CONVENTIONAL AND DISTRIBUTED MODE LOUDSPEAKER ARRAYS FOR THE APPLICATION OF WAVE-FIELD SYNTHESIS TO VIDEOCONFERENCE

Jose J. Lopez, Basilio Pueo and Maximo Cobos

Institute for Telecommunications and Multimedia Applications (iTeAM)
Technical University of Valencia, Spain

ABSTRACT

This paper presents two prototypes for the application of Wave-Field Synthesis (WFS) to high realism 3D videoconference systems based on the concept of virtual window. To provide the acoustic window, we have chosen WFS as the spatial sound rendering technique that generates a very realistic sound field using loudspeaker arrays. On one hand, a conventional cone loudspeaker array has been employed in conjunction with an autostereoscopic plasma display. On the other hand, a multiexcited Distributed Mode Loudspeaker (DML) in the form of Multiactuator Panel (MAP) has been also employed as an alternative technology to the cone loudspeakers. Because of its low visual profile and the fact that the vibration is low enough to be imperceptible to the human eye; they can be integrated into a room interior and simultaneously used as projection screens. Until now, the small size of current prototypes did not allow its use for big immersive environments, but in this paper we present one of the largest MAP reported.

Index Terms— WFS, DML, MAP, videoconference

1. INTRODUCTION

Videoconference systems have been around the market for long time. Their more ambitious aim has always been to avoid the need of having physical presence of people for carrying out meetings. However, their impact in the market has not been as important as many people expected. The reason for that is, essentially, that the sense of realism was far away from the expected. In order to improve this, the main efforts actually are focused on making the feeling of being there as real as possible. The goal is to achieve a video screen that appears to be a virtual window to the other side of the conference.

The concept of virtual windows should include image and sound. In both cases, the human perception system is able to obtain a 3D sense of the space using two eyes and two ears. That is why an acoustic window should not only provide a realistic image perception of the space, but spatial sound as well.

Two main technologies are usually employed to produce stereoscopic images: the systems where the user must wear special glasses (polarized, shuttered or anaglyph), and the autostereoscopic displays that provide 3D perception without the need for special glasses or other headgear, [1]. In this paper we will present and analyze two types of prototypes combining 3D image and spatial sound.

The simplest and best known method for providing spatial sound is stereo, which is able to position a source in the space using a pair of loudspeakers. However, only a listener located in a middle position between the loudspeakers is able to localize correctly the source, otherwise, the localization accuracy is severely degraded. On the other hand, multichannel surround sound systems (5.1, 6.1, 7.1), well established in the cinema industry, are not suitable for videoconference. This is due to the fact that their main objective is the reproduction of special effects in movies, and the rear loudspeakers would not add any significant contribution in a meeting. In [2], Wave-Field Synthesis was proposed as a spatial sound system for videoconference. The extension of the sweet spot offered by WFS is completely suitable for live-size videoconference with multiple participants.

This paper is structured as follows: in Section 2, an introduction to WFS is reported and in Section 3, the WFS prototype for autostereoscopic displays is described. In Section 4, a review of DML technology and its multiexcited extension for WFS, known as MAP, is carried out. In Section 5, a large MAP prototype which is simultaneously used as WFS system and projection screen is presented. Finally, the conclusions of our work are discussed in Sect. 6.

2. WAVE FIELD SYNTHESIS

Wave Field Synthesis (WFS) is a spatial reproduction technique that, by analogy to holography and in base to the Huygens principle, reproduces an acoustic field inside a volume from the stored signals recorded in a given surface. Huygens principle tells that the wave front radiated by a source behaves like a distribution of sources that are in the wave front, named secondary sources. WFS was first proposed with application to 3D sound by Berkhout [3]. The synthetic wave front is created by loudspeaker arrays that substitute the individual secondary sources. The ideal situation would be when an area, which is completely surrounded by loudspeakers, is fed with signals that create a volumetric velocity proportional to the particle velocity normal component of the original wave front.

The main advantage of these systems is that the acoustic scene has no sweet spot. When listeners move inside the listening area, the sound pressure level changes also in a realistic way according to its relative position to the virtual source, fig. 1. However, the recreation of a true natural wave field can be fulfilled with certain restrictions. Huygens principle needs to be discretized in practice, which means that an infinite continuous secondary source distribution is replaced by a number of finite arrays of equidistant discrete loudspeakers. The lack of continuity leads to a maximum usable frequency, known as spatial aliasing frequency, whereas the finiteness of the array causes some truncation effects. These loudspeaker arrays can be manufactured with typical dynamic transducers or by placing several exciters in a single DML. These approaches are introduced in the next sections.

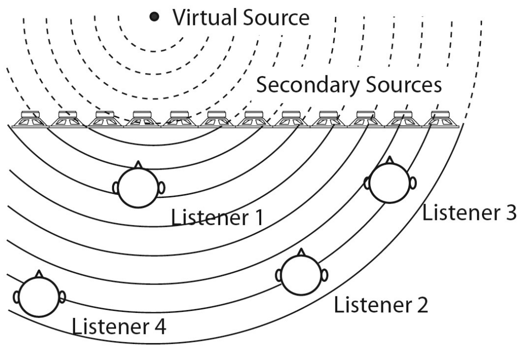

Fig. 1. Several listeners in an extended area perceive a virtual source with spatial fidelity.

3. CONVENTIONAL LOUDSPEAKER ARRAY FOR AUTOSTEREOSCOPIC DISPLAY

Autostereoscopic displays provide 3D perception without the need for special glasses. They are available commercially in different sizes and, although they are not as big as current video projectors, it is common to find 42'' and 50'' models in the market. In this prototype, the Philips 42" 3DWOW autostereoscopic 3D display has been employed. The slanted multi-view lenticular lens technology affords full brightness and full contrast and allows multiple users to view 3D content at the same time, within a large comfort zone. Moreover, a true color representation is ensured by the lenticular lens technology.

Two arrays have been used simultaneously, one above the display and the other below in order to cause the sensation that the sound comes from the display itself, Fig. 2. Each array consists of 12 loudspeakers with a separation distance of 18 cm. The array is longer than the display, which allows reducing the edge effects at the end of the array and enables the possibility of including bigger displays in the future.

Fig. 2. Prototype combining two WFS arrays of cone loudspeaker with a 42" autostereoscopic screen.

The signals corresponding to each source and their exact position are required in WFS for rendering the acoustic wave field, Fig 1. As several speakers are usually found in a videoconference, it is needed to capture the sound of each of them independently and to estimate their positions. Next, we will propose two different systems for that purpose, although only the first one has been implemented in the prototype. Figure 3a shows how the sound of each participant is captured by means of spot-mics. In order to know their correct position and synthesize their voices in the correct location, computer vision techniques are utilized for detecting target faces with the same camera

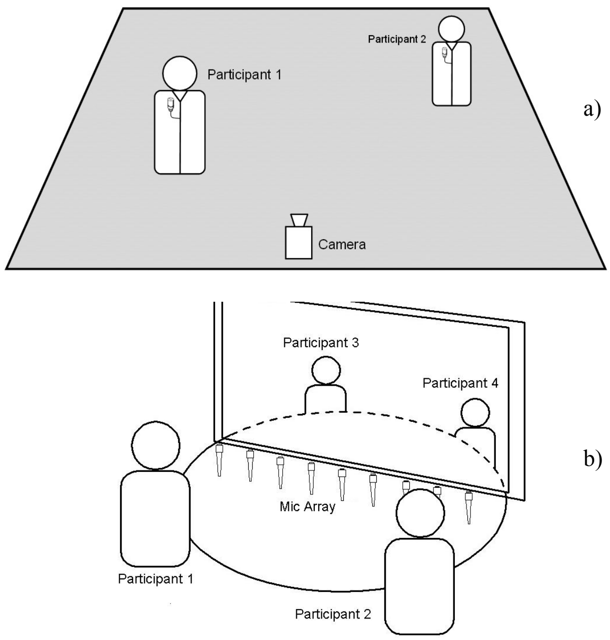

Fig. 3. Sound capture system, a) using spot-mics and the support of a videocamera and b) using a microphone array and beamforming.

used for video transmission. In the other case, Fig. 3b, a microphone array provides a way of separating and positioning the signals of the speakers. This second alternative using beamforming and DOA estimation will be implemented in a future prototype.

The system has been tested in real time showing an efficient performance and good sound localization in the face of each speaker. However, for improving the sense of realism and to exploit the potential of WFS, bigger screens will be required in the future. A 50’’ prototype will be tested soon and larger autostereospic displays will be available in the future.

4. DISTRIBUTED MODE LOUDSPEAKERS

The DML essentially consists of a thin, stiff panel that vibrates in a complex pattern over its entire surface by means of a electro-mechanic transducer called exciter. The exciter is normally a moving coil device, which is carefully positioned and designed to excite the natural resonant modal structure of the panel optimally. In Fig. 4, a graphical representation of a DML is presented, which shows panel, exciter and housing.

DMLs are panels of finite extent deploying bending waves. The DML relies on the optimization of its eigenmodes to produce a modal density that is sufficiently high to give the impression of a continuous spectrum [4]. The excitation of bending waves on panels results in sound radiation with distinct qualities with regard to the pistonic motion of typical dynamic loudspeakers. A traditional loudspeaker acts for the most part of its radiation as a phase coherent radiator, and thus, it has a correlated output. However, the uncorrelated output of a DML produces an omnidirectional directivity response over the major part of the audio frequency band [5]. In addition to this, DML sources produce reflections that are less correlated to the direct sound than those radiated from piston sources and thus, constructive and destructive interference of sound is minimized [6] .

One of the practical advantages of DMLs is their ease to mount directly on the wall surface. Besides, they are light-weight loudspeakers with a small back housing that can be get unnoticed as part of the decoration. Since the panel surface can be large and the vibration is low enough to be imperceptible to the human eye, they can be integrated into a room interior and simultaneously used as projection screens [2]. In this way, image and sound are fully integrated for multimedia applications. Furthermore, the cost of DMLs is generally lower than that of dynamic loudspeakers on baffles. These features make DMLs very suitable for WFS reproduction, which will be introduced in the next section.

Encouraged by the positive results on sound localization, the applicability of single-exciter DMLs for WFS reproduction was tested for the first time in [7], reporting that individual panels reconstructed the wave field

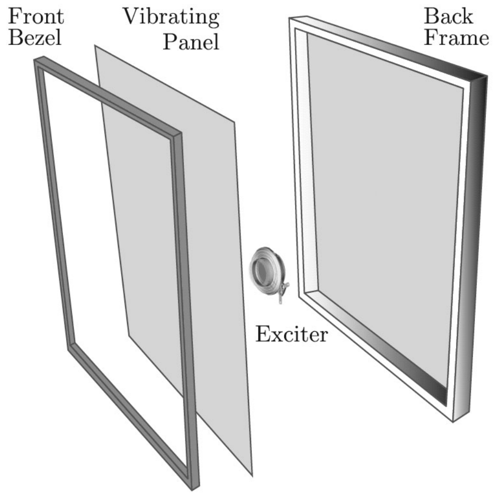

Fig. 4. Block diagram of a Distributed Mode Loudspeakers with only one exciter (wiring is omitted).

correctly. However, the secondary sources spacing required by the WFS algorithm to acquire a reasonable useful bandwidth, forced the size of panels to be very low. This conferred DMLs weak bass response due to the lack of excited modes in the low frequency region. In [8], Boone proposed to extend the DML technology to a panel with multiple exciters, each driven with a different signal. Such a configuration would act as a WFS array if every exciter on the panel would excite only a small part of the panel around the exciter position, which was experimentally confirmed in [9]. Since exciters in a DML operate by converting electrical signals into mechanical movement which is applied to the panel, these panels are also known in the technical literature as Multiactuator Panels (MAP). There are some benefits for MAPs to be used in WFS reproduction. They can easily be integrated into a living room because of its low visual profile. Furthermore, the vibration of the surface is almost negligible so that it can be used as projection screens [10].

5. LARGE MAP FOR WFS AND PROJECTION

The well-known 3D displays that require the viewer to wear special glasses present two different images in the same display plane. The glasses select which of the two images is visible to each of the viewer’s eyes. Technologies for this include polarization, shuttering or anaglyph. In this prototype we selected the shuttering technology were a double frame-rate is employed (left and right eye emitted alternatively) in combination with a shutter glasses that block the opposite image. The projector employed was an InFocus DepthQ working at 120 Hz with DLP technology.

For the projection screen a large MAP was especially designed and built, fig. 4, to meet the demands of immersive audio applications. For that purpose, it includes a horizontal line of exciters composed of 13 exciters with 18 cm spacing, presenting an aliasing frequency of approximately 1 kHz.

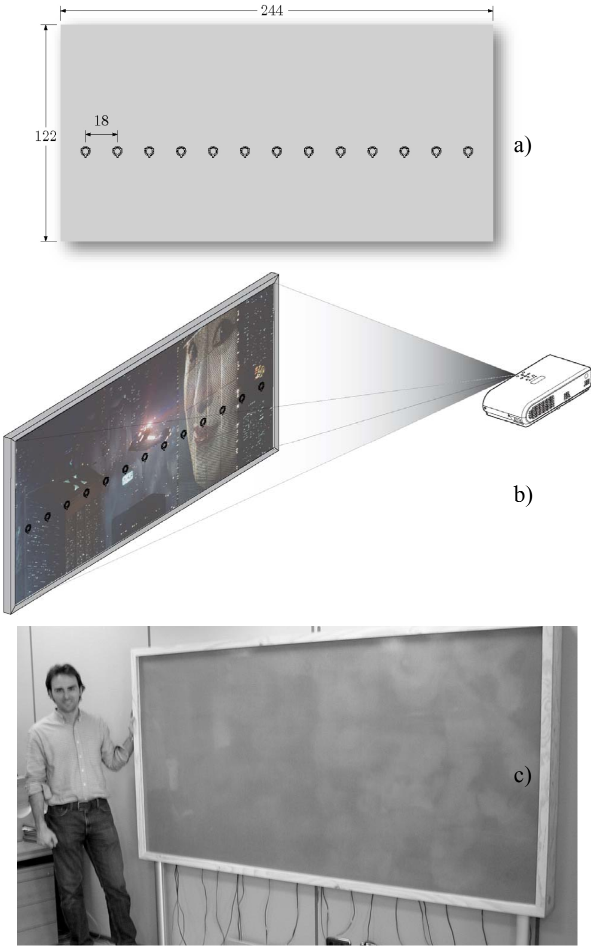

Fig. 5. Large MAP, a) block diagram and measures, b) employment in conjunction with a projector, c) photograph of the resulting prototype panel assembled and ready for use.

The panel is a sandwich of polyester film bonded to an impregnated paper honeycomb 5 mm thick using a thermoplastic adhesive (cell size = 4.8 mm). Its bending rigidity is 4.23 and 2.63 Nm in the x and y directions respectively and has an areal density of 0.51 kg/m2. Due to its size, frequencies until 100 Hz can be reproduced successfully. More about the acoustic performance and audio quality of this panel was analyzed and previously presented by the authors in [11].

6. CONCLUSIONS

Two prototypes for the application of Wave-Field Synthesis to high realism 3D videoconference systems have been presented, built and tested successfully. They provide and acoustic window that generates a very realistic sound field using loudspeaker arrays. The first one uses conventional cone loudspeaker arrays and has been employed in conjunction with an autostereoscopic plasma display. The use of two arrays, one above the display and the other below, has shown a good performance, causing the sensation that the sound comes from the display itself. However, the small size of the screen, only 42” did not take profit of the full potential of WFS. On the other hand, a multiexcited DML in the form of MAP has been also employed as an alternative technology to the cone loudspeakers. In this case, the big size of the screen in conjunction with the realistic sound provided by the WFS produced a better sense of immersion. The main improvement with respect to smaller MAPs laid on the useful bandwidth for on and off axis radiation. However, the use of supporting subwoofers is still advisable for a natural sound reproduction. Until now, the small size of current prototypes did not allow its use for real and big immersive environments, but in this paper we presented one of the largest MAP reported.

7. ACKNOWLEDGEMENTS

This work has been partially supported by the Spanish Ministry of Education and Science under project TEC2006-13883-C04-01 and by the by the Spanish Administration agency CDTI, under project CENIT-VISION 2007-1007.

8. REFERENCES

[1] N.A. Dodgson, “Autostereoscopic 3D Displays”, *Computer*, vol 38, no. 8, pp. 31 – 36, Aug. 2008
[2] W. de Brujin and M. Boone, “Application of Wave-Field Synthesis in life-size videoconferencing,” in *114th Conv. Audio Eng. Soc*., Amsterdam, Netherlands, Mar. 2003.
[3] A. Berkhout “A holographic approach to acoustic control,” *J. Audio Eng. Soc*., vol.36, no.12, pp.977–995, 1988.
[4] J. W. Panzer and N. Harris, “Distributed-mode loudspeaker radiation simulation,” in *105th Conv. Audio Eng.* Soc., no. 4783, San Francisco, USA, Sept. 1998.
[5] J. A. Angus, “Distributed mode loudspeaker polar patterns,” in *107th Conv. Audio Eng. Soc*., no. 5065, New York, USA, Sept. 1999.
[6] N. Harris, V. Gontcharov, and M. Hawksford. “Measurement and simulation results comparing the acoustic of various direct radiators in the presence of a dominant specular reflection”. in *109th Conv. Audio Eng. Soc.*, number 5215, Los Angeles, USA, Sept. 2000
[7] M. Boone and W. de Brujin, “On the applicability of distributed mode loudspeaker panels for Wave Field Synthesis based sound reproduction,” in *108th Conv. Audio Eng. Soc.*, Paris, France, Feb. 2000.
[8] M. Boone, “Multi-actuator panels (MAPs) as loudspeaker arrays for Wave Field Synthesis,” *J. Audio Eng. Soc*., vol. 52, no. 7–8, pp. 712–723, July 2004.
[9] M. Kuster, D. de Vries, D. Beer, and S. Brik, “Structural and acoustic analysis of multiactuator panels,” *J. Audio Eng. Soc*., vol. 54, no. 11, pp. 1065–1076, Nov. 2006.
[10] G. Thiele, H. Wittek, and M. Reisinger, “Potential Wavefield Synthesis applications in the multichannel stereophonic world,” in *AES 24th International Conf. on Multichannel Audio*, Banff, Canada, June 2003.
[11] B. Pueo, J. Escolano, J.J. López and G. Ramos, “On Large Multiactuator Panels for WFS Applications,” in *124th Conv. Audio Eng. Soc*., Amsterdam, May. 2008.