PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Experiment on audio-visual mutual effect on subjective impression in architectural space by HMD VR display

Takane TERASHIMA[1]; Chiharu HIRAOKA[2]; Xue GONG[3]; Yasunobu TOKUNAGA[4];
Ayumi ISHIKAWA[5];

[1-3] Mie University, Japan

[4] National Institute of Technology, Maizuru College, Japan

[5] National Institute of Technology, Gifu College, Japan

ABSTRACT

The purpose of this study is to clarify the influence of visual information on subjective impression for room sound field. Authors have carried out subjective experiments on it using room impulse response signal as auditory stimulus and interior panorama picture as visual stimulus, and found that reverberance/clarity for room sound field is affected by information sensible from room volume, dimensions, geometry or finishing etc. Visual stimulus of interior panorama picture is projected on flat screen in these experiments, but visual experiences were not always same as those in real room. In this report, new experiments using head mounted display instead of flat screen in presenting visual stimulus were conducted. Results were compared with previous experiment of flat screen, and the presentation methods of visual stimuli were inspected. It was found that subjective dimensions for room were affected by reverberance of auditory impression in the new experiment unlike previous one. Influence of visual information was clarified in the close condition to reality.

Keywords: Visual, Auditory, Mutual Effect, Cross-modality

1. INTRODUCTION

The purpose of this paper is to clarify the mutual effects of audio and visual impression in room interior and sound field. Authors have carried out subjective experiments using room impulse response signal through loudspeakers and interior panorama picture projected on a flat screen in anechoic chamber [1~4]. It is found that reverberance/clarity for room sound field is affected by visual information such as room volume, dimensions, geometry or finishing etc., whereas width/brightness/hardness etc. of room interior attribute seems to be little affected by auditory information such as reverberation time etc. Visual stimulus of interior panorama picture (QuickTime VR type) projected on flat screen was used in these experiments, but situations of visual experiences are different from real room to some extent. In this report, new experiments using Head Mounted Display (VR Headset, HMD) instead of flat screen in presenting visual stimulus were conducted. Results were compared with previous experiments, and the difference of presentation methods of visual stimuli were inspected. Then, all the results of experiments are compared and integrated, so that the properties of mutual effect of audio and visual in architectural room could be clarified again.

2. SUBJECTIVE EXPERIMENT

Subjective experiments about cross-modality in architectural space were carried out by audio-visual stimuli consisted of internal panorama pictures and impulse response sound of rooms.

[1] tera@arch.mie-u.ac.jp
[2] chiharu@e.arch.mie-u.ac.jp
[3] gongxue@e.arch.mie-u.ac.jp
[4] tokunaga@maizuru-ct.ac.jp
[5] ishikawa@gifu-nct.ac.jp

2.1 Stimulus

Table 1. Rooms and their acoustic properties (500Hz)

Name	Type/Attribute	Volume [m^3]	Rev. Time (T_{30}) [s]	Clarity (C_{80}) [dB]
Cl	Class Room	750	0.74	9.2
Ha	Hall	12000	1.45	0.1
Gy	Gymnasium	6000	3.29	3.2

Figure 1. Room "Cl"

Figure 2. Room "Ha"

Figure 3. Room "Gy"

Visual stimuli were made from panorama internal pictures taken from 3 rooms listed in Table 1. Overview pictures of room interiors are shown as Figure 1-3.

Table 2. Reverberation Times of auditory stimuli [s] (500Hz)

Stimuli Name	subscript					
	1	2	3	4	5	6
cl	0.23	0.74	1.21	1.76	2.25	2.74
ha	0.62	0.95	1.47	1.92	2.41	2.85
gy	0.78	1.34	1.82	2.28	3.33	4.78

*Gray cell means original (actual) reverberation time of the room.

Models of sound files were made by Digital Signal Processor (DSP) to approximate to impulse responses measured in these 3 rooms, then auditory stimuli were made from these models and modified in reverberation times by DSP operation. Reverberation times of auditory stimuli were listed in Table 2.

In experiment, these auditory and visual stimuli were provided to subjects separately (in single-mode) or combinedly (in multi-mode). And visual stimuli were presented to subjects through screen projection or HMD in which they could look around internal of room. Auditory stimuli were presented to subjects through headphone or loudspeaker in anechoic chamber and they could hear the room response.

2.2 Method

Table 3. Presentation Method including previous experiment

Method	Visual stimuli	Auditory stimuli
SL*	Screen projection	Loudspeaker in anechoic chamber
SH		Headphone
VL	Head Mount Display (VR Headset)	Loudspeaker in anechoic chamber
VH		Headphone
RH	(Real Space)	Headphone

*reported in previous paper [1~4]

Figure 4. HMD with Headphone (Method "VH")

Subjects were presented with stimuli in 4 modes (audio/visual x single/multi) per each method. Including previous experiments, all the methods of presentation are shown in Table 3. The situations of experiments VH, SL and VL are shown in Figure 4, 5 and 6. The results of experiments by presentation method SL had been previously reported [1~4], then in this report the results of SH, VL, VH, and RH are described in addition to those of SL. (RH means that subjects evaluate in real space with headphone.) The number of subjects were 10. Subjects were asked to answer rating in 7 scales (-3, -2, -1, 0, +1, +2, +3) for every stimulus in evaluation items listed in Table 4. Subjective impressions of all experiments were compared in terms of difference between methods and modes.

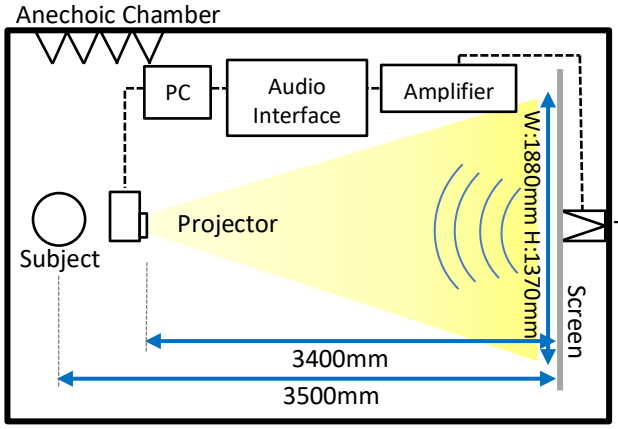

Figure 5. Presentation Method "SL"

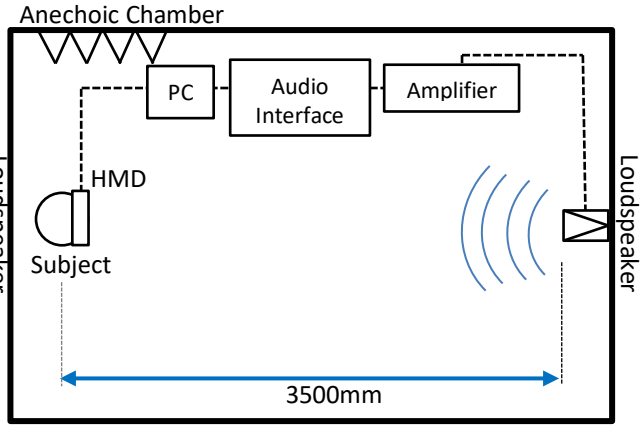

Figure 6. Presentation Method "VL"

Table 4. Subjective evaluation Items

Mode		Item	-3 <----------Adjective----------> +3		
Audio-Visual (Multi)	Visual (Single)	Brightness	Dark	···	Bright
		Warmth	Cold	···	Warm
		Width	Narrow	···	Wide
		Dimension	2-dimensional	···	3-dimensional
		Familiarity	Strange	···	Familiar
		Complexity	Simple	···	Complex
		Ceiling Height	Low	···	High
		Hardness	Soft	···	Hard
	Audio (single)	Length of Reverberation	Short	···	Long
		Fullness of Reverberation	Poor	···	Full
		Clarity of Acoustics	Dim	···	Clear

3. RESULTS AND DISCUSSIONS

3.1 Auditory impression

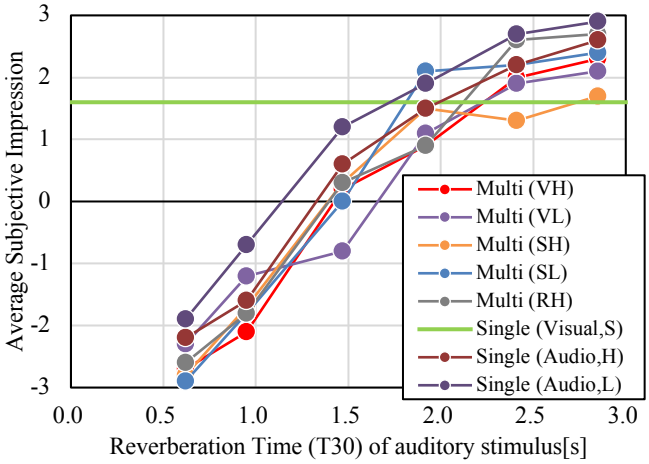

Figure 7. "Length of Reverberation" (Ha)

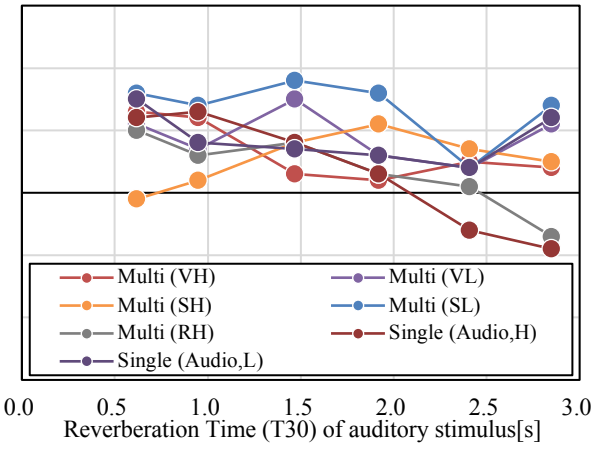

Figure 8. "Clarity of Acoustics" (Ha)

Subjective evaluations about auditory impression in different methods and modes for the same stimulus are compared.

Figure 7 and 8 show the results of "Length of Reverberation"/ "Clarity of Acoustics" for space "Ha". In Figure 7, at any methods, average subjective impression, "Length of Reverberation" tends

to increase with reverberation time of stimulus. Impression for auditory stimuli by loudspeaker reproduction in audio single-mode at short reverberation time are a little higher than other methods, but there is no significant difference of impression for the same stimulus between methods. Expected impression from visual stimulus only (visual single-mode S) is about 1.6 corresponding to reverberation time 2.0. In Figure 8, as a whole, auditory impressions "Clarity of Acoustics" seem to decrease with reverberation times of auditory stimuli, but in method VL and SL in multi-mode, and method L in audio single-mode, auditory impressions behave regardless of reverberation times. It seems to be difficult for subjects to judge Clarity by listening impulse response signals radiated from loudspeaker in this experiment and subjective responses were scattered. With regard to SH, Clarity is maximum at near reverberation time 2.0s, but the reason cannot be explained.

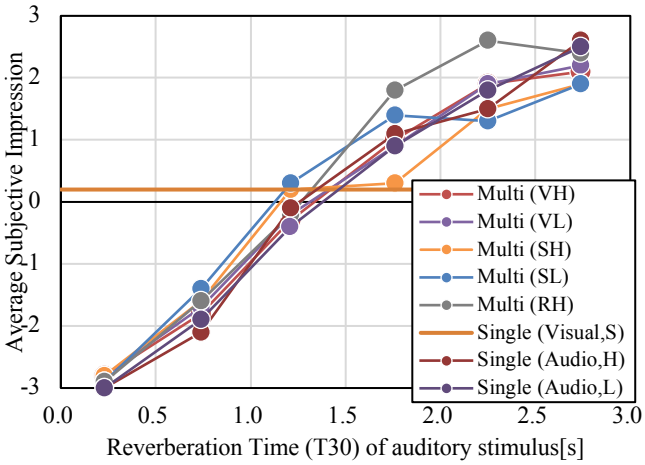

Figure 9. "Length of Reverberation" (Cl)

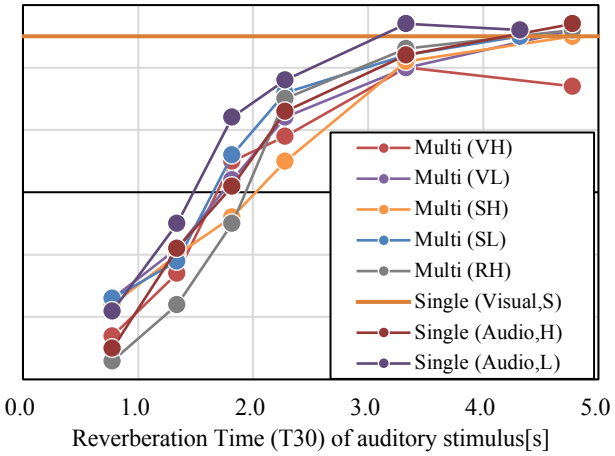

Figure 10. "Length of Reverberation" (Gy)

Figure 9 and 10 show the results of "Length of Reverberation" for room "Cl" and "Gy". In Figure 9, at any methods, auditory impression tends to increase with reverberation time and there is no significant difference of impression for the same stimulus between methods except for method RH (real room with headphone), in which impressions for auditory stimuli at reverberation times longer than 1.5s are higher than other method. Expected impression from visual stimulus (visual single-mode S) is about 0.2 corresponding to reverberation time 1.2s, that is fairly longer than actual reverberation time. In Figure 10, similarly to Figure 9, t any methods, auditory impression tends to increase with reverberation time and there is no significant difference of impression between methods. Expected impression from visual stimulus is about 2.5 corresponding to reverberation time 3.3s over, that corresponds to actual reverberation time on the whole.

3.2 Visual impression

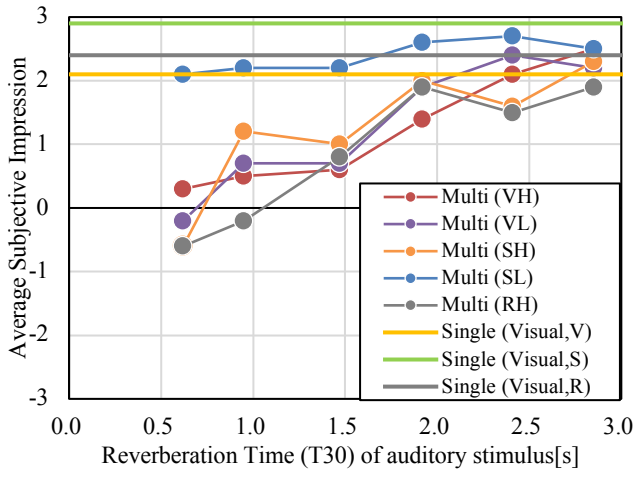

Figure 11. "Width" (Ha)

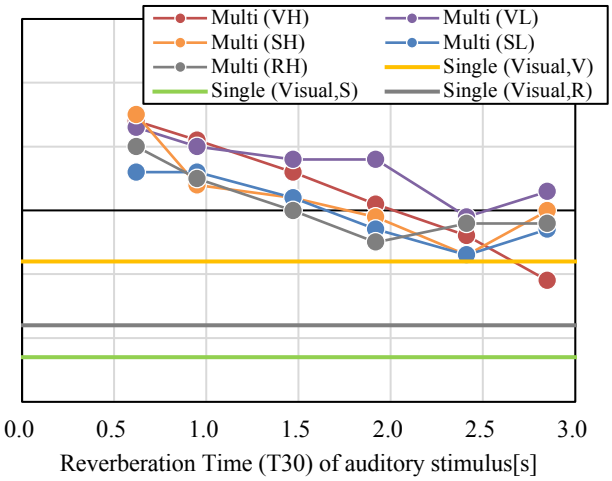

Figure 12. "Hardness" (Ha)

Figure 11 shows the result of "Width" and Figure 12 shows the result of "Hardness" for space "Ha".

In Figure 11, visual impression about width of room increased with reverberation time of auditory stimulus in all the method of experiment except for SL. There is no significant difference between modes for the same auditory stimulus. “Width” by method SL changes little and is almost constant near the visual single-mode values regardless of reverberation time. It is supposed that when visual information is provided on screen and auditory information is provided by loudspeaker, audio-visual integration about room dimension would not occur and visual impression of width might not be influenced by auditory impression. In Figure 12, “Hardness” is also influenced by auditory information and tends to decrease with reverberation time. There is no significant difference between modes for the same auditory stimulus. As a whole, “Hardness” in visual single-mode is lower than multi-mode.

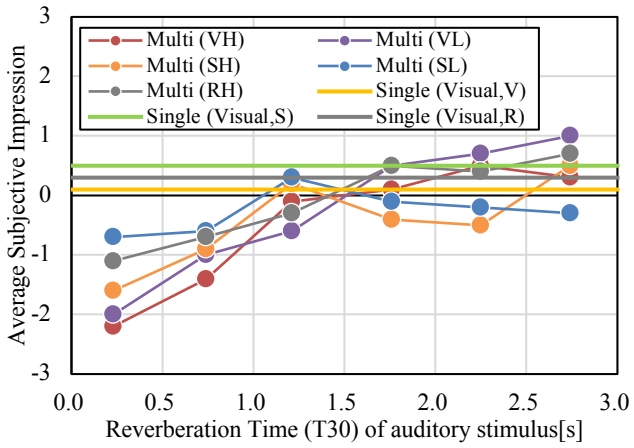

Figure 13. “Width” (Cl)

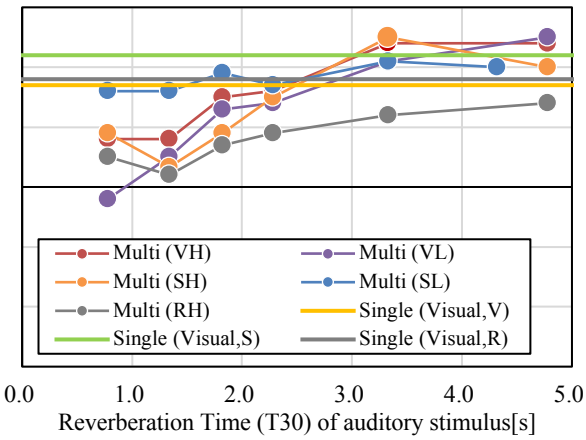

Figure 14. “Width” (Gy)

Figure 13 shows the result of “Width” for space “Cl” and Figure 14 for space “Gy”. Like Figure 11, visual impression about width of room increased with reverberation time in all the method of experiment except for SL. Width by method SL changes little and is almost constant near the visual single-mode values regardless of reverberation time.

3.3 Difference between multi-mode and single-mode values

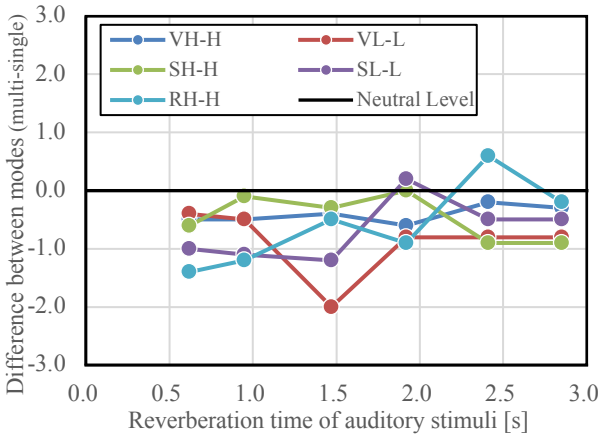

Figure 15. Difference between modes ("Length of Reverberation", Ha)

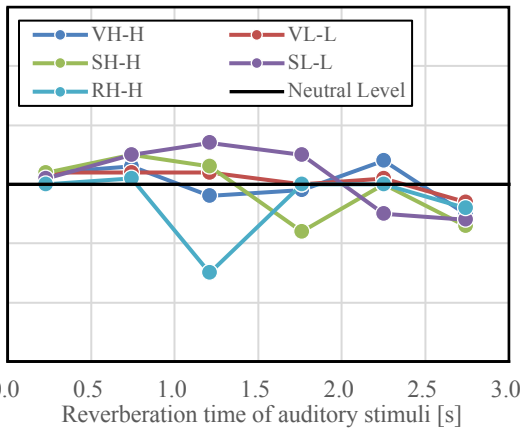

Figure 16. Difference between modes ("Length of Reverberation", Cl)

Differences of impression between multi- and single-mode values are shown in Figure 15-18. In these figures, vertical axis means impression level in multi-mode relative to single-mode, and negative value means that multi-mode value is smaller than single-mode. In room “Ha” and “Gy”, which have large volume, most of plotted points for auditory impression “Length of Reverberation” have negative values at most methods and stimuli. It is supposed that auditory image came to have a kind of boundary limitation by being provided with concrete internal image of visual stimuli in multi-mode, while auditory image could spread widely as free space without boundary in single-mode regardless of reverberation time. In small room “Cl”, plotted points are located around neutral level and auditory impression seems to be little influenced by visual information except for

stimulus cl_3 at RH-H. In this room, a short reverberation times auditory impression has consistency with visual information but points shift a little in positive or negative direction in a range of audio-visual integration. The reason why the evaluation for cl_3 in multi-mode (real room and headphone presentation) shift largely to negative area cannot be explained.

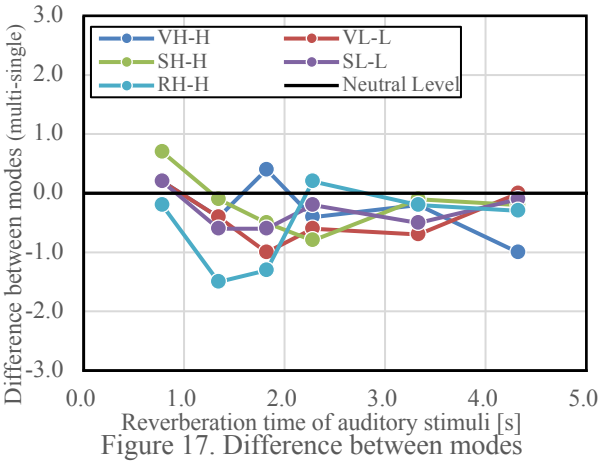

Figure 17. Difference between modes ("Length of Reverberation", Gy)

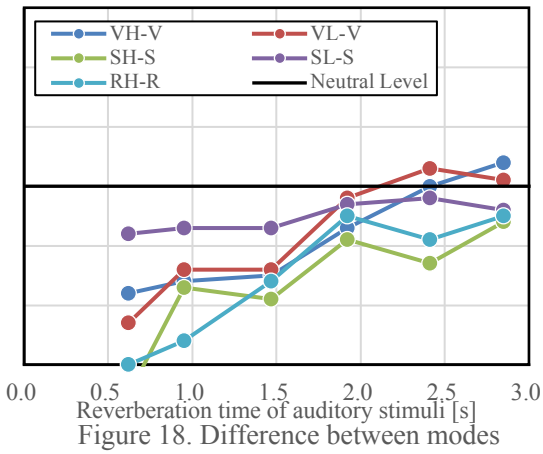

Figure 18. Difference between modes ("Width", Ha)

In Figure 18, visual impression of "Width" is clearly influenced by auditory information; reverberation times, and in room "Ha" impression shift in negative direction largely especially at short reverberation times except for SL-L. It seems that at long reverberation times over 2.0s, which is longer than actual room, visual impression has consistency with auditory information. In SL-L, the difference of impression almost constant at a little negative value for all reverberation times. It is possible that in that case subjects cannot integrate visual and audio information, and the combination of presentation methods, screen projection and loudspeaker production, which has been previously used in our experiments, might be unsuitable in our studies.

4. CONCLUSIONS

- In subjective experiments about audio and visual cross-modality, screen projection and HMD, loudspeaker and headphone were cross-combinedly used as methods of stimulus presentation and the results were compared.
- Almost all the methods of experiment behave similarly for the same stimuli except for the method SL; screen projection and loudspeaker reproduction combination multi-mode.
- It is confirmed that auditory impression in terms of reverberance is influenced by visual information by other methods of stimuli presentation than previous method.
- The method SL showed the different behavior from other methods for visual impressions "Width", and visual impression seems to be little influenced by auditory stimuli. It is supposed to be unsuitable in our studies.

REFERENCES

1. Y. Takeyama, A. Ishikawa and T. Terashima; Relationship of auditory impression for room space with discordance among auditory and visual sense Part II, Proceedings of Annual Research Meetings of AIJ, Tokai Chapter, Vol.56, pp.321-324 (2018)
2. Y. Takeyama, A. Ishikawa and T. Terashima; Comparison of stimulus presentation methods in experiments to study on subjective impression for audio and visual of interior, Summaries of Technical Papers of Annual Meetings of AIJ, Environmental Engineering I, pp.141-142 (2018)
3. A. Ishikawa and T. Terashima; Study on the effect of sense of incongruity caused by combination with visual information in subjective reverberance for indoor sound field in architectural space, Proc. of the Fifth International Conference on Human-Environment System (ICHES2016 Nagoya), 20030, p.10 (2016)
4. A. Ishikawa, T. Terashima and Y. Tokunaga; Effect of Visual Stimulus on Subjective Impression of Indoor Sound Fields with Various Reverberation Times, Proceedings of internoise2014, the 43th International Congress on Noise Control Engineering, 133 (2014)