A New Digital Audio System Design to Cope with High Audio Scene (HAS) Encoding Technology for All Round Applications

IP Ralph W.L.
Department of Mechanical Engineering
The University of Hong Kong
ralphip@hku.hk

IP Percy W.H.
Technology Innovation Ltd.
Cable TV Tower
info@tech-inno.com

Abstract

A brand new audio signal processing technology namely "High Audio Scene (HAS)" has been developed to compensate the inadequacies of Dolby Digital formats that widely used in daily life audio processing systems. The challenging of using Dolby systems is the expected acoustic conditions including building environment and audio configuration facilities should be satisfied. The coded multiple channels signal from HAS encoders are from ordinary stereo sources without the expensive set up in source recording. The quality of the regenerated sound is compatible to Dolby Surround EX and less acoustically controlled environment is needed. The audio systems configured either in 2.1-, 3.2- or 5.2-channel HAS format can generate a VAST and STRONG VOCAL audio effect similar to the hearing in a natural environment. However, less preserved power is need because of using interference characteristics can intensify the sound sources to many times. The technology is expected to be widely applied for different environment including classrooms, dwells, corridors, auditoriums, cinemas, and concert halls without much building re-revolution work.

1. Introduction

Loudspeakers commonly refer to the type of audio transducers using electro-mechanical principles to generate a sound wave by piston oscillation effects. The generated oscillations will be transformed into audio waves under different environmental settings before the reaching to the audiences. The most typical settings that always concerned have included closed areas, open spaces, reflective surfaces, absorption surfaces and geometrical shape of the acoustic environment in the reproduction of high fidelity (Hi-fi) [1] sound wave. The propagating direction of the audio waves generated from loudspeakers is coaxial to the voice coil oscillation axis, and the wave propagation envelop is governed by the conical shape or flat diaphragm of the speakers, thus, the active zone of the sound waves generated from the speaker types is always limited by the constraints of direction and envelop following Stokes' law [2, 3] of attenuation.

The other known physical phenomena in sound propagation also play equally important roles for audio engineers to design the high-end audio systems. The major concerned wave propagation properties have included echo, reflection/adsorption and interference. However, few available audio system designs can fully avoid the negative effects in wave propagation for the task of reproducing the Hi-fi results from the original sound sources. The commonly used solutions in the regeneration of a Hi-fi sound output from high-end audio systems are to enhance the environment. For example, mineral wool, glass wool and micro perforated plates are used to make acoustic panels for auditoriums, halls, lecture theatres, libraries, meeting rooms and recording rooms to reduce echoes and uncontrollable reflections. Many strategies for building interior structure design have been investigated for the purposes of providing the best environment in the appreciation of musical performance and listening of human speeches. Therefore, architectural acoustics has been a hot scientific topic for noise control in the building design engineering. The demands of applying architectural acoustics are not only for the design of opera houses and concert halls. More widely, noise suppression is critical in the design of multiple unit dwellings, domestic apartments, teaching classrooms and business premises that may affect significantly of noise from the nearby pubs, bars and road-side traffic. As a result, investment for the establishing of an acoustic environment is always larger than the audio equipment cost for Hi-fi sound listening is expected.

2. Dolby Digital technology

The idea of Dolby Digital [4] is to create up to 6 or 8 discrete channels of sound wave. This sound processing technique was initially used in film industry but now has been popularized in audio/video systems including CD, LD, VCD, DVD and Blue-ray players. Claimed results of audio enhancement using Dolby Digital processing techniques are the feeling of rich, enveloping surround sound, and adding realism to the sound sources. The surround sound encoding/decoding technology is the reigning standard in Dolby Digital processors that compels up to 7.1-channel soundtracks for an enveloping audio experience to the audiences.

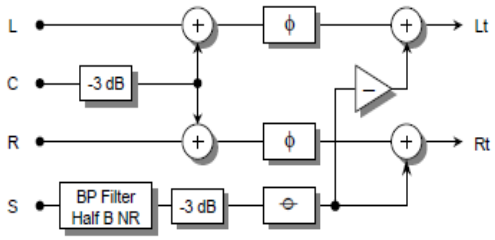
Source:www.dolby.com/us/en/consumer/technology/home-theater/dolby-digital.html

Figure 1. The speaker arrangement for Dolby 5.1-channel surround systems

In a 5.1-channel system as shown in Figure 1, the most elaborate mode in common usage involves five channels for normal-range speakers (20 Hz – 20,000 Hz) including right front, center, left front, right rear and left rear, and one channel (20 Hz – 120 Hz allotted audio) for a subwoofer driven low-frequency effects (LFE). In the playback of films/video images, the left and right channels primarily reproduce musical swells and accurately communicate movement across the screen. The centre channel mainly reproduces dialogue and human speeches at the front central position. The two surround channels (Left rear surround and Right rear surround) produce sound envelops for the creation of surrounding feeling. The LFE channel— ".1" in 5.1-channel system delivers deep, powerful bass effects. Thus, audiences will be engaged in a surround sound perception environment when experiencing Dolby Digital technology.

2.1. Technology of Dolby surround encoding

In the coding process, Dolby Digital divides the source spectrum into narrow frequency bands using mathematical models derived from the characteristics of the ear, and analyzes each band to determine the audibility of those signals. In the producing of the surround audio effect in the 5.1-channel format, Dolby Surround [5] encoding processors record 4-channel (Left, Centre, Right and Surround) sound sources to produce a 2-channel (Left total, Right total) surround soundtracks as illustrated in Figure 2.



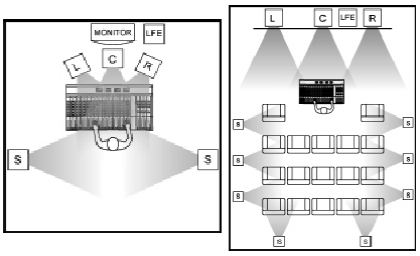
L: Left, C: Centre, R: Right, S: Surround, Lt: Left total, Rt: Right total

Source: 5.1-Channel Production Guidelines, Dolby System

Figure 2. Block diagram of the 5.1-channel Dolby Surround encoding algorithms

2.2. Recording and playing of audio sources coded in Dolby Surround format

Due to multiple channels are needed for audio encoding in the 5.1- or 7.1-channel Dolby process, the required conditions in this production method are not only for artists and audio engineers, the design and building of recording facilities are equally important. The facilities have included the layout, size, shape and acoustic design of recording rooms. Two typical design of a Dolby Surround recording room are illustrated in Figure 3.

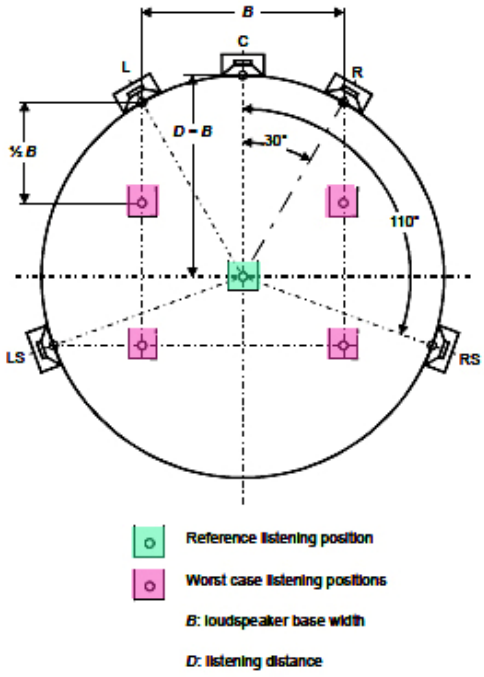


Source: 5.1-Channel Production Guidelines, Dolby System

Figure 3 Layout of Dolby Surround recording room design for (a) standard and (b) cinema configuration

In order to regenerate the sound effect accurate in the playback of soundtracks coded in 5.1- or 7.1-channel Dolby Surround format, a listening room layout designed for the best perception referring to the standard of ITU-R [6] is illustrated in Figure 4. Aside

from the alignment of sound source on the same horizontal plan, specific geometry of the speaker position is recommended. The Centre speaker directly in front, Left/Right speakers are 30° from center (forming a 60° angle) and the Surround speakers are at 110° off center.



Source: International Telecommunication Union

Figure 4. Recommended playback layout for 5.1-channel Dolby Surround audio based on the standard of ITU-R

3. High Audio Scene (HAS) technology

HAS is a new digital audio signal processing algorithm to create 2.1-, 3.2- and 5.2-channel digital sound data from stereo (2-channel) audio sources. Thus, no 4-channel input source is needed same as Dolby Surround encoding method illustrated in Figure 2. HAS can encode most of the stereo soundtracks no matters recorded in analog or digital format. One of an essential purposes for the development of HAS format is to fill the inadequacy of Dolby Digital 5.1- or 7.1-channel encoding for many stereo sounds hold in old fashion storage devices including audio cassette tapes, VHS, V-8 and Hi-8. On the other hand, HAS provides a relatively economical audio encoding method to avoid the necessity of using expensive recording facilities as shown in Figure 3. Figure 5 illustrates the algorithm of HAS for 5.2-channel encoding.

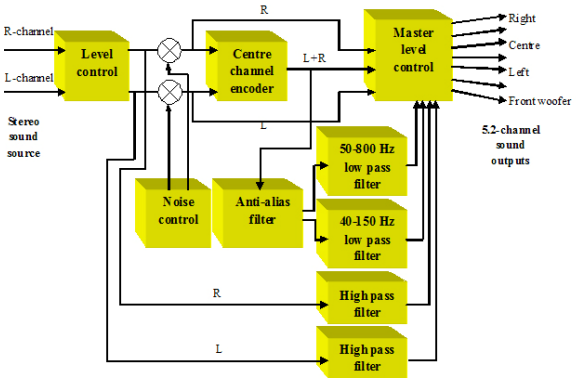


Figure 5. HAS encoding algorithms for 5.2-channel audio systems

HAS consists of five discrete full range main channels (L, C, R, L aux and R aux) and two band-limited LFE channels for added bass, i.e. " .2-channel". The five main channels deliver full frequency bandwidth from 80 Hz to 25 kHz, the two LFE channels deliver limited bandwidth from 50 to 800 Hz (front woofer) and from 10 to 150 Hz (rear subwoofer). In comparison with Dolby 5.1-channel surround format, HAS needs two high frequency auxiliary main channels (L/R aux) and an additional LFE channel. The frequency ranges of the two audio encoding systems are listed in Table 1.

Table 1. Bandwidth of different channel speakers for Dolby and HAS systems

Speaker channel	Dolby 5.1-channel surround	HAS 5.2-channel
Left, Right, Centre	80 Hz – 20 kHz	80 Hz – 20 kHz
L total, R total	80 Hz – 20 kHz	
Subwoofer	10 Hz – 120 Hz	
Front woofer		50 Hz – 800 Hz
Rear subwoofer		10 Hz – 150 Hz
L/R auxiliary		20 kHz – 25 kHz

Because of HAS uses two separated LFE channels in bass management and the new L/R auxiliary channel for high frequency signals regeneration, a more natural effect is therefore produced. A testing result has shown the output effect from a HAS 5.2-channel sound signals could be a distinguishable human voice standing out from a noisy background.

4. Technology development of HAS

The study presented here is an upfront research on the development of new digital audio encoding models to fill the inadequacy in the regeneration of natural, seamless, detail and subtleties of sound quality. It has no duplication tasks of enhancing the feeling of stereo, surrounding, cinematic experience, thrilling and exciting in the encoding of audio signals, because of a full range of Dolby product can provide these features. However, a high equipment cost together with a minimum acoustic environment is needed to achieve these special feeling. Due to inherent limitation of the chosen material in the design of electro-magnetic type piston speakers, only 1/3 of sound power can be delivered to the audiences. As a result, most of Hi-fi audio systems use high power loudspeakers to preserve a full range of sound spectrum. For example, a general purpose home theatre needs 50 - 100 W front and surround speakers, and 75 - 200 W woofers or sub-woofers. The total power preserved to a 5.1-channel home theatre system, e.g. SONY Blue-ray™ HTS [7] can be as large as 700 W that equivalent to 1/3 motor power of a domestic front loading washing machine. This is a critical issue to be improved in the design a green electrical product. Therefore, a key technical requirement in the development of the new audio system is to keep the minimum of input power, and maintain a compatible output power to other similar design. The tasks in the development of the HAS audio processing technology, and an implementation of the technology onto multiple channel audio systems will be discussed as follows.

4.1. Minimum input power

One of an important design feature of HAS audio systems is for all round applications including the implementing onto portable computers, e.g. laptops and notebooks. The power could be used for these electronic devices is usually limited to 5 W, thus, small size and low power speakers are needed. Figure 6 shows conceal type speakers are installed in notebook computers and internally connected to the sound driver cards.

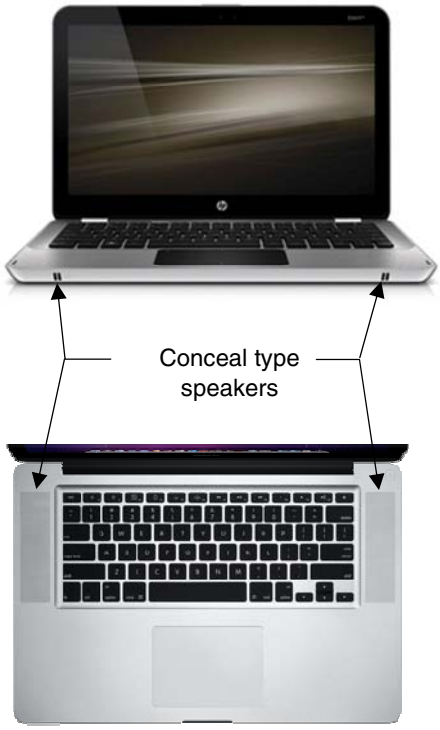


Figure 6. The conceal speakers installed on notebook computers

Most of these conceal speakers cannot provide a full range of sound spectrum that causes a poor results for music listening. A solution recommended in this research which also a special feature of the speaker design for HAS systems is to use the forced vibration principle for the enlargement of each sound channel. In the new design, two common piston type loudspeakers are used and mounted on a reflective surface as illustrated in Figure 7.

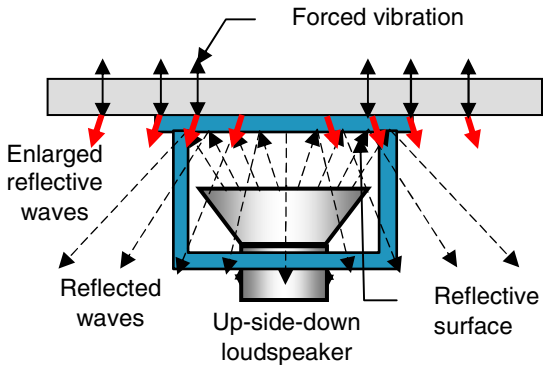


Figure 7. Enlargement of sound amplitude by force vibration method from reflective waves

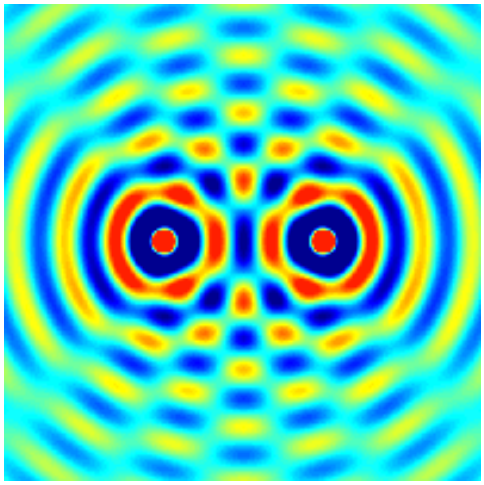
The reflective sound path design can magnify the sound source amplitude up to three times by the

principle of forced vibration. When the sound waves hit on the reflective surface, some of them are reflected and propagated, and some of them are conducted to the computer case and create a forced vibration result. The amount of vibration will depend upon properties of the case, such as materials, texture and surface roughness. In general speaking, glossy surface and hard material will produce higher enlarged reflective waves from vibration. When the best reflection is achieved, less output power from the speakers is required. The tested input power for the speakers used on notebook and laptop computers can be less than 3 W each. For this low power consumption design, the USB format can be used for both of signals and power transmission.

For the restricted reason of two conceal type speakers are only used on a notebook computer, a limit sound enhancement effect could be produced even the HAS encoding technique is applied for this special 2.0-channel audio system. However, the technology has given a minimum improvement of sound properties.

4.2. Intensify of outputs by interference

Using wave reflection properties and inducted force vibration to enlarge the wave propagation intensity could be a special feature of the new audio system design. The second investigated strategy for further intensify of the sound amplitude is to use the phenomenon of interference. Interference can be created from two wave sources traveling along different sound paths as illustrated in Figure 8.



Source: http://en.wikipedia.org/wiki/Interference

Figure 8. Patterns of the circular wave interference created from two point sources

The occurred interference can be constructive to magnify the sources, or destructive to diminish the sources. In order to create a constructive interference for the purpose of intensifying the outputs, the crucial factors of the source distance, phase angle and source frequency have been studied. Results from the measurement of a 2-speaker system that the loudness is near tripled when the distance between the speakers are from 0.6 m to 2 m. The testing results also show that interference can be achieved for the best positioning of the speakers according to their frequent brands. It is an important finding to enable HAS the ability to create the STRONG VOCAL effect in the encoding of human voices/speeches from an ordinary 2-channel audio source. Due to the interference effect can intensity the sound source, the expected speaker power is less than 5 W for a floor area of 15 m x 15 m. This system configuration can also be suitable for the environment without proper acoustic control, such as classrooms and general assembly halls.

4.3. Create a VAST effect by reflection and interference

A nature and wide perception feeling namely VAST is expected to be generated from the new audio wave encoding method. It is not the same as the Dolby Surround effect by creating an encirclement sound envelop to the audiences. The configuration of 3.2- and 5.2-channel systems adopting HAS encoding technology is mainly for the creation of VAST and strong vocal feeling for songs, speeches and music listening. The said feeling may not be created from the traditional 5.1-channel Dolby audio systems because they may have less interference from the surround speakers. The new system configuration is to place a woofer at the front of the audiences, and a subwoofer next to the rear walls as illustrated in Figure 9.

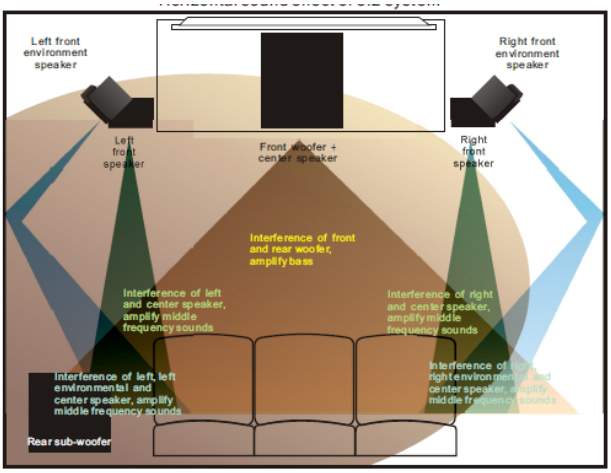


Figure 9. The sound envelop simulation for the 5.2-channel HAS audio system design

In the design, the reflected sound waves from the Left, Centre, and Right speakers will meet the sound waves created by the woofers and reflected from the rear walls. Interference could be occurred when a carefully positioning of the speakers and the woofers/subwoofers. The surround sound envelop generated from the traditional 5.1-channel Dolby systems is shown in Figure 10 for comparison.

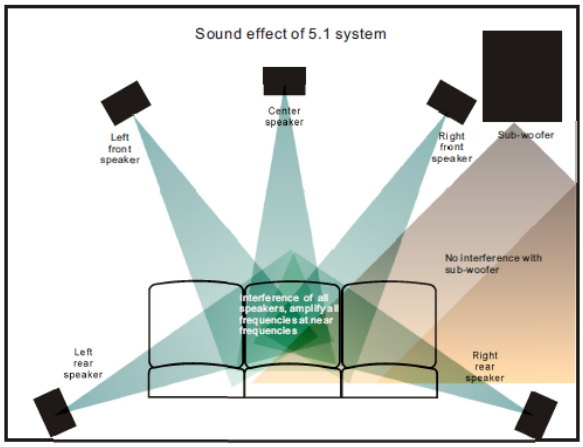


Figure 10. The sound envelop simulation for the 5.1-channel Dolby audio system design

The estimated speaker power for 3.2- and 5.2-channel configuration systems using HAS encoding algorithm could not be more than 10 W and 20 W respectively.

5. Generic algorithm development to cope with special environmental factors

The sound generation and propagation from the new audio system design is mainly by reflection and interference. Thus, some of inherence problems in the use of traditional Hi-fi units can be minimized, such as the golden position and acoustic facility as mentioned in Section 2. However, the new system design may not totally remove problems introduced from the environment, such as the room shape and wall surface conditions. It easily understandable that high reflection will come from the bare walls and irregular shape buildings may destruct regular reflection patterns. While the 5.2-channel systems are used for song playing or for film watching, the strong vocal effect for human voices may not be fully interpreted in a concrete wall room. It is because the highly reflective wall surface causes too much of echoes and sound distortion. In the development of model algorithms for 3.2- and 5.2-channel systems, several extreme causes should be considered to make the encoding technology to be more generic.

6. Conclusion

The sound sources encoded in Dolby format will create a limitation for the best position of Hi-fi perception results according the principle of ITU-R. In addition, the environment design and facility quality for recording and playback have also played important roles in the regeneration of good sound quality. High facility cost is also paid for the Hi-fi production environment for the reasons of controlling sound wave propagation problems. In opposite, the new developed sound encoding algorithm – HAS adopted the common wave properties of reflection and interference to assist the generation of VAST and STRONG VOCAL effects. The proposed system configuration either in 3.2- or 5.2-channel format can generate an audio effect similar to the hearing in a natural environment. In addition, the two channel encoding formats are customized for songs, films, operas and symphonies to regenerate and realism effects that can compatible with Dolby Digital Surround Ex. However, less preserved power is need for the audio systems configured under HAS format because of the interference characteristic can intensify the sources to many times.

References

[1] H.A. Hartley, “Audio Design Handbook”, *Gernsback Library*, New York: 1958.
[2] DIN 45500 High Fidelity Audio Equipment and Systems, Deutsches Institut für Normung.
[3] BS EN 61305 Household High-Fidelity Audio Equipment and Systems, BSI Standards, 1995.
[4] Dolby Laboratories, Inc. www.dolby.com/index.html

[5] 5.1-Channel Production Guideline, Issue 1, Dolby Laboratories Inc, 2000
[6] ITU, International Telecommunication Union, http://www.itu.int/ITU-R/index.asp?category=information&rlink=rhome&lang=en
[7] SONY, Blue-ray Disc™ Home Theatre System http://www.sonystyle.com/webapp/wcs/stores/servlet/ProductDisplay?catalogId=10551&storeId=10151&langId=-1&productId=8198552921665532064#specifications