We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



19

Sound Image Localization on Flat Display Panels

Gabriel Pablo Nava, Yoshinari Shirai, Kaji Katsuhiko, Masafumi Matsuda, Keiji Hirata and Shigemi Aoyagi
NTT Communication Science Labs
Japan

1. Introduction

In recent years, we have experienced rapid advancements on solid-state luminous displays which have lead to their introduction in many, if not all, multimedia applications. Moreover, former light projectors used in immersive audiovisual spaces have been succeeded by large LCD panels which have overcome requirements such as dark room, occlusion-free sightline between projector and screen, spacious installation, etc. But on the other hand, while the screens used with light projectors allow a relatively free placement of loudspeakers (e.g. behind the screen) to provide realistic sound spatialization, flat LCD displays impose new challenges from the acoustical point of view. One of them is to provide positional correspondence between the video images on the display and their sound image while allowing multiple users to properly localize the sound images on the LCD panels. To achieve this, *auditory displays* have been traditionally implemented with conventional stereo loudspeakers installed at the sides of the video display panel as shown in Fig. 1. However, a fundamental problem with such setups is that correct sound image localization is achieved only by listeners positioned at the symmetrical axis of the stereo array. Listeners standing at the asymmetrical areas tend to perceive a sound image shifted towards their closest loudspeaker. This problem has been extensively studied and attributed to the so-called *precedence effect*: from multiple sound sources radiating similar sound intensities, a listener tends to localize a single sound image close to the nearest source, given that the differences of arrival times of the sounds reaching his ears are between about 1 and 50 ms (Gardner, 1968). As demonstrated by a number of precise experiments (Rakerd, 1986), the precedence effect provides important psychoacoustical cues to the process of sound localization in humans as well as in other species (Litovsky et al., 1999), nevertheless, it represents a fundamental problem in stereo reproduction systems that attempt realistic sound spatialization over a wide area. For this reason, several approaches to expand the area of accurate sound image perception have appeared in the literature. Some of them have coped the problem as to that of developing loudspeakers with radiation characteristics that help to equalize the sound from both channels over a broad area, while others focus on pre-processing the audio signals (either in the analog or digital domain) to create an spatialization effect on the listeners. In the first scenario, Bauer reported a loudspeaker arrangement aiming the optimum trade off between the position of the listener and the balance of acoustic energy distributed from the stereo channels (Bauer, 1960). He concluded that stereo loudspeakers whose radiation pattern follows the cosine law (such as that of the front lobe of an acoustic dipole), together

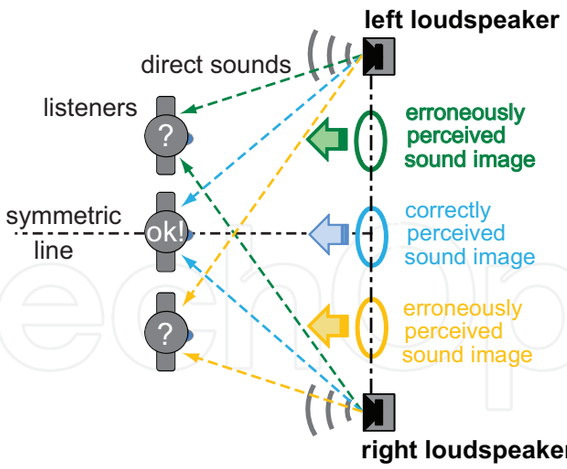


Fig. 1. Traditional stereo loudspeaker setup. Listeners at the asymmetrical areas tend to perceive a sound image shifted towards their closest loudspeaker direction.

with their placement facing in counter fire, can achieve balance of the stereo channels not only at the symmetrical line but also within the (circular) area that encloses the two acoustic dipoles. On the contrary, Kates, based on his model that takes into account the directional characteristics of the human ear (Kates, 1980), debated that the loudspeakers should be more directional (rather than with a spherical radiation) to target an specific audience area. This idea was further supported by Davis who employed arrays of multi-loudspeaker drivers per stereo channel to achieve the desired radiation pattern over the audio band (Davis, 1987). In his implementations, Davis used crossover networks to deliver specific-band signals to each loudspeaker driver of determined frequency response so that the overall interaction of the individual loudspeakers would lead the sought directionality. In the second scenario, but still aiming at the directional radiation issue, Aoki *et. al.* proposed a 6-drivers stereo loudspeaker array (three drivers per channel) capable of maintaining in-head sound image localization within the seating area (a rectangular table), (Aoki & Koizumi, 1987). Each loudspeaker driver of the triplets are excited by either delayed, inverted and/or amplified versions of the original stereo signals. Thus, when the audio signals are radiated, the unwanted acoustic energy at specific areas is suppressed by cancelation of the opposed-phase waves. A more recent approach following similar ideas was introduced by Ródenas *et. al.* in what they called a *position-independent* stereo system (Ródenas et al., 2003). In contrast to the system of Aoki *et. al.*, the position-independent implementation used pairs of loudspeaker drivers in each channel. Ródenas et. al. further proposed psycoacoustic models to achieve optimum directivity of the loudspeakers through their excitation by signals filtered with optimal FIR filters. Another example of binaural reproduction based on Wiener filters was reported by Kim and Wang, (Kim & Wang, 2003). Other approaches, although suited to a single listener, take advantage of the computational power of recent computers to incorporate video tracking systems that provide real-time information of the listener's position (with respect to the loudspeakers) to adjust the audio signals that should be delivered by the stereo loudspeakers in order to render a sound image at the recognized position, (examples: Gardner (1997); Kyriakakis et al. (1998); Merchel & Groth (2009)). Yet, there is another tendency of recent audio reproduction systems which is not limited to stereo channels but rather requires large numbers of collinearly arranged loudspeakers to provide audio spatialization to an specific

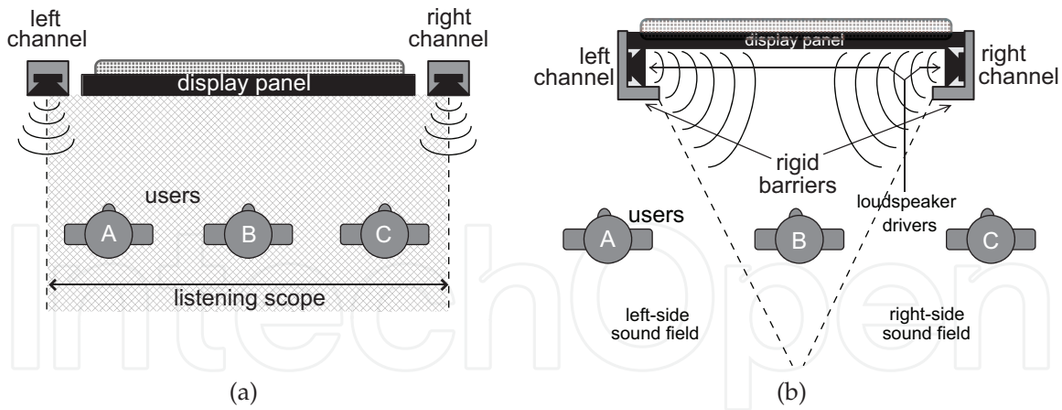


Fig. 2. (a) Conventional stereo loudspeakers with a display panel. (b) The new loudspeaker design with L-like rigid barriers.

audience. Their principle is based on the assumption that any given wave field can be accurately reconstructed by the superposition of elementary spherical waves radiated from a number of individually driven sound sources (loudspeakers). The most representative technique of this approach is the Wave Field Synthesis, or WFS (Berkhout et al., 1993), and its successful performance in applications including joint video and audio/speech has been substantiated by experimental work with setups in which the large array of loudspeakers can be easily hidden behind the thin projection screens (Melchoir et al., 2003; Werner & Boone, 2003). As Werner points out, accurate audio spatialization reproduction enhances the naturalness of the interaction among the participants in a teleconference, leading to a more productive collaborative work (Werner & Boone, 2003). From that perspective, WFS has probed to provide a high quality auditory interface for the target audience. Nevertheless, as modern teleconference-collabotarion systems tend to include solid-state luminous displays large enough (usually forming complete walls) to immerse the users, the loudspeaker placement and the acoustic influence of the flat hard displays becomes a serious problem to straightforwardly implement sound spatialization with existent methods and conventional loudspeakers. At this level, the loudspeaker design and placement seem to play such an important role as that of the audio signal processing involved to render a sound image at a given position. Nonetheless, relatively less research has been reported in this field. It is, therefore, the aim of this chapter to introduce a novel loudspeaker design for flat hard display panels on which localization of sound images is desired. Its simple, yet effective, principle is demonstrated with numerical simulations, objective and subjective experiments, and real-time user-interactive applications.

2. Loudspeakers for flat display panels

2.1 An L-like design to improve sound image localization

Conventional stereo loudspeakers for video displays are typically mounted as independent units placed at the sides of the display panel, facing straight to the audience. In spite of the existent stereo spatialization techniques, the listening area achieved by this arrangement is delimited by the separation of the stereo loudspeakers as illustrated in Fig. 2(a). On the other hand, some acoustic engineers (Bauer, 1960; Kates, 1980), based on experimental work, recommend an angular placement of the loudspeakers facing counter-fire into the target

audience. Such loudspeaker angulation results in a an expansion of the overall radiation pattern of the stereo array. Following a similar idea, in the novel design proposed here, the loudspeaker drivers are, in addition, embedded to the interior of L-like rigid structures (referred as *rigid barriers* hereafter) which at the same time are attached to the lateral edges of the display panel. These rigid barriers act as attenuators of sound intensity as a function of the listening direction. In this way, the loudspeaker drivers are, indeed, off the plane of the display surface and pointing in counter-fire into its interior. Fig. 2(b) shows an sketch of the proposed loudspeaker arrangement for flat displays. With such configuration, the design aims to mechanically redistribute the radiated acoustic energy so as to balance the sound strength at the off-symmetrical axis. To achieve that, the design relies on two basic assumptions:

1. the display panel is flat and hard, and
2. monaural sound signals are to be rendered on the panel.

Note that, in principle, compliance of the above statements allows applicability of the design to any display panel of a given size. Nevertheless, the dimensions of the rigid barriers have to be optimized because of the dependency on operative frequency, the size of the panel and the coverage area. Such issue is addressed in further paragraphs.

Let us consider an example where a monaural sound image at the center of the display panel of Fig. 2(b) is desired. Moreover, suppose that the audio signals to drive the left and right channels are identical, i.e. zero phase and equal amplitude. At the symmetrical point B (or along the symmetrical axis), the acoustic signals radiated by the loudspeakers are observed with the same (or nearly the same) time and intensity and through a direct path from both channels. Therefore, as in a conventional stereo loudspeaker setup, a listener at B is able to perceive a central sound image on the display. At the asymmetrical positions A and C, the sound from the closest loudspeaker arrives first through indirect paths (by diffraction on the edges), but its intensity is attenuated by the rigid barrier so as to prevent the masking effect of the sound from the opposite channel (which arrives through a direct path). Note also that, with this configuration, the acoustic energy from one channel is fundamentally radiated to the front of the panel and to the opposite-side field. Thus, the hypothesis is that as the sound from the farthest loudspeaker remains unmasked, the effect of precedence is expected to be reduced as function of the listening position, leading to an improvement of sound image localization at the asymmetrical areas. With the aid of numerical simulations, this hypothetical supposition will be further demonstrated.

2.2 Optimization of the L-like design

Although the simplicity of the L-like design makes it easy to implement, the dimensions of the rigid barriers are not fixed and should be optimized for a particular case. If for example, the sound of an audio-visual application is to be spatialized over an effective listening area defined by the coverage angle θ from the surface of the display panel of size h on Fig. 3(a), then the size δ of the L-like rigid barriers has to be optimized to achieve the desired level of sound intensity radiated over a grid of N (receiving) field points. Therefore, the optimum size $\delta_{mboxopt}$ is the one that minimizes the following objective function within a number β of

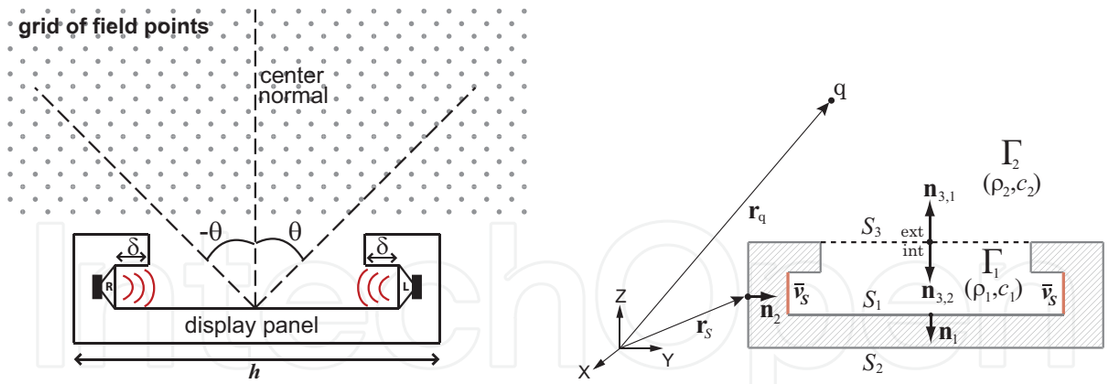


Fig. 3. (a) Optimization of L-like rigid barriers for a target listening area. (b) Theoretical model for the numerical optimization.

discrete frequencies in a given audible frequency band, as

$$\delta_{\text{opt}} = \min_{\delta \in \mathbb{R}^+} \frac{1}{\beta} \sum_{k=1}^{\beta} G(f_k, \delta) \qquad (1)$$
$$\text{subject to} \quad 0 < \delta \leq \max \ ,$$

where

$$G(f_k, \delta) = 10 \ \log_{10} \left[\frac{1}{N} \sum_{q=1}^{N} \left(\left| p_q^{\text{des}}(f_k, \theta) \right|^2 - \left| p_q^{\text{sim}}(f_k, \delta) \right|^2 \right| \right] \ . \qquad (2)$$

Here, p_q^{des} and p_q^{sim} are the actually desired and the numerically simulated sound pressures at the q-th field point. Alternatively, the sound field p_q^{des} can be further represented as

$$p_q^{\text{des}}(f_k, \theta) = p_q^{\text{des}}(f_k) \cdot W(\theta) \ , \qquad (3)$$
$$0 \leq W(\theta) \leq 1 \ ,$$

i.e., the sound pressure p_q weighted by a normalized θ-dependent function $W(\theta)$ to further control the desired directivity. Thus, G measures the average (in dB's) of the error between the (squared) magnitude of p_q^{des} and p_q^{sim}.

The acoustic optimization is usually performed numerically with the aid of computational tools such as finite element (FEM) or boundary element (BEM) methods, being the latter a frequently preferred approach to model sound scattering by vibroacoustic systems because of its computational efficiency over its FEM counterpart. BEM is, indeed, an appropriate framework to model the acoustic display-loudspeaker system of Fig. 3(a). Therefore, the adopted theoretical approach will be briefly developed following boundary element formulations similar to those described in (Ciskowski & Brebbia, 1991; Estorff, 2000; Wu, 2000) and the domain decomposition equations in (Seybert et al., 1990).

Consider the solid body of Fig. 3(b) whose concave shape defines two subdomains, an interior Γ_1 and an exterior Γ_2, filled with homogeneous compressible media of densities ρ_1 and ρ_2,

where any sound wave propagates at the speed c_1 and c_2 respectively. The surface of the body is divided into subsegments so that the total surface is $S = S_1 + S_2 + S_3$, i.e. the interior, exterior and an imaginary auxiliary surface. When the acoustic system is perturbed with a harmonic force of angular frequency ω, the sound pressure p_{q} at any point q in the 3D propagation field, is governed by the Kirchhoff-Helmholtz equation

$$\int_S \left(p_S \frac{\partial \Psi}{\partial \mathbf{n}} - \Psi \frac{\partial p_S}{\partial \mathbf{n}} \right) dS + p_{\mathrm{q}} = 0 \,, \tag{4}$$

where p_S is the sound pressure at the boundary surface S with normal vector $\mathbf{n}$. The Green's function Ψ is defined as $\Psi = e^{-jkr}/4\pi r$ in which $k = \omega/c$ is the wave number, $r = |\mathbf{r}_S - \mathbf{r}_{\mathrm{q}}|$ and $j = \sqrt{-1}$. Moreover, if the field point q under consideration falls at any of the domains Γ_1 or Γ_2 of Fig. 3(b), the sound pressure p_{q} is related to the boundary of the concave body by

• for q in Γ_1:

$$C_{\mathrm{q}}\, p_{\mathrm{q}} + \int_{S_1+S_3} \left(\frac{\partial \Psi}{\partial \mathbf{n}}\, p_{S_1+S_3} - \Psi\, \frac{\partial p_{S_1+S_3}}{\partial \mathbf{n}} \right) dS =$$

$$C_{\mathrm{q}}\, p_{\mathrm{q}} + \int_{S_1} \left(\frac{\partial \Psi}{\partial \mathbf{n}_1}\, p_{S_1} + j\omega\rho_1\, \Psi\, v_{S_1} \right) dS + \int_{S_3} \left(\frac{\partial \Psi}{\partial \mathbf{n}_{3,1}}\, p^{\mathrm{int}}_{S_3} + j\omega\rho_1\, \Psi\, v^{\mathrm{int}}_{S_3} \right) dS = 0 \tag{5}$$

• for q in Γ_2:

$$C_{\mathrm{q}}\, p_{\mathrm{q}} + \int_{S_2+S_3} \left(\frac{\partial \Psi}{\partial \mathbf{n}}\, p_{S_2+S_3} - \Psi\, \frac{\partial p_{S_2+S_3}}{\partial \mathbf{n}} \right) dS =$$

$$C_{\mathrm{q}}\, p_{\mathrm{q}} + \int_{S_2} \left(\frac{\partial \Psi}{\partial \mathbf{n}_2}\, p_{S_2} + j\omega\rho_2\, \Psi\, v_{S_2} \right) dS + \int_{S_3} \left(\frac{\partial \Psi}{\partial \mathbf{n}_{3,2}}\, p^{\mathrm{ext}}_{S_3} + j\omega\rho_2\, \psi\, v^{\mathrm{ext}}_{S_3} \right) dS = 0 \tag{6}$$

Note that in the latter equations (5) and (6), the particle velocity equivalent $\partial p/\partial \mathbf{n} = -j\omega\rho\, v$ has been used. Thus, p_{S_i} and v_{S_i} represent the sound pressure and particle velocity on the i-th surface S_i. The parameter C_{q} depends on the solid angle in which the surface S_i is seen from p_{q}. For the case when q is on a smooth surface, $C_{\mathrm{q}} = 1/2$, and when q is in Γ_1 or Γ_2 but not on any S_i, $C_{\mathrm{p}} = 1$.

To solve equations (5) and (6) numerically, the model of the solid body is meshed with discrete surface elements resulting in a number of L elements for the interior surface $S_1 + S_3$ and M for the exterior $S_2 + S_3$. If the point q is matched to each node of the mesh (collocation method), equations (5) and (6) can be written in a discrete-matrix form

$$\mathbf{A}_{S_1}\, \mathbf{p}_{S_1} + \mathbf{A}^{\mathrm{int}}_{S_3}\, \mathbf{p}^{\mathrm{int}}_{S_3} - \mathbf{B}_{S_1}\, \mathbf{v}_{S_1} - \mathbf{B}^{\mathrm{int}}_{S_3}\, \mathbf{v}^{\mathrm{int}}_{S_3} = 0 \,, \tag{7}$$

and

$$\mathbf{A}_{S_2}\, \mathbf{p}_{S_2} + \mathbf{A}^{\mathrm{ext}}_{S_3}\, \mathbf{p}^{\mathrm{ext}}_{S_3} - \mathbf{B}_{S_2}\, \mathbf{v}_{S_2} - \mathbf{B}^{\mathrm{ext}}_{S_3}\, \mathbf{v}^{\mathrm{ext}}_{S_3} = 0 \,, \tag{8}$$

where the $\mathbf{p}_{S_i}$ and $\mathbf{v}_{S_i}$ are vectors of the sound pressures and normal particle velocities on the elements of the i-th surface. Furthermore, if one collocation point at the centroid of the each element, and constant interpolation is considered, the entries of the matrices $\mathbf{A}_{S_i}, \mathbf{B}_{S_i}$, can be

computed as

$$a_{l,m} = \begin{cases} \int_{s_m} \dfrac{\partial \Psi(\mathbf{r}_l, \mathbf{r}_m)}{\partial \mathbf{n}_l}\, ds & \text{for } l \neq m \\ 1/2 & \text{for } l = m \end{cases}$$

$$b_{l,m} = -j\omega\rho_k \int_{s_m} \Psi(\mathbf{r}_l, \mathbf{r}_m)\, ds\,, \tag{9}$$

where s_m is the m-th surface element, the indexes $l = m = \{1, 2, \ldots, L \text{ or } M\}$, and $k = \{1, 2\}$ depending on which subdomain is being integrated.
When velocity values are prescribed to the elements of the vibrating surfaces of the loudspeaker drivers (see $\overline{v}_S$ in Fig. 3(b)), equations (7) and (8) can be further rewritten as

$$\mathbf{A}_{S_1}\, \mathbf{p}_{S_1} + \mathbf{A}^{\text{int}}_{S_3}\, \mathbf{p}^{\text{int}}_{S_3} - \hat{\mathbf{B}}_{S_1}\, \hat{\mathbf{v}}_{S_1} - \mathbf{B}^{\text{int}}_{S_3}\, \mathbf{v}^{\text{int}}_{S_3} = \bar{\mathbf{B}}_{S_1}\, \bar{\mathbf{v}}_{S_1}\,, \tag{10}$$

$$\mathbf{A}_{S_2}\, \mathbf{p}_{S_2} + \mathbf{A}^{\text{ext}}_{S_3}\, \mathbf{p}^{\text{ext}}_{S_3} - \hat{\mathbf{B}}_{S_2}\, \hat{\mathbf{v}}_{S_2} - \mathbf{B}^{\text{ext}}_{S_3}\, \mathbf{v}^{\text{ext}}_{S_3} = \bar{\mathbf{B}}_{S_2}\, \bar{\mathbf{v}}_{S_2}\,, \tag{11}$$

Thus, in equations (10) and (11), the $\widehat{\mathbf{v}}_{S_i}$'s and $\overline{\mathbf{v}}_{S_i}$'s denote the unknown and known particle velocities, and $\widehat{\mathbf{B}}_{S_i}$'s and $\overline{\mathbf{B}}_{S_i}$'s their corresponding coefficients.
At the auxiliary interface surface S_3, continuity of the boundary conditions must satisfy

$$\rho_1\, p^{\text{int}}_{S_3} = \rho_2\, p^{\text{ext}}_{S_3} \tag{12}$$

and

$$\frac{\partial p^{\text{int}}_{S_3}}{\partial \mathbf{n}_{3,1}} = -\frac{\partial p^{\text{ext}}_{S_3}}{\partial \mathbf{n}_{3,2}}\,, \quad \text{or} \quad , -j\omega\rho_1\, v^{\text{int}}_{S_3} = j\omega\rho_2\, v^{\text{ext}}_{S_3} \tag{13}$$

Considering that both domains Γ_1 and Γ_2 are filled with the same homogeneous medium (e.g. air), then $\rho_1 = \rho_2$, leading to

$$p^{\text{int}}_{S_3} = p^{\text{ext}}_{S_3} \tag{14}$$

and

$$-\, v^{\text{int}}_{S_3} = v^{\text{ext}}_{S_3} \tag{15}$$

Substituting these interface boundary parameters in equations (10) and (11), and rearranging into a global linear system of equations where the surface sound pressures and particle velocities represent the unknown parameters, yields to

$$\begin{bmatrix} \mathbf{A}_{S_1} & \mathbf{0} & \mathbf{A}^{\text{int}}_{S_3} & -\hat{\mathbf{B}}_{S_1} & \mathbf{0} & -\mathbf{B}^{\text{int}}_{S_3} \\ \mathbf{0} & \mathbf{A}_{S_2} & \mathbf{A}^{\text{ext}}_{S_3} & \mathbf{0} & -\hat{\mathbf{B}}_{S_2} & \mathbf{B}^{\text{ext}}_{S_3} \end{bmatrix} \begin{bmatrix} \mathbf{p}_{S1} \\ \mathbf{p}_{S2} \\ \mathbf{p}_{S3} \\ \hat{\mathbf{v}}_{S1} \\ \hat{\mathbf{v}}_{S2} \\ \mathbf{v}^{\text{int}}_{S3} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{B}}_{S_1}\, \bar{\mathbf{v}}_{S_1} \\ \bar{\mathbf{B}}_{S_2}\, \bar{\mathbf{v}}_{S_2} \end{bmatrix}. \tag{16}$$

Observe that the matrices $A's$ and $B's$ are known since they depend on the geometry of the model. Thus, once the vibration $\bar{\mathbf{v}}_{S1}$ of the loudspeakers is prescribed, and after equation (16) is solved for the surface parameters, the sound pressure at any point q can be readily computed by direct substitution and integration of equation (5) or (6). Note also that, a

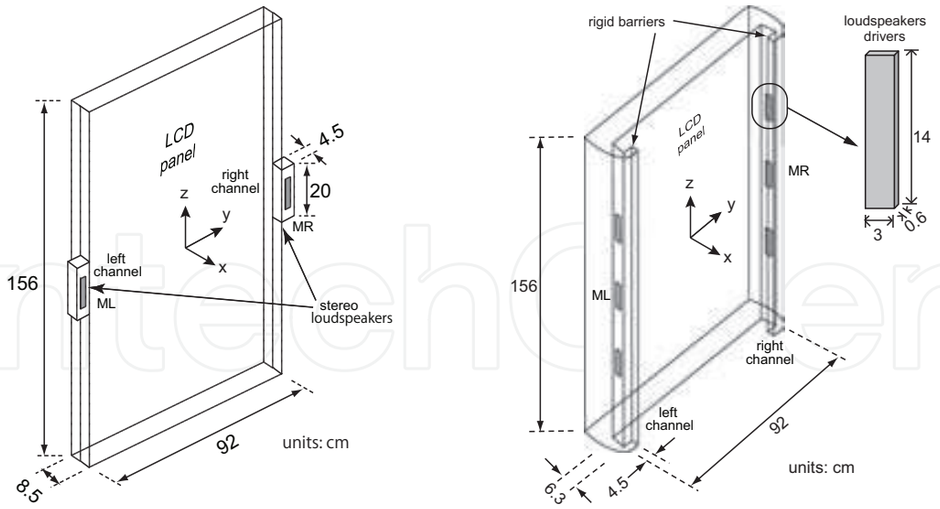


Fig. 4. Conventional stereo loudspeakers (a), and the L-like rigid barrier design (b), installed on 65″ flat display panels.

multidomain approach allows a reduction of computational effort during the optimization process since the coefficients of only one domain (interior) have to be recomputed.

3. Sound field analysis of a display-loudspeaker panel

3.1 Computer simulations

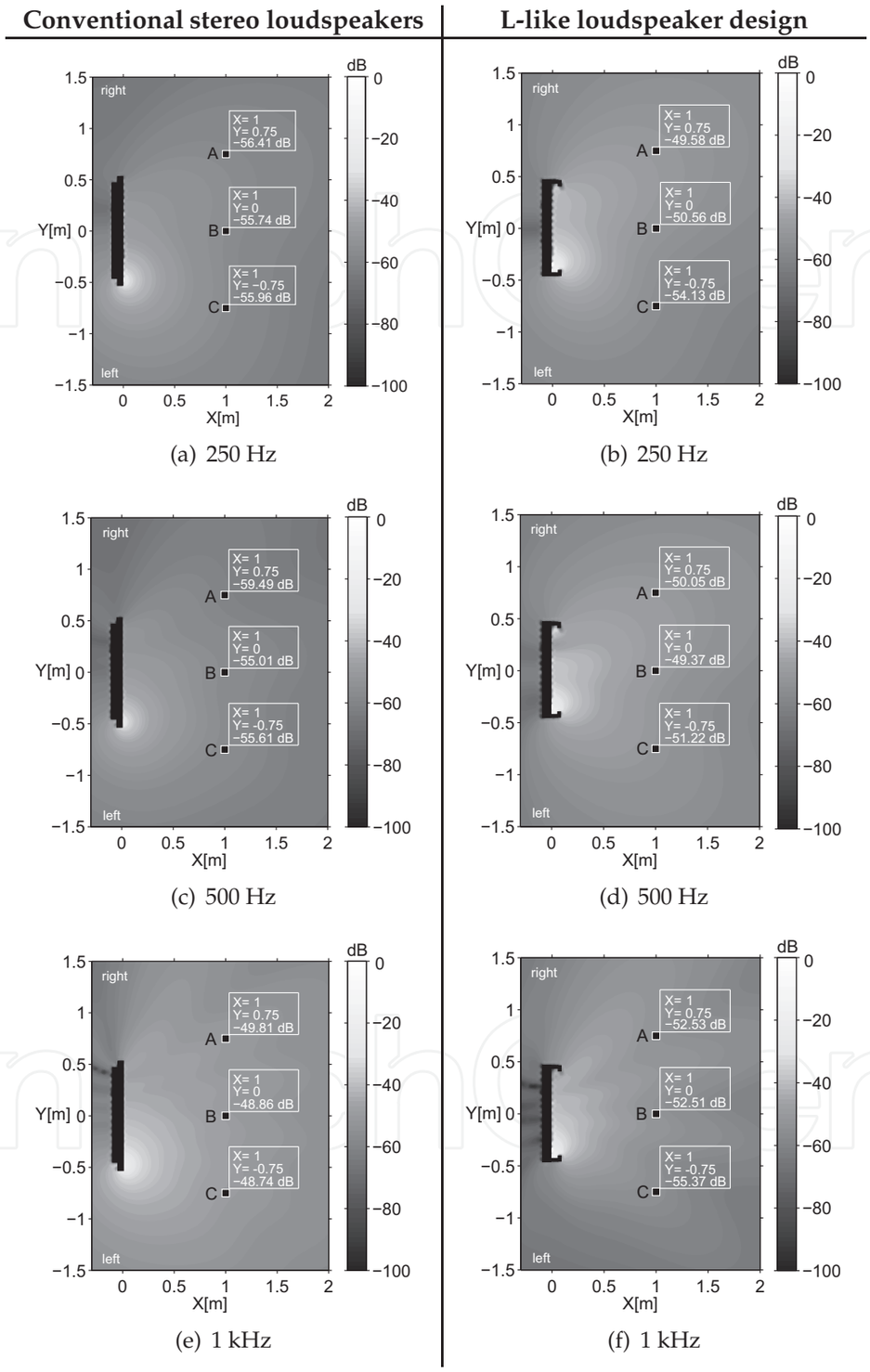
3.1.1 Numerical model

Recent multimedia applications often involve large displays to present life-size images and sound. Fig. 4(a) shows an example of a vertically aligned 65-inch LCD panel intended to be used in immersive teleconferencing. To this model, conventional (box) loudspeakers have been attached at the lateral sides to provide stereo spatialization. A second display panel is shown in Fig. 4(b), this is a prototype model of the L-shape loudspeakers introduced in the previous section. The structure of both models is considered to be rigid, thus, satisfying the requirements of flatness and hardness of the display surface.

In order to appreciate the sound field generated by each loudspeaker setup, the sound pressure at a grid of field points was computed following the theoretical BEM framework discussed previously. Considering the convention of the coordinate system illustrated in Figs. 4(a) and 4(b), the grid of field points were distributed within -0.5 m $\leq x \leq$ 2 m and -1.5 m $\leq y \leq$ 1.5 m spaced by 1 cm. For the numerical simulation of the sound fields, the models were meshed with isoparametric triangular elements with a maximum size of 4.2 cm which leaves room for simulations up to 1 kHz assuming a resolution of 8 elements per wavelength. The sound source of the simulated sound field was the left-side loudspeaker (marked as ML in Figs. 4(a) and 4(b)) emitting a tone of 250 Hz, 500 Hz and 1 kHz, respectively for each simulation. The rest of the structure is considered static.

3.1.2 Sound field radiated from the flat panels

The sound fields produced by each model are shown in Fig. 5. The sound pressure level (SPL) in those plots is expressed in dB's, where the amplitude of the sound pressure has been



(a) 250 Hz (b) 250 Hz

(c) 500 Hz (d) 500 Hz

(e) 1 kHz (f) 1 kHz

Fig. 5. Sound field generated by a conventional stereo setup (left column), and by the L-like loudspeaker design (right column) attached to a 65-inch display panel. Sound source: left-side loudspeaker (ML) emitting a tone of 250 Hz, 500 Hz, and 1 kHz respectively.

normalized to the sound pressure p_{spk} on the surface of the loudspeaker driver ML. For each analysis frequency, the SPL is given by

$$\mathrm{SPL} = 20 \log_{10} \frac{|p_{\mathrm{q}}|}{|p_{spk}|} \tag{17}$$

where a SPL of 0 dB is observed on the surface of ML.
In the plots of Figs. 5(b), 5(d) and 5(f), (the L-like design), the SPL at the points A and B has nearly the same level, while point C accounts for the lowest level since the rigid barriers have effectively attenuated the sound at that area. Contrarily, Figs. 5(a), 5(c) and 5(e), (conventional loudspeakers) show that the highest SPL level is observed at point C (the closest to the sounding loudspeaker), whereas point A gets the lowest. Further note that if the right loudspeaker is sounding instead, symmetric plots are obtained. Let us recall the example where a sound image at the center of the display panel is desired. When both channels radiate the same signal, a listener on point B observes similar arrival times and sound intensities from both sides, leading to a sound image perception on the center of the panel. However, as demonstrated by the simulations, the sound intensities (and presumably, the arrival times) at the asymmetric areas are unequal. In the conventional stereo setup of Fig. 4(a), listeners at points A and C would perceive a sound image shifted towards their closest loudspeaker. But in the loudspeaker design of Fig. 4(b), the sound of the closest loudspeaker has been delayed and attenuated by the mechanical action of the rigid barriers. Thus, the masking effect on the sound from the opposite side is expected to be reduced leading to an improvement of sound image localization at the off-symmetry areas.

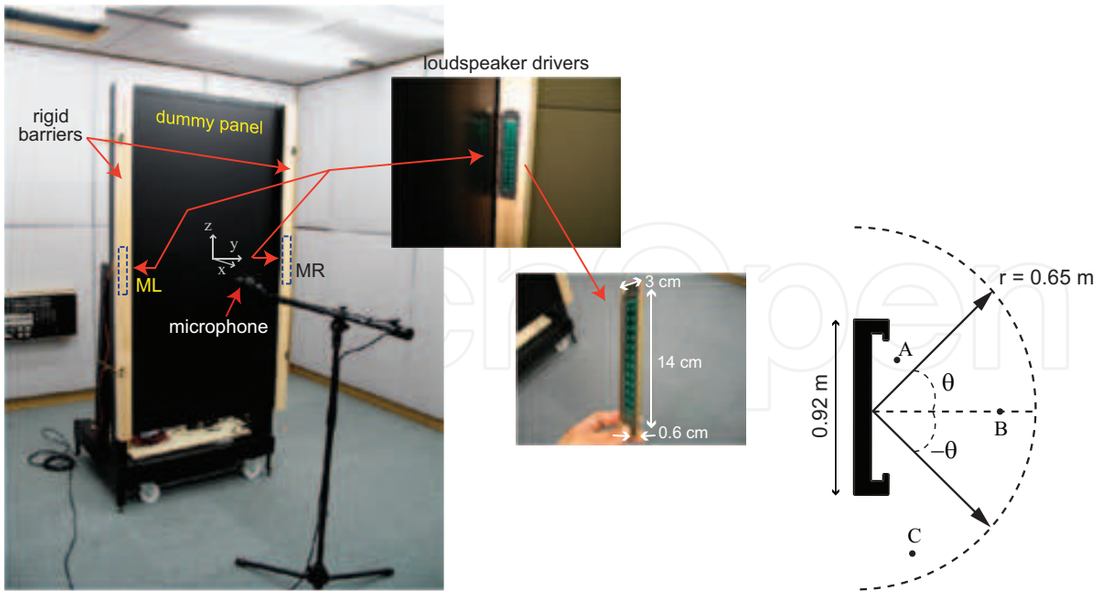
3.2 Experimental analysis

3.2.1 Experimental prototype

It is a common practice to perform experimental measurements to confirm the predictions of the numerical model. In this validation stage, a basic (controllable) experimental model is desired rather than a real LCD display which might bias the results. For that purpose, a flat dummy panel made of wood can be useful to play the role of a real display. Similarly, the rigid L-like loudspeakers may be implemented with the same material. An example of an experimental prototype is depicted in Fig. 6(a) which shows a 65-inch experimental dummy panel built with the same dimensions as the model of Fig. 4(b). The loudspeaker drivers employed in this prototype are 6 mm-thick flat coil drivers manufactured by FPS Inc., which can output audio signals of frequency above approximately 150 Hz. This experimental prototype was used to performed measurements of sound pressure inside a semi-anechoic room.

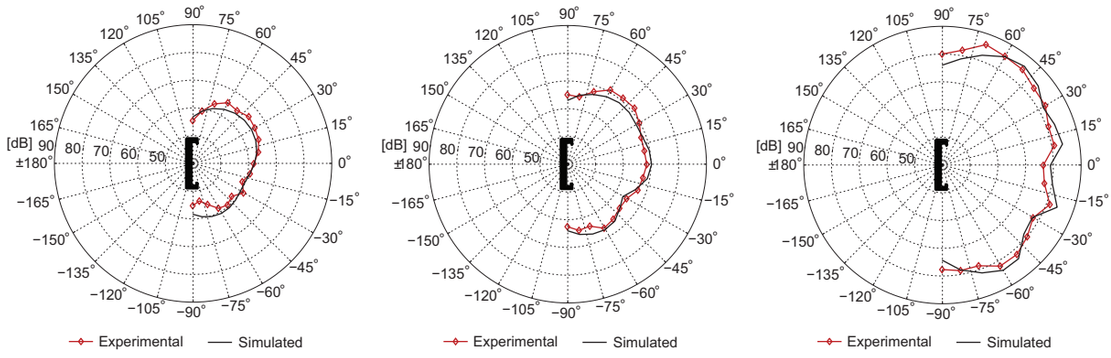
3.2.2 Sound pressure around the panel

The sound field radiated by the flat display panel has been demonstrated with numerical simulations in Fig. 5. In practice, however, measuring the sound pressure in a grid of a large number of points is troublesome. Therefore, the first experiment was limited to observe the amplitude of the sound pressure at a total of 19 points distributed on a radius of 65 cm from the center of the dummy panel, and separated by steps of 10^{o} along the arc $-90^{\mathrm{o}} \leq \theta \leq 90^{\mathrm{o}}$ as depicted in Fig. 6(b), while the left-side loudspeaker ML was emitting a pure tone of 250 Hz, 500 Hz and 1 kHz respectively.



(a) (b)

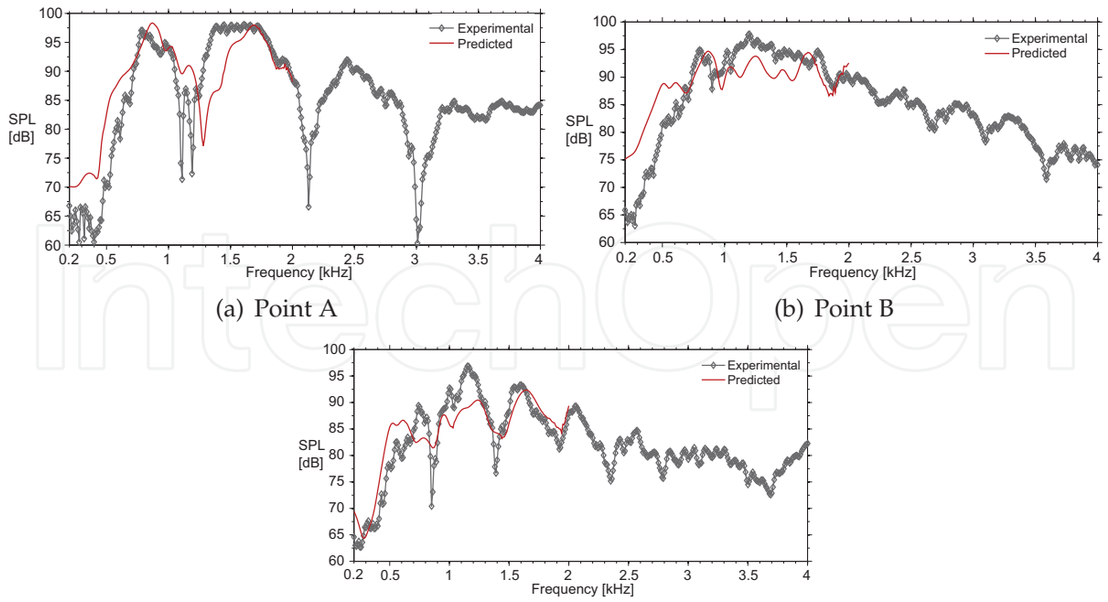
Fig. 6. (a) Experimental dummy panel made of wood resembling a 65-inch vertically align LCD display. (b) Location of the measurements of the SPL generated by the dummy panel.



(a) 250 Hz (b) 500 Hz (c) 1 kHz

Fig. 7. Sound pressure at three static points (A, B and C), generated by a 65-inch LCD panel (Sharp LC-65RX) within the frequency band 0.2 - 4 kHz.

The attenuation of sound intensity introduced by the L-like rigid barriers as a function of the listening angle, can be observed on the polar plots of Fig. 7 where the results of the measurements are presented. Note that the predicted and experimental SPL show close agreement and also similarity to the sound fields of Fig. 5 obtained numerically, suggesting that the panel is effectively radiating sound as expected. Also, the dependency of the radiation pattern to the frequency has been made evident by these graphs, reason why this factor is taken into account in the acoustic optimization of the loudspeaker design.



(a) Point A (b) Point B

(c) 1 kHz

Fig. 8. Sound pressure level at a radius of 65 cm apart from the center of the dummy panel.

3.2.3 Frequency response in the sound field

A second series of measurements of SPL were performed at three points where, presumably, users in a practical situation are likely to stand. Following the convention of the coordinate system aligned to the center of the panel (see Fig. 6(a)), the chosen test points are $A(0.25, 0.25)$, $B(0.5, 0.0)$ and $C(0.3, -0.6)$ (in meters). At these points, the SPL due to the harmonic vibration of both loudspeakers, ML and MR, was measured within the frequency band 0.2–4 kHz with intervals of 10 Hz. For the case of the predicted data, the analysis was constrained to a maximum frequency of 2 kHz because of computational power limitations. The lower bound of 0.2 kHz is due to the frequency characteristics of the employed loudspeaker drivers.

The frequency response at the test points A, B and C, are shown in Fig. 8. Although there is a degree of mismatch between the predicted and experimental data, both show similar tendencies. It is also worth to note that the panel radiates relatively less acoustic energy at low frequencies (approximately below 800 Hz). This highpass response was originally attributed to the characteristics of the experimental loudspeaker drives, however, observation of a similar effect in the simulated data reveals that the panel, indeed, embodies a highpass behavior. This feature can lead to difficulties in speech perception in some applications such as in teleconferencing, in which case, reenforcement of the low frequency contents may be required.

4. Subjective evaluation of the sound images on the display panel

The perception of the sound images rendered on a display panel has been evaluated by subjective experiments. Thus, the purpose of these experiments was to assess the accuracy of the sound image localization achieved by the L-like loudspeakers, from the judgement of a group of subjects. The test group consisted of 15 participants with normal hearing capabilities whose age ranged between 23 and 56 years old (with mean of 31.5). These subjects were

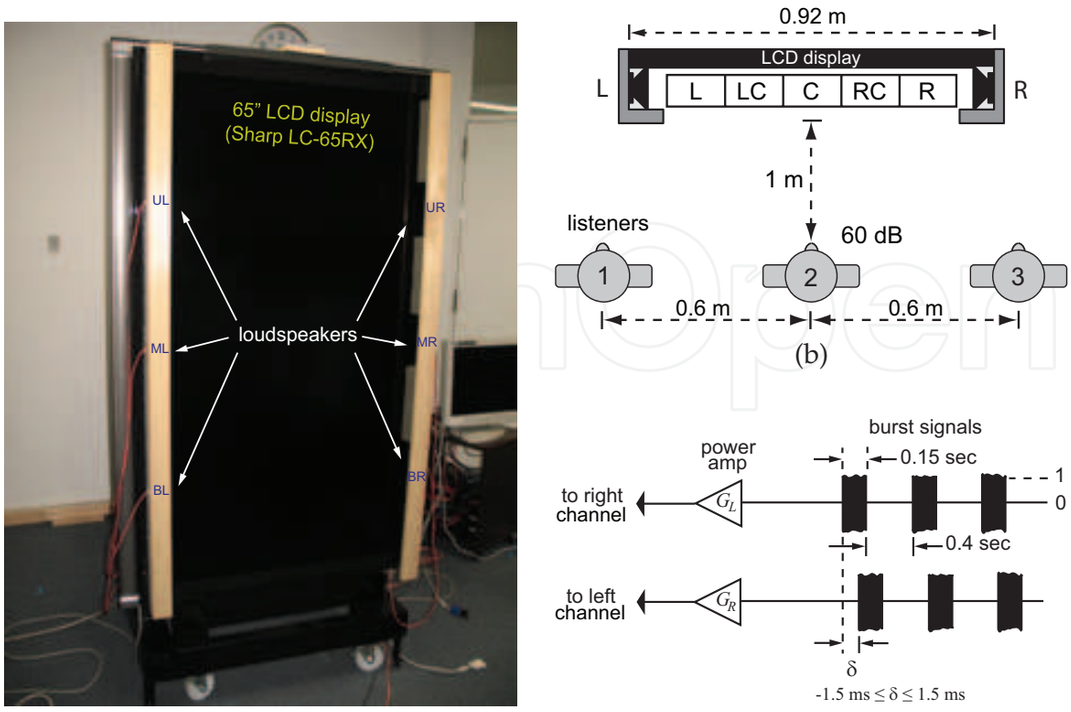


Fig. 9. 65-inch LCD display with the L-like loudspeakers installed (a). Setup for the subjective tests (b). Broadband signals to render a single sound image between ML and MR on the LCD display.

asked to localize the sound images rendered on the surface of a 65-inch LCD display (Sharp LC-65RX) which was used to implement the model of Fig. 9(a).

4.1 Setup for the subjective tests

The 15 subjects were divided into groups of 3 individuals to yield 5 test sessions (one group per session). Each group was asked to seat at one of the positions 1, 2 or 3 which are, one meter away form the display, as indicated in Fig. 9(b). In each session, the participants were presented with 5 sequences of 3 different sound images reproduced (one at a time) arbitrarily at one of the 5 equidistant positions marked as L, LC, C, RC, and R, along the line joining the left (ML) and right (MR) loudspeakers. At the end of each session, 3 sound images have appeared at each position, leading to a total of 15 sound images at the end of the session. After every sequence, the subjects were asked to identify and write down the perceived location of the sound images.

To render a sound image at a given position, the process started with a monaural signal of broadband-noise bursts with amplitude and duration as specified in Fig. 9(c). Therefore, to place a sound image, the gain G of each channel was varied within ($0 \leq G \leq 1$), and the delay δ between the channels was linearly interpolated in the range -1.5 ms $\leq \delta \leq 1.5$ ms. In such a way that, a sound image on the center corresponds to half the gain of the channels and zero delay, producing a sound pressure level of 60 dB (normalized to 20 μP) at the central point (position 2).

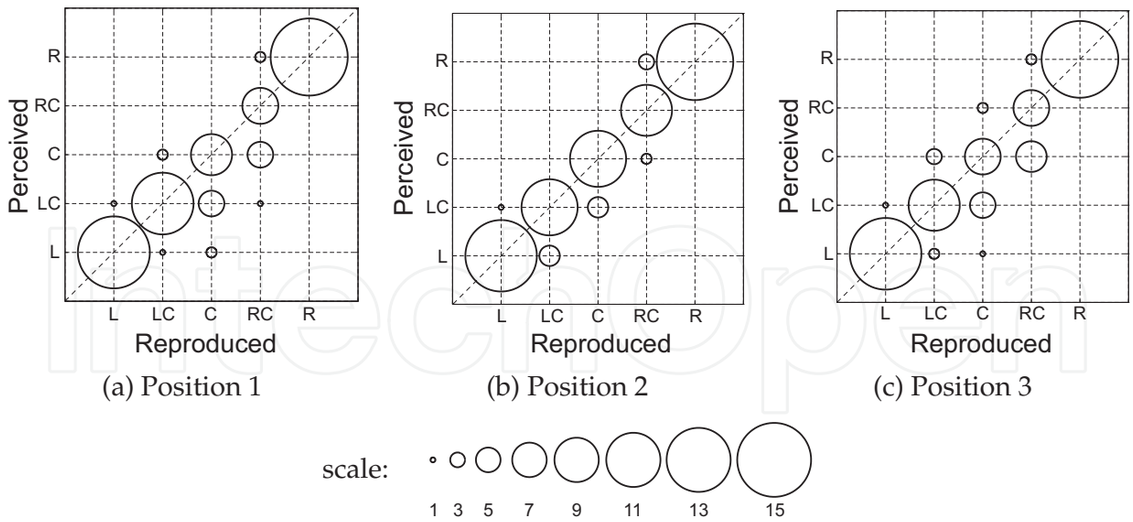


Fig. 10. Results of the subjective experiments.

4.2 Reproduced versus Perceived sound images

The data compiled from the subjective tests is shown in Fig. 10 as plots of *Reproduced* versus *Perceived* sound images. In the ideal case that all the reproduced sound images were perceived at the intended locations, a high correlation is visualized as plots of large circles with no sparsity from the diagonal. Although such ideal results were not obtained, note that the highest correlation between the parameters was achieved at Position 2 (Fig. 10(b)). Such result may be a priori expected since the sound delivered by the panel at that position is similar to that delivered by a standard stereo loudspeaker setup in terms of symmetry. At the lateral Positions 1 and 3, the subjects evaluated the sound images with more confusion which is reflected with some degree of sparsity in the plots of Figs. 10(a) and (c), but yet achieving significant level of correlation. Moreover, it is interesting to note the similarity of the correlation patterns of Figs(a) and (c) which implies that listeners at those positions were able to perceive similar sound images.

5. Example applications: Multichannel auditory displays for large screens

One of the challenges of immersive teleconference systems is to reproduce at the local space, the acoustic (and visual) cues from the remote meeting room allowing the users to maintain the sense of presence and natural interaction among them. For such a purpose, it is important to provide the local users with positional agreement between what they see and what they hear. In other words, it is desired that the speech of a remote speaker is perceived as coming out from (nearby) the image of his/her face on the screen. Aiming such problem, this section introduces two examples of interactive applications that implement multichannel auditory displays using the L-like loudspeakers to provide realistic sound reproduction on large display panels in the context of teleconferencing.

5.1 Single-sound image localization with real-time talker tracking

Fig. 11 of the first application example, presents a multichannel audio system capable of rendering a remote user's voice at the image of his face which is being tracked in real-time by video cameras. At the remote side, the monaural signal of the speech of a speaker (original

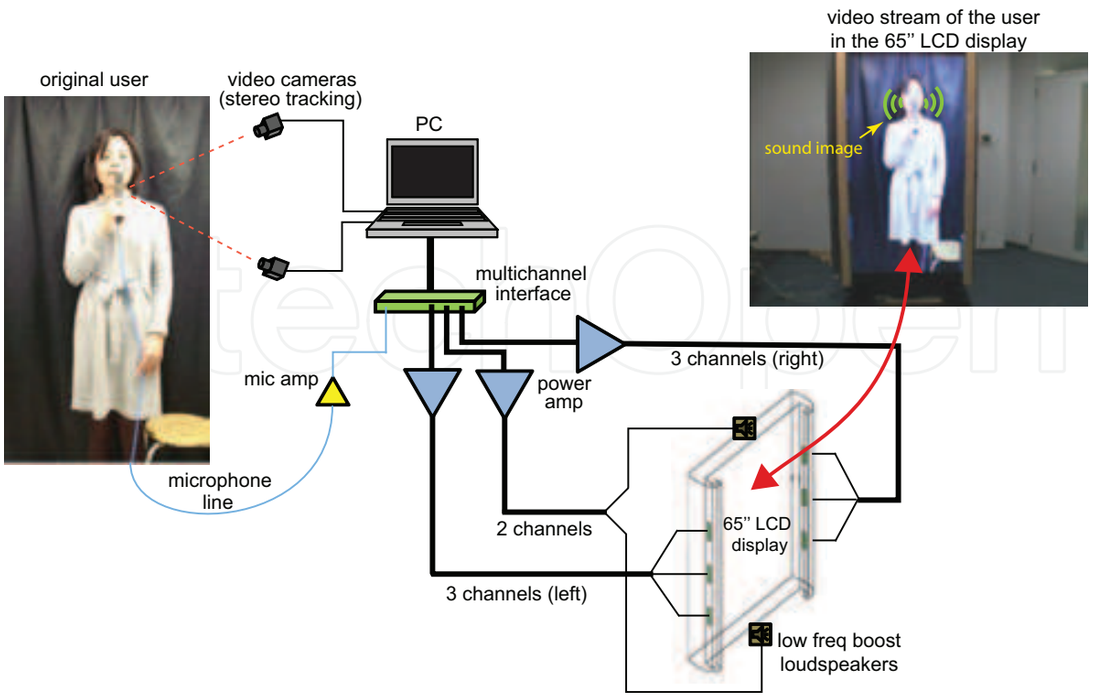


Fig. 11. A multichannel (8 channels) audio system for a 65-inch LCD display, combined with stereo video cameras for real-time talker tracking.

user, Fig.11) is acquired by a microphone on which a visual marker was installed. The position of the marker is constantly estimated and tracked by a set of video cameras. This simple video tracking system assumes that the speaker holds the microphone close to his mouth when speaking, thus, the origin of the sound source can be inferred. Note that for the purpose of demonstration of the auditory display, this basic implementation works, but alternatively it can be replaced by current robust face tracking algorithms to improve the localization accuracy and possibly provide a hands-free interface.

In the local room (top-right picture of Fig. 11), while the video of the remote user is being streamed to a 65-inch LCD screen, the audio is being output through the 6-channel loudspeakers attached to the screen panel. In fact, the 65-inch display used in this real-time interactive application is the prototype model of Fig. 4(a) plus two loudspeakers at the top and bottom to enforce the low frequency contents. Therefore, the signal to drive these booster loudspeakers is obtained by simply lowpass filtering (cut off above 700 Hz) the monaural source signal of the microphone. As for the sound image on the surface of the display, once the position of the sound source (i.e. the face of the speaker) has been acquired by the video cameras, the coordinate information is used to interpolate the sound image (left and right, and up/down), thus, the effect of a moving sound source is simulated by panning the monaural source signal among the six lateral channels in a similar way as described in section 4.1. The final effect is a sound image that moves together with the streaming video of a remote user, providing a realistic sense of presence for a spectator in the local end.

5.2 Sound positioning in a multi-screen teleconference room

The second application example is an implementation of an auditory display to render a remote sound source on the large displays of an immersive teleconference/collaboration room

known as t-Room (see ref. to NTT CS). In its current development stage, various users at different locations can participate simultaneously in a meeting by sharing a *common virtual space* recreated by the local t-Room in which each of them is physically present. Other users can also take part of the meeting by connecting through a mobile device such a note PC. In order to participate in a meeting, a user requires only the same interfaces needed for standard video chat through internet: a web camera, and a head set (microphone and earphones). In Fig. 12 (right lower corner), a remote user is having a discussion from his note PC with attendees of a meeting inside a t-Room unit (left upper corner). Moreover, the graphic interface in the laptop is capable of providing full-body view of the t-Room participants through a 3D representation of the eight t-Room's decagonally aligned displays. Thus, the note PC user is allowed to navigate around the display panels to change his view angle, and with the head set, he can exchange audio information as in a normal full-duplex audio system. Inside t-Room, the local users have visual feedback of the remote user through a video window representing the note PC user's position. Thus, this video window can be moved to the remote user's will, and as the window moves around (and up/down) in the displays, the sound image of his voice also displaces accordingly. In this way, local users who are dispersed within the t-Room space are able to localize the remote user's position not only by visual but also by audible cues.

The reproduction of sound images over the 8 displays is achieved by a 64-channel loudspeaker system (8 channels per display). Each display is equipped with a loudspeaker array similar to that introduced in the previous section: 6 lateral channels plus 2 low frequency booster channels. As in the multichannel audio system with speaker tracking, the sound image of the laptop user is interpolated among the 64 channels by controlling the gain of those channels necessary to render a specific sound images as a function of the video window position. Non-involved channels are switched off at the corresponding moment. For this multichannel auditory display, the position of the speech source (laptop user) is not estimated by video cameras but it is readily known from the laptop's graphic interface used to navigate inside t-Room, i.e., the sound source (face of the user) is assumed to be nearby the center of the video window displayed at the t-Room side.

6. Potential impact of the sound image localization technology for large displays

As display technologies evolve, the future digital environment that surrounds us will be occupied with displays of diverse sizes playing a more ubiquitous role (Intille, 2002; McCarthy et al., 2001). In response to such rapid development, the sound image localization approach introduced in this chapter opens the possibility for a number of applications with different levels of interactivity. Some examples are discussed in what follows.

6.1 Supporting interactivity with positional acoustic cues

Recent ubiquitous computing environments that use multiple displays often output rich video contents. However, because the user's attentive capability is limited by his field of vision, user's attention management has become an issue of research (Vertegaal, 2003). Important information which is displayed on a screen out of the scope of the user's visual attention may just be missed or not realized on time. But on the other hand, since humans are able to accurately localize sound in a 360^o plane, auditory notifications represent an attractive alternative to deliver information (e.g. (Takao et al., 2002)). Let us consider the specific example of the video interactivity in t-Room. Users have reported discomfort when using

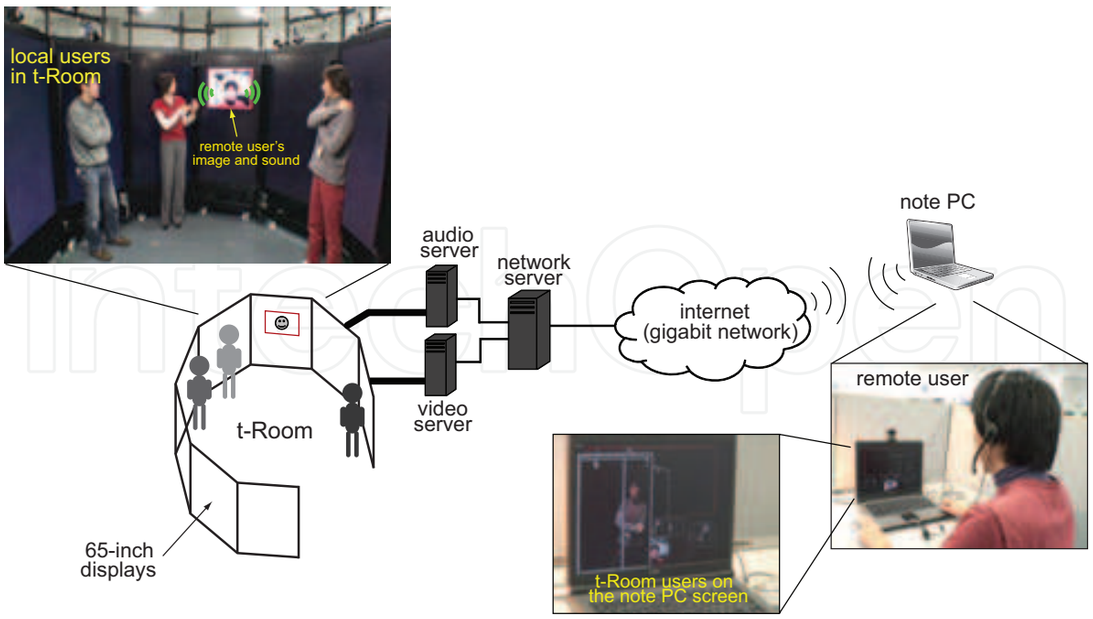


Fig. 12. Immersive teleconference room (t-Room) with a multichannel (64 channels) auditory display to render the sound images of remote participants on the surface of its large LCD displays.

the mouse pointer which is often visually lost among the eight surrounding large screens. This problem is even worsened as users are free to change their relative positions. In this case, with the loudspeaker system introduced in this chapter, it is possible to associate a subtle acoustic image positioned on the mouse pointer to facilitate its localization. Another example of a potential application is in public advertising where public interactive media systems with large displays have been already put in practice (Shinohara et al., 2007). Here, a sound spatialization system with a wide listening area can provide information on the spatial relationship among several advertisements.

6.2 Delivering information with positional sound as a property

In the field of Human Computer Interaction, there is an active research on the user-subconscious interactivity based on the premise that humans have the ability to subconsciously process information which is presented at the background of his attention. This idea has been widely used to build not only ambient video displays but also ambient auditory displays. For example, the whiteboard system of Wisneski et al. (1998) outputs an ambient sound to indicate the usage status of the whiteboard. Combination of musical sounds with the ambient background has been also explored (E. D. Mynatt & Ellis, 1998).

In an ambient display, the source information has to be appropriately mapped into the background in order to create a subtle representation in the ambient (Wisneski et al., 1998). For the case of an auditory ambient, features of the background information have been used to control audio parameters such as sound volume, musical rhythm, pitch and music gender. The controllable parameters can be further extended with a loudspeaker system that in addition allows us to position the sound icons according to information contents (e.g. depending on its relevance, the position and/or characteristics of the sound are changed).

6.3 Supporting position-dependent information

There are situations where it is desired to communicate specific information to a user depending on his position and/or orientation. This occurs usually in places where the users are free to move and approach contextual contents of his interest. For example, at event spaces such as museums, audio headsets are usually available with pre-recorded explanations which are automatically playbacked as the user approaches an exhibition booth. Sophisticated audio earphones with such features have been developed (T. Nishimura, 2004). However, from the auralization point of view, sound localization can be achieved only for the user who wears the headset. If a number of users within a specific listening field is considered, the L-like loudspeaker design offers the possibility to control the desired audible perimeter by optimizing the size of the L-like barriers to the target area and by controlling the radiated sound intensity. Thus, only users within the scope of the information panel listen to the sound images of the corresponding visual contents, while users out of that range remain undisturbed.

7. Conclusions

In this chapter, the issue of sound image localization with stereophonic audio has been addressed making emphasis on sound spatialization for applications which involve large flat displays. It was pointed out that the effect of precedence that occur with conventional stereo loudspeakers setups represents an impairment to achieve accurate localization of sound images over a wide listening area. Furthermore, some of the approaches dealing with this problem were enumerated. The list of the survey was extended with the introduction of a novel loudspeaker design targeting the sound image localization on flat display panels. Compared to existent techniques, the proposed design aims to achieve expansion of the listening area by mechanically altering the radiated sound field through the attachment of L-like rigid barriers and a counter-fire positioning of the loudspeaker drivers. Results from numerical simulations and experimental tests have shown that the insertion of the rigid barriers effectively aids to redirect the sound field to the desired space. The results also exposed the drawbacks of the design, such as the dependency of its radiation pattern with the dimensions of the target display panel and the listening coverage. For such a reason, the dimensions of the L-like barriers have to be optimized for a particular application. The need for low-frequency reenforcement is another issue to take into account in applications where the intelligibility of the audio information (e.g. speech) is degraded. On the other hand, it is worth to remark that the simplicity of the design makes it easy to implement on any flat hard display panel.

To illustrate the use of the proposed loudspeaker design, two applications within the framework of immersive telepresence were presented: one, an audio system for a single 65-inch LCD panel combined with video cameras for real-time talker tracking, and another, a multichannel auditory display for an immersive teleconference system. Finally, the potentiality of the proposed design was highlighted in terms of sound spatialization for human-computer interfaces in various multimedia scenarios.

8. References

Aoki, S. & Koizumi, N. (1987). Expansion of listening area with good localization in audio conferencing, *ICASSP '87*, Dallas TX, USA.

Bauer, B. B. (1960). Broadening the area of stereophonic perception, *J. Audio Eng. Soc.* 8(2): 91–94.

Berkhout, A. J., de Vries, D. & Vogel, P. (1993). Acoustic control by wave field synthesis, *J. Acoustical Soc. of Am.* 93(5): 2764–2778.

Ciskowski, C. & Brebbia, C. (1991). *Boundary Element Methods in Acoustics*, Elsevier, London.

Davis, M. F. (1987). Loudspeaker systems with optimized wide-listening-area imaging, *J. Audio Eng. Soc.* 35(11): 888–896.

E. D. Mynatt, M. Back, R. W. M. B. & Ellis, J. (1998). Designing audio aura, *Proc. of SIGCHI Conf. on Human Factors in Computing Systems*, Los Angeles, US.

Estorff, O. (2000). *Boundary Elements in Acoustics, Advances and Applications*, WIT Press, Southampton.

Gardner, M. B. (1968). Historical background of the haas and/or precedence effect, *J. Acoustical Soc. of Am.* 43(6): 1243–1248.

Gardner, W. G. (1997). *3-D Audio using loudspeakers*, PhD thesis.

Intille, S. (2002). Change blind information display for ubiquitous computing environments, *Proc. of Ubicomp2002*, Göterborg, Sweden, pp. 91–106.

Kates, J. M. (1980). Optimum loudspeaker directional patterns, *J. Audio Eng. Soc.* 28(11): 787–794.

Kim, S.-M. & Wang, S. (2003). A wiener filter approach to the binaural reproduction of stereo sound, *J. Acoustical Soc. of Am.* 114(6): 3179–3188.

Kyriakakis, C., Holman, T., Lim, J.-S., Hong, H. & Neven, H. (1998). Signal processing, acoustics, and psychoacoustics for high quality desktop audio, *J. Visual Com. and Image Represenation* 9(1): 51–61.

Litovsky, R. Y., Colubrn, H. S., Yost, W. A. & Guzman, S. J. (1999). The precedence effect, *J. Acoustical Soc. of Am.* 106(4): 1633–1654.

McCarthy, J., Costa, T. & Liongosari, E. (2001). Unicast, outcast & groupcast: Toward ubiquitous, peripheral displays, *Proc. of Ubicomp2001*, Atlanta, US, pp. 331–345.

Melchoir, F., Brix, S., Sporer, T., Roder, T. & Klehs, B. (2003). Wave field synthesis in combination with 2D video projection, *24th AES Int. Conf. Multichannel Audio, The New Reality*, Alberta, Canada.

Merchel, S. & Groth, S. (2009). Analysis and implementation of a stereophonic play back system for adjusting the "sweet spot" to the listener's position, *126th Conv. of the Audio Eng. Soc.*, Munich, Germany.

NTT CS The future telephone: *t-Room*, NTT Communication Science Labs. http://www.mirainodenwa.com/e_index.html.

Rakerd, B. (1986). Localization of sound in rooms, III: Onset and duration effects, *J. Acoustical Soc. of Am.* 80(6): 1695–1706.

Ródenas, J. A., Aarts, R. M. & Janssen, A. J. E. M. (2003). Derivation of an optimal directivity pattern for sweet spot widening in stereo sound reproduction, *J. Acoustical Soc. of Am.* 113(1): 267–278.

Seybert, A., Cheng, C. & Wu, T. (1990). The resolution of coupled interior/exterior acoustic problems using boundary element method, *J. Acoustical Soc. of Am.* 88(3): 1612–1618.

Shinohara, A., Tomita, J., Kihara, T., Nakajima, S. & Ogawa, K. (2007). A huge screen interactive public media system: mirai-tube, *Proc. of 2th international Conference on Human-Computer interaction: interaction Platforms and Techniques*, Beijin, China, pp. 936–945.

T. Nishimura, Y. Nakamura, H. I. H. N. (2004). System design of event space information support utilizing cobits, *Proc. of Distributed Computing Systems Wrokshops*, Tokyo, Japan, pp. 384–387.

Takao, H., Sakai, K., Osufi, J. & Ishii, H. (2002). Acoustic user interface (aui) for the auditory displays, *Communications of the ACM* 23(1-2): 65–73.

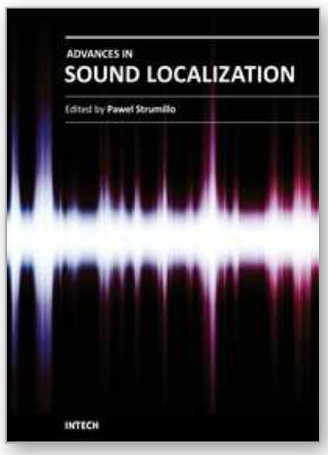
Vertegaal, R. (2003). Attentive user interfaces, *Communications of the ACM* 46(3): 30–33.

Werner, P. J. & Boone, M. M. (2003). Application of wave field synthesis in life-size videoconferencing, *114th Conv. of the Audio Eng. Soc.*, Amsterdam, The Netherlands.

Wisneski, C., Ishii, H. & Dahley, A. (1998). Ambient displays: Turning architectural space into an interface between people and digital information, *Proc. of Int. Workshop on Cooperative Buildings*, Darmstadt, Germany, pp. 22–32.

Wu, T. (2000). *Boundary Element Acoustics, Fundamentals and Computer Codes*, WIT Press, Southampton.

IntechOpen



Advances in Sound Localization
Edited by Dr. Pawel Strumillo

ISBN 978-953-307-224-1
Hard cover, 590 pages
Publisher InTech
Published online 11, April, 2011
Published in print edition April, 2011

Sound source localization is an important research field that has attracted researchers' efforts from many technical and biomedical sciences. Sound source localization (SSL) is defined as the determination of the direction from a receiver, but also includes the distance from it. Because of the wave nature of sound propagation, phenomena such as refraction, diffraction, diffusion, reflection, reverberation and interference occur. The wide spectrum of sound frequencies that range from infrasounds through acoustic sounds to ultrasounds, also introduces difficulties, as different spectrum components have different penetration properties through the medium. Consequently, SSL is a complex computation problem and development of robust sound localization techniques calls for different approaches, including multisensor schemes, null-steering beamforming and time-difference arrival techniques. The book offers a rich source of valuable material on advances on SSL techniques and their applications that should appeal to researches representing diverse engineering and scientific disciplines.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

Gabriel Pablo Nava, Yoshinari Shirai, Kaji Katsuhiko, Masafumi Matsuda, Keiji Hirata and Shigemi Aoyagi (2011). Sound Image Localization on Flat Display Panels, Advances in Sound Localization, Dr. Pawel Strumillo (Ed.), ISBN: 978-953-307-224-1, InTech, Available from: http://www.intechopen.com/books/advances-in-sound-localization/sound-image-localization-on-flat-display-panels

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen