A Study on the Acoustic Characteristics of Large Spaces

M.Kaitea and H.Tachibanab

a Taisei Technology Center, 344-1, Nase-cho, Totsuka-ku, Yokohama 245-0051, Japan
b Institute of Industrial Science, Univ. of Tokyo, 4-6-1, Komaba, Meguro-ku, Tokyo 153-0041, Japan

In order to obtain a guideline for acoustic design of such large spaces as sports’ domes and arenas, acoustic measurements were performed in several large spaces constructed in Japan. From the results, acoustic characteristics of large spaces were investigated.

INTRODUCTION

In recent years, a lot of such large spaces as sports’ domes and arenas for musical events have been designed and constructed in Japan. In these large spaces, poor speech audibility is often experienced, which is caused by such acoustic problems peculiar to large spaces as long reverberation, long pass echo and focusing of reflections. At present, however, any design guideline has not been established for these large spaces and the traditional method based on the concept of optimum reverberation time is still being used.

As a basic study on this problem, we performed acoustic measurements in three large spaces constructed in Japan, in which the “directional room impulse responses” were measured using a sharp directional microphone in addition to the measurements using omni-directional microphone. From the results, such room acoustic parameters as $C_{80}$ (clarity) and $t_s$ (center time) were obtained in each direction. In the spaces with movable ceiling system, the effects of the change of room volume and the reflection from the membrane roof were investigated.

MEASURING METHOD

Although in almost all cases PA system is used in this kind of large spaces, the measurements were performed by the usual room acoustic measurement techniques. That is, the sound source signals were radiated from a dodecahedral omni-directional loudspeaker system and room impulse responses were measured at several points through omni-directional and directional microphones. For the impulse response measurement, MLS and TSP methods were used. Figure 1 shows the directions in which the directional microphone was set.

The measurements were performed in the following three large spaces:

(a) “S-stadium” for American football game and big rock concerts; room volume $V=740,000$ m$^3$,

(b) “S-arena” for musical events and such sports events as basketball with movable divided ceiling (ceiling height: 20~30m); $V=190,000$~$320,000$ m$^3$,

(c) “K-dome” for various kinds of sport events for citizen with a movable roof made of single membrane (5,200m$^2$); $V=600,000$ m$^3$.

RESULTS AND DISCUSSION

From the impulse responses measured through the omni-directional microphone, the values of clarity $C_{80}$, center time $t_s$ and reverberation time $T_{60}$ were calculated. $T_{60}$ and $C_{80}$ were obtained by the integrated impulse response method. From the results measured through the directional microphone, $C_{80}$, $t_s$, and $t_0$ in each direction were calculated. In this paper, the results of these values in 500Hz octave band are presented.

(1) The changes of parameters by the difference of measurement position

Figure 2 shows the spatial distribution of the values of omni-directional and directional $C_{80}$ measured at
Concerning the room acoustic characteristics in large spaces, the results of the field measurements performed in three spaces are reported. In this study, the applicability of the “directional impulse response” was tried to obtain the acoustic information in each direction. The effect of the change of room volume (ceiling height) and that of the existence of the membrane roof were also investigated and it has been confirmed that the reverberation condition can be controlled by these movable systems.
Acoustic Design of Stations of Underground

Chr. A..Shirjetsky a, Stanislav A. Kostarev b

a Research Institute of Building Physics, 21, Lokomotivnij Proezd, 127238, Moscow, Russian Federation.
b Tunneling association of Russia, Sasdovaya-Spasskaya,21,107217,Moscow,Russian Federation.

The problem of acoustic design of the halls for the passengers on the metro stations are presented in the combinations of actions, which are included the sound absorption's lining of the station room and the creating of the sound reinforcement system. The goal of this actions is the achievement of the acceptable level of summary noise in the hall and the good articulation of transmitted speech signal. In paper there are system of methods of step-to-step design for the decision of tasks of noise control, creating in hall station the optimal reverberation time and the optimal amplification of speech's information.

The maintenance of conditions of acoustic comfort at stations of underground demands performance of a complex of measures, which purpose is the achievement of optimum conditions sound transmission of information speech, for what syllable intelligibility should be not less than 80 % (i.e. not less than 0,6 on an index of an articulation) [1], at a level admitted of a background noise till a curve of limiting spectra –75 accepted as normative levels at stations of underground of Russia [2]. The analysis of an acoustic regimes of stations of the underground shows, that the decision of the specified task sound problems is possible only at the following sequence of operations of acoustic designing (block diagram of acoustic design see in a fig. 1).

FIGURE 1. The block diagram of a sequence of operation designing of underground stations.
1. Representation architectural–planning and constructive dates of room.
2. Analysis of noise regime of the room for the coordination its with normative levels.
3. 1-nd correction of the project.
4. Acoustic design of room for the coordination with zones of optimum for reverberation time and other acoustic parameters.
5. 2-nd correction of the project.
6. Electroacoustic design of room adhered to acoustic and noise regimes.
7. Conclusion of the dates on all acoustic parameters of room and requirements to characteristics of the electroacoustic equipment and engineers equipment.

On the 1- st stage the analysis of the noise phons of the initial project of station of underground is made, basic which the account of maximum levels of noise in a room of station under the known formula for diffuse of a field is presenting the sum of a direct sound and reverberation component: of noise;

$$L_n = 10 \log \left( \frac{\sum_{i=1}^{N} \Omega W_i}{4 \pi r^2} + \frac{1}{B_0} \sum_{i=1}^{N} W_i \right),$$  \hspace{1cm} (1)

where $L_n$ - 1-oct, levels of sound power everyone of source of noise considered in a range of frequencies 63-8000 Hz in dB relation to a threshold level $W_0 = 10^{-12}$;

$\Omega$ - factor of an orientation of i-st source of a sound;

$N$ - total quantity of sources of a sound (at counter movement of trains i=2);

$B_0$ - room constant of a station under the initial project;

$r$ - distance of a researched point of a field up to i-st source of a sound.
Being set known levels of sound power of sources of a sound, their orientation and required distributions of allowable noise levels makes i-st correction of a premise room of station, so that average factor of a sound absorption of a premise room corresponded of a required constant premise room

$$\alpha = \frac{B_1}{B_1 + S} ,$$  \hspace{1cm} (2)

where S - common area of a premise room of station.

The following stage of acoustic designing represents control account of reverberation time under the known formula of Eyring, with the purpose of check of a volume optimum of station of underground Topt \((V)\), with the account transmission of the speech information (see fig. 2).

![FIGURE 2. Recomended reverberation time on frequencies 500-2000 Hz for halls of stations jf underground](image)

In case of insufficiency common fond of a sound absorption for achievement Topt in a wide range of frequencies, there is the 2 -nd correction of design of project with the additional increase of fond of a sound absorption of station. The result of the 2- nd correction is the control account of spectral distribution of levels of a background noise \(L_b(f)\) on corrected value of a room constant.

On the following stage is the electroacoustic design of sound reinforcement of system of station, which main conditions are: 1) maintenance on all sound field the good relation of a signal /noise appropriate of the basic dynamic range (up to 90 %) speech signal; 2) maintenance in all points of acoustical perception of speech of the relation of a direct sound to diffuse field not less than 3-4.

At the account of the given nonuniformity of a field \(\pm \Delta L\), the minimally necessary level sound field can be defined under the formula:

$$L_{\text{min}} (f) \geq L_b (f) + D + \Delta L ,$$ \hspace{1cm} (3)

The second condition requires use in long, disproportional rooms of station \(s\) of underground the zone distributed system of loudspeakers, so that the maximal removal of the passenger from a sound source did not exceed distance

$$R_{\text{max}} \leq 0.14 \sqrt{B(f)D_j(\theta)} ,$$

where \(D_j(\theta) = \Omega \Phi^2(\theta)\) parameter of an orientation of a source \(\Omega\) - is a factor of concentration of source, \(\Phi(\theta)\) - its polar diagram of an orientation on angle \(\theta\) the received point of station. The account of a sound field on parameters \(\Omega\) and \(\Phi(\theta)\) is made separately for three ranges of frequencies: low (125-250 Hz) , average (500-1000 Hz) and high (2000-4000 Hz). The calculation of transfer of the speech information determining on in average and, especially, high frequencies for which is obligatory the achievement of non-uniformity of a sound field in limits \(\Delta L < \pm 3dB\).

Besides the common circuit disorilution of loudspeakers should be chosen so that the time of a delay of signals between them did not exceed 25-30 ms in any point of a sound field. The submitted here technique of acoustic designing of halls of stations of underground in complete volume is stated in “a Manual…” [3].

REFERENCES

2. CH 2.2.4/2.1.8.562-96 “Noise on workplaces, in rooms of inhabited, public buildings and in the territory of inhabited building”.
Guidelines for Predicting Acoustic Characteristics in Subway Stations

Jian Kang\textsuperscript{a} and Raf J. Orlowski\textsuperscript{b}

\textsuperscript{a} School of Architecture, University of Sheffield, Western Bank, Sheffield, S10 2TN, UK
\textsuperscript{b} Arup Acoustics, St. Giles Hall, Pound Hill, Cambridge CB3 0AE, UK

In long spaces such as subway stations classical room acoustic theory is not applicable since the assumption of a diffuse field does not hold with the extreme dimensional condition. In recent years considerable work has been carried out on acoustics of long spaces. This paper briefly reviews the previous theoretical work and then presents acoustic design guidelines based on the design and research work for the new subway stations in Hong Kong. Theoretical models were applied for predicting reverberation and speech intelligibility of public address system. In parallel, acoustic scale models were developed to investigate the effectiveness of strategic architectural acoustic treatments. The prediction showed good agreement with measurement.

**INTRODUCTION**

In subway stations most platform and concourse spaces are very long in one dimension and relatively short in the other two, and that consequently the classical formulae are invalid for calculating reverberation time (RT) in these types of spaces. In recent years considerable theoretical and practical work has been carried out on acoustics of long spaces \cite{1}. This paper starts with a brief review of relevant theoretical work. It then outlines an acoustic design manual for the Hong Kong Mass Transit Railway Corporation (MTRC) new subway stations \cite{2}.

**THEORY**

A series of theoretical and computer models has been developed for acoustics of long spaces \cite{1}. This includes a theoretical/computer model for geometrically reflecting boundaries using the image source method, a computer model for diffusely reflecting boundaries using the radiosity method, a theoretical/computer model for calculating acoustic indices from multiple sources, a method to predict the temporal and spatial distribution of train noise in subway stations, and a practical method for predicting acoustic indices in long enclosures, especially the speech intelligibility of multiple loudspeaker public address (PA) systems in subway stations. Using the theories special characteristics of long enclosures have been studied. The variation in reverberation along the length has been investigated and it is shown that RT generally increases with increasing source-receiver distance. In the case of multiple sources, it has been demonstrated that the reverberation is dependent on the number and position of sources. Comparison has been made between sound fields resulting from geometrical and diffuse boundaries, and considerable difference has been shown. Overall, it is suggested that the long enclosure theories should be applied if the length is greater than six times of the width and the height.

The design guidelines described in this paper are mainly based on the theoretical model using the image source method \cite{3}, especially a simplified formula for calculating RT in long enclosures, known as Kang/Orlowski equation. The formula applies typically to subway stations with a RT in the approximate range of 1-2s \cite{1}. The calculation is made as a function of source-receiver distance:

\[
RT_d = \frac{150}{850M - 10\log\left[\frac{d}{d + 850}\left(1 - \alpha\right) W H + \frac{1}{425}\right]}
\]

where $d$ (m) is the source-receiver distance, $W$ (m) and $H$ (m) are the width and height of enclosure, $\alpha$ is the average absorption coefficient of boundaries, and $M$ (dB/m) is the air absorption coefficient.

The ideal conditions for the application of this formula are as follows: rectangular enclosure, length approximately greater than six times the width and height, geometrically reflecting boundaries, absorbent end walls, uniform and angle-independent boundary absorption coefficient, a single point source at the center of cross-section, receiver at the center of cross-section, source-receiver distance is greater than width and height, and decay curves are approximately linear.

The measured RT results from a number of subway stations as well as from scale models of MTRC stations were compared to predictions using the formula. Good agreement was found.

**REFERENCES**

\textsuperscript{1} \textsuperscript{2}
DESIGN GUIDELINES

The objective of the acoustic design manual is to enable acoustic engineers and designers to assess the amount and disposition of acoustic absorption likely to be required to achieve specified reverberation times on platforms and concourses [2]. In addition, guidelines are given for the distribution of loudspeakers to maximize speech intelligibility. The guidelines are based on the calculation using the above formula, coupled with detailed scale model measurements of generic station types of the MTRC.

In the design manual all spaces in MTRC stations are categorized into one of three fundamental types, namely Type ‘A’ spaces – long; Type ‘B’ spaces - well-proportioned; and Type ‘C’ spaces - complex. For Type ‘B’ spaces the classic theories can be used, while Type ‘C’ spaces should be analyzed with the help of scale models. The guidelines presented in this paper relate only to Type ‘A’ spaces, which may be further divided into one of six sub-categories, three relate to typical station concourse areas, and three to typical platform areas. The dimensional criteria for platform areas are detailed in Table 1.

Table 1. Sub-categorization of Type ‘A’ platforms.

<table>
<thead>
<tr>
<th>Space type</th>
<th>Dimensions (m)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>&gt;6W 7-9 5-6</td>
<td>Kwai Fong</td>
</tr>
<tr>
<td>P2</td>
<td>&gt;6W 16-20 3.5-4.5</td>
<td>Tseung Kwan O</td>
</tr>
<tr>
<td>P3</td>
<td>&gt;6W 2.5-3.5 2.5-3.5</td>
<td>North Point</td>
</tr>
</tbody>
</table>

As an example, design guidance for Type P2 is outlined below. The target mid-frequency RT is 1.2s, and the target rapid speech transmission index (RASTI) is 0.45, with 0.5 over 90% of the floor area. A target RT spectrum is also specified.

The mid-frequency RT in a Type P2 platform may be estimated as a function of the quantity and distribution of absorptive treatment in the space using the prediction chart given in Figure 1. The absorptive treatment is assumed to be mineral wool or its acoustic equivalent of approximate thickness of 75mm. The absorption coefficients of other boundaries are assumed to be in the range 0.05 to 0.1 across the frequency range 125-4kHz.

The optimum distribution of absorptive treatment comprises three longitudinal bands distributed evenly across the central soffit downstand area, as illustrated in Figure 2. If deviations from the optimized treatment configuration are unavoidable, guidelines are given to minimize localized adverse effects on speech intelligibility. The deployment of absorptive treatment should be based on a rectangular grid formed from single loudspeaker grid cells, defined as the plan area bounded by the smallest repeatable loudspeaker arrangement within and constituting the array. The absorptive treatment may deviate from the optimum distribution provided alternative treatment locations can be identified within the cell to render the total treatment area unchanged. Limits for discontinuities are also recommended.

CONCLUSIONS

Based on substantial theoretical and scale modeling work, precise design guidelines have been developed for subway stations. It has been proved that the guidelines are useful for practical design.

REFERENCES

Simulation of the Reverberant Space in the Multichannel Audio Using the Convolution Method

A. Czyzewski, A. Kornacki, G. Szwoch, B. Kostek

Technical University of Gdansk, Sound & Vision Engineering Department, 80-952 Gdansk, Poland

The convolution method is commonly used to simulate the reverberant space by convolving monophonic or stereophonic sounds with the impulse responses of the room. In this paper, application of this method to the multichannel audio is proposed. The impulse responses of the real room were recorded. Each of the audio channels was obtained using the convolution of the adequate room impulse response with monophonic source sound. The results of the convolution were then combined and encoded as the multichannel surround audio in the format 5.1. The time and spectral analyses of the resulting sounds, as well as the listening tests were performed. The results of these experiments are presented and discussed in the paper. The presented method allows one to simulate the acoustical conditions of the room where the monophonic audio was acquired. Possible applications of this method include advanced Internet teleconferencing, in which the bandwidth requirements may be decreased by transmitting only monophonic sounds and the impulse responses of the room instead of the whole multichannel audio.

INTRODUCTION

The convolution of audio signal with the impulse response of the room is commonly used in order to simulate the acoustics of the room in the recording or during playback in another room. The impulse response of a room may be either recorded using the high amplitude impulse signal or simulated by the computer modeling system. Spatialization of the sound using the multichannel techniques is now getting widespread [1,2,3]. The aim of the experiments described in this paper is to simulate the acoustic properties of the recording hall using the convolution of monophonic audio signal with the multichannel impulse response of the hall. This procedure may be useful in multichannel recording and internet teleconferencing systems [4].

EXPERIMENTS

The aim of the preliminary experiments was to simulate the acoustical properties of the auditory hall situated at the Technical University of Gdansk, Poland. The diagram of the hall is shown in Fig. 1. The measured reverberation time of the hall is about one second. The impulse responses of the hall were recorded using the gunshot as a sound source and five microphones. The omnidirectional condenser Neumann KM-83 microphones were used to record the direct signal as well as signals reflected from the room walls.

FIGURE 1. The auditory room: x - sound source, o - microphones

The sound source and microphones were situated as shown in Fig. 1. Four of the five microphones were placed in the corners of the room, about two meters from the walls in order to reduce the influence of the sound waves reflection from the walls. The signals received by the microphones were recorded using the mixing console and a multitrack recorder. The recorded sound signals were digitized and stored on a computer hard disk in the *.wav format (48 kHz, 16 bit). Several recordings were made and their results were averaged in order to reduce the influence of acoustical disturbances.
In the next stage, the recorded impulse responses of the room were convolved with the monophonic audio test signals. In order to examine the influence of the convolution on the quality of sounds of different kind, three test signals were used: a dialogue, a short piece of piano music and a set of percussive sounds. Two approaches to the convolution were made. The first one was based on the mathematical convolution formula. The Fast Fourier Transforms of both the test signal and impulse response were calculated, using the Mathematica computer system. The resulting spectra were multiplied and the final result of the convolution was obtained by computing the Inverse FFT of the result of this multiplication [1]. The second method used the SEKD Samplitude 2496 software with Room Simulator module. The second approach was much faster, it does not, however, give any control over the convolution algorithm, allowing only changing the values of predefined parameters.

The results of the convolution for the same test signal were combined to form the multichannel 5+1 audio signal, using the SEKD Samplitude 2496 audio editor. In the center channel the source test signal mixed with the result of convolution obtained for the signal recorded by center microphone was placed. The remaining four channels (front-left, front-right, rear-left, rear-right) contained the results of convolving the test signal with the impulse response recorded by a given microphone. Next, the combined signal was encoded to the AC3 format using the Astarte a.PACK encoder, in order to use the resulting sounds in the listening tests.

RESULTS

The time and spectral analysis of the sound signals obtained by the convolution were performed. Fig. 2 shows the exemplary time domain plots of the piano music piece convolved with the impulse responses of the room. Analysis of the time- and frequency-domain plots proved that results of the convolution with each recorded impulse response differ from each other.

The preliminary listening tests were performed in the listening room (4.5 × 5 m) using the multichannel 5+1 audio system. During the tests, no significant amount of noise or other distortions introduced into signal during convolution was noticed. Persons taking part in tests described the processed sound as more spatial and similar to the sound perceived in the auditory room where impulse responses were recorded.

CONCLUSIONS

A method of spatializing monophonic audio signal was presented. The results of the preliminary experiments showed that convolving the signal with multichannel impulse response allows one to simulate the acoustical properties of the recording room. At the next stage of the research, impulse responses of different rooms will be recorded and used for convolution. The systematic subjective listening tests will be performed by the group of experts who will compare the quality of audio signals recorded in different rooms and obtained using the convolution.

REFERENCES

2. J. Klepko, 5-Channel Microphone Array with Binaural Head for Multichannel Reproduction, 103rd AES Convention, New York 1997, preprint No. 4541

ACKNOWLEDGEMENTS

The research is sponsored by the Committee for Scientific Research, Warsaw, Poland. Grant No. 8 T11D 00218.
On the disagreement between speech transmission index (STI) and speech intelligibility

H. Onaga\textsuperscript{a}, Y. Furue\textsuperscript{b} and T. Ikeda\textsuperscript{a}

\textsuperscript{a}Faculty of Science and Engineering, Kinki University, Kowakae 3-4-1, Higashi-osaka, 577-8502 Japan
\textsuperscript{b}Faculty of Engineering, Fukuyama University, Sanzo, Gakuen-cho 1, Fukuyama, 729-0292 Japan

STI contradicts the generally accepted concept of the importance of early energy for speech intelligibility because energy concentration due to strong reflections with any delay time increases the value of STI \cite{1}. The intelligibility test results indicate clear disagreement with STI. STI cannot be considered to correspond to intelligibility.

\textbf{THEORETICAL EXAMINATION ON STI}

Schroeder\cite{2} showed that the modulation transfer function (MTF) is the absolute value of the Fourier transform of the squared impulse response $h^2(t)$ divided by its total energy.

$$m ( F ) = \frac{\left| \int_{0}^{\infty} h^2 ( t ) e^{ - j 2 \pi F t } d t \right|}{\int_{0}^{\infty} h^2 ( t ) d t} \quad (1)$$

It is clear from the definition of MTF that an impulse response $h(t)$ and its inversion in the time axis $h(t_c - t)$ have the same MTF value. Which means MTF (also STI) equally evaluates early energy and late energy.

At first, we examine STI in single reflection fields. The MTF of single reflection fields is written as

$$m ( F ) = \sqrt{1 + \frac{2 \eta \cos (2 \pi F t)}{1 + \eta \cos (2 \pi F t)}} \quad (2)$$

In single reflection fields, MTF has the same value for the power ratio $h_1$ and $1/h_1$, and for delay times $t_1$ and $-t_1$. The inverted arrival order of direct and reflected sound causes no change in the MTF value. The calculated STI is shown in Fig. 1 as a function of the delay time of a single reflection. The STI takes the same value for positive and negative delay times and relative levels of reflection.

Next, we examine the series of constant level reflections of 1 ms intervals with duration 100 ms is divided into five equal portions in the time domain (A1, A2, A3, A4, and A5), and one of the five portions are amplified by L dB. The calculated STI curves are plotted in Fig. 2 as a function of relative amplitude and show that enhanced reflections cause an increase in the STI regardless of the delay time.

\textbf{INTELLIGIBILITY IN SINGLE REFLECTION FIELDS}

Intelligibility tests were carried out in single reflection fields. The speech source material was an anechoic recording of a male speaker (4.9 mora/second speech rate). The material used was the 4-mora Japanese word lists composed by Sakamoto et al. \cite{3}. Sound fields were synthesized using a digital delay machine. The test sounds were produced through loudspeakers located in front of the subject in an anechoic room at a A-weighted...
FIGURE 3. Word intelligibility for single reflection field. The numbers in the figure show delay time of the reflection. The level of the reflection is -6dB.

FIGURE 4. Word intelligibility for single reflection field. The numbers in the figure show relative level of the reflection. The delay time of the reflection is 80 ms.

SPL of 62dBA. The Hoti-spectrum-noise is generated simultaneously as a masking noise at 68dBA. The subjects were 30 elderly (aged 63-92 years) persons.

The mean intelligibility scores are plotted in Fig. 3 and 4. From these Figures, we can see that the condition of positive delay time always show higher intelligibility than negative delay time and that of negative relative levels than positive relative levels. Both of which means stronger primary sound causes higher intelligibility than stronger secondary sound in spite of the same STI value.

INTELLIGIBILITY IN MULTIPLE REFLECTION FIELDS

Sound fields were synthesized using a digital delay machine and a digital reverberator. Each sound field has the same direct sound and reverberation, however a series of 5 strong reflections was initiated at different delay times(Fig. 5). The subjects were 28 elderly (aged 56-79 years) and 9 young (aged 18-22 years) persons.

The mean intelligibility is shown in Fig. 6. Intelligibility is not a monotonic function of the STI. Though, a tendency can be observed for longer delay times to cause lower intelligibility.

CONCLUSION

The STI evaluates the energy concentration or dispersion in the time domain regardless of the delay time, which contradicts the generally accepted concept of the importance of early reflection for speech intelligibility. The intelligibility test results reconfirm the importance of early energy, and indicate clear disagreement between the STI and intelligibility. STI cannot be considered to correspond to intelligibility.

ACKNOWLEDGEMENT

The authors would like to thank the NTT Basic Research Laboratories and Prof. Y. Suzuki of Tohoku Univ. for providing the new word lists. Part of this study was supported by a research grant from the Kajima Foundation.

REFERENCES