Effet of transient signal length on cross-correlation functions in a room

Hirofumi Yanagawa
Department of OMDD, Pioneer Electronic Corporation, 26104-chome, Hanazono, Tokorozawa-shi, Saitama, 359 Japan

Yoshio Yamasaki and Takeshi Itow
School of Science and Technology, Waseda Univrtsity, 4-1, Ohkubo3-chome, Shinjuku-ku, Tokyo 160 Japan

(Received 29 May 1987; accepted for publication 18 July 1988)

An investigation wasmade of cross-correlation functions with transient signals between two points, which correspond to both earas of a listener in a room. The interaural cross correlation is closely related to the subjective impressions of sound fields. The study attempted to account in a comprehensive way for the combined effects that initial reflected and reverberation sounds from music or other transient signals have on such impressions. To this end, cross-correlation functions of the rise and fall of the sound field from transient signals were derived from the impulse responses at two points in the hall.These results were combined with image-sources distribution patterns derived by the closely located four-point-microphone method; then, a comparative explanation was made of the changes with transient time duration of the cross-correlation functions. Good agreement was found between chang with time in experimentally derived maximal cross-correlation function values and the changes with time in image-sources distribution of the sound field.

PACS numbers: 43.55.Br,43.55.Hy,43.66.Pn

INTRODUCTION

Spatial impression of the sound field is of two kinds: that produced by initial reflected sound and that produced by succeeding reflected sound (diffuse sound). Sound image spaciousness [1] is mainly conveyed by the former component, while the latter component contributes so-called subjective diffuseness.[2]

Various studies have examined the spatial impression of sound fields. These have shown that interaural cross-correlation coefficients,[3,4] or interaural cross-correlation functions,[5,6] are major elements affecting spatial impression.

As an evaluation metbod, cross-correlation analysis has been found applicable not only to original sound fields, such as are produced in concert halls, but also to reproduced sound fields.[7-9]

Numerous studies have treated these two components separately. However, for an integrated evaluation of subjective impressions of sound fields created by music or other transient signals, it is necessary to find out how interaural cross-correlation functions change over time. In a previous study, the present authors carried out measurements of changes with time in correlation coefficients between two points in a sound field during the reverberation process (transient sound fields).[10] In that study, we indicated that the diffuseness of sound fields during the reverberation process is closely linked to correlation coefficients. In the study, reported here, a more-detailed investigation was made of changes with time in cross-correlation functions between two points. This was done by indirectly measuring impulse responses in the sound field, instead of by direct measurement.

Further, the closely located four-point-microphone method was used to determine image-sources distribution pattens. In this method, impulse responses were collected at four points in very close proximity and were processed as digital sigmals. These results show that the initial reflected sounds of the sound field can be considered as sounds arriving from image sources. Transient sound fields were accordingly divided into the rise and fall of the sound field. Correspondence was determined between the image-sources distribution and the changes with time of the cross-correlation functions during this rise and fall and the changes were explained in physical terms.

1. CROSS-CORRELATION FUNCTIONS OF TRANSIENT SIGNALS

When a signal $S(t)$ radiates in a room, the convolution between the sound pressures $P_1(t)$ and $P_2(t)$ at two points in the room and the impulse responses $h_1(t)$ and $h_2(t)$ can be expressed by

$$P_1(t) = \int_{-\infty}^{\infty} S(\gamma) h_1(t-\gamma) d\gamma , \tag{1}$$

$$P_2(t) = \int_{-\infty}^{\infty} S(\lambda) h_2(t-\lambda) d\lambda . \tag{2}$$

Subjective diffuseness of the sound field is especially pronounced in the reverberation process after the sound has ceased being radiated at its source. It is thus convenient to conceive of the radiation and termination of signals as proceeding in steps.

The rise and fall of a sound field can be defined by

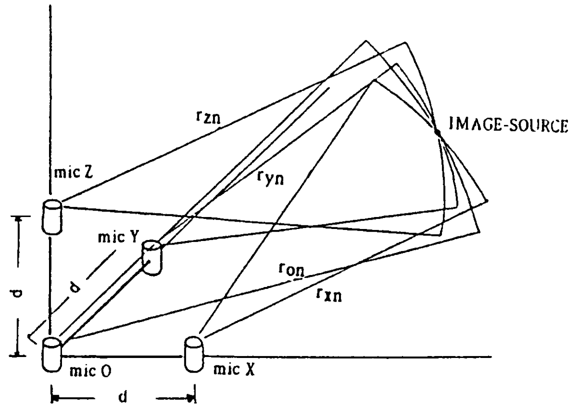

FIG.1. An image-source location. The location of oan image source is the intersection of sphere surfaces. The radii are r_{on}, r_{xn}, r_{yn}, and r_{zn}, respectively.

introducing the concept of an image-sources distribution. The rise of the sound field starts when the signal $S(t)$ is emitted simultaneously $(t=0)$ at the actual sound source and at all image sources and lasts until the signal reaches the receiving point. Likewise, the fall of the sound field begins when the sound source stops radiating. The direct sound dies out first, follwed by a gradual process of image sources dying out in their order of proximity to the point of sound reception. Incorporating these classifications clearly in

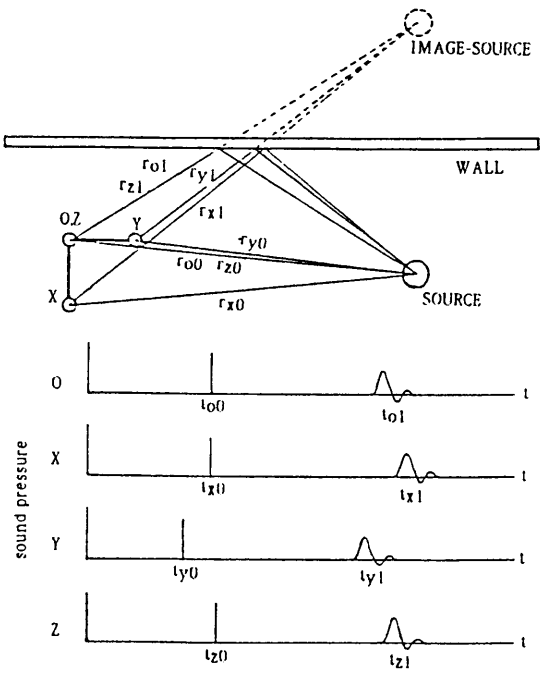

FIG.2. A model of impulse responses observed at four-point microphones.

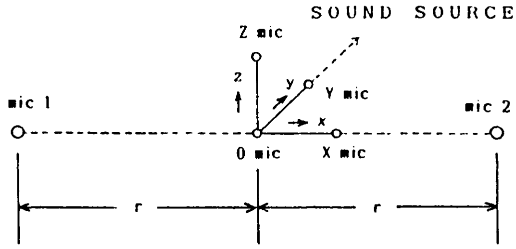

FIG.3. Microphone placement. Microphones X,Y,Z and O : closely located four-point microphone. Microphones 1 and 2 are used for cross-correlation function measurements, r_c =16cm.

Eqs.(1) and (2), the follwing equations are derived by using the time window function $W(t)$. These equations assume that in the rise of the sound field, signals are emitted at

$$P_{d1}(t)=\int_{-\infty}^{\infty}S(\gamma)W(\gamma)h_1(t-\gamma)d\gamma\,, \tag{3}$$

$$P_{d2}(t)=\int_{-\infty}^{\infty}S(\gamma)W(\gamma)h_2(t-\gamma)d\gamma\,. \tag{4}$$

$$\begin{aligned}&W(\gamma)=1,\quad 0<\gamma<t,\\&W(\gamma)=0,\quad \text{elsewhere}.\end{aligned} \tag{5}$$

Likewise, assuming that at the fall of the sound field signals stop radiating at $t=0$, the time window $W(t)$ in Eqs.(3) and (4) is as follows:

$$\begin{aligned}&W(\gamma)=1,\quad -\infty<\gamma<0,\\&W(\gamma)=0,\quad \text{elsewhere}.\end{aligned} \tag{6}$$

The cross-correlation function $\Phi_{d12}(t,\tau)$ for each of these cases can be expressed as

$$\Phi_{d12}(t,\tau)=\left\langle P_{d1}(t)P_{d2}(t+\tau)\right\rangle, \tag{7}$$

where the terms within the angular brackets are ensemble averages. If the signal $S(t)$ is white noise, then since the autocorrelation functions become delta functions, Eq. (7) can be rewritten as

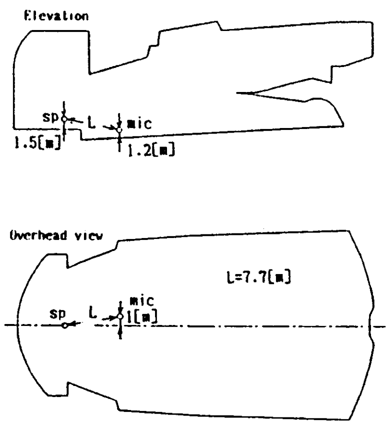

FIG.4. Location of sound source and sound receiving point (Ohkuma Hall).

$$\Phi_{d12}(t,\tau) = S^2 \int_{-\infty}^{\infty} h_1(t-\gamma) h_2(t+\tau-\gamma) W(\gamma) d\gamma, \tag{8}$$

$$\langle S(t)S(t+\tau)\rangle = S^2\delta(\tau).$$

In the rise of the sound field,

$$W(\gamma)=1, \quad 0<\gamma<t, \qquad W(\gamma)=0, \quad \text{elsewhere.} \tag{9}$$

Also, in the fall of the sound field,

$$W(\gamma)=1, \quad -\infty<\gamma<0, \qquad W(\gamma)=0, \quad \text{elsewhere.} \tag{10}$$

In other words, rather than seeking to determine the ensemble average of white noise in a repeated, discontinuous way, the changes with time of correlation functions can be derived from impulse responses.

As time elapses in a risings sound field the range of inte gration expands from the starting point of the impulse response toward the tail end of the waveform. Similarly, in a falling sound field there is a gradual diminishment from this starting point. Changes with time in correlation functions can accordingly be estimated in the following way.

In the rise of the sound field [see Eqs. (8) and (9)], at first a high correlation is shown, with maximal values close to 1. This is due to direct sound. These maximal values decrease as time elapses, along with the decreasing directionality and intensity of succeeding reflected sounds.

This phenomenon can be viewed in terms of the image-sources distribution derived by the closely located four-point-microphone method. Assume a sphere whose radius increases as time elapses in proportion to the speed of the sound, with the sound reception point as the center. When the actual sound source is included in the sphere (i.e., when direct sound arrives), the maximal correlation function value is 1. Thereafter, as the radius of the sphere increases, each time a new image source becomes encompassed within the sphere the maximal correlation value changes, until eventually, it reaches a steady-state value. When the transient

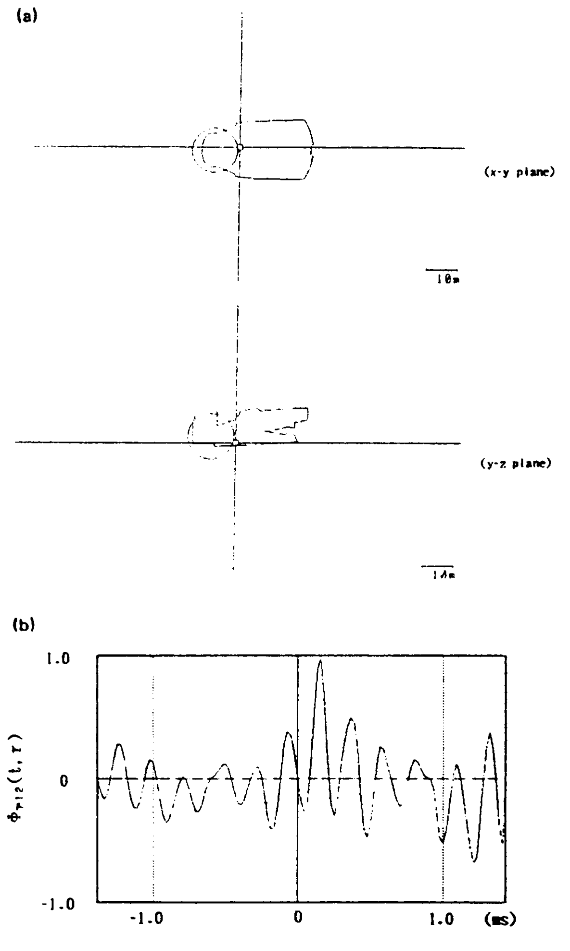

FIG.5. (a) Image-sources distribution , rise of sound field (t :0-3ms). **(b)** Cross-correlation function [see Eqs. (8) and (9)], rise of sound field (t :0-3ms).

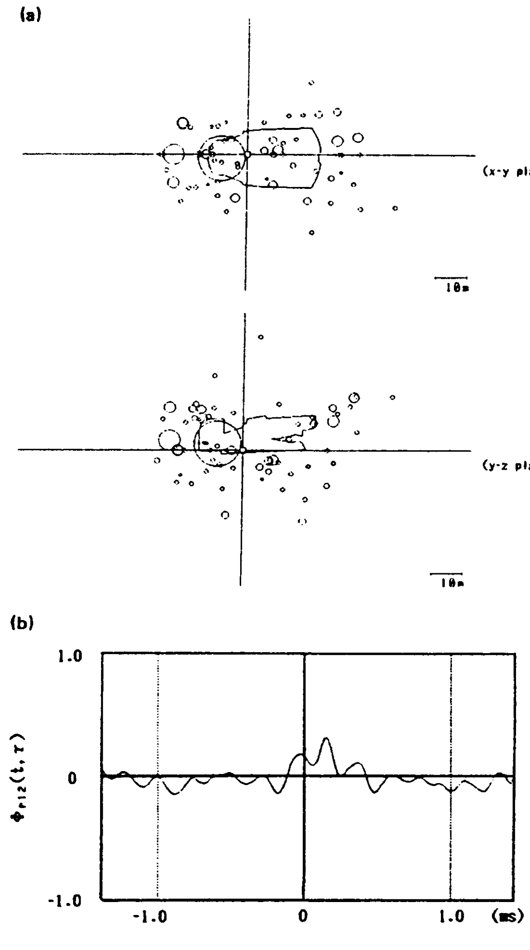

FIG.6 (a) Image-sources distribution, rise of sound field (t :0-100ms). **(b)** Cross-correlation function [see Eqs. (8) and (9)], rise of sound field (t :0-100ms).

ignal length becomes longer, the radius should also be larger.

In the fall of the sound field [see Eqs. (8) and (10)], the correlation functions are determined by the sound from image sources outside the above-mentioned sphere. Initially, when the sound terminates at its source, direct sound from the actual sound source stops contributing to the cross-correlation functions and the maximal correlation values drop below the steady-state value reached in the rising sound field. Thereafter, as time elapses, the phereincreases in radus; after a number of image sources radiating initial reflected correlation functions those of a practically diffuse sound field. When the ransients signal length becomes shorter, the radius should also be larger.

In the soecialcase when the transient signal is short amd constant, the sound field does no reach steady state. This condition can be expressed only by the time window $W(t)$ in Eq.(8):

$$W(\gamma)=1, \quad -\infty<\gamma<T, \qquad W(\gamma)=0, \quad \text{elsewhere.} \tag{11}$$

wherer T is the duration time of the signal.

Seen in terms of the image-sources distribution, the cross-correlation functions are determined by the sound from image sources inside a spherical shell, which has a sound reception point as its center and whose radius expaInds in proportion to the speed of sound c. The thickness of the spherical shell equals to cT.

To summarize the three cases discussed above, the

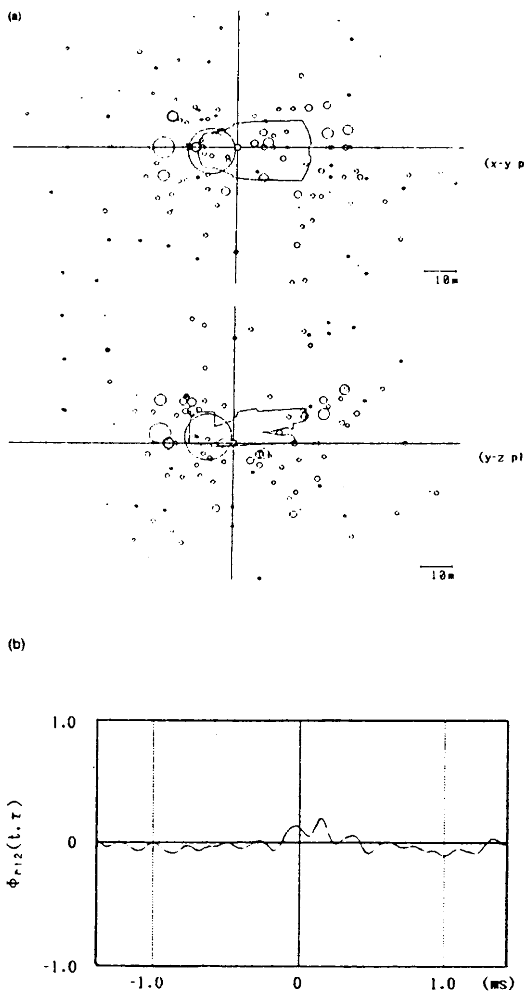

FIG.7 (a) Image-sources distribution, rise of sound field (t :0-500ms). **(b)** Cross-correlation function [see Egs. (8) and (9)], rise of sound field (t :0-500ms).

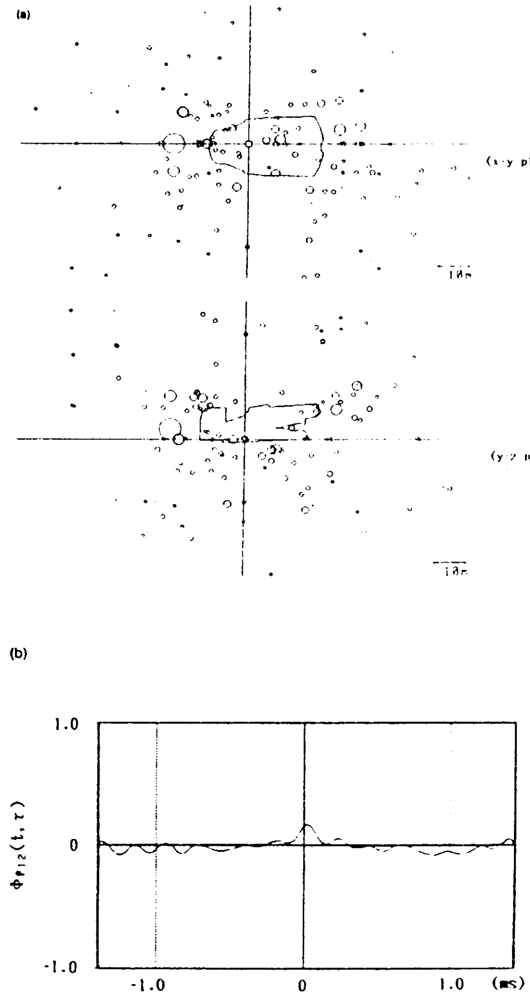

FIG.7 (a) Image-sources distribution, fall of sound field (t :3-500ms). **(b)** Cross-correlation function [see Egs. (8) and (9)], fall of sound field (t :3-500ms).

crossss-correlation functions of a transient signal in a room are cquivalent to those found when white noise is radiated stably from the sound source and/or the image sources within a shell or inside or outsidde the sphere, whose radius increases as time elapses.

II. CLOSELY LOCATED FOUR-POINT-MICROPHONE METHOD [11]

Cross-correlation functions are estimated from the image-sources distribution, as stated in Sec. I. The image-sources distribution in a concert hall can be derived from the impulse responses by using the closely located four-point-microphne method, which has been recently developed.

Suppose that a sampled-impulse response is obtained at a listening point. We assume here that each of observed datum in the response should be representative of a reflected sound that is coming to the listening point from an image source. Thus we can caleulate the distance between the image position and the listening point.

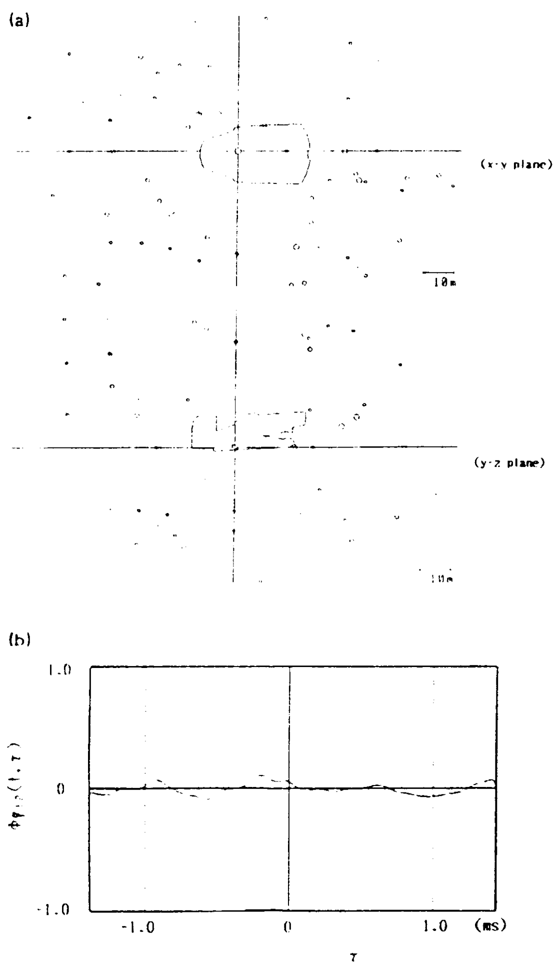

FIG.9. (a) Image-sources distribution, fall of sound field (t :100-500ms). **(b)** Cross-correlation function [see Eqs.(8) and (10)], fall of sound field (t :100-500ms).

We can identify theoretically the position of the image-sources in a three-dimensional space if we obtain the impulse responses using four microphones that are located close to each other at the listening point. We also assume that all the reflections follow geometric acoustics theory. The effects of diffraction or more complex wave behaviors are still being studied.

Figure 1 demonstrates this scheme, where observing points are located on the rectangular coordinate axes and at the origin.

Now r_{on}, r_{xn}, r_{yn}, and r_{zn} are the distances between an image source and each of the observation points X, Y, Z , and the origin O . From spherical equations, the coordinate of an image source for each reflection (X_n, Y_n, Z_n) is

$$\begin{aligned} X_n &= (d^2 + r_{on}^2 - r_{xn}^2)/2d, \\ Y_n &= (d^2 + r_{on}^2 - r_{yn}^2)/2d, \\ Z_n &= (d^2 + r_{on}^2 - r_{zn}^2)/2d. \end{aligned} \tag{12}$$

Figure 2 shows a simple model and its impulse responses at each observation point.

III. MEASUREMENT OF THE IMAGE-SOURCES DISTRIBUTION AND OF CROSS-CORRELATION FUNCTIONS IN A TRANSIENT SOUND FIELD

Measurements were made of cross-correlation functions and sound field measurements were made using the closely located four-point-microphone method. This was done in order to find the correlatin between changes with time in cross-correlation functions and the image-sources distribution. Sound field measurements were made in the Ohkuma Lecture Hall of Waseda University in Japan. This hall has a capacity of 5420 m^3 and a reverberation time of 1.3s.

The sound source consisted of a pair of 20-cm-diam full-range loudspeakers without cabinets, attached together face to face. The input signal fed in phase to each loudspeaker was a series of rectangular pulses at random intervals; however, no interval was shorter than the reverberation time of the hall. The amplitude of the pulses was 100 V and the pulse width was $5\mu s$.

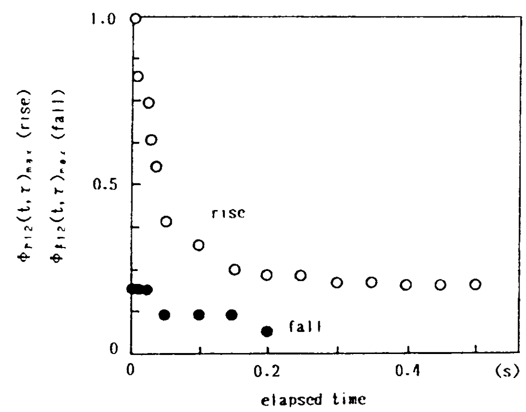

FIG.10. Changes with time in maximal cross-correlation function values during rise or fall of sound field.

Two microphones were used for the cross-correlation function measurements. These were placed side by side, 16 cm apart and facing the sound source from the measuring point of the closely located four-point-microphone method (see Fig. 3). The distance between the two microphones $(2r_e = 32\text{cm})$ was determined to approximate the cross-correlation function between both ears of a listener in a diffuse sound field.[5] The sound source was located at the center of the stage, at a height of 1.5m, while the sound reception points were at the front center of the hall, at a height of 1.2m above the floor. The distance of the sound reception points from the sound source was 7.7m(see Fig.4).

The impulse responses of the hall were fed into an analog-to-digital converter and digitally processed by a personal computer. Because of the random interval of the signal, the periodic noise was also reduced after averaging the impulse responses.

Changes with time in the cross-correlation functions and in image-sources distribution during the rise of the sound field in the hall are shown in Figs.5-7. The image-sources distributions shown in the figure are projections of image sources on the floor or side surfaces. The center of the circle showing the image source represents its location, while the area of the circle represents its power. The range of the distribution shown here is up to approximately 200ms from the time of direct sound arrival. The number of image sources increases with elapsed time; as can be seen clearly, this is accompanied by a decrease in the maximal values of cross-correlation functions.

The characteristics in the fall of the sound field are shown in Figs.8 and 9. Although less pronounced than in the rise of the sound field, a decrease in the maximal cross-correlation function values can be seen, as opposed to those values found during the steady state.

Changes with time in the maximal cross-correlation function values within $\tau = 1.0\text{ms}$ are shown in Fig. 10 for the rise and fall of the sound field, respectively. In Ohkuma Hall a practically steady state is reached in approximately 150ms.

IV. CONCLUSIONS

First, during the rise and fall of a sound field, changes with time in image-sources distribution were investigated by the closely located four-point-microphone method. Considering the sound radiating from actual and images sources as white noise, which is emitted in the sound field and ceases radiating in steps, the cross-correlation functions between two points in the sound field during these transient states can be determined from impulse responses.

These cross-correlation functions are not only functions of relative time differences, but also of elapsd times, with the beginning of an impulse response serving provisonally as the time origin. changes in cross-correlation functions in the rise and fall of the sound field can be explained in terms of changes in the distribution of image sources. In the rise of the sound field these are due to the inclusin of more and more image sources within a sphere whose center is the sound reception point and whose radius increases with time in proportion to the speed of sound. Similarly, changes during the fall of the sound field are due to image sources remaining on the outside of the sphere. ln cases where the transient signal length is relatively short and constant, the cross-correlation functions depend in the same way on image sources within a spherical shell centering around the sound reception point

The present authors now believe that the cross-correltion function of a transient signal should be closely related to spatial impressions in a concert hall. The psychological aspect of this point is now being studied.

ACKNOWLEDGMENTS

The authors thank Shiro Ise and Nobumasa Seiyama for their assistance in the many experiments.

1 M.Barron and A.H.Marshall,"Spatial impression due to early Inteal reflection in concert halls; The derivation of a physical mea sure," Vib. 77, 211-232 (1981) .

2 P.Damaske,"Subjective Untersuchung von Schallfeldern,"Acustica **19**, 199-213 (1967/68).

3 T.Anazawa, H.Yanagawa, and T.Itow,"On correlation coefficients of both ears and ‘feeling of wideness,'" IEICE Jpn. Tech. Rep. No. EA70-13 (1970)(in Japanese).

4 J. Blauert and W.Lindemann, "Auditory spaciousness; Some further psychoacoustics analyses," J.Acoust.Soc. Am.**80**, 533-542 (1986)

5 Y.Ando and Y.Kurihara, "Nonlinear response in evaluating the subjective diffuseness of sound field," J.Acoust.Soc.Am.**80**, 833-836 (1986)

6 M.R.Schroeder, D.Gottlob, and K.F.Siebrasse, "Comparative study of European concert halls: Correlation of subjective preference with geomet ric and acoustic parameters," J.Acoust.Soc.Am.**56**, 1195-1201 (1974)

7 H.Yanagawa, H.Higashi, and S.Mori, "Interaural correlation coefficients of the dummy head and the feeling of wideness," Acoust.Soc.Jpn.Tech.Rep.No.1135-1 (1976)(in Jpanese)

8 M.Tohyama and A.Suzuki, "lnteraural cross-correlation coeflieients in reproduced sound fields," JAS J.**4**, 3-13 (1985)(in Jpanese)

9 Y.Hirata, "Improving stereo at L.F.," Wireless World 89, 60-62 (1983)

10 T.Anazawa, H.Yanagawa, and T.Itow, "On transient correlation coefli cients of sound field," Acust. Soc. Jpn. Tech. Rep. No.AA 71-3 (1971)(in Japanese)

11 K.Endoh, Y.Yamasaki, and T.Itow, "Grasp and developmentof spatial information in a room by closely located four-point microphone method."ICASSP **86**, 909-912 (1986)