results show that 15 observers perceived the sound images moved upward from a position slightly higher than the loudspeaker-2 for the ascending stimuli and the sound images moved downward for the descending stimuli. The rest two observers perceived opposite movement to the former observers. They were the same observers as those who perceived direction opposite to the other observers in the experiment-1.

Fig.5 shows the results of the 15 observers for the stimuli ascending in pitch. The ordinate indicates the averaged perceived heights of the sound images produced by the first and the last segments of the stimulus, respectively, from the center of the loudspeaker-1.

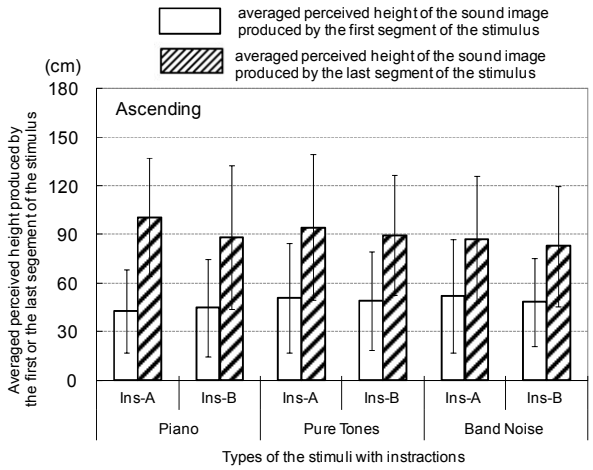

Figure 5.- The averaged perceived heights of the sound images produced by the first and the last segments of the stimuli ascending in pitch. Only one loudspeaker from which the stimuli were output, was set up just in front of the observers. For the empty boxes, the hatched boxes, Ins-A, Ins-B and the vertical bars, see the caption in Figure 3.

Despite only one loudspeaker in front, the most observers perceived the sound images moved upward from a position slightly higher than the loudspeaker-2 from which the stimuli were output. The perceived first and the last locations of the sound images for the first and the last segments, respectively, were approximately 15-20cm lower than those in the experiment-1. The observers were required to answer a card number at which they perceived the sound images were localized in order to compare the results with those of the experiment-1. As the distance between the adjacent cards is 30cm, that an averaged perceived height for the first segments of the sounds is 45cm, for example, can be recognized that half of the observers perceived the sound images for the first segments around the card numbered 3, and the rest perceived around the card numbered 2.

Fig.6 shows the results of the 15 observers for the stimuli descending in pitch. The most observers perceived the sound images moved downward to a position slightly higher than the loudspeaker-2. The difference between the highest position of the sound images and the lowest for piano-tones was approximately 60cm, while that for one-third octave noises was less than 30cm. The averaged positions of the sound images for the last segments in the experiment-2 were approximately 15cm lower than those in the experiment-1. Only for the piano tones, the perceived positions of the sound images for the first segments were as high as those in the experiment-1.

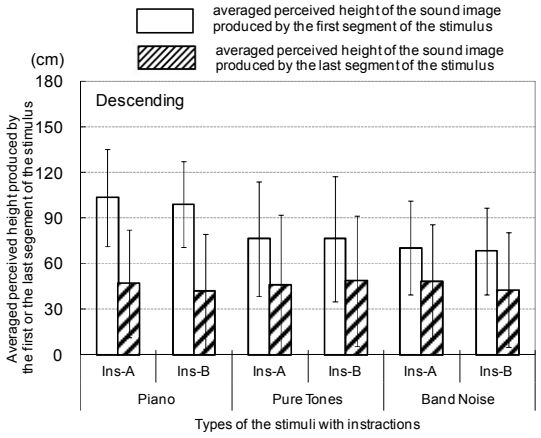

Figure 6.- The averaged perceived heights of the sound images produced by the first and the last segments of the stimuli descending in pitch. For other details, see the caption in Fig.3.

DISCUSSION

Vertical sound localization of tones with constant pitch

Prior to the experiment-1 and -2, Agatsuma and Miyasaka[1] performed experiments of vertical sound localization using the stimuli consisting of pure tones with three types of constant frequencies of 1, 4 and 8 kHz and one-octave band noises with the center frequencies of 1, 4 and 8kHz for 28 observers. A stimulus was radiated randomly through one of the seven loudspeakers perpendicularly set up as shown in Fig.1 except the highest loudspeaker.

Fig.7 shows a part of the results for the pure tones in the case that they were radiated from the loudspeaker-2. Fig.8 shows a part of the results for the one-octave band noises radiated from the loudspeaker-2. In both figures, the abscissa indicates the card number which was attached to the corresponding loudspeaker from which the sound was heard, while the ordinate indicates the total number of the observers.

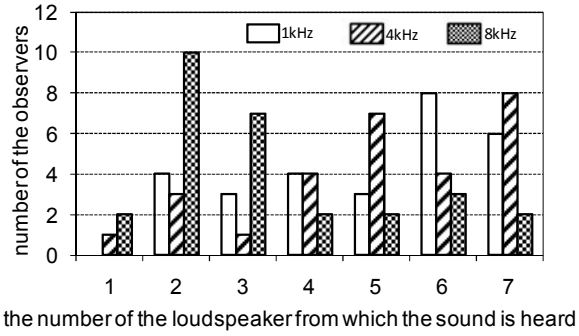

Figure 7.- Relationship between the number of the observers and the identified loudspeaker number when the pure tones are radiated from the loudspeaker-2.

For 1kHz-tone in Fig.7, there were only four observers (14%) who perceived the sound images at the loudspeaker-2 from which the real sounds were output. The rest observers (86%) perceived the sound images higher than the loudspeaker-2. For 4kHz-tone, there were also only three observers (11%) who perceived the sound images at the loudspeaker-2. For 8 kHz-tone, ten observers (36%) perceived the sound images at the loudspeaker-2. These results indicate that the sound images are approximately perceived at equal to or higher positions than the position of the loudspeaker-2. However, it is not clear that the higher the tones are in pitch, the higher the sound images are located.

For one-octave band noises with the center frequency of 1, 4 and 8 kHz in Fig.8, most observers perceived the sound images at equal to or 30cm higher than the position of the loudspeaker-2. However, it is also not clear that the higher the center frequencies of the noises are, the higher the sound images are located. These results imply that it will be difficult for observers to perceive the locations of the sound images along their fundamental frequencies when they are presented randomly in pitch.

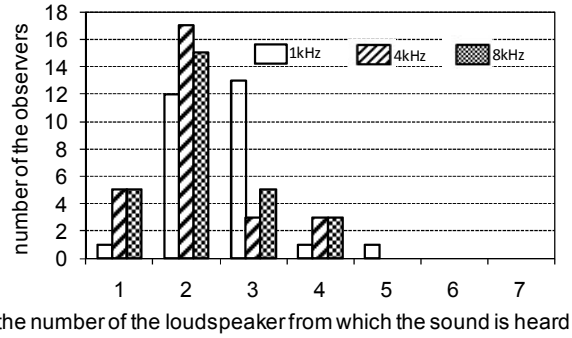

Figure 8.- Relationship between the number of the observers and the identified loudspeaker number when the one-octave band noises are radiated from the loudspeaker -2.

Vertical sound localization of tones with constant pitch when they are moved upper or lower

Agatsuma and Miyasaka performed another experiments of vertical sound localization using the one-third octave band noises with the center frequencies of 1, 4 and 8kHz as function of moving speed. For each center frequency, the stimuli were moved smoothly from lower positions to higher positions and in reverse direction. The observers were required to point the locations of the sound images at the starting points of them and at the end points.

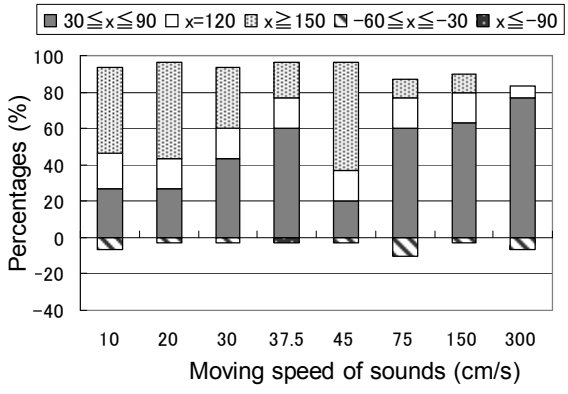

Figure 9.- Psychological moving distances of sounds which were moved from the loudspeaker-1 to -5 with the various moving speeds shown in the abscissa. The ordinate indicates the percentages of observers who perceived the moving distances. The filled boxes, for example, indicate the percentages of the observers who perceived the sounds moved between 30cm and 90cm.

Fig.9 shows one of the results for one-third octave band noises as a function of moving speeds of the sounds. The sounds were moved from the loudspeaker-1 to 5 with various moving speeds.The abscissa indicates the moving speed (cm/s) of the stimuli. The ordinate indicates the percentages of observers who perceived the moving distances. The filled gray boxes, for example, indicate the percentages of the observers who perceived the sounds moved between 30cm and 90cm.

The results show that over 90% of the observers perceived that the sound images moved upward for all the stimuli when they were moved upward irrespective of the movins speeds. These results imply that observers perceive the sound images move smoothly from low to high although the center frequencies of the noises are kept constant.

Vertical sound localization of piano-tones with ascending and descending frequencies over three octaves

In the experiment-1 and -2, the fundamental frequencies of the stimuli were ascending or descending. When these sounds will successively be concatenated over 3 octaves, how the observers perceive the sound images? In order to investigate the question, an additional experiment was performed. Only piano-tones were used. The first half of the stimulus was the ascending tone sequences, while the latter was the descending sequences, both of them were the same as used in the experiment-1. The stimulus was moved smoothly upward from the loudspeaker-1 to the loudspeaker-4 for the first half and successively downward from the loudspeaker-4 to the loudspeaker-1 for the last half of the stimulus. The observers were required to answer the highest location of the sound images by the loudspeaker-number.

Fig.10 shows the results of the experiment. Only five observers answered that the highest location of the sound images was at the loudspeaker-4 from which the sound segment with highest fundamental frequency (C7) was output. Approximately more than half observers perceived the maximum locations 90cm no less higher than the location of the loudspeaker-4. These results imply that the locations of the sound images will depend on a series of stimuli with ascending and/or descending fundamental frequencies.

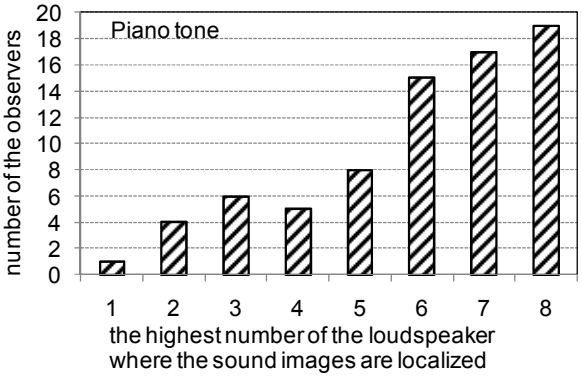

Figure 10.- The highest number of the loudspeaker where the sound images are localized for piano-tones. The first half of the stimulus was the ascending tone sequences, while the latter was the descending sequences, both of them were the same as used in the experiment-1. The stimulus was moved smoothly upward from the loudspeaker-1 to the loudspeaker-4 for the first half and successively downward from the loudspeaker-4 to the loudspeaker-1 for the last half of the stimulus.

CONCLUSION

As already mentioned in INTRODUCTION, the purpose of this study is to clear whether the sound image will move or not for sounds with fundamental frequencies varying from low to high or from high to low when they are presented through a fixed loudspeaker. The results show that the sound images are perceived to move upward or downward with successive increase or decrease of the fundamental frequencies, respectively, although it is difficult for observers to identity the correct locations of the sound images when the stimuli with constant fundamental frequencies are presented randomly at the loudspeakers perpendicularly set up. These results imply that effective vertical movements may be realized, for example, by transfer sounds played by bass instruments smoothly to those by soprano instruments.

ACKNOWLEDGEMENTS

The author express thanks to Mr.K.Iwata and Mr.H.Miyauchi who assisted the author to perform the experiments and gather the experimental data. A part of this work was supported by the Assistance Grants of HBF (Hoso Bunka Foundation).

REFERENCES

1 Y.Agatsuma and E.Miyasaka, “Perceptual movement of sounds fed through multiway loudspeakers set up perpendicularly”, 19th International Congress on Acoustics, Madrid, Spain, (2007).

2 K. Kurozumi, “Movement of a phantom sound image produced from loudspeakers placed in a vertical plane”, Journal of Acoustical Society of Japan (in Japanese), pp 359-360 (1990)

3 K. Kurozumi, “Direction of sound image produced form two loudspeakers placed in a vertical plane”, Journal of Acoustical Society of Japan (in Japanese), pp 437-438 (1989)

4 S. Ferguson and D. Cabrera, “Vertical Localization of Sound from Multiway Loudspeakers”, Journal of Audio Engineering Society Vol. 53, No. 3, pp 163-173 (2005 March)