

immersive 3D sound for cinema

Authors:

Toni Mateos, R&D Director at **imm sound**, Director of Audio Research at Barcelona Media, and Associate Lecturer at University Pompeu Fabra

Vicente López, Executive Chairman at **imm sound**, Vicepresident of Barcelona Media, and Professor of Computer Science at Universitat Pompeu Fabra

Alex Artigas, Marketing Manager at **imm sound**

Acknowledged contributions by:

Marc Orts, Re-recording Mixer

Oriol Tarragó, Sound Designer

Introduction

Cinema sound is on the verge of a revolution, possibly comparable to the transition from silent films to talkies in the late 20s. The new change is towards immersion. Up to now, the only benefit sound witnessed in the shift from analog to digital was a mere change of physical support, and the consequent elimination of lossy compression from cinema soundtracks. There is, however, a whole new world of possibilities that full exploitation of digitalization offers, in terms of listening experience, in terms of freedom in the creative processes and in terms of simplifying and rationalizing the workflow in the production-distribution-exhibition chain. This white paper describes a technology for immersive 3D sound production that breaks with the traditional proposals for simply adding more and more audio channels, and sets the ground for the immediate present and future of cinema sound.

Even with only two ears [1], our real life auditory experience is fully three-dimensional. We constantly hear, even unconsciously, sounds arriving from multiple directions, and we are capable of perceiving complex properties, like proximity, size and shape. Sound systems that fail in delivering this full experience often result in a spectator's lack of immersion, lack of believability, and absence of the illusion of being there.

The narrative, creative and immersive use of sound in films has therefore always been limited by the available technology. As an example, sounds could only be located in the screen during the decades of mono and stereo, a fact with obvious creative limitations. In this respect, the standardization of 5.1 surround systems over the last 20 years has only helped filling part of the hole; the two extra surround channels are normally redeemed secondary, and used only for limited purposes [2].

Proposals to overcome these aesthetical limitations by means of using more loudspeakers have existed for more than 20 years. Examples are the 7.1, 9.1, 10.1, 11.1, 13.1 layouts, and the notable proposals by great audio experts, like the 10.2 by THX's founder Tomlinson Holman (mid 90s [3]), and the 22.2 by K. Hamasaki, from the giant Japanese broadcaster NHK (early 2000s [4]).

All proposals for simply adding more and more audio channels have nonetheless earned little attention from the cinema industry. Without the corresponding advance in technology, such proposals become less and less scalable, as more channels imply more complexity and, therefore, less creativity and lengthier more expensive post-production.

imm sound's technology breaks complexity at its root, by eliminating the concept of audio channel. The technology allows film production, post-production and delivery to be 100% independent of the loudspeaker layouts where content is to be exhibited. **imm sound's** 3D soundtracks are always reproduced fully adapted to every cinema's specific layout, thus tackling all present and future configurations at once. With one single shot, all aforementioned proposed multi-loudspeaker layouts become part of **imm sound**.

On the other hand, this simple yet powerful paradigm opens an unlimited world of creative possibilities, by freeing sound designers and mixers from the hassle of creating while keeping in mind a large set of audio channels. Otherwise frequent processes like sending a stereo track to the top-front-left and mid-back-center channels are simply eradicated from sound design. Creative professionals need only concentrate on sound and space. It's that simple.

Ultimately, and most importantly, **imm sound's** technology allows the cinema industry to deliver compelling immersive experiences that rival real life.



The impact of immersive sound

Manuel Faria, CEO of the innovative sound design company Indigo sometimes explains this anecdote to stress the often undervalued impact of sound in the audience: "For a few weeks, cinemas in Portugal showed a 3 minute sound experience we created with immersive sound, but without any image. It narrated an airplane accident from the point of view of one of the passengers. The audience would listen to it with all lights off. It was then followed by Avatar, with almost 3 hours of action, shooting and explosions. Yet, cinema owners received several complaints from spectators that declared "they had not gone to the cinema to witness such aggressive and drastic situations our piece had put them in! You can even feel good sound pulsating through your stomach!"

Technology

As introduced above, the key feature of **imm sound**'s technology is the independence of content creation from the loudspeaker layouts used in exhibition. One single imm 3D soundtrack is distributed to venues with a different number and location of loudspeakers, and **imm sound**'s processors adapt the soundtrack to each layout, producing a 100% tailored sound experience.

This is accomplished by removing the concept of audio channel in all processes in sound post-production, mixing, and distribution. State of the art and in-house developed algorithms ensure that all decisions regarding sound spatialization are performed without reference to any concept of channel. Algorithms include not only object-based audio processing but also higher-order Ambisonics and perceptual-based spatial sound processing.

All such powerful 3D sound related algorithms are performed by **imm sound**'s Immersive Audio Workstation, or IAW. The IAW integrates transparently with all professional audio workstations, like Pro Tools, and all professional mixing consoles. The workflow is therefore hardly changed for sound recorders, designers and mixers, who operate as always in all processes not related to sound spatialization.

Figure 1 shows one of the perhaps more representative examples of the creative freedom brought by the change of paradigm. On the left, a mono track is spatialized in a traditional 11.1 multi-channel setup, by routing it to a top-front-left, mid-front-left and top-rear-left channels. On the right, the same track is spatialized using the IAW, without any reference to channels, simply by locating it in space and assigning a suitable apparent size.

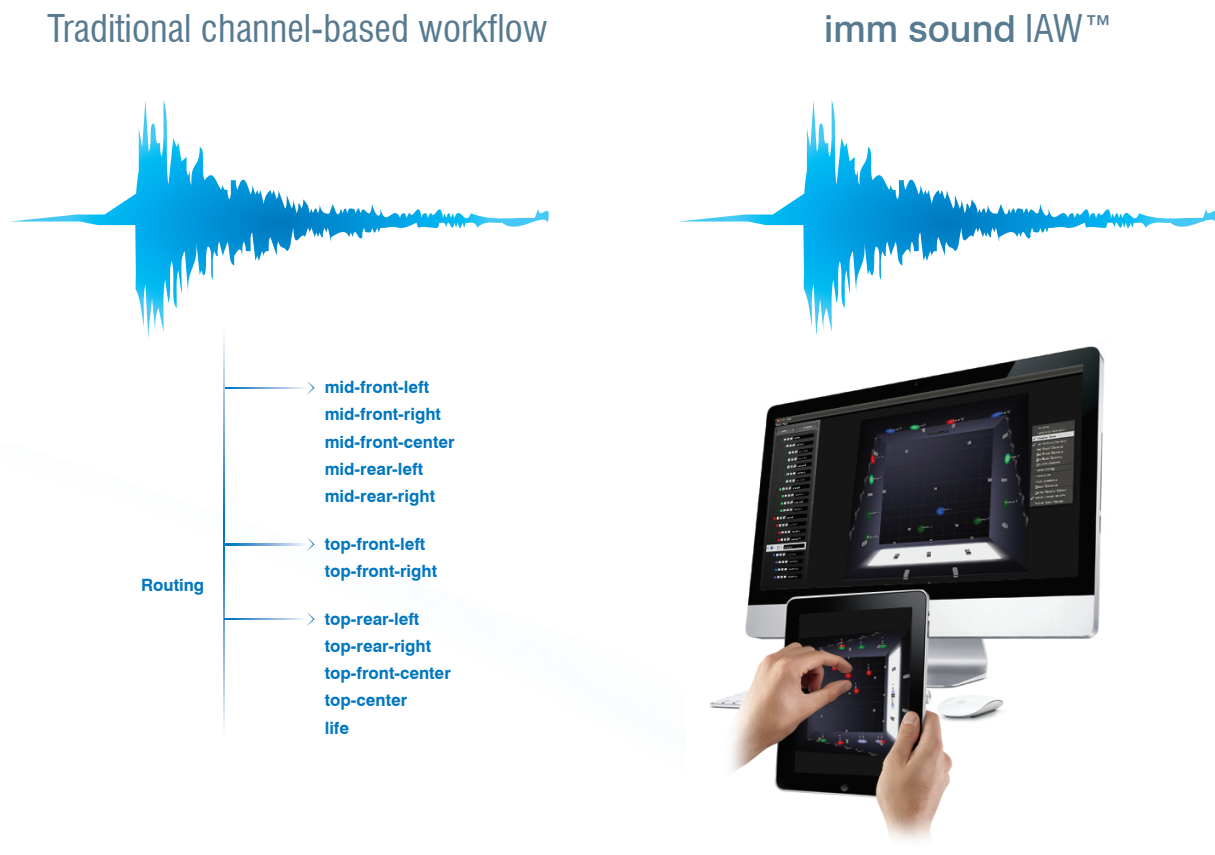


Figure 1: Spatialization of a single sound in a traditional channel-based (left) and in a channel-free workflow (right), the latter using a tactile interface provided by the IAW, whereby sounds are simply located in space, and assigned a given apparent size.

It might be worth taking this example a bit further. The soundtrack resulting from the channel-based decisions (left figure) cannot be reproduced in a straight forward manner in any loudspeaker layout other than the 11.1 it was created for. Any automatic conversion to say, a 5.1 format, will suffer from well-known problems of downmix processes (see downmix section below), including wrong relative volumes of different elements in the soundtrack, and comb filtering. In contrast, the resulting imm 3D soundtrack can be straightforwardly reproduced in any layout, or converted to the corresponding multi-channel format, from stereo and 5.1, to 14.1 or 23.1, with absolute respect to the original dynamics, and absence of downmix problems.

Further creative processes that are much simpler and powerful in channel-free workflows are the use of 3D reverberations and 3D ambient sounds, which often provide realism and naturalness hard to create purely in post-production.

We shall conclude this section describing the final steps of distribution and playback, which follows a fully DCI compatible process [3]. imm 3D soundtracks are mastered and exported to a digital support, like a hard-drive. Alongside, 5.1 and 7.1 versions can be created and monitored automatically from the imm 3D soundtrack.

The size of imm 3D soundtracks is typically about twice the size of a 5.1 soundtrack, which is, in any case, a small fraction of the size of the movie images. In alternative multi-channel based proposals, extra channels are folded into 5.1/7.1 streams by means of lossy compression, which should be unacceptable in the digital era. **imm sound** uses a 100% lossless format.

Digital Cinema Packages (DCPs) are authored as usual, including both the 5.1/7.1 versions and a reference to the **imm sound** version. Both the DCP and the imm 3D soundtrack are distributed in one single physical support. In cinemas without **imm sound**'s processors, ingestion is only in the cinema server, and reproduction is in standard 5.1/7.1 mode. In **imm sound** certified cinemas, ingestion is due also in the **imm sound** 3DSP processor, which plays the 3D soundtrack slaved to the cinema server (fig. 2).

Upon ingestion of imm 3D soundtracks, **imm sound**'s 3DSP audio processors, which are configured to match with the specific loudspeaker layout of each venue, produce a 100% tailored local version to provide the best possible experience.

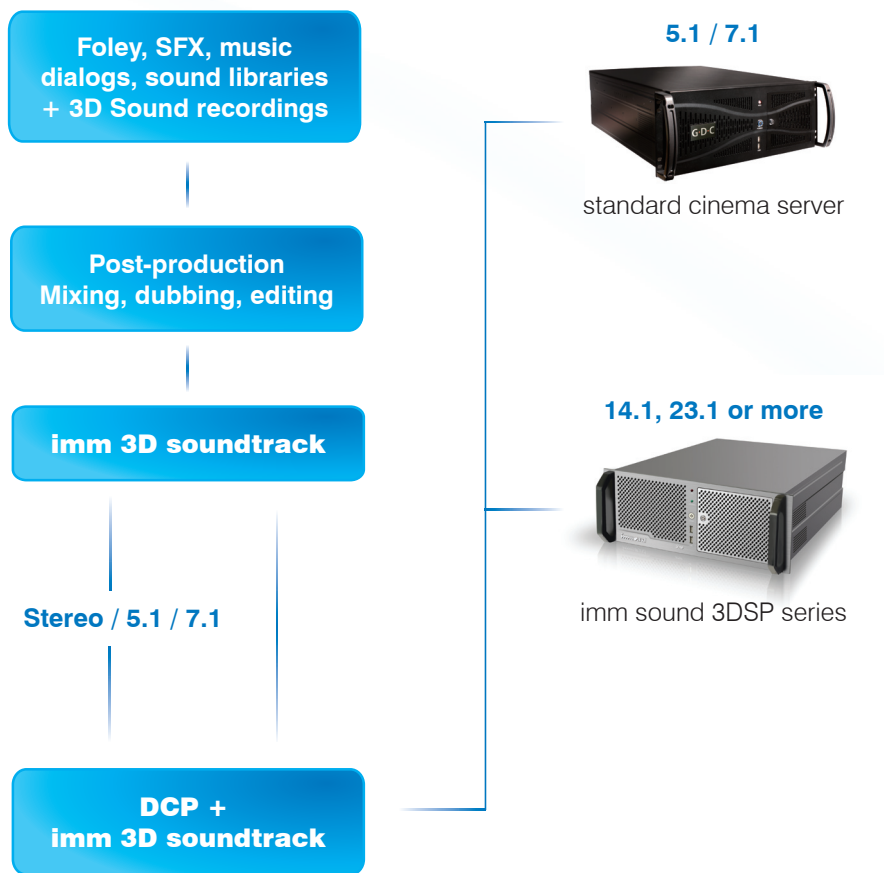


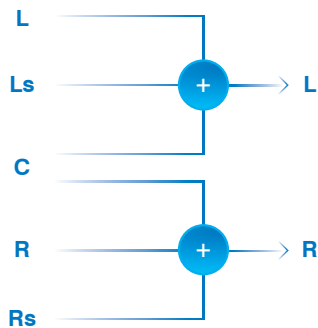
Figure 2: One single hard-drive, containing both the DCP with the 5.1/7.1 versions and the imm 3D soundtrack, is distributed and ingested in the cinema server. In **imm sound** certified cinemas, the same hard-drive is ingested in the **imm sound** processor too, which is permanently slaved to the cinema server.

Downmix is a problem of the past

Downmix is the process of generating a version of a soundtrack with a smaller number of channels than the original. A relevant example to cinema is the downmix of a 5.1 soundtrack to stereo, where, for example, the final left channel is a weighted combination of the original left-surround, left and center channels (fig. 3). It is well-known that such mixing of channels produces severe degradations of the resulting audio [6]: the original balance between different sounds is not respected, and audible artifacts like comb filtering appear in an uncontrolled manner. For this reason, different versions of traditional multi-channel soundtracks are always performed manually, in a lengthy and heavily supervised process. Downmix is a problem inherent to multi-channel soundtracks.

In contrast, imm 3D soundtracks do not contain any information about channels. They can generate any multi-channel version with absolute respect to the original dynamics, and absence of downmix problems. With **imm sound**, downmix becomes a problem of the past.

Typical downmix process from 5.1 to stereo



Absence of any downmix process

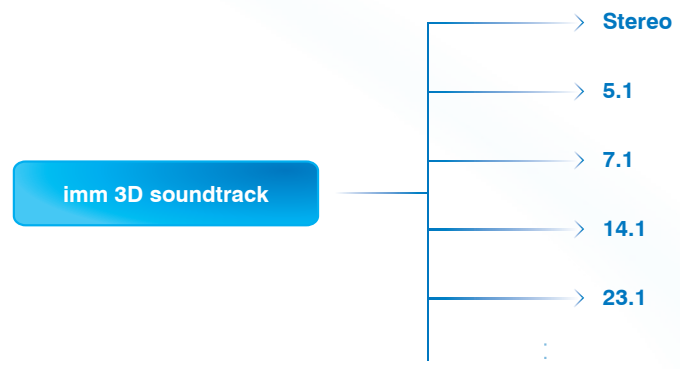


Figure 3: Left, typical downmix process from 5.1 to stereo, whereby the resulting signals are weighted sums of the originals, and exhibit unwanted variations in dynamic levels and comb filtering. Right: absence of any downmix process or artifact in the generation of multi-channel soundtracks out of an original imm 3D soundtrack

Recommended loudspeaker layouts

Given that imm 3D soundtracks can be reproduced fully exploiting any loudspeaker layout, the question arises as to which layouts should be considered ideal for the cinema industry, where only the highest quality is admissible. For the first time, this question can be posed only in terms of immersion, aesthetics and narrative considerations, not in terms of how many audio channels fit in a given physical support, or how many channels can creative professionals cope with.

While universal consensus is probably impossible, let us rephrase the question in terms of objective factors that every spectator can appreciate, in order to fix a minimum set of goals that every layout should satisfy.

The first goal should be that spectators perceive a significant change in their experience. This can be mostly achieved by exploiting all parts of the soundscape not available in 5.1 systems, especially, the top layers. The second goal should be that sound designers forget about restrictions; they should not bear in mind issues like “I’d better not move this sound of a dragon flapping too much above the spectators, because there is only one independent channel in the ceiling, and therefore, low angular resolution”.

Both goals point towards similar requirements: the top layers should be sufficiently populated with individually fed loudspeakers. In average theatres, the top of the screen, and the top of the back wall, are seen at about 25° elevation angle from the best seat. This leaves a huge space on top of the audience. There are proposals for installing plenty of loudspeakers in the ceiling, but fed by one single independent signal, normally referred to as the voice of God, mimicking the diffuse concept that surround channels in 5.1 systems have. This offers poor resolution, and remains far from fulfilling both of the aforementioned goals: smooth sound motion in the large area above the audience is not possible.

imm sound's expertise, gained through extensive tests in large variety of layouts, carried out for more than 7 years, and confirmed in professional cinema theatres operating worldwide 365 days a year, led to several conclusions:

- The use of the diffuse surround concept, by which one independent signal is fed to many loudspeakers should be avoided, as it destroys proper sound localization. Note that, of course, it is always possible to use independently fed loudspeakers as a traditional diffuse line, if the sound engineer wants. **imm sound**'s worldwide installers are already offering ceiling loudspeaker solutions with wide angular dispersion, light weight, and easy mounting (fig. 4).
- Not less than three independent channels should be located directly in the ceiling over the audience.



Figure 4: Example of a ceiling loudspeaker solution used in various **imm sound** theatres around the globe, with wide angular dispersion, light weight, and easy mounting.

All worldwide **imm sound** certified cinemas enjoy imm 3D soundtracks reproduction in a variety of layouts which are tailored based on the exhibitor specific needs; examples include the minimum recommended 14.1 or the typical 23.1 premium layout.

All certified layouts fulfill the aforementioned goals by minimal addition of loudspeakers on top of standard 5.1/7.1 systems, ensuring full 5.1/7.1 compatibility. 5.1/7.1 soundtracks are reproduced through the standard loudspeakers, with standard equalization, as if all extra installed loudspeakers did not exist.

The added value of imm sound in cinema

The content industry is completing the transition to the digital era, for which **imm sound** technology is fully designed. Besides enabling tools to create and manipulate sound in a new way, it simplifies and rationalizes the distribution process, and delivers to the audience an immersive experience without precedent in the analogue era.

In the present scenario, where movies are enjoyed in all kind of devices and places, the industry needs to adopt approaches in which the quality of the experience is the best possible for every case without the need to produce a different product for every scenario. **imm sound** technology has been designed to make this possible. One single imm 3D soundtrack needs to be produced, with the guarantee that it will be decoded so as to deliver the best sound experience that is possible in the infrastructure available to the spectator. Either in a large theatre with dozens of loudspeakers, at home with only a few, or on a tablet with headphones, the spectator enjoys the film knowing that the soundtrack offers the best possible sound for the infrastructure at hand. **imm sound** decoding technology is the keystone upon which a new way of producing, distributing and enjoying the movies can be build.

The cinema industry as a whole benefits from the better quality of experience that is possible with **imm sound** technology. But all members of the industry chain can take advantage of this opportunity. Exhibitors get a differentiating feature that brings spectators to their screens. Movie theatres deliver a unique immersive experience. Producers stop generating different products for different layouts. Sound mixers finally concentrate on sound, not in channels, and create wonderful soundscapes in a highly intuitive way. Creativity is enhanced by the use of an enriched cinematic language which can help telling stories better, and in novel ways. Everybody wins.

Producers

Several essential novel features of **imm sound**'s technology are relevant when planning a movie in terms of budget and distribution. First, one single imm 3D soundtrack tackles all formats, available and future. Thus, one single version will sound perfectly adapted, and maximally exploiting the possibilities of any conceivable reproduction layout: stereo, 5.1, 7.1, 11.1, 14.1, 23.1... All at the same time. Within one single DCP.

Second, the freedom that channel-independent post-production brings to sound designers enables them to concentrate on creativity, not on channel-based decisions. Besides, one single post-production streamline leads to both the imm 3D soundtrack and the stereo/5.1/7.1 versions, without any loss of quality due to downmix proceses, which affect other traditional multi-channel proposals. As a result, post-production time and movie's total sound budget do not increase, and might even slightly reduce.

Third, **imm sound** is the ideal companion to also exploit markets beyond cinema. Home users greatly benefit from the freedom of locating loudspeakers anywhere they choose. As it is well known, home environment limitations turn into only a small percentage of users actually locating home 5.1 systems properly. The expectations for traditional twice-as-many loudspeakers proposals do not seem viable in this respect, a fact that might worry producers about the quality of their movies when reproduced at home. With **imm sound**, home users can locate as many loudspeakers as desired, and anywhere it suits their needs. Producers are sure that their content will playback 100% adapted to every user's choice! This feature might be key for massive deployment of 3D sound systems at home.

Post-production studios

imm sound provides a unique tool, the Immersive Audio Workstation, or IAW, for efficient and creative 3D sound post-production. It seamlessly integrates with all major audio workstation, like Pro Tools, and with all professional mixing consoles. The IAW takes care of all purely 3D sound related processes, and provides intuitive tactile interfaces for sound spatialization, allowing engineers to concentrate only in the creativity, not in channels.

With the IAW, limitations in 5.1 or other traditional multi-channel systems are overcome. Coherent and continuous sound motion is now possible along every direction about the audience, sounds can have well-defined depth, proximity, and it is possible to create rich atmospheres and three-dimensional reverbs. The use of 3D recordings and upmix from stereo or 5.1 to 3D are novel powerful possibilities offered by the IAW. (see next page for further details)

The IAW also allows, with a simple switch, monitoring of the mix in a wide variety of formats, like 5.1, 7.1 and all other **imm sound** recommended layouts, including 14.1 and 23.1. The IAW exports a final master or separate stems as imm 3D soundtracks/stems, or to any standard multi-channel format. Only one hard-drive is needed for distribution to all cinemas. Those without **imm sound** compatible audio processors will reproduce 5.1/7.1, as standard.

3D recordings

While dialogues can continue benefitting from close and boom mikes, there is a whole world of interesting soundscapes that deserve proper 3D recordings: rich ambient sounds, reverberations, and even 3D-recorded foley, like airplanes or trains passing by. The freedom provided by **imm sound's** channel-free technology allows the use of all existing, and soon-to-arrive, microphone solutions.

Unlike in channel-based approaches, the possibilities are endless. The choice of solution can be based on aesthetics, narrative intention and logistics, not on the number of channels for which the content is prepared. It is possible to use any combination of standard microphone arrays, custom extensions that enable capturing the sound above the listener's plane, and inherently 3D microphones, like the Soundfield [7] or the Eigenmike [8].

The workflow from capture to post-production does not require any change. Recordings are time aligned in Pro Tools, and seamlessly recognized and processed by the Immersive Audio Workstation.

3D upmix

3D upmix is a real-time conversion of any stereo or 5.1 source to **imm sound's** channel-free 3D format. It analyses the original signals and, via a physical and psycho-acoustical inference process, generates full 3D soundscapes while respecting one of the most sacred laws of cinema: dialogue, direct sound of instruments, etc. always remain in the screen.

When used in post-production, it helps creating compelling ambiances and reverbs out of stereo or 5.1 stems or recordings; it also allows obtaining an initial version of a 3D soundtrack, on top of which engineers can concentrate and polish the important scenes.

When used in cinema theatres, it is the perfect companion for alternative content, especially when broadcasted in stereo.

Exhibitors

Exhibitors from around the globe have already been enjoying **imm sound** in fully operative commercial installations since early 2010. On one hand, they are well aware that the impact of immersive 3D sound on the audience is large, because unlike in the change from analog to digital, there is not a single spectator that leaves the audience without having noticed the new experience. On the other hand, they benefit from the following advantages of the technology:

- **Compatibility.** It is fully compatible with 5.1/7.1 systems; **imm sound** recommended layouts always include a standard 5.1/7.1 layout as sub-systems. Thus **imm sound** certified venues can function in standard 5.1/7.1 mode when required.
- **Flexibility.** Exhibitors have the freedom to choose layouts according to their needs, including venue architectural restrictions and budget. **imm sound** provides technical support and ensures the quality of the theatre. Both new venues and refits are currently under operation worldwide.
- **Affordability.** All **imm sound** certified layouts add the minimal number of loudspeakers to the standard surround configuration. In traditional proposals for only one or two ceiling channels, many loudspeakers need to be placed to distribute sound. Thanks to **imm sound**'s concept, which allows for full exploitation of sound localization above the audience, different manufactures have already provided new solutions of highly dispersive loudspeakers tailored for ceiling mounting. This facilitates installation and, perhaps paradoxically, reduces the number of required loudspeakers, and thus, the final budget.



Automatic EQ systems

While 5.1 cinemas are still largely calibrated manually, venues with larger number of independent channels desperately call for automation of part of the calibration process. **imm sound**'s cinema processor provides the highest quality state of the art algorithms for automatic equalization of up to 120 channels, based on precise FIR and IIR filters, which greatly simplifies the cinema installation process.

Perfect alignment of all loudspeakers in a 3D layout produces the feeling that a continuous sound canvas surrounds the audience, that loudspeakers become transparent. It also ensures content creators and distributors that films will sound in cinemas as they sound in studios.

Universal compatibility

By breaking with the concept of channels, **imm sound** takes a privileged place in the cinema sound standards ecosystem. On the one hand, imm 3D soundtracks can be reproduced fully adapted to any present or future loudspeaker layout. On the other hand, by being based on the high standard that sounds can be precisely located at any direction about the audience, any traditional multi-channel soundtrack can be faithfully reproduced in **imm sound**'s certified cinemas, via a process known as remapping, implemented in **imm sound**'s cinema processors.

Compatibility with all present and future multi-loudspeaker formats and configurations is at the heart of **imm sound**'s technology.

Conclusion

imm sound is a new paradigm. It tackles the root of the problems that have impeded immersive 3D sound's deployment, despite the many proposals for multi-channel layouts that have coexisted for almost 20 years. It exploits the power of the digital era to provide a truly new solution, not just a way to embed more audio channels in a physical support.

For the first time, content creation is made independent of the number of loudspeakers and their position. One single imm 3D soundtrack exploits the best of all present and future loudspeaker layouts, ensuring always the highest quality and fidelity.

The impact is at all levels. In post-production, creativity is maximized by freeing designers and mixers from having to cope with 10-20 channels to base their decisions. The Immersive Audio Workstation integrates with all major studio software and hardware, enabling the engineers to produce efficiently and concentrate on the creative processes.

In distribution, one single DCP is needed for all venues, regardless of whether they have **imm sound** compatible audio processors or not. In theatres, it provides freedom to exhibitors to install layouts that best suit their needs and venue specificities, while providing full 5.1/7.1 compatibility, and requiring a minimal addition of loudspeakers and amps; even less than traditional proposals for only one or two distributed ceiling channels.

imm sound makes a contribution to the cinematic language that may enhance the cinema experience in a way that can not be predicted today. We can not foresee how creative professionals will take advantage of the possibility of immersing the spectator in the middle of any imaginable soundscape. Nor can we predict how far communication with the spectator will improve with spatial sound, but we are sure that the cinema experience will be better.

imm sound's immersive sound technology is a breakthrough. It is ahead of other proposals by at least a couple of years. It may take some time until it is fully adopted by the industry, but it is clear that **imm sound** technology is leading the way towards the 21st century cinema sound. Immersive sound is much more than 3D sound, if understood as a simple addition of loudspeakers in the ceiling; it is about quality of the spectator experience, about natural soundscapes that rival real life. **imm sound** technology has been designed to make possible, efficient and inexpensive this evolution of cinema sound.

References

[1] **Spatial hearing: the psychophysics of human sound localization**, Jens Blauer, MIT press, 1997.

[2] **Dolby surround mixing manual**,

http://www.dolby.com/uploadedFiles/zz-_Shared_Assets/English_PDFs/Professional/44_SurroundMixing.pdf

[3] **The Number of Audio Channels**, Holman, Tomlinson. AES 100th Convention, 1996, paper 4292.

[4] **5.1 and 22.2 multi-channel sound productions using an integrated surround sound panning system**, Hamasaki, Kimio; Hatano, Wataru; Hiyama, Koichiro; Komiyama, Setsu; Okubo, Hiroyuki. AES 117th Convention, 2004, paper 6226.

[5] **Digital Cinema Initiatives specification**,

<http://www.dcinovies.com/specification/>

[6] **Comparison of Quality Degradation Effects Caused by Limitation of Bandwidth and by Down-mix Algorithms in Consumer Multichannel Audio Delivery Systems**, Zielinski, SK, Rumsey, F and Bech, S, AES 114th Convention, 2003, paper 5602.

[7] **The Soundfield microphone**,

<http://soundfield.co.uk/>

[8] **The Eigenmike microphone**,

http://www.mhacoustics.com/mh_acoustics/Eigenmike_microphone_array.html

imm sound®

Diagonal 177
E-08018 Barcelona

Tel (+34) 93 485 3880
Fax (+34) 93 309 3188

info@immsound.com
www.immsound.com