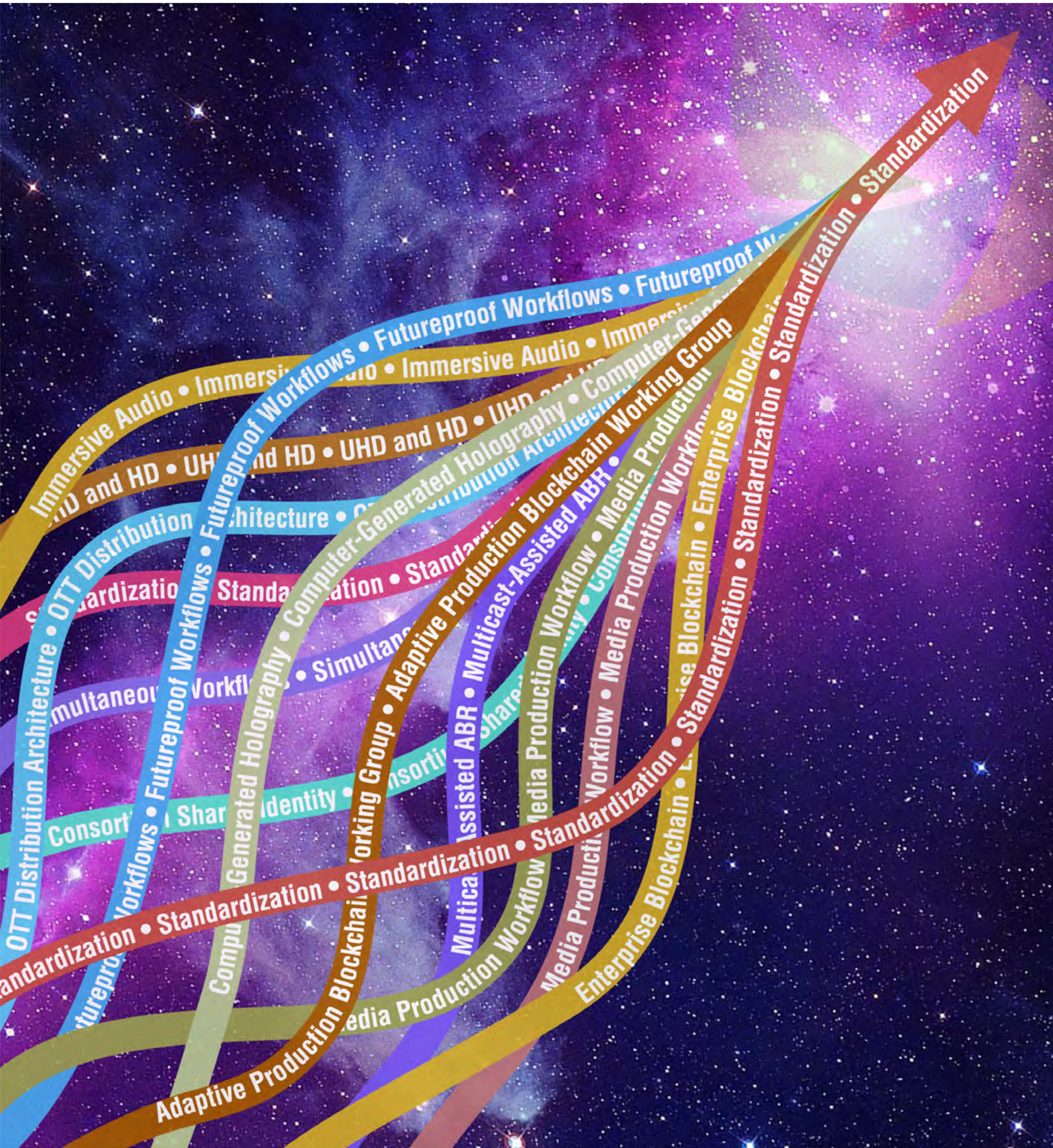
SMPTE

MOTION IMAGING JOURNAL

Covering Emerging Technologies for the Global Media Community



Immersive Audio: Futureproof Workflows for the Real World

By Brian A. Vessa

Introdução:

Falar de interoperabilidade deveria ser um mantra para as novas tecnologias. Infelizmente a masterização e distribuição de áudio imersivo é feita, atualmente, usando fluxos de trabalho proprietários, limitando sua interoperabilidade e aceitação final. Tentando mudar isso, este artigo propõe um avanço com o SMPTE ST 2098-2 Bitstream de áudio imersivo (do inglês IAB), já que é importante que os fluxos de trabalho sejam tão eficientes e interoperáveis para que se possa criar conteúdo à prova de futuro. Também é crítico que os equipamentos, que esperamos possam processar e reproduzir este formato, sejam interoperáveis. Claro que como bem dizem os autores "ainda estamos no estágio inicial de implantação do IAB" e, por isso, os fluxos de trabalho adotados agora podem moldar como a indústria irá avançar no futuro próximo. Boa leitura, bom ano e bem vindos ao futuro!

Tom Jones Moreira

Abstract

Mastering and distributing immersive audio is currently done using proprietary workflows, limiting its interoperability and ultimate acceptance. To move forward with the SMPTE ST 2098-2 Immersive Audio Bitstream (IAB) in an eloquent manner, it is important that the workflows be as efficient and interoperable as possible while creating futureproofed content. It is also critical that the equipment we expect to process and play IAB is interoperable. We are at the crucial beginning stage in the rollout of IAB, and the workflows that are adopted now will shape how the industry moves forward.

The key is to plan, budget, and create for the big picture: long-term delivery of both feature and television content to multiple distribution channels rather than the immediate, short-term window. Designing and mixing immersive audio at the beginning rather than adding it later makes the product better and is more efficient. This allows the content to be delivered to any market and saves money in the long run.

Futureproofing is the key to building a valuable and versatile library. Shooting and finishing in 4K/HDR, designing and mixing in immersive audio, and authoring these into Interoperable Master Format (IMF) are key to the longevity and salability of the library.

Interoperability is the key to efficient distribution, and it must be nurtured and finessed to realize its full potential. To achieve full interoperability, content creators and distributors need to make their business case and push manufacturers for the adoption and full implementation of standards.

To move forward with IAB in an eloquent manner, it is important that the workflows be as efficient and interoperable as possible while creating futureproof content. It is also critical that the equipment we expect to process and play IAB is ready for the task when content is delivered. We are at the crucial beginning stage in the rollout of IAB. Workflows that are adopted now will shape how the industry moves forward.

Keywords

Audio, automation, Digital Cinema Packages (DCP), futureproof, immersive audio, Immersive Audio Bitstream (IAB), Interoperable Master Format (IMF), mixing, workflow

Introduction

Immersive sound adds the third dimension of height to the consumer listening experience and allows for pinpoint control of individual components in the soundscape via audio objects (clips of pulse code modulation (PCM) audio with associated metadata to direct their playback). Since its introduction a decade ago, it has grown in stature and advanced technologically to the point where the majority of "tentpole" titles are mixed natively in immersive sound. It is also utilized by gaming and virtual reality (VR) applications.

Mastering and distributing immersive audio is currently done using proprietary workflows, limiting its interoperability and ultimate acceptance. To address this problem, SMPTE standardized the ST 2098-2 Immersive Audio Bitstream (IAB) in 2018.[1] Additional standards have been written that standardize its use in Digital Cinema Packages (DCPs) and the Interoperable Master Format (IMF). A number of codecs have been developed to carry it to the consumer for a variety of applications.

To move forward with IAB in an eloquent manner, it is important that the workflows be as efficient and interoperable as possible while creating futureproof content. It is also critical that the equipment we expect to process and play IAB is ready for the task when content is delivered. We are at the crucial beginning stage in the rollout of IAB. Workflows that are adopted now will shape how the industry moves forward.

This article examines current issues related to the creation, mastering, distribution, and playback of immersive audio and proposes real-world workflows to address them. The reality of interoperability while achieving consistent and quality playback that accurately brings the artistic intent to the consumer is addressed, proposing steps to achieving these in the short and long term.

Current Top-Level Workflows

Top-Level Feature Workflows

To create efficient IAB workflows, it is important to understand the basic workflow processes already at play. **Figure 1** shows a typical top-level workflow for creating a 5.1 theatrical feature soundtrack mix. The release sound format target (5.1) is given up front and all elements are created to that format. The "Original Version" (OV) 5.1 printmaster (composite mix) is created from the final mix stems and packaged into a DCP, containing image, audio, text files, and a composition play list (CPL) that plays them in sync. It is then distributed to cinemas. In the process, a "Music and Effects" (M&E) mix is also created in the same format that removes the dialogue and "fills" the production effects that are lost when doing so with Foley and specific sounds. This M&E is used to create dubbed language mixes, which are packaged into DCPs for release in multiple territories.

The OV theatrical mix and dubbed languages are then available for "downstream" use. Ideally, a "nearfield" 5.1 home theater mix is performed to optimize for the home environment. A Left total-Right total (LtRt) home theater printmaster (two channel mix containing encoded surround information) is also created. These are conformed to a long-form video of the movie and packaged with it, creating a "video master." This video master with a 5.1 and LtRt OV mix can then be used to create a number of deliverables for physical (BD, DVD) and digital (streaming and broadcast) distribution. The dubbed language mixes are used to create a Foreign Language Master (FLM), which is the source for physical and digital distribution internationally. Additional languages are often created to service territories that did not receive a theatrical release.

Top-Level Television Workflows

The television workflow is similar to the home theater part of the feature workflow. **Figure 2** shows a typical top-level workflow for creating 5.1 and LtRt television soundtrack mixes. The broadcast sound format target (5.1) is given up front and all elements are created to that format. The OV mixes are packaged into a video master to create the first-market broadcast deliverables. An M&E is also created for dubbing, and the dubbed languages go into an FLM to create the international broadcast deliverables. In some cases, the content may get a theatrical release, in which case a theatrical printmaster and DCP are also created. This part of the workflow is shown as dashed lines.

One significant difference between feature and television audio is the dynamic range and overall level restrictions specified by the broadcast industry (both terrestrial broadcast and streaming). While this initially grew out of the limitations of analog broadcast, today's restrictions require that the level and dynamic range of the program content be more similar to commercials, so that when going from program to commercial and back to program, the level change is not jarring. By specifying a loudness target that all programs must meet, switching from one channel to another channel maintains a similar level. The governing specifications are Advanced Television Systems Committee (ATSC) A/85 (-24 LKFS) in the U.S. and European Broadcast Union (EBU) R128 (-23 LUFS) in Europe. Some broadcasters have their own specifications.

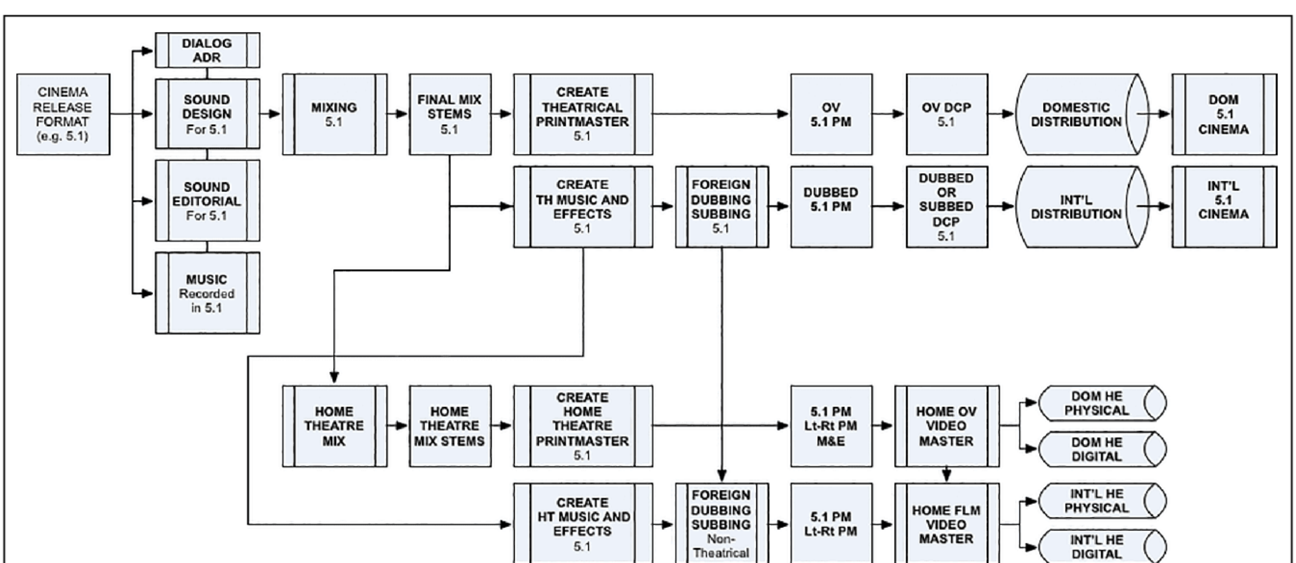


FIGURE 1. Top-level feature soundtrack workflow.

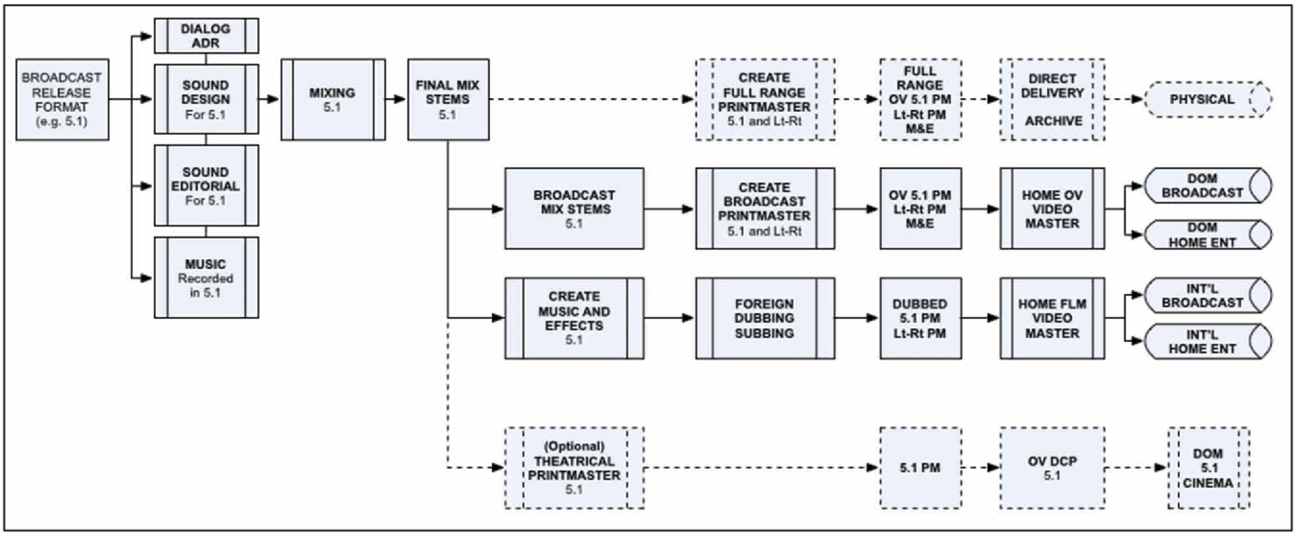


FIGURE 2. Top-level television soundtrack workflow.

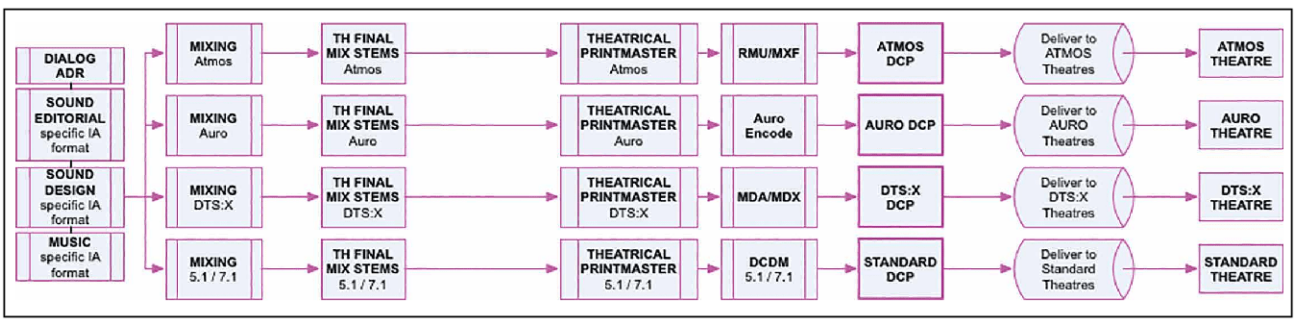


FIGURE 3. The challenge of multiple immersive audio mixes and deliverables.

While this works well in its context, the content's dynamic range and impact are compromised. Though unfortunately rare in today's tight schedules, a full-range mix should be created for archive and physical distribution, and the broadcast mix created in parallel or immediately afterward. The full-range workflow is shown as dashed lines in **Fig. 2**.

Enter Immersive Audio

Feature Immersive Audio Workflows for Cinema

When immersive sound systems began to spring up in cinemas a decade ago and a desire to deliver to these systems emerged, the theatrical audio workflow had to adapt quickly. Each brand of immersive sound system required its own mix and specific DCP to play the immersive audio soundtrack. Sound designers, scoring mixers, and re-recording mixers are needed to learn these new formats and how to best create for them. And it was still necessary to create a 5.1 and 7.1 mix for standard theaters, with corresponding DCPs. **Figure 3** illustrates the challenges.

The bottom line is that greenlighting an immersive audio soundtrack meant a very painful and expensive experience both for the creatives and the distributors—a big impediment to wide adoption.

As a result, a number of workflows have been tried to gain efficiency. **Figure 4** shows an example of a current workflow, where the movie is mixed in one immersive sound format (Atmos in this example) and then converted (remixed/downmixed) to create the other formats. Though more efficient than **Fig. 3**, it still presents challenges to post schedules and does not really help distribution, since multiple DCPs are still required.

Feature Immersive Audio Workflows for the Home

Immersive sound for the home is an exciting recent entry into the consumer marketplace. Both feature and television immersive audio content are now available on physical disc. A number of services have begun to stream it as well, creating a new market and demand. ATSC 3.0 provides for the transmission of immersive audio using AC-4 or MPEG-H.

Consumer audio manufacturers offer many ways to play immersive audio at home. There are a number of full-blown home theater systems with individual loudspeakers for the true Audiophile that sound amazing. Competitively priced soundbars are now available that do a reasonable job of conveying immersive sound for the average consumer, especially with the addition of rear loudspeakers. Headphone technology is making tremendous strides and has become an excellent way to

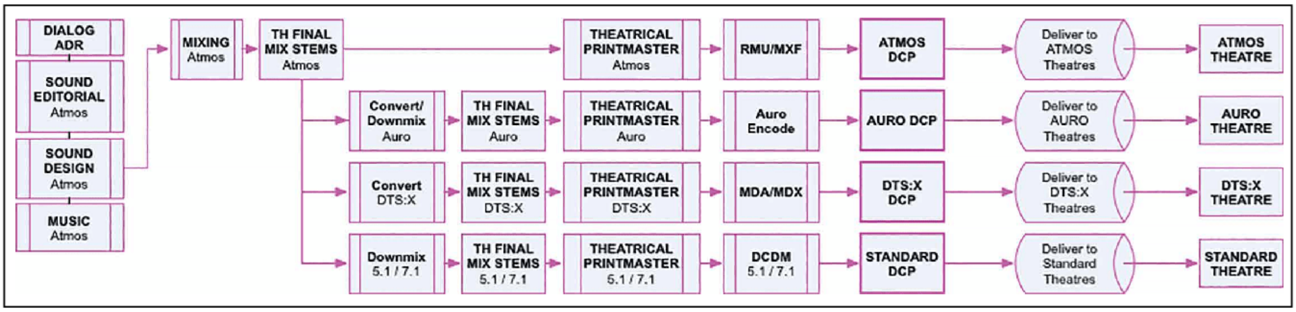


FIGURE 4. Example immersive remix workflow.

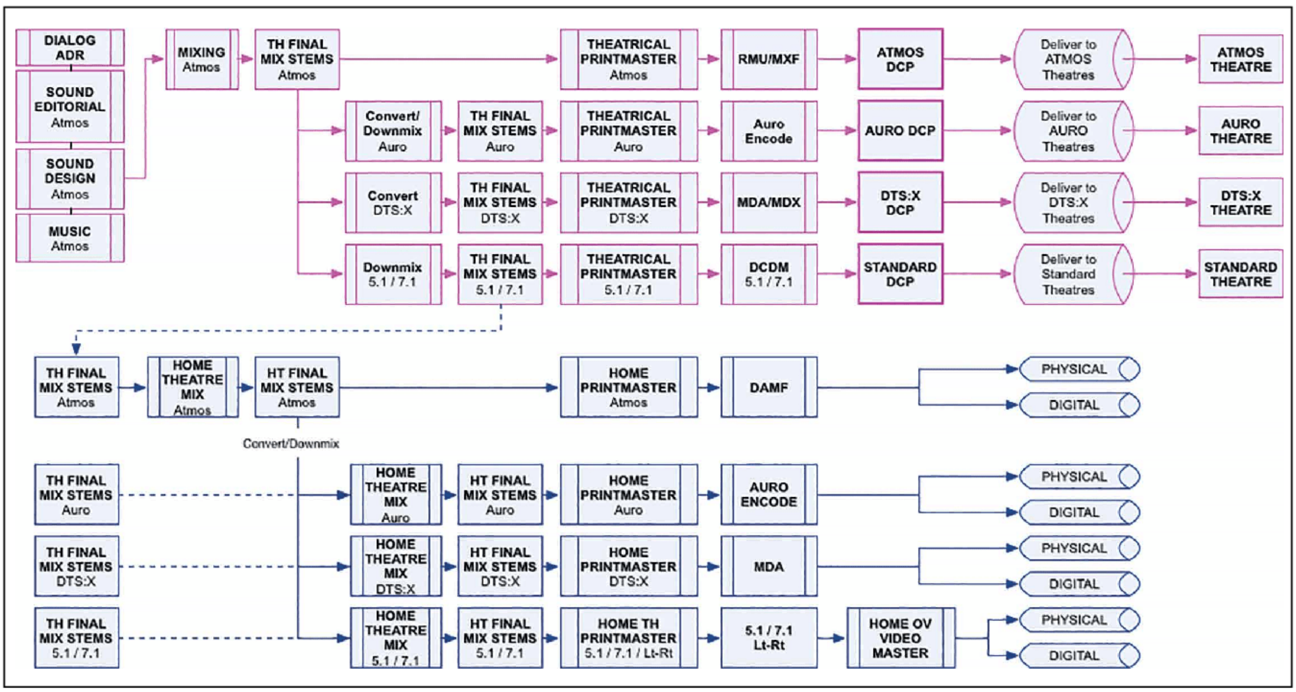


FIGURE 5. Current immersive mixing workflow for cinema and home.

enjoy immersive sound from any device. There is also emerging technology that conveys immersive sound directly through the loudspeakers in a computer or mobile device that shows great promise.

Creating and distributing immersive audio for the home has its own technical and distribution challenges. It still requires separate mixes and deliverables for each brand of immersive sound format, which are distinct from the theatrical deliverables. It still must be conformed to long-form video, but carrying it in the same video master has been impossible until recently. Therefore, it must be delivered separately. **Figure 5** illustrates the current workflow for cinema and home. Each home mix format is created either by using the theatrical stems in each format (dashed lines) or by remixing from one home mix format to the other, as shown in the theatrical workflow.

Television Immersive Audio Workflows

Creating immersive audio mixes for television content has been an after-the-fact affair until recently, but native immersive audio mixing is now finding its way into television workflows. Mixing, mastering, and delivering television immersive audio content face essentially the same challenges as features and can be solved similarly.

IAB

The multiple proprietary deliverables for immersive audio have been an impediment to its adoption and distribution. Thanks to dedicated efforts by a number of industry audio engineers, SMPTE published ST 2098-2 IAB in 2018.[1] This standard defines an interoperable delivery bitstream that can carry immersive audio. Companion standards for digital cinema have also been published that define how IAB is to be wrapped into an MXF track file (ST 429-18)[2] and packaged into a DCP (ST 429-19).[3]

Using these standards, a single IAB DCP can be created and delivered to any theater with an immersive sound system capable of reading and playing the IAB bitstream. This is a definite game-changer.

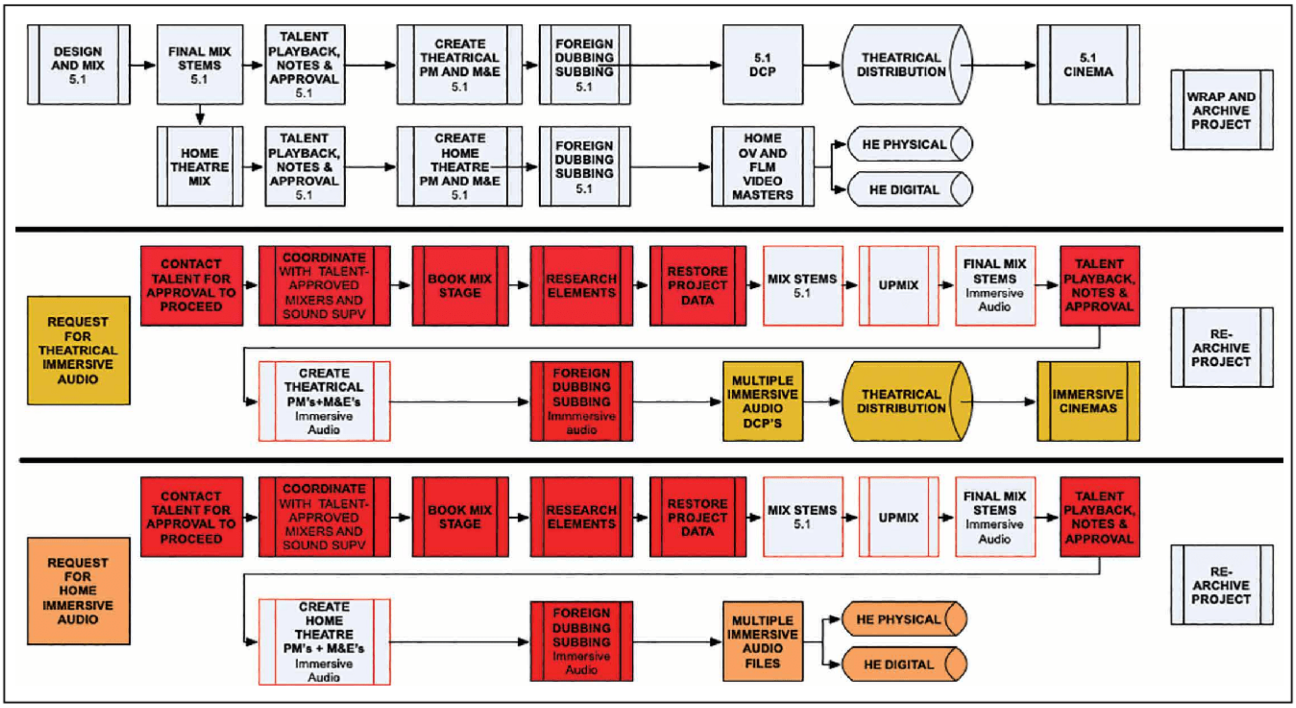


FIGURE 6. Short-sighted first window.

With IAB, there can be one immersive audio mix, one deliverable, and playout in multiple theaters with various immersive sound systems. All of the technical elements are in line for the IAB DCP rollout, but the industry has some challenges. Discussion and proposed solutions follow.

IMF

The Interoperable Master Format (IMF) is the brainchild of engineers from Hollywood studios who wished to have a file-based replacement for videotape as a mastering medium. Using DCP as a springboard, they created a specification for a digital video package (DVP), which was then brought to SMPTE, standardized in 2013, and renamed IMF. It is also a collection of image, audio, and text files with a CPL to play them in sync, but is far less constrained than DCP as it can have many different files and CPLs. Each CPL plays a unique combination of the assets, called a composition. The IMF can be expanded as assets for a title grow.

This is a very flexible mastering system and an excellent source from which to make deliverables. If desired, a subset of assets and a CPL to play them can be packaged into an Interoperable Master Package (IMP) and delivered to a facility or "Business to Business" (B2B) partner.

In addition to being able to carry "standard" audio, such as 5.1 and LtRt, IMF is the perfect solution for mastering and delivering IAB, thanks to SMPTE publishing ST 2067-201 in 2019.[4] IAB can now be part of a title's IMF assets and delivered or transcoded as needed to create multiple deliverables.

Short-Sighted First Window

Now that we have looked at the current workflows and introduced IAB and IMF, how can we use these new tools to improve workflows and gain efficiency?

The first step is to recognize and address the short-sighted first window that encumbers current workflows. The short-sighted first window is driven solely by the deliverables required by the first market release and does not take into account the bigger picture of the lifecycle of content delivery. **Figure 6** shows how this workflow operates, and why it sets itself up for a lot of pain that could have been avoided.

It begins with the noble goal of creating the deliverables on the delivery list for the first air or screen, but when a market requests immersive audio later, it costs more to create (often read: prohibitive), talent needs to approve, scheduling is a hassle…so, it either does not happen or happens haphazardly and is much more difficult to do. It generally does not give as good a result since the sound design was not done with immersive audio in mind and can only be expanded so much in an up-mix.

There is a similar situation in the picture world, where the image is shot and/or finished to the first window delivery requirement (usually 2K or HD) and then later, it is determined that there is a market for 4K/UHD and HDR. It costs a lot to create these after the fact (often read: prohibitive), talent needs to approve, scheduling is a hassle…and this is if it is even possible at all based on how the show was shot and what files were delivered. Similarly, it generally does not give as good a result.

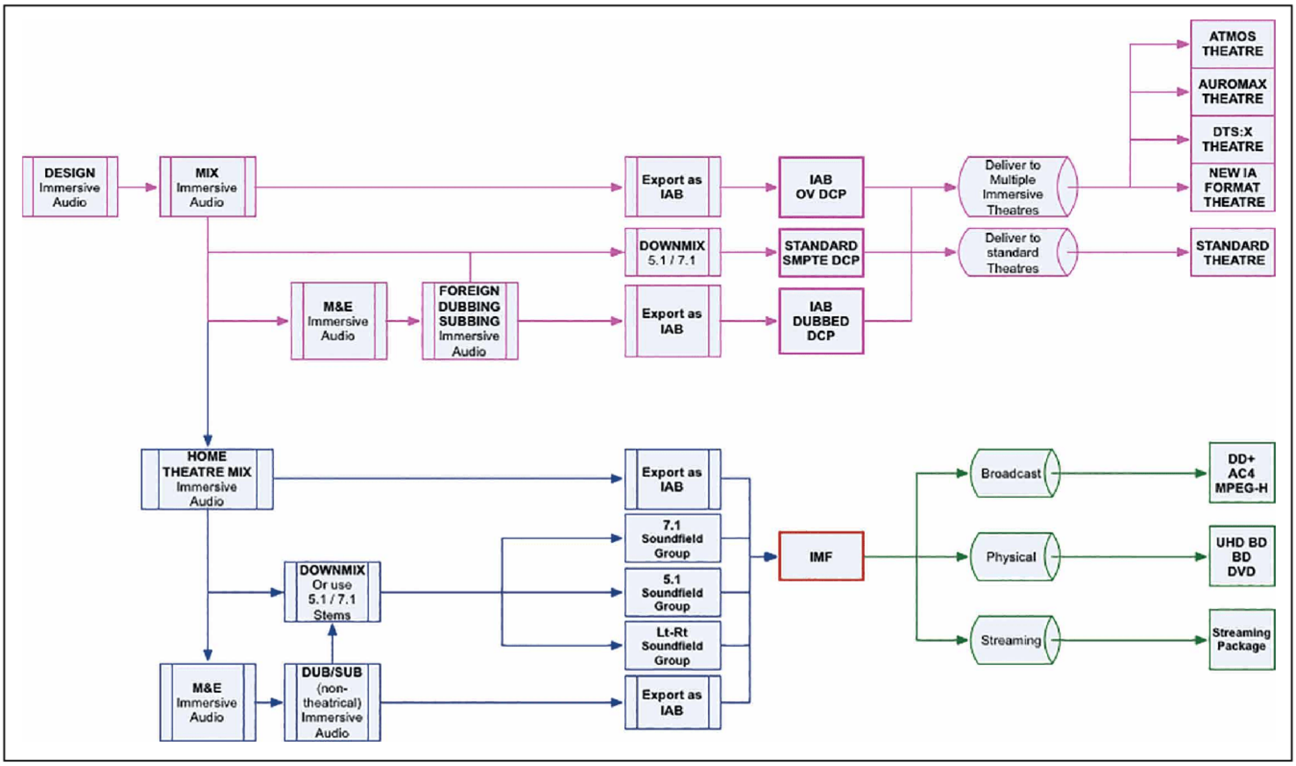


FIGURE 7. Native future-forward path workflow for features.

Futureproofing

The key to creating a sustainable and efficient immersive audio workflow is to design with the future in mind. These days, there is no extra cost to design and natively mix immersive audio like there was ten years ago. Sound design and mixing should be undertaken with the ultimate goal of an immersive audio release, even if there is not one planned for the first delivery window.

For features, the ultimate goal is no longer the cinema release, but all future markets. For television content, the ultimate goal is not the first window broadcast, it is also all future markets.

Therefore, the sound design, mixes, and deliverables for those future markets should be done while the production is active and the materials at hand to futureproof. While it may appear to be more complicated and expensive to do this "extra" work up front, it is actually far simpler and cheaper in the long run.

Future-Forward Path

The future-forward path is where the production and post-production workflows are planned for the long-term distribution of the content and all phases of the creation move in lock step toward that ultimate goal.

Figure 7 illustrates the "native" future-forward workflow for features. The theatrical sound design and mixing are immersive, and 5.1 and 7.1 are created from the immersive mix. This yields an IAB DCP and standard DCP. The workflow goes directly into a home immersive mix, which produces home IAB. The home 5.1, 7.1, and LtRt are created either from the home immersive mix or from the theatrical 5.1/7.1 stems using the home immersive mix automation. Together, these yield an IMF with IAB and standard audio tracks, ready to use as a source for servicing any and all downstream markets in a very time-efficient manner.

This workflow fosters the highest quality immersive experience for the consumer. As an added bonus, the standard mixes are more enveloping, since they inherit the additional audio detail and holographic sensation from the immersive mix.

Figure 8 illustrates the native future-forward path for television content. In this workflow, both the sound design and mixing are immersive, and the 5.1 and LtRt mixes are created from the immersive mix. To futureproof, the mixes are first created full range and then another or parallel pass is done to create the A/85 or R128 mixes for first-window broadcast. If a theatrical release is planned, DCPs are created from the full range mixes.

All audio is then put into IMF, ready to use as a source for servicing any and all future markets.

The native path has many advantages, yielding the highest quality audio in multiple formats with maximum time efficiency.

Figure 9 illustrates a compromise alternative path in case an immersive mix room is not available for the initial mix. A television workflow is shown, but it can also be used for features.

With this path, the sound design is immersive but a standard mix is performed first. This mix is laid out

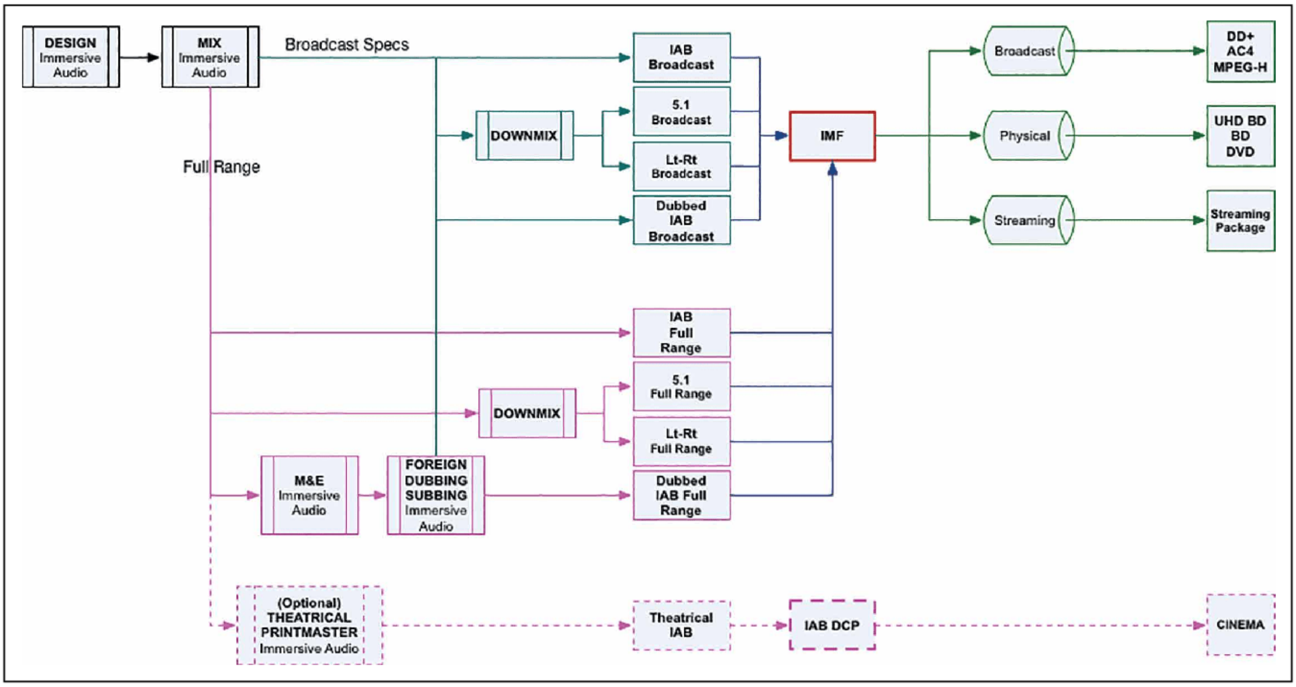


FIGURE 8. Native future-forward path workflow for television.

with an immersive mix in mind, often with additional stems that can go directly into objects. This yields 5.1 and LtRt mixes both in full range and A/85 or R128 and also yields full range immersive-ready stems. The project then moves on directly to an immersive mix stage that completes the immersive mix in full range and A/85 or R128. All audio is then put into an IMF, ready to use as a source for servicing any and all future markets.

For the alternative workflow to be effective and efficient, it is important that the sound design is immersive and flows seamlessly through the standard mixing process to the creation of the immersive mix, which should be performed as close to the timeframe of the standard mix as possible.

Interoperability in the Real World

When SMPTE or any other standards organization creates standards, they are designed to be used in the real world. We work hard to ensure that equipment manufacturers can actually implement these standards, and most standards committees have representatives from that sector for this very reason. But once the standard is published, it is still up to the manufacturers to ultimately choose how or if they implement it. Resources, development cycles, product vision, market share, and politics all play their part in the implementation decisions. In some cases, standards have been written but never fully implemented, and true interoperability was never achieved.

Interoperability in the real world needs cooperation and business justification as well as finessing and nurturing to come to fruition. Many of the things we take for granted that "just work" were hard-fought wins and business agreements.

The promise of IAB is to make one immersive mix, play it in any theater with any brand of immersive sound system, and create any flavor of immersive audio deliverable from it. This level of interoperability is highly desired and is achievable.

As usual, getting there means traveling through some bumps in the road before hitting smooth highway.

Cinema IAB Rollout

Digital cinema relies on a very constrained implementation of standards to interoperate as well as it does. New features have been slow to happen, as no one wants to rock the boat and risk a dark screen. As we have seen with the very slow conversion from INTEROP DCP to SMPTE DCP, even though software and firmware updates have been available for some time, theaters have not been keen to update gear that "is working fine as is." Also, there has been reticence by manufacturers to adopt all of the features of SMPTE DCP. So, a constrained set of features was codified by the Intersociety Digital Cinema Forum (ISDCF), and test DCPs were created to check interoperability. SMPTE RDD 52 "SMPTE DCP Bv2.1 Application Profile"[9] specifies this constrained SMPTE DCP.

In light of this, the SMPTE DCP rollout has made significant strides, but it is still ongoing.

The Cinema IAB rollout and adoption of IAB DCP face similar challenges, but fortunately the number of theaters requiring updates is considerably more manageable. As with SMPTE DCP, the SMPTE ST 2098-2 IAB standard is designed to be forward-looking and thus exceeds any current implementation. And some

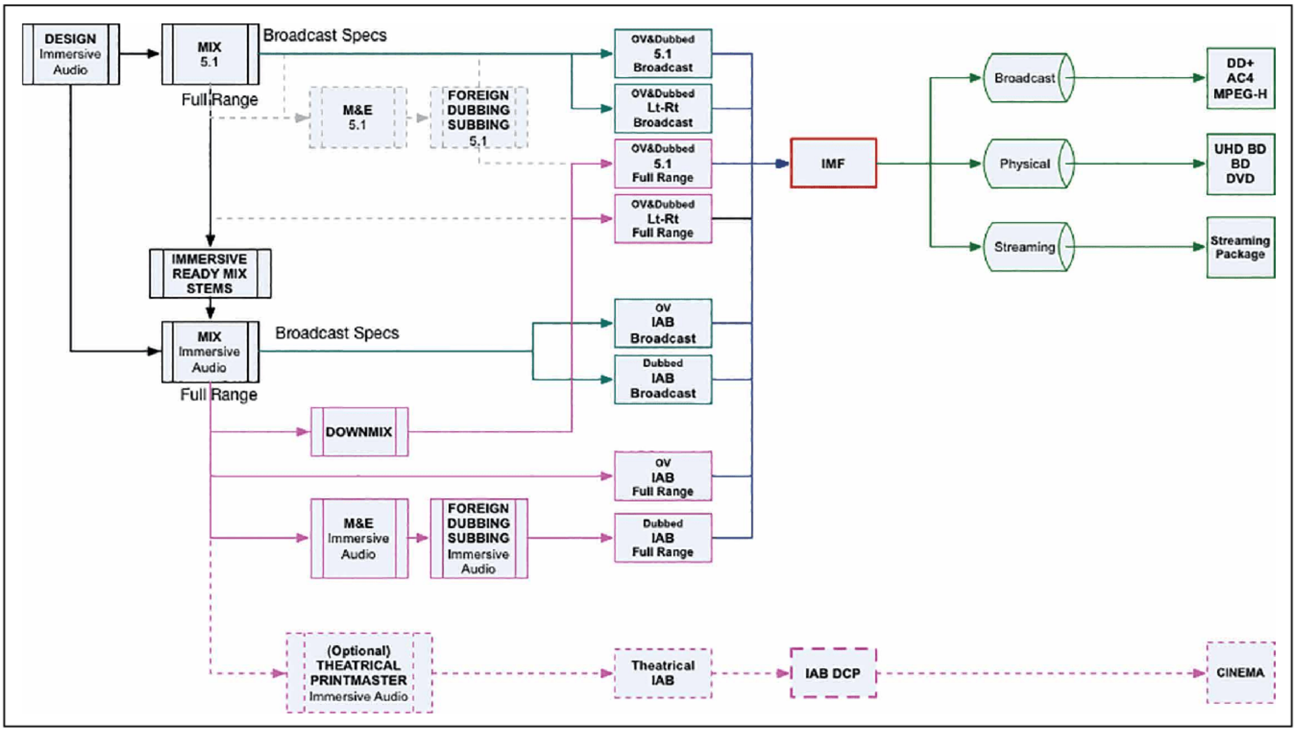


FIGURE 9. Alternative future-forward path workflow for television.

manufacturers are reticent to do the updates necessary to adopt all of its features.

To facilitate the Cinema IAB rollout, a constrained set of IAB features has been codified by ISDCF into SMPTE RDD 57 "ST 2098-2 IAB and Packaging Constraints: IAB Application Profile 1."[10] This constrains IAB to the features implemented in current Atmos systems, an approach that was agreed by all the manufacturers in this space. IAB Profile 1 gives creation tools and playback devices a defined target to achieve critical initial interoperability and a big green light for the rollout of IAB DCP.

To move from proprietary DCP deliveries to IAB DCP, (1) manufacturers need to issue updates for theater systems and mixing tools, (2) exhibition and mix stages have to be willing to install those updates, and (3) studios and other content providers need to feel comfortable that there will be no dark screens.

The good news is that most of the installed base of equipment in immersive theaters and mix stages is already IAB Profile 1 capable, and the major mastering facilities can already create IAB DCP. There remain around 1200 immersive theaters that need upgrades to be able to play an IAB Profile 1 DCP. The rollout has been hampered by the Covid-19 pandemic but is looking promising for early 2022.

IAB IMF Rollout

IAB IMF adoption is moving more slowly than cinema, as IMF itself is not yet integral to all content workflows. The IMF User Group, a similar body to ISDCF in the home entertainment world, is bringing together parties from diverse sectors that have a mutual interest in utilizing IMF in their workflows. The UG has created a number of best practices which are helping IMF to become a worldwide standard.

Many manufacturers in the IMF space have implemented or plan to implement IAB, but this is taking time, as other IMF features are also on their development roadmap. Several manufacturers successfully utilized beta software at the February 2020 SMPTE IMF plugfest to author and play IAB IMF, and several transcoders are now available that can read IAB and create deliverables. The advent of IAB Profile 1 is also important in the adoption of IAB in IMF.

Having IAB in IMF is a big win for mastering and B2B distribution. Currently, the multiple proprietary immersive file formats and codecs for home and broadcast are handled individually as separate deliveries outside of IMF. The integration of IAB into IMF means that one mezzanine package can be created that contains both standard and immersive audio, and any audio deliverable can be created from a single master.

Attaining Full Interoperability

The key to full interoperability is involvement. Users (e.g., content providers and distributors) need to work diligently with manufacturers to encourage the adoption and implementation of standards every step of the way. The cinema plugfests organized by ISDCF and the IMF plugfests organized by SMPTE are two

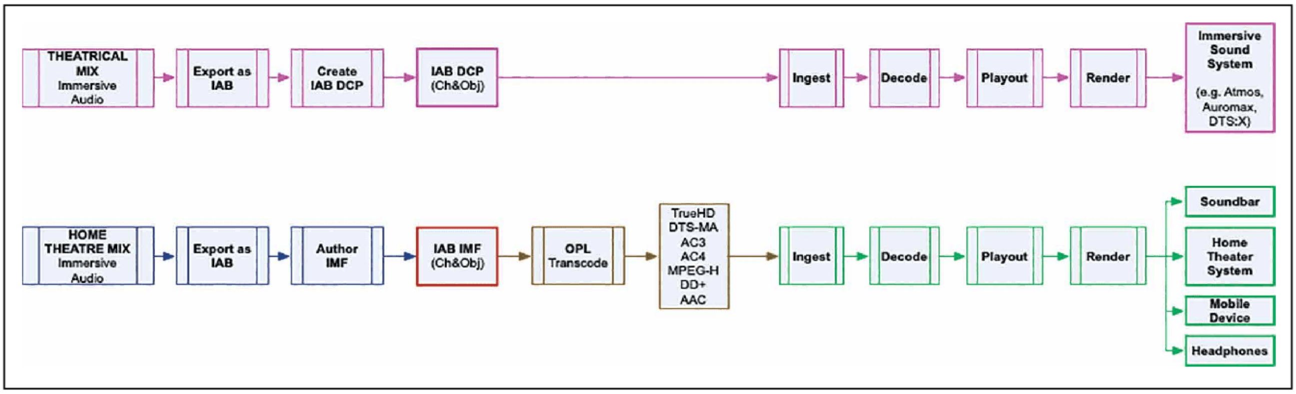


FIGURE 10. Delivering immersive audio with channels and objects.

excellent venues to continue the march toward full interoperability. The work that ISDCF and the IMF UG are doing to bring together the user communities and create best practices is fostering excellent cooperation and shared knowledge, but it is ultimately up to the users to push the use cases and beat the drum for full interoperability.

The business case for full interoperability is obvious. The efficiency and cost savings in creation and distribution, coupled with the increased availability of immersive content, is a win-win.

Playability and Translation

Having come up with this interoperable way to convey immersive audio, which can be delivered many ways and played through many sound systems, how can we guarantee that what the consumer hears is anything like what we mixed? The short answer is we cannot, but we can take measures to ensure that everything is there to allow it to happen.

Even 5.1 mixes, consisting of six known channels, do not sound the same in every sound system. Nor do proprietary immersive mixes played on multiple systems of that same brand. There are always some differences in the playback systems and the rooms.

Immersive sound systems have the challenge of rendering objects as well as reproducing channels. We have been dealing with channels for some time and have created a fair amount of consistency in their reproduction, but rendering objects is a much newer concept. Each sound system has its own approach.

The key component is the immersive audio renderer (what I like to call the "magic renderer"). This is a "black box" with a ton of DSP that has been programed with proprietary secret sauce. It interprets the channels and objects in realtime and directs them into the attached playback system. This could be a large theater sound system, a small theater sound system, a home theater system, a soundbar, a TV, a pad, a phone, or a set of headphones. That is a big variety!

Delivering Effectively

There are two basic concepts for the delivery of immersive audio. The first concept, illustrated in **Fig. 10**, is to mix with channels and objects and then carry those all the way through to the playback system, where the renderer does its magic and creates the best immersive sound it can.

A prime example of this concept is the cinema workflow. The IAB DCP contains channels and objects. The theater playback server feeds them to the renderer, which creates the playback channels that then go to the cinema processor and loudspeakers. The renderer knows where the loudspeakers are in the auditorium and renders accordingly. Since this is also how the immersive sound system works on a mixing stage, a reasonable translation is quite possible.

For home delivery, the IAB is transcoded into proprietary codecs that carry the channels and objects to home systems. The home playback system will have one or more licensed chips to decode and feed the channels and objects to its renderer. These are then rendered into playback channels and fed to the consumer's playback system, which as noted could be anything from a home theater system to a phone. Thus, the translation of the immersive mix to the consumer is potentially quite variable and depends greatly on how a given renderer renders the channels and objects.

The second concept, illustrated in **Fig. 11**, is to mix with channels and objects and then "prerender" these into a set of channels based on known loudspeaker configurations. This Soundfield Group of prerendered channels is then labeled per SMPTE ST 377-4,[5] ST 377-41[6] and ST 377-42,[7] packaged and delivered. A theatrical example of this method is IMAX 12.0.

This method can work well for consumer applications, since the delivery of channels is relatively straightforward through established pipelines. It can also lessen translation variables.

For example, 7.1.4 is a standard immersive soundfield configuration for the home. The exact loudspeaker setup and angles are defined in ITU BS.2159-8.[8] Excellent translation can be obtained by creating a 7.1.4 home mix

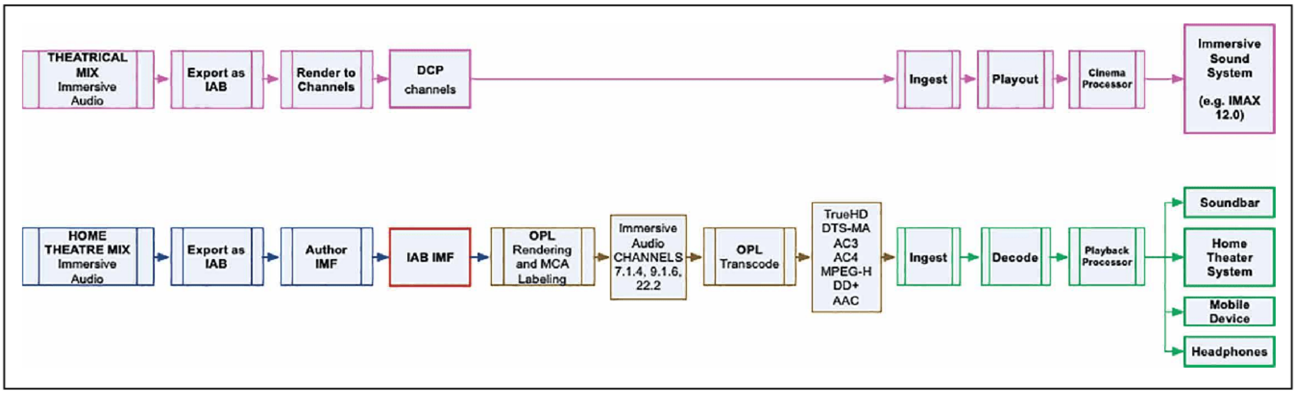


FIGURE 11. Delivering immersive audio as rendered channels.

using channels and objects and putting the resulting IAB into IMF. To make deliverables, a defined OPL is used to prerender the IAB using 7.1.4 as the target, yielding a 12-channel 7.1.4 soundfield group. This can be easily carried, coded, and decoded and will yield very good results if played into a home 7.1.4 system or smaller format device.

When considering how to deliver immersive audio, choose the best delivery system for the pipeline and expected playback system.

Futureproof Workflow Recommendation

We have explored a number of concepts and workflows related to the creation and distribution of immersive audio. **Figure 12** brings these together and illustrates a recommended futureproof workflow. Immersive audio is used in both the cinema and home workflows, parenting all immersive and standard format audio elements. A single IAB IMF is the source for multiple downstream deliverables directed by output profile lists (OPLs). OPLs are created as soon as a client is onboarded and are called each time the client requests deliverables. This fosters customization, automation, and repeatability.

Whether the delivery is DCP, IMF, codecs, or channels, having futureproofed immersive audio content ready is key. It also allows for more lucrative sales into multiple markets because the sales team knows what is available and can offer top shelf, right now.

Conclusion

We have over a decade of immersive audio experience in our industry. It has matured and is here to stay. Immersive sound adds significant realism and "viscerality" to a soundtrack. Most tentpole features and a number of television shows are natively mixed in immersive

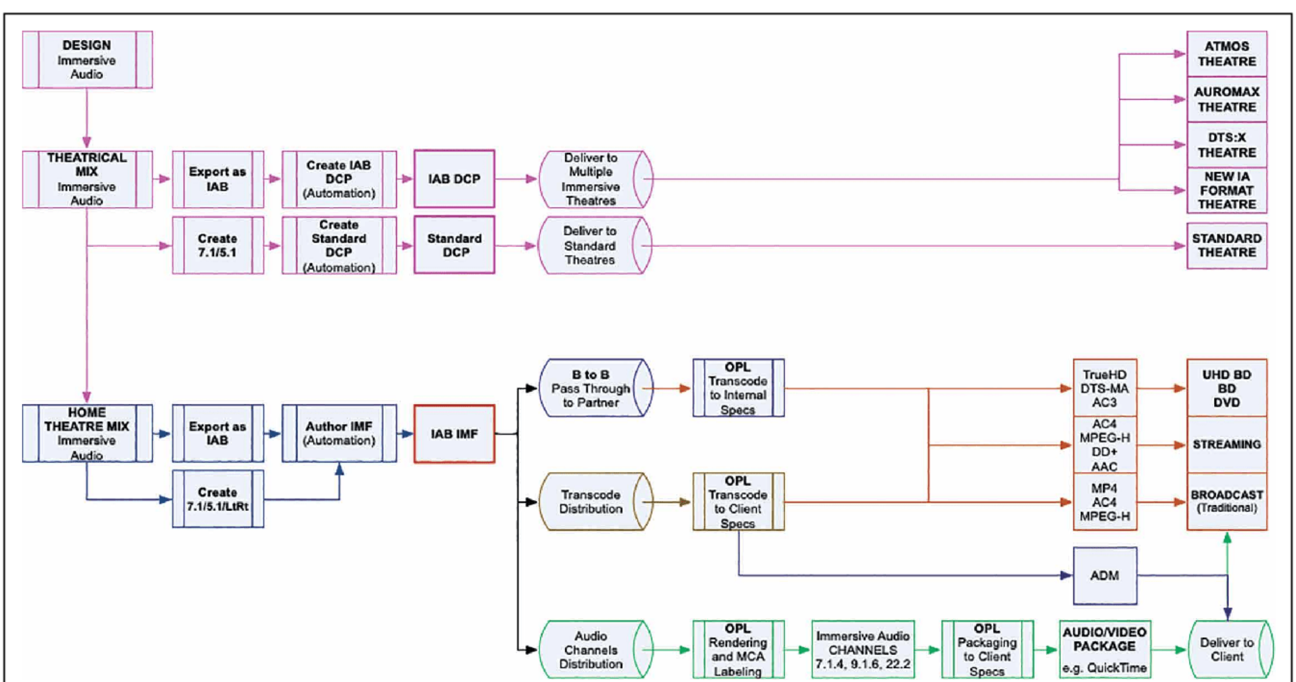


FIGURE 12. Futureproof workflow recommendation with automation and OPL.

sound. It is no longer more expensive to create than standard audio and has many tangible benefits, but its use has been hampered by short term, inefficient workflows, and proprietary deliverables. This article has examined the immersive audio workflows that have evolved for feature and television content, noting their inefficiencies and demonstrating how they can be streamlined to create for today and the future at the same time.

We are on the cusp of interoperable, single-inventory delivery using the ST 2098-2 IAB with standardized IAB DCP and IAB IMF. By using efficient and interoperable creation and distribution workflows, immersive audio is cost-effective and ultimately highly profitable. Creating the soundtrack natively in immersive audio gives a higher quality product for every audio format and opens the door for sales into any future market.

Key Takeaways

1) Plan, budget, and create for the big picture.
    - Create for the long-term delivery of both feature and television content to multiple distribution channels rather than the immediate, short-term window.
2) Designing and mixing for immersive audio from the beginning rather than adding it later….
    - Makes for a higher quality product.
    - Is more time-efficient.
    - Is more cost-efficient.
    - Allows immersive content to be delivered to any market at a moment's notice.
3) Futureproofing is the key to building a valuable, versatile, and salable library.
    - Shooting and finishing in 4K/HDR.
    - Designing and mixing in immersive audio.
    - Authoring all into IMF to create a single mezzanine master.
4) Interoperability is the key to efficient distribution and a reliable consumer experience.
    - Must be nurtured and finessed to realize its full potential.
    - Content creators and distributors must make the business case for it.
    - Manufacturers must be encouraged to adopt and implement standards to achieve full interoperability.
5) Automation is the key to efficient distribution; interoperability is the key to automation.
6) While it may appear to be more expensive to futureproof up front, it is actually far cheaper in the long run.

References

1. SMPTE, ST 2098-2:2021, "Immersive Audio Bitstream Specification."
2. SMPTE, ST 429-18:2019, "D-Cinema Packaging-Immersive Audio Track File."
3. SMPTE, ST 429-19:2019, "D-Cinema Packaging-DCP Operational Constraints for Immersive Audio."
4. SMPTE, ST 2067-201:2019, "Immersive Audio Bitstream Level 0 Plug-in."
5. SMPTE, ST 377-4:2021, "MXF Multichannel Audio Labeling Framework."
6. SMPTE, ST 377-41:2021, "MXF Multichannel Audio Controlled Vocabulary."
7. SMPTE, ST 377-42:2021, "MCA Label Controlled Vocabulary."
8. International Telecommunication Union-Radiocommunication (ITU-R), BS 2159-8 (2019), "Multichannel sound technology in home and broadcasting applications."
9. SMPTE, RDD 52:2020 "D-Cinema Packaging-SMPTE DCP Bv2.1 Application Profile."
10. SMPTE, RDD 57:2021 "ST 2098-2 Immersive Audio Bitstream and Packaging Constraints: IAB Application Profile 1."

About the Author

Brian A. Vessa is a dedicated audio professional with over 35 years of experience in the industry. After attending the University of California Los Angeles (UCLA) Engineering School, Los Angeles, CA, USA, he became a recording engineer, producing albums and recording orchestras. He transitioned into film as a music editor and sound editor, became a rerecording mixer at Cannon Films, Los Angeles, CA, and Metro-Goldwyn-Mayer (MGM), Los Angeles, CA, and then handled audio restoration at NT Audio, Santa Monica, CA, USA. He was hired by Sony Pictures Entertainment, Culver City, CA, USA, in 1998, and today is their executive director of digital audio mastering and representative to Digital Cinema Initiatives (DCI). He is a member of the Academy Sound Branch, SMPTE, and the Audio Engineering Society (AES). He is the founding chair of the SMPTE TC-25CSS Technology Committee on Cinema Sound Systems and serves on many SMPTE and AES committees. He was a member of the original committee that conceived and created the Interoperable Mastering Format (IMF), helping write a fully detailed standard for IMF prior to it being brought to SMPTE. He has been instrumental in standardizing the Immersive Audio Distribution Bitstream (IAB) for Cinema and IMF as well as a host of related standards. He is a past chair of the DCI Technical Committee and has written many audio specifications, including a white paper on near-field mixing for home theater that has been widely adopted.