inter.noise 2014
MELBOURNE AUSTRALIA
16-19 NOVEMBER

Effect of Visual Stimulus on Subjective Impression of Indoor Sound Fields with Various Reverberation Times

Ayumi ISHIKAWA[1]; Takane TERASHIMA[2]; Yasunobu TOKUNAGA[3]

[1, 2] Mie University, Japan

[3] Maizuru National College of Technology, Japan

ABSTRACT

In architectural and urban space, we are always exposed to multi-modal stimuli of visual information and sound fields in various scenes of everyday life. The purpose of this study is to clarify relationships between subjective impressions for vision (size, shape, colour, design etc.) and auditory (sound field as a structure of reflections) of indoor/outdoor space, and to acquire information which contributes to architectural design or acoustic design. In this study, laboratory experiments were carried out, in which subjective impression for indoor sound fields were measured when auditory stimuli with different reverberation times and visual stimuli of indoor VR pictures were presented simultaneously. In our previous papers [1~4], we compared measured values of single-modal presentation (auditory stimuli only) with those of multi-mode (visual and auditory stimuli), and could approximately understand the influence of visual stimuli on auditory impression. In this paper, it is mainly described that additional experiments were carried out under newly arranged conditions, so that the psychological influence of incongruity between visual and auditory stimulus on subjective impression for sound field could be clarified.

Keywords: Multi and Single-modal, Visual and Auditory stimuli, Subjective impression
I-INCE Classification of Subjects Number(s): 51.1

1. INTRODUCTION

In this study, we consistently had carried out experiments to understand the effect of presenting multimodal information of visual and auditory on the subjective impression of sound fields. In this paper, results of a series of three experiments are reported. In experiment I (Ex.I), various subjective impressions for multimodal presentation of combined visual stimuli and auditory stimuli with different reverberation time (RT) were measured. In experiment II (Ex.II), similar experiments as Ex.I were carried out, but auditory stimuli are synthesized and controlled with RT. Experiment III (Ex.III) is consisted of the following 3 kinds of subjective evaluation experiments in the laboratory in order to clarify the effect of visual information on auditory impression in various real architectural spaces.

- Ex.III-1. Subjective visual impressions and predicted RT were measured when only visual stimuli (virtual reality (VR) images of 3 kinds of real architectural spaces) were presented to subjects.
- Ex.III-2. Subjective auditory impressions were measured when only auditory stimuli (3 sound fields of the same architectural spaces as Ex.III-1 and 6 sound fields synthesized with RT at 0.5 s interval on personal computer) were presented to subjects.
- Ex.III-3. Subjective auditory impressions and the sensual degree of incongruity were measured when multimodal stimuli (27 combinations of visual stimuli, the same as Ex.III-1 and auditory stimuli, the same as Ex.III-2) were presented to subjects.

The results of these experiments were analyzed statistically and the effect of visual information on subjective auditory impression was discussed.

[1] ayumi@e.arch.mie-u.ac.jp
[2] tera@arch.mie-u.ac.jp
[3] tokunaga@maizuru-ct.ac.jp

2. EXPERIMENTAL OUTLINE

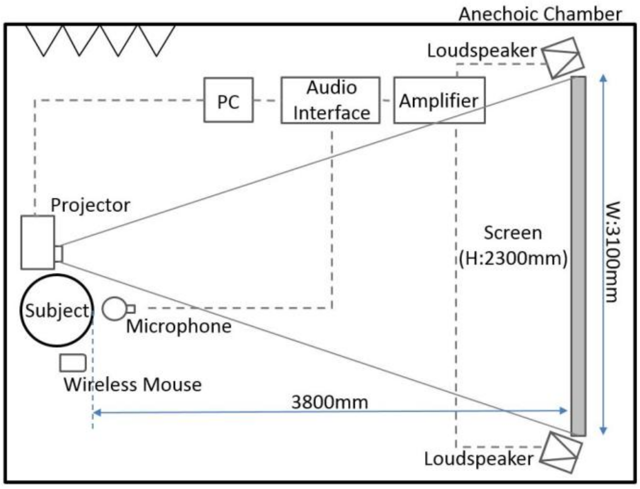

Figure 1 - System outline of Ex.I and II

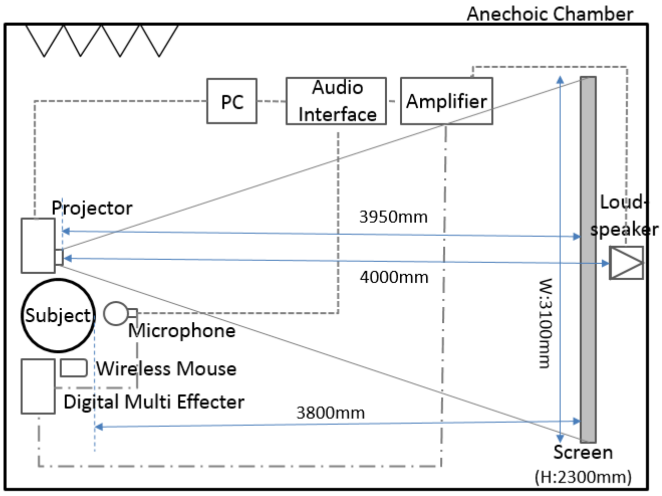

Figure 2 - System outline of Ex.III

2.1 Ex.I and II

The experimental system was built, in which visual stimuli by VR images and auditory stimuli by superimposed sound field were presented to subjects individually or simultaneously in anechoic chamber, and subjective impressions for stimuli were measured. As shown in Figure 1, in anechoic chamber, VR image was projected on the screen, and a subject was able to overlook 360 degrees of inner view of each architectural space by mouse operation manually. A microphone was installed in front of the subject, and the sound, which he spoke or clapped, was collected through microphone and processed in real time convolution with impulse response (IR) of auditory stimulus through personal computer (PC). Sound field of auditory stimulus was reproduced stereophonically from loudspeakers. 10 subjects with normal eyesight and hearing ability participated in these experiments.

<u>Procedure</u>

Table 1 - Acoustical properties (based at 500 Hz) of IRs used as auditory stimuli at Ex.I and II

	IR	T_{sub} [s]	EDT [s]	C_{80} [dB]
	Based IR	3.56	3.11	-1.8
Ex. I	1	0.52	0.53	11.5
	2	1.05	1.21	3.9
	3	1.56	1.55	1.9
	4	2.03	1.83	-1.0
	5	2.52	2.22	-0.7
	6	3.13	2.73	-1.5

	IR	T_{sub} [s]	EDT [s]	C_{80} [dB]
	Based IR	3.56	3.11	-1.8
Ex. II (III)	7	0.80	0.99	6.0
	8	1.09	1.26	2.7
	9	1.42	1.41	2.8
	10	1.73	1.57	-0.6
	11	2.17	2.06	-0.6
	12	3.06	2.46	-1.8

Table 2 - Use and acoustical properties (based at 500 Hz) of spaces used as visual stimuli at Ex.I and II

Space	Use	T_{sub} [s]	EDT [s]	C_{80} [dB]
a	Class room	0.68	0.65	8.8
b	Restaurant	0.94	7.23	1.1
c	Corridor	1.57	1.42	5.6
d	Entrance	1.94	1.99	4.8
e	Gymnasium	3.07	2.72	3.2

The common procedure of Ex.I and II are summarized as follows. First, only auditory stimuli were presented (single-mode), and subjects were asked to answer the subjective auditory impression for each stimulus about “Duration of Reverberation”, “Reverberance” and “Clarity” in 7-step category scales (-3 ~ +3 or 0 ~ +7). Next, auditory and visual stimuli were simultaneously presented (multi-mode), and subjects were asked to answer subjective impression about auditory and visual. Questions about auditory were the same as single-mode and those about visual are 14 adjective pairs about vision (e.g. "Brightness", "Complexity" and "Capacity"). And additionally subjects were asked to answer sensual degree of incongruity between sound and image in 4-step category scales from 0 (not detected) to +3 (most detected).

Stimulus

Five existing spaces "a" ~ "e" were adopted as materials of visual stimuli for experiment, which are spaces used daily by a large number of general public, and have different inner view design and indoor acoustical properties. Acoustical properties (based at 500 Hz) of auditory stimuli are listed in Table 1. RTs of auditory stimuli in Ex.I were changed at 0.5 s interval by editing waveform of basic IR on PC. RTs of auditory stimuli of Ex.II were also changed by editing waveform so that the distance of subjective RT between stimuli should be equal on the basis of the result of subjective evaluation "Duration of Reverberation" in Ex.I. The information of visual stimuli is shown in Table 2.

2.2 Ex.III

As shown in Figure 2, Ex.III was the similar to Ex.I and II was built, but with a digital Multi-effecter (YAMAHA SPX990J) was added to that. Only 1 loudspeaker installed in the center behind the screen, and sound fields were reproduced monophonically from that. 10 subjects (different from Ex.I and II) with normal eyesight and hearing participated in this experiment.

Procedure of Ex.Ⅲ-1

Table 3 - Use and acoustical properties (based at 500 Hz) of spaces used as visual stimuli at Ex.III

Space	Use	T_{sub} [s]	EDT [s]	C_{80} [dB]
A	Class room	0.68	0.65	8.80
B	Hall (multi-purpose use)	1.50	1.71	0.07
C	Gymnasium	3.07	2.72	3.20

Table 4 - Question items of Ex.III-1

Category	Items	Pair of Adjective	
Color perception	Brightness	Whitish	Blackish
	Saturation	Vivid	Sordid
	Hue	Warm	Cool
	Color	Monochrome	Polychrome
	Luminosity	Light	Dark
Space perception	Solidity	Solidity	Flat
	Complexity	Complex	Simple
	Openness	Open	Closed
	Area	Large	Small
	Softness	Soft	Hard
	Intimacy	Intimate	Unfamiliar

First, only VR images as visual stimuli shown in Table 3 were presented randomly in single-mode, and subjects overlook inner views of each architectural space by the same manner of Ex.I and II. Next, subjects were asked to identify and determine expected RT from visual information, to adjust RT values by switch operation on the digital-multi-effecter by themselves while listening to the processed sound through the digital-multi-effecter. Finally, subjects were asked to answer the subjective visual impression in 7-step category scales (-3 ~ +3). Question items are shown in Table 4. Moreover, they were asked to answer whether to have experiences with the spaces in visual stimuli in 4-step ("Have used it", "Probably have used it", "Probably have not used it", and "Never used it").

Procedure of Ex.Ⅲ-2

Table 5 - Question items of Ex.III-2

Items	Pair of Adjective	
Duration of Reverberation	Long	Short
Reverberance	Rich	Poor
Clarity	Clear	Vague

Only sound fields as auditory stimuli were presented randomly in single-mode. There were 9 stimuli, 3 sound fields of real architectural space used as visual stimulus in Ex.III-1 and 6 sound fields synthesized on PC (the same stimuli as Ex.II). Subjects listen to sound fields reproduced by convoluting sound through microphone with auditory stimuli, and were asked to answer subjective auditory impressions in 7-step category scales (-3 ~ +3). Question items are shown in Table 5.

Procedure of Ex.Ⅲ-3

Three visual stimuli used in Ex.III-1 and 9 auditory stimuli used in Ex.III-2 were combined into 27 audio-visual stimuli combinations (multi-modal stimuli). These were presented randomly to subjects. They were asked to answer the subjective auditory impression for multi-modal presentation (visual and auditory simultaneously). Question items were the same as Ex.III-2. Additionally, they were asked to answer the sensual degree of incongruity for the combination of sound fields and images in 7-step category scales (-3: much incongruity ~ +3: no incongruity).

3. RESULT & DISCUSSION

3.1 Result of Ex.I

Figure 3 shows the average values of the subjective evaluation of “Duration of Reverberation”. Subjects were able to distinguish the difference between stimuli because most of them evaluated according to the order of RT duration, in both modes. At long RT (>=2.0 s), the order of subjective values and RTs do not agree and subjects seem not to be able to distinguish the difference between stimuli. And in most cases, values in multi-mode are also higher than single-mode. However, subjects do not necessarily feel that the combination of real auditory and visual condition of each space is most suitable. Properties of “Reverberance” were similar as “Duration of Reverberation” in most aspects.

Figure 4 shows the average values of subjective evaluation of “Clarity”. It seems that the values in multi-mode are lower than single-mode in many cases. Particularly at long RT, the influence by visual stimuli is large and the values greatly fall down from single-mode value.

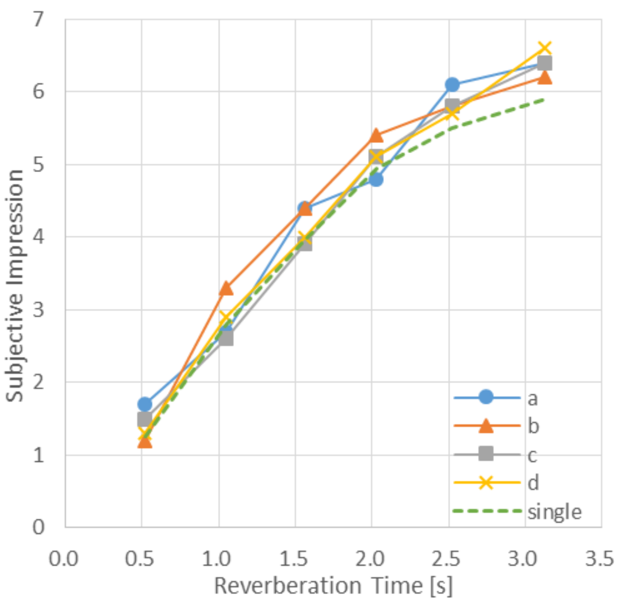

Figure 3 – Average values of “Duration of Reverberation” (Ex.I)

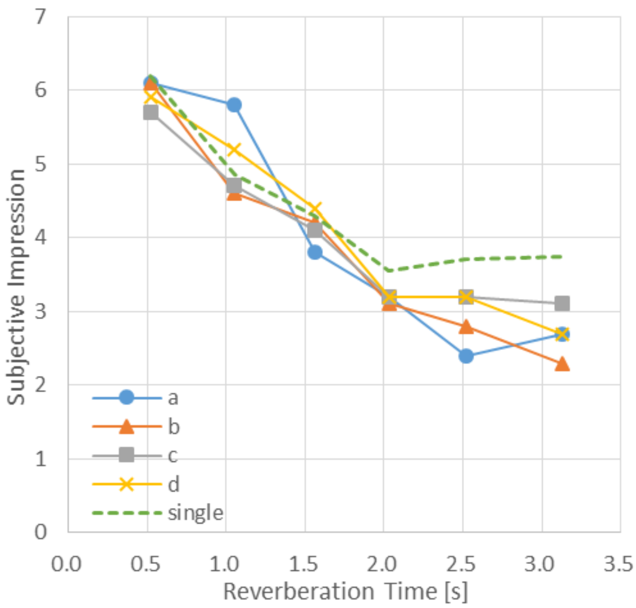

Figure 4 – Average values of “Clarity” (Ex.I)

3.2 Result of Ex.II

The properties of results from Ex.II were mostly the same as Ex.I. About "Duration of Reverberation" at stimuli in both modes, the relationship of subjective values and RT was almost linear. The longer RT becomes, the higher subjects evaluate in many cases of "Reverberance", but the relationship of values and RT is less linear for RT than in "Duration of Reverberation". In addition, each auditory stimulus was evaluated by the order of RT duration in single-mode, but the value for auditory stimulus "6" (RT=3.06 s) is lower than stimulus "5" (RT=2.17 s) in multi-mode except the combination with visual stimulus "c". As RT of auditory stimuli became long, "Clarity" became small, and it means the evaluation is almost contrary to "Duration of Reverberation". At long RT stimuli, there is nonlinearity between values and RT.

<u>**Change of Subjective Evaluation between Modes**</u>

Figure 5 and 6 show the ratio of the average value of "Duration of Reverberation" and "Clarity" in multi-mode to single-mode. The ratio is higher than 1.0 in most multi-modal stimulus for "Duration of Reverberation" (and "Reverberance"), i.e. subjective impression about reverberation increase when auditory stimulus is added with visual information. In addition, when the stimuli with longer RT than auditory stimulus "4", the variance of ratios in visual stimuli became smaller, and ratios seem to approach 1.0. However, there seems to be no difference in tendency to change by visual stimulus.

On the other hand in "Clarity", there is a certain amount of multi-modal stimuli that are rated less than 1.0. When the stimuli with smaller RT than auditory stimulus "5", ratios seem to be near 1.0, and in the stimuli with longer RT, ratios vary greatly by the visual stimulus.

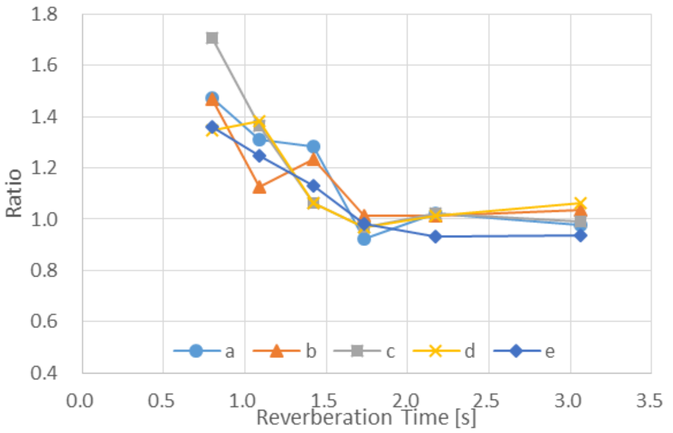

Figure 5 – Ratio of "Duration of Reverberation" in multi-mode to single-mode (Ex.II)

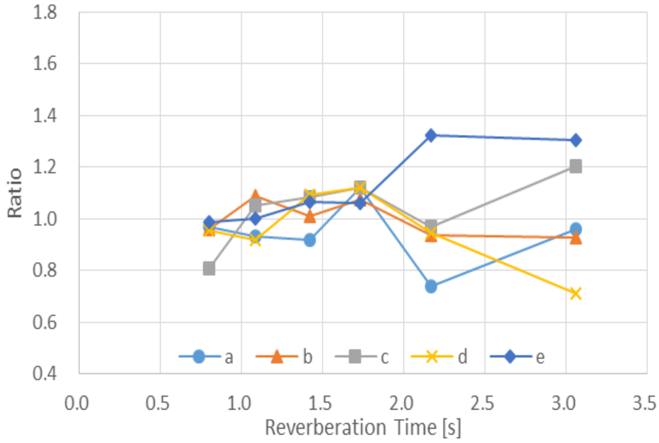

Figure 6 – Ratio of "Clarity" in multi-mode to single-mode (Ex.II)

3.3 Discussion about Result from Ex.I and II

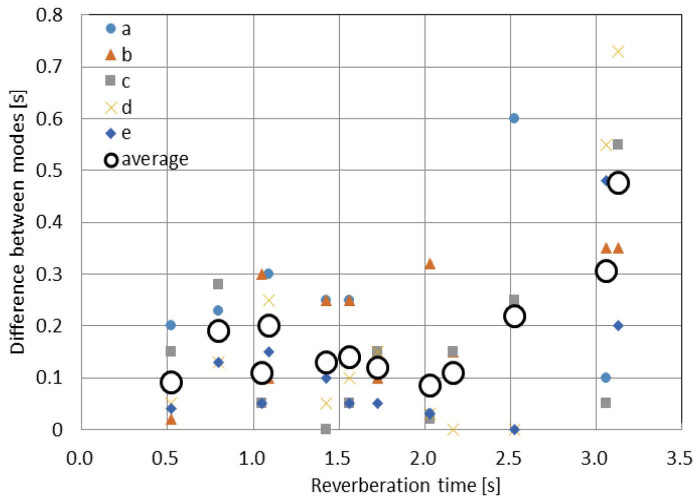

Figure 7 – Difference of SRT between single-mode and multi-mode (Ex.I and II)

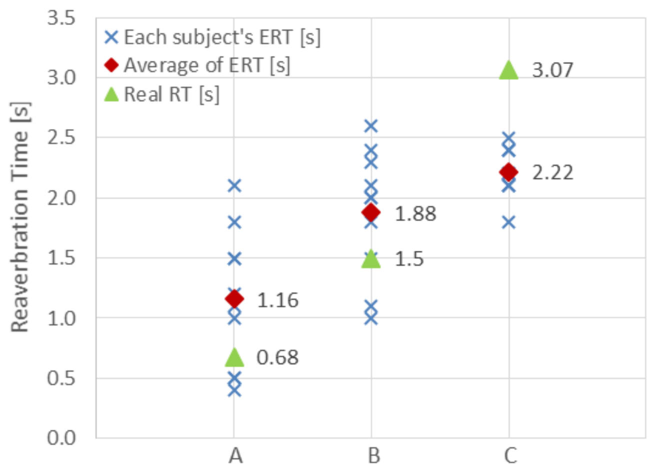

Figure 8 – Expected reverberation time of each space (Ex.III-1)

Subjective values of RT in "Duration of Reverberation"

Subjective values of RT (so to speak, Subjective RT, SRT) were derived from Ex.I (Figure 3) and Ex.II, by converting the subjective values of "Duration of Reverberation" in single-mode into RT values graphically. Figure 7 shows the difference of SRT between single-mode and multi-mode. When auditory stimulus is the shortest (RT=0.52 s), difference of SRT between modes is around 0.1 s and 0.2 s at maximum. On the other hand, when auditory stimulus is the longest (RT=3.13 s), difference between modes is around 0.5 s. As a whole, there is a tendency that the longer RT of auditory stimulus is, the larger the difference between modes is. And also the distribution of difference becomes wider.

Sense of Incongruity between Auditory and Visual Stimuli

It was expected that subjects feel less incongruity for the multi-modal stimuli in which RT values of auditory stimuli are nearly equal to those of real space in visual stimuli, but it was not always valid.

3.4 Result and Discussion of Ex.III-1

Experience and Expected Reverberation Time

Questionnaire results asking subjects of their experience of using the real space in visual stimulus or not, the ratio of answer "(Probably) has used it" is 80% for space "A", 100% for "B", 90% for "C" and it seems that most of subjects have experienced real spaces in visual stimuli.

Figure 8 shows the results of measuring expected reverberation time (ERT) for each visual stimulus of each subject when only visual stimuli (3 images of spaces) were presented. The error between average ERT and the original RT was -0.48 s for "A" (classroom), -0.38 s for "B" (hall) and +0.85 s for "C" (gymnasium). Average ERT for "A" and "B" exceeded original RT, and the error was small, but the values of ERT varied widely. On the other hand, average ERT of "C" was greatly smaller than original RT, and the error was quite large, but variance of ERT values was small. It seems that it is hard for subjects to predict RT value of gymnasium in which they have few opportunities to concentrate to listen to sound in comparison with classroom or hall. In addition, it seems that average ERT of "C" was judged smaller than original RT because inner volume of the "C" is about half of "B".

Results of examination for difference between population means of ERT for visual stimuli shows that significant difference is detected at 1 % level between "A" and "B", and "A" and "C". Between "B" and "C", of which the ERTs were similar, there was not significant difference.

Subjective Impression for Visual Stimulation

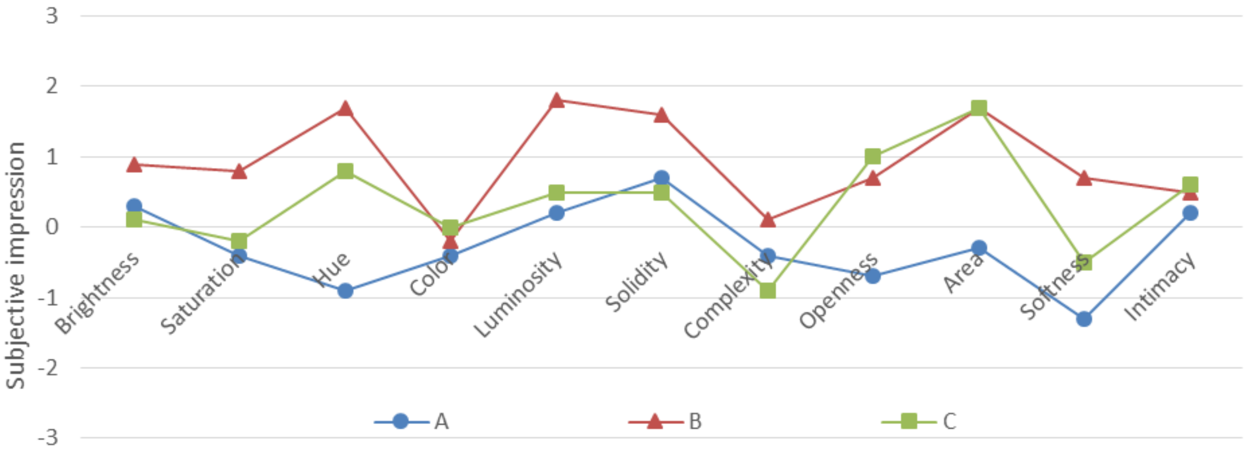

Figure 9 − Average values of visual impression (Ex.III-1)

Figure 9 shows the profiles of average values of subjective impression for visual stimuli. As a whole, the average value is low and the standard deviation is small for "A". For "B", the average value for color perception such as "Color", "Hue" and "Brightness" is particularly high. The subjective values of "Brightness" were expected to be approximately equal, because physical illuminance on the screen of visual stimuli was controlled equally in experimental conditions, but the subjective value for "Luminosity" of "B" was a little high. It seems to be due to the influence of other color perception. In addition, standard deviation of "B" was larger than "A" and "C" as a whole. About "C", the standard deviation of color perception was relatively small, but that of space perception such as "Solidity", "Complexity" and "Openness" was large. In addition, "Area" of "B" was evaluated to be almost the same as "B", but the real volume of "C" is about half of "B". It seems that to the volume of "B" is supposed to be smaller than real because "B" has complicated interior while "C" is a simple rectangular room.

As a result of difference examination of the population means between visual stimuli in each questionnaire item, significant difference was detected in only "Hue" in color perception. In addition, significance was also detected in "Color" between "A" and "B" at 5 % level, in "Brightness" between "A" and "C", "B" and "C" at 1 % level. In space perception, significant difference was detected in "Solidity" between "B" and "C", in "Openness" and "Area" between "A" and "B", "A" and "C", and in "Softness" between "A" and "B", "B" and "C" at 1 % or 5 % level.

3.5 Result and Discussion of Ex.III-2

Subjective Impression for Auditory Stimulation

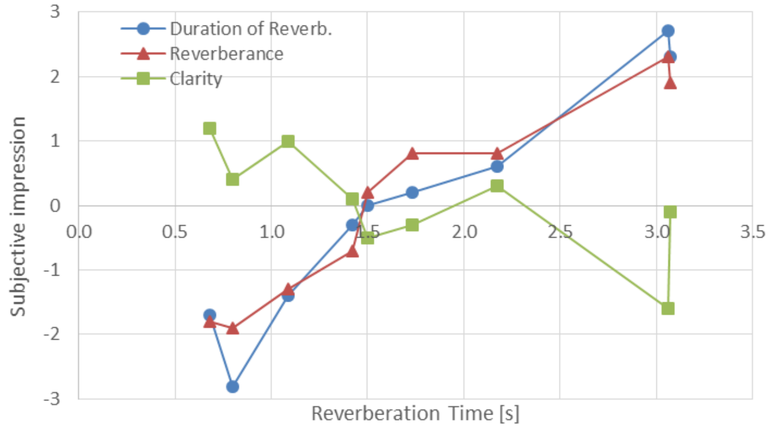

Figure 10 − Change of average value of auditory impression (Ex.III-2)

Figure 10 shows the change of average values of subjective impression for auditory stimuli in relation with RT. The changes of "Duration of Reverberation" and "Reverberance" are similar and almost contrary to "Clarity". In these 2 items, subjective values are almost in proportion to RT values of auditory stimuli. On the other hand, "Clarity" is in inverse proportion to RT in perspective, but does not necessarily depend on RT. Between auditory stimuli "9", "b", "10" and "11" (RT=1.50 ~ 2.17 s), the change of subjective value for each questionnaire item is small. However, auditory stimuli with short RT ("a", "7" and "8") and with long RT ("11" and "12") have large variation in subjective impression.

As a result of difference examination of the population means between auditory stimuli in each questionnaire item, the significant difference was detected between most auditory stimuli in "Duration of Reverberation" and "Reverberance" at 1% or 5% level. It is natural that the significance was not detected between auditory stimuli with close RT. On the other hand, in "Clarity", the significant difference was detected between only 7 combinations of stimuli with fairly large differences in RT. It is more difficult for subjects to distinguish the "Clarity" difference between stimuli than "Duration of Reverberation" and "Reverberance" by presentation of only auditory stimulation.

3.6 Result and Discussion of Ex.III-3

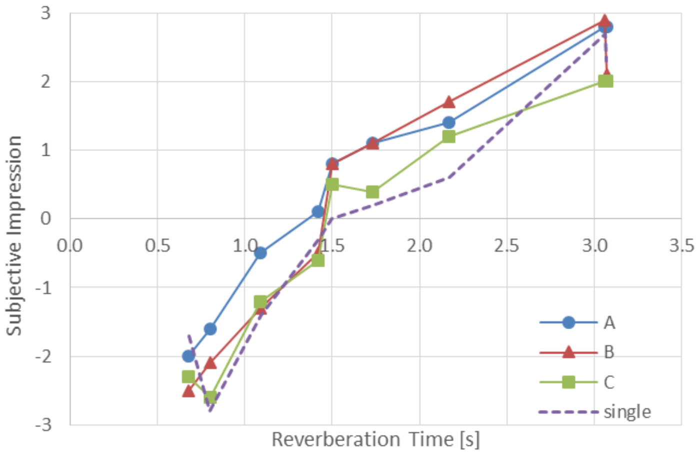

Figure 11 − Average values of "Duration of Reverberation" in multi-mode (Ex.III-3)

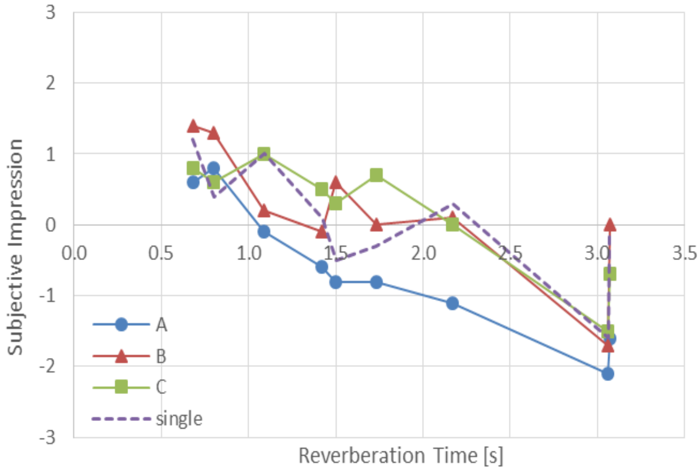

Figure 12 − Average values of "Clarity" in multi-mode (Ex.III-3)

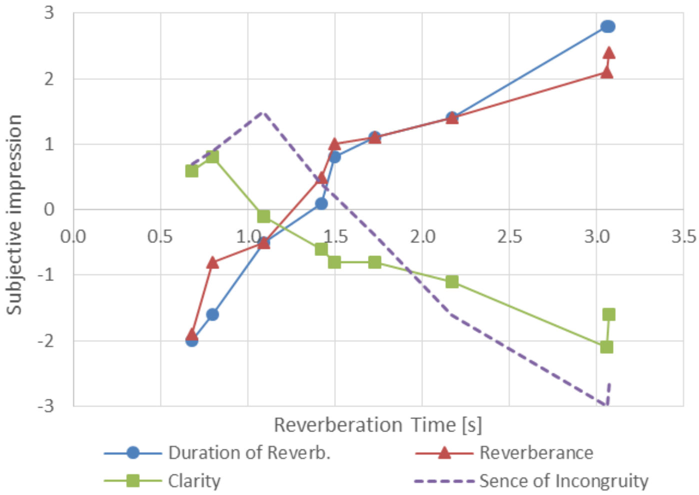

Figure 13 – Average values of “A” in multi-mode (Ex.III-3)

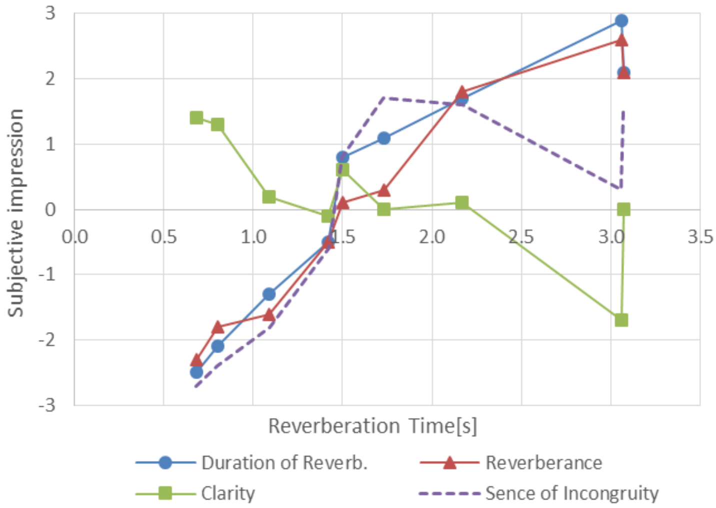

Figure 14 – Average values of “B” in multi-mode (Ex.III-3)

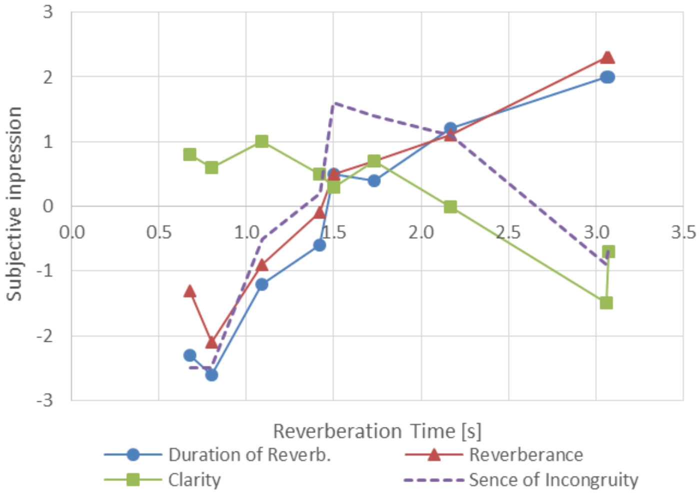

Figure 15 – Average values of “C” in multi-mode (Ex.III-3)

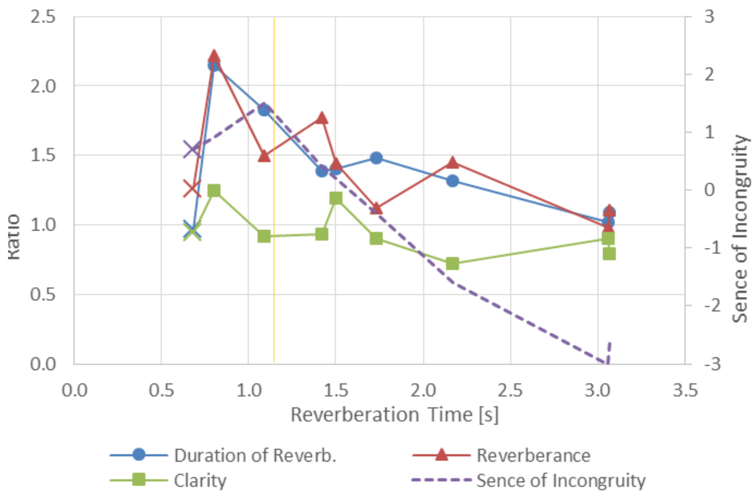

Figure 16 – Relationship of Ratio and Sense of Incongruity for “A” (Ex.III-3)

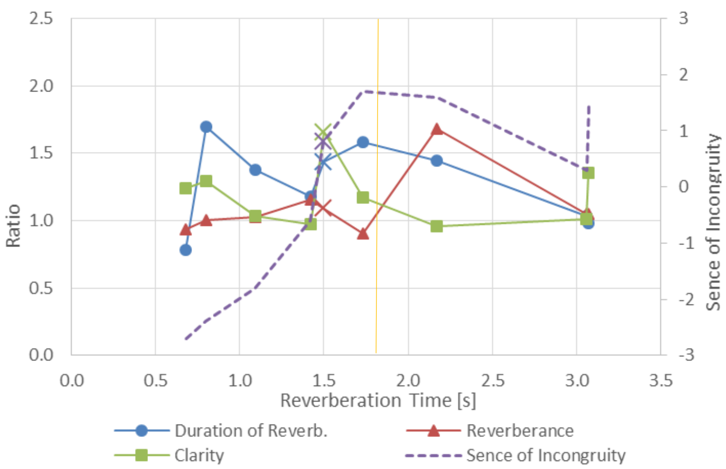

Figure 17 – Relationship of Ratio and Sense of Incongruity for “B” (Ex.III-3)

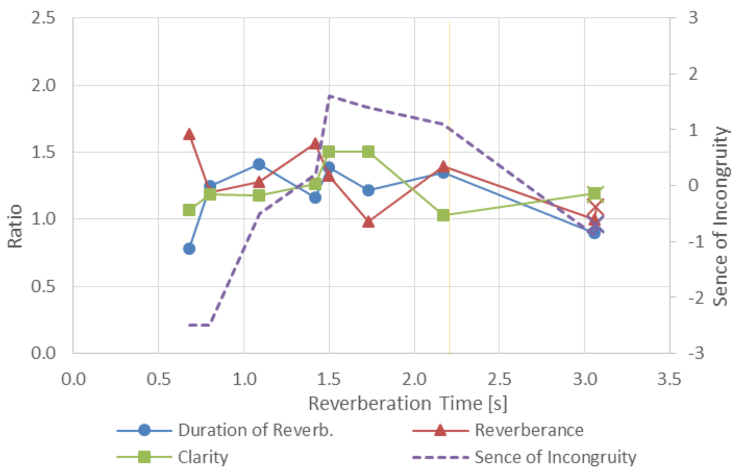

Figure 18 – Relationship of Ratio and Sense of Incongruity for “C” (Ex.III-3)

Subjective Impression for Audio-Visual Stimulation (Multi-mode)

Figure 11 and 12 shows the profiles of average values of auditory impression when auditory and visual stimuli were simultaneously presented to subjects. In these figures, the profile of auditory single-mode values (dashed line) from Ex.II is added. A similar tendency was seen in “Duration of Reverberation” and “Reverberance” regardless of visual stimuli, and multi-mode values are slightly higher than single-mode. Multi-mode presentation makes subjects feel reverberation more long and rich than single-mode. In Figure 12, a tendency of “Clarity” was approximately opposite to “Duration of Reverberation” and “Reverberance”, but the range of change by RT is smaller. For multi-mode presentation with visual stimulus of “A” (classroom), the values are lower than single-mode, while for “B” (hall) and “C (gymnasium), values in multi-mode are almost equal to single-mode.

Sense of Incongruity

Figure 13 ~ 15 shows the change of auditory impressions by RT for each visual stimulus and the profile of average values of “Sense of Incongruity” in multi-modal presentation. In Figure 13, as “A” has long ERT (=1.16 s) than original real RT (=0.68 s), it seems that sense of incongruity became high in RT >= 1.5 s. In Figure 14 for “B”, sense of incongruity became low in RT >= 1.5 s. If RT is short, sense of incongruity become high, but longer RT than original tends to be permitted. In Figure 15 for “C”, as ERT was shorter than original RT, sense of incongruity for combinations with short RT of auditory stimulus (such as “b” or “10”) is lower than the combination of “C” and “c” (real situation).

Ratio of Subjective Evaluation in Multi-mode to Single-mode

We expressed the ratio of average subjective values in multi-mode presentation (Ex.III-3) to single-mode (Ex.III-2) in order to consider the influence on auditory impression by visual information in detail.

About “Duration of Reverberation”, because the ratios are more than 1.0 between RT=0.80 ~ 2.17 s in most cases, from this it is understood that subjects feel reverberation longer by multi-modalization. On the other hand, there were little changes of impression by multi-modalization when RT of auditory stimulus was extremely long (“12” or “c”) or short (“a”). In addition, the range of change for “A” was the largest, and for “C” the smallest. It is assumed that subjects feel reverberation longer particularly for the space with short RT. Furthermore, as a result of difference examination of population means between multi-mode and single-mode (only auditory presentation), significant difference was detected between auditory stimulus “8” and audio-visual stimulus “A-8” at 5 % level.

About “Reverberance”, because almost all ratios are more than 1.0, subjects feel reverberation richer in multi-mode than single-mode in most cases. The tendency of influence by multi-modalization is similar to that of “Duration of Reverberation”. In addition, as a result of difference examination of population means between modes, significant difference was detected between “7” and “A-7”, and between “A-7” and “B-7” at 5 % level.

In “B” and “C”, “Clarity” became higher by multi-modalization because almost all ratios were more than 1.0. On the other hand, in “A”, “Clarity” was spoiled by multi-modalization, because almost all ratios were slightly less than 1.0. In addition, the range of change in values was small in all space compared with “Duration of the Reverberation” and “Reverberance”. Furthermore, as a result of difference examination of population means between modes, significant difference was detected between “11” and “A-11” at 5 % level.

Relationship of Ratio and Sense of incongruity

Furthermore, Figure 16 ~ 18 show the ratio of subjective auditory impressions in multi-mode to single-mode for each visual stimulus (solid line) and the change of average values of “Sense of Incongruity” (dashed line). In these figures, ERT (vertical line) and original RT (marker “x”) of each space are also shown.

In “A”, “Sense of Incongruity” is low in RT range of 0 to ERT + 0.34 s (RT + 0.8 s), and it becomes large from ERT + 0.57 s (RT + 1.1 s). It is natural that sense of incongruity grows large if RT is longer than ERT. On the other hand, in RT <=1.5 s any combination of visual and auditory stimuli tend to be permitted. In addition, change of ratio about “Duration of Reverberation” and “Reverberance” is similar, but that about “Clarity” often decreases by multi-modalization.

In “B”, “Sense of Incongruity” is high in RT range of 0 to ERT -0.46 s (RT -0.08 s), and that decreases from ERT -0.38 s (RT ± 0 s). In contrast with “A”, if RT is shorter than “b” (original IR for “B”), “Sense of Incongruity” is large, and the longer reverberation than that tends to be permitted. In addition, “Duration of Reverberation” and “Reverberance” decreased by adding extremely short IR (“a” etc.). There is a tendency that there is no change or increase by adding IR except that. On the other hand, “Clarity” tends to increase generally.

In "C", "Sense of Incongruity" is high in RT range of 0 to ERT -1.13 s (RT -1.98 s), and that disappears from ERT -0.8 s (RT -1.65 s) to ERT -0.05 s (RT -0.9 s), but is high again from ERT + 0.8 s (RT -0.01 s). If RT is longer than ERT, subjects seem to feel "Sense of Incongruity" even if it is original IR ("c") or RT is too short (shorter than 1.09 s).

4. SUMMARY

Experiments were carried out in which subjective auditory impressions were measured for single and multi-mode presentation by auditory and visual stimulation. When visual stimulation was added to auditory stimulation, the influences on subjective auditory impression and the sensual degree of incongruity were measured and analyzed especially in relation with RT.

About Ex.I and II

- Average subjective values of "Duration of Reverberation" and "Reverberance" in multi-mode became higher than single-mode under conditions of auditory stimulation within RT=1.5 s. The tendency of fluctuation by visual stimulation was seen in "Clarity" under conditions of auditory stimulation with RT>=2.0 s.
- The large change of average subjective values was observed at the border of RT=1.5 ~ 2.0 s.

About Ex.III

- In the space where one has little opportunity to concentrate on hearing sounds, the gaps between ERT and original real RT become large, and the subjective auditory impressions vary widely.
- The determination for "Duration of Reverberation" and "Reverberance" were accomplished in conformity with acoustic physical quantity, but that of "Clarity" was not necessarily. In particular, in the case of IR edited on a PC, auditory impression seems for synthesized IR to be different from real room IR even if RT of these IR are similar.
- If the gap between ERT and original RT is large, "Sense of Incongruity" becomes large. In this experiment, the tendency that auditory impression was influenced by visual impression was seen according to conventional knowledge, because people primarily make decisions based on visual information.
- "Duration of Reverberation" was evaluated as longer, and "Reverberance" was evaluated as richer by multi-modalization in almost all visual stimuli. On the other hand, for "Clarity", evaluations vary by visual stimulation and it seems to be greatly influenced by ERT.

Overall

The association of subjective impression of auditory and visual varies differently for every stimulus, but some relationships prove to be clear. Particularly, it was found that the subjective auditory impression changes by visual information, and it may be related with sense of incongruity existed in a combination of multi-modal stimulation.

As future works, we will try to relate physical quantities of visual stimulation to auditory impression, and will determine the most influent parameter of visual quantity. Useful design knowledge from this study will be derived if we can clarify a relationship between physical quantity of visual stimulation and auditory stimulation.

REFERENCES

1. Wani Y. "Effect of information on subjective impression for sound field in architectural space". Proc. International Conference of Acoustics 2013; June 2013, Montreal, Canada 2013: 3pNSc19.
2. Wani Y. et al. "Effect of combination on subjective impression for visual and auditory stimulus in architectural space". Proc. Architectural Institute of Japan 2013; August 2013, Hokkaido, Japan 2013: 40041, pp. 89-90.
3. Wani Y. et al. "Effect of reverberation time on subjective impression for visual and auditory stimulus in architectural space" Proc. Architectural Institute of Japan (Tokai) 2014; February 2014, Nagoya, Japan 2014: 304, pp. 393-396.
4. Ishikawa A. et al. "Effect of visual stimulus and reverberation time on subjective impression of indoor sound fields" Proc. Architectural Institute of Japan 2014; September 2014, Kobe, Japan 2014: 43440.
5. Iwamiya S. "Multimodal Communication of Music and Image" Kyusyu University Publication; 2000.
6. Densil C. et al. "AUDITORY VERSUS VISUAL SPATIAL IMPRESSION: A STUDY OF TWO AUDITORIA" Proc. ICAD 04-Tenth Meeting of the International Conference on Auditory Display 2004; July 2004, Australia 2004.