

LATEST DEVELOPMENTS IN LOW BIT-RATE AND HIGH-QUALITY MULTI-CHANNEL AUTOMOTIVE AUDIO

MARC GAYER¹ AND ROBERT BLEIDT²

¹ *Fraunhofer Institute for Integrated Circuits, Am Wolfsmantel 33, 91058 Erlangen, Germany*

marc.gayer@iis.fraunhofer.de

² *Fraunhofer USA Digital Media Technologies, 100 Century Center Court, Suite 504, San Jose, CA 95112, USA*

robert.bleidt@dmf.fraunhofer.org

In 2006, the ISO/MPEG standardization group finalized the MPEG Surround specification, which allows for efficient and backward compatible compression of high-quality multi-channel sound using parametric coding techniques. Since then the MPEG Surround standard has been included in a number of digital broadcast standards as the multi-channel audio codec of choice and its use in electronic music distribution and media players is underway. This paper summarizes the technical background of MPEG Surround and discusses its use in digital broadcast systems with a special focus on automotive applications such as car radio receivers, rear seat entertainment, and portable music player docking stations in the car.

INTRODUCTION

With the increasing availability of high quality and fully discrete multi-channel reproduction systems in cars there is a desire of car audio equipment manufacturers and radio broadcasters to improve their offerings with new features that utilize these surround loudspeaker setups. The MPEG Surround audio coding standard [1] provides low bitrate and high quality compression of multi-channel audio content and avoids the need for a transmission bandwidth consuming simulcast through the use of a backward compatible downmix. It has been standardized or is currently under standardization for a number of broadcast systems that are relevant for the automotive sector.

This paper discusses the technical background and features of MPEG Surround and provides information on the status of its adoption and applicability in numerous broadcast scenarios. The usage of MPEG Surround in electronic music distribution to portable music players and digital storage media, as well as a MPEG Surround car docking station design with multi-channel decoding capability for this content will be discussed. Implementation aspects are explained and a possible combination of multi-channel audio coding with stereo-to-surround upmixing technologies is presented.

1 MPEG SURROUND TECHNICAL BACKGROUND

This chapter provides a brief description of the MPEG Surround technology, which has been discussed in more detail in previous AES publications [2][3][4][5][6].

1.1 Concept of Spatial Audio Coding

MPEG Surround is based on a principle called Spatial Audio Coding (SAC), which exploits the perceptual inter-channel irrelevance in multi-channel audio signals to achieve higher compression rates. These inter-channel dependencies can be captured in terms of spatial cues, which are parameters describing the spatial image of the multi-channel audio signal. Spatial cues typically include level/intensity and phase differences as well as measures of coherence/correlation between channels and they can be represented in an extremely compact way.

Figure 1 shows the basic working principle of an MPEG Surround encoder. The multi-channel audio signal is fed into the MPEG Surround encoder module which extracts the spatial cues and at the same time generates a downmix that usually is a backward compatible mono- or stereophonic signal. MPEG Surround also offers the possibility to use an artistic downmix, reflecting the choices of a sound engineer, instead of the automatic downmix generated by the MPEG Surround encoder. While the downmix signal can also be transmitted as uncompressed PCM it is desirable for low-bitrate applications such as broadcast systems to use a compression audio codec to achieve a bandwidth efficient transmission of the audio content. The exact interaction of MPEG Surround with audio codecs such as MPEG-4 AAC-LC or HE-AAC [15] and MPEG Layer-2 [16] will be discussed in later chapters of this paper. The compression of the downmix signal can be done by an unaltered audio codec with the ability to reuse existing implementations. The MPEG Surround

spatial cues are transmitted in a low-bitrate side channel, e.g. the ancillary data portion of the downmix bitstream.

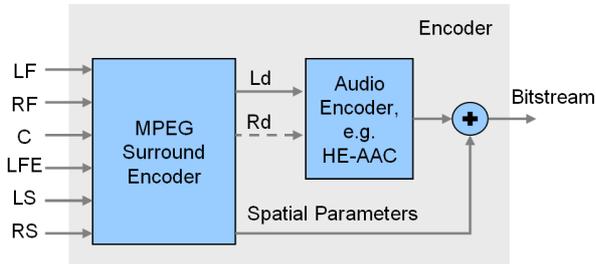


Figure 1: MPEG Surround encoder principle

The bitstream containing MPEG Surround data is then transmitted via broadcast or storage media to a decoder as pictured in Figure 2. The mono or stereo downmix is decoded by the core audio decoder and used by the MPEG Surround decoder module together with the spatial parameters to generate a psychoacoustically discrete multi-channel audio signal.

A legacy decoder, which is not capable of making use or detecting the spatial parameters that are hidden in a backward compatible way in the bitstream, will just decode the mono or stereo downmix. Therefore the simulcast of the multi-channel and the downmix signal is not required with MPEG Surround.

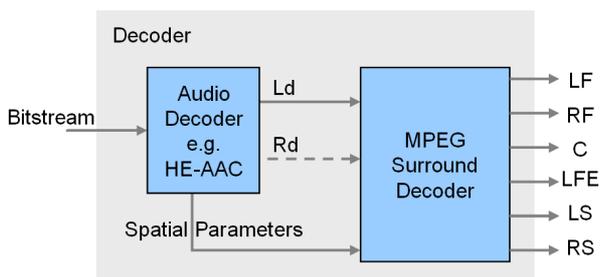


Figure 2: MPEG Surround decoder principle

1.2 Bitrate vs. Audio Quality Scalability

To increase its usability in as many applications as possible, MPEG Surround covers a broad range both in terms of spatial side information rate and multi-channel audio quality. Figure 3 shows the high flexibility of the MPEG Surround operation point between transparency on the one hand and very low spatial parameter bitrate on the other hand.

For digital broadcast or internet streaming applications the focus will usually be on low bitrates which consequently means low spatial parameter bitrates. MPEG Surround offers good multi-channel audio quality at spatial side info rates of only 2 to 5 kbit/s. It even offers a non-guided mode, also called 'enhanced matrix mode', where no spatial side information needs

to be transmitted at all and which still provides better quality than legacy matrix-based surround systems [7]. This allows the use of MPEG Surround even in systems where the additional spatial side information can not be transmitted and also in analog systems.

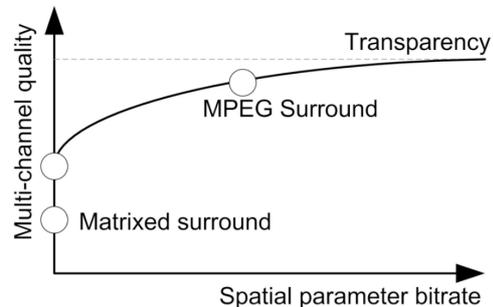


Figure 3: Rate/distortion scalability

Further scalability in the low-bitrate range from 2 kbit/s up to about 32 kbit/s can be achieved by sophisticated spatial parameter choice, parameter frequency resolution, time resolution, and quantization resolution on the encoder side.

In electronic music distribution where for most users audio quality is of higher importance than small file sizes, the system can reach perceptual transparency by increasing the spatial side information rate through the use of residual coding, which is a wave-form coding extension that codes the error signal originating from the limits of the purely parametric model.

1.3 Binaural Rendering Mode

MPEG Surround supports a binaural rendering mode to experience a virtual multi-channel loudspeaker setup over conventional stereo headphones [3]. This operation mode is of particular interest for automotive rear-seat entertainment applications or for mobile devices such as portable music players or mobile TV receivers.

Two different use cases are supported. In the first case, referred to as "3D", the transmitted stereo downmix is converted to a 3D headphone signal at the encoder side, accompanied by spatial parameters. On the decoder side legacy stereo devices will automatically render a 3D headphone output. If the same bit stream is decoded by an MPEG Surround decoder, the 3D downmix can be converted to a standard multi-channel output for loudspeaker playback using the transmitted spatial parameters.

In the second, probably more common use case, a conventional MPEG Surround downmix and spatial parameter bitstream is decoded using the so-called "binaural decoding" mode. Here the binaural synthesis is applied at the decoder side using efficient algorithms that avoid a processing power consuming intermediate decoding of the signal to the full multi-channel representation. Instead, the binaural output is generated

by a 2-in, 2-out matrix operation which applies the transmitted spatial parameters and the HRTF (Head Related Transfer Function) data in one step and thus saves a significant amount of processing power and data memory.

1.4 MPEG Surround Profiles and Levels

Technologies such as MPEG Surround are fully standardized in ISO/MPEG documents [1]. Other standards that make use of MPEG Surround such as DAB+, DMB, Digital Radio Mondiale or others that will be mentioned in chapter 2 of this paper usually only give a reference to the applicable MPEG standards document and, if required, provide additional remarks. To simplify this procedure MPEG Surround ‘profiles’ and ‘levels’ have been pre-defined, which take into account that MPEG Surround is an extremely flexible technology which can cover a multitude of application scenarios with diverse requirements with regard to functionality, available spatial side information rate and processing power, number of audio channels, and so on. Currently a single “Baseline MPEG Surround Profile” is available, which contains the following tools, which are partly explained in this paper and in more detail in the referenced literature:

- artistic downmix functionality
- matrix compatibility
- enhanced matrix mode decoding
- temporal shaping
- residual coding (not including artistic downmix residuals)
- binaural decoding
- 3D audio decoding
- low power decoding

This profile is further subdivided into six hierarchical levels that come with increasing number of output channels, range of sampling rates, bandwidth of the residual signal, and computational complexity when advancing to a higher level. The following table lists the supported tree configurations, maximum number of audio output channels, and availability of residual coding for each level.

Level 1 allows for stereo output of multi-channel content with and without the use of binaural rendering techniques. It is ideally suited for mobile devices that are equipped with stereo loudspeakers or headphones and exhibits very low computational complexity. This level includes the ability to upmix from a mono based operation to stereo loudspeaker reproduction, the 5-1-2 mode with 5-1-5 streams. Levels 2 and 3 additionally offer discrete 5.1 output for loudspeakers. Levels 4 and 5 extend to 7.1 output, while Level 6 can handle up to 32 channels and operate at up to 96 kHz sampling rate.

Level	Tree configurations	Max. output chan.	Max. Fs [kHz]	Residual coding
1	515, 525, 727	2.0	48	n/a
2	515, 525, 727	5.1	48	n/a
3	515, 525, 727	5.1	48	yes
4	515, 525, 727	7.1	48	yes
5	515, 525, 757, 727	7.1	48	yes
6	515, 525, 757, 727, arbitrary tree extension	32 incl. LFE	96	yes

Table 1: Levels of the MPEG Surround Baseline Profile

2 ADOPTION OF MPEG SURROUND IN DIGITAL BROADCAST STANDARDS FOR AUTOMOTIVE USE

This chapter provides an overview on the adoption of MPEG Surround in various already deployed broadcast systems or ones that are currently under standardization. For a number of systems the use of MPEG Surround has already been defined in the respective standards. For other systems the applicability of the MPEG Surround coding algorithm has been shown in trial implementations or it is inherent due to the use of an appropriate core codec such as MPEG HE-AAC and the backward-compatible bitstream transport feature of MPEG Surround.

The overall bitrate that is required for multi-channel audio transmission when using MPEG Surround in a low-bitrate, i.e. non-residual, mode is the same as the typical stereo bitrates of the respective broadcast standards since in these cases MPEG Surround will operate in a mode that as a typical example only requires around 3-5 kbit/s for the spatial side information.

2.1 DAB (Digital Audio Broadcasting)

DAB, developed by EUREKA Project 147 and first deployed in 1995, uses MPEG-1 Layer-2 audio coding and operates in the terrestrial Band III (174-230MHz) and L Band (1452-1492MHz). DAB has seen widespread use with locally varying success and acceptance in Europe, Canada, and China. Typical audio bitrates are 128 to 192 kbit/s for stereo 48 kHz content. The combination of the Layer-2 audio codec with

MPEG Surround has recently been added to DAB under the name "DAB Surround". DAB Surround enables the backward-compatible transmission of high quality multichannel audio for DAB, which is a rather old but well established system that has millions of receivers in use.

2.2 DAB+

Using the exact same transmission and channel coding techniques as traditional DAB and also incorporating all data services known in DAB, the new DAB+ standard replaces the MPEG-1 Layer-2 audio codec with the much more efficient MPEG-4 HE-AAC v2 codec. The AAC core audio frame length in DAB+ is 960 samples, which leads to a frame length of 1920 in HE-AAC mode. Additional techniques such as superframing, Reed Solomon and Fire Codes as well as Cyclic Redundancy Checks provide means for error protection and correction and assure backward compatibility to the DAB data stream formatting and framing. Typical audio bitrates for HE-AAC v2 as used in DAB+ are only one third of the ones used for Layer-2 coding and subsequently the number of transmitted programs can be tripled in DAB+ compared to DAB. Standardized in 2007 in ETSI TS 102 563, MPEG Surround is the only option for multi-channel audio coding for DAB+ and the MPEG Surround Levels 2 and 3 of the Baseline Profile will be used in DAB+.

DAB+ services will first be established in Europe and the Australian Government has announced DAB+ as the official digital radio standard in the country.

2.3 DRM (Digital Radio Mondiale) and DRM+

DRM, first deployed in 2003 and standardized in ETSI ES 201 980, is a digital radio broadcast system for transmission frequencies below 30 MHz, i.e. in terrestrial long, medium, and short wave AM bands. Currently under standardization, DRM+ extends the possible transmission range and bandwidth and will operate in all the broadcasting bands below 120 MHz. All data services used in DAB can be used also in DRM under the restriction of a much lower bit rate. Currently DRM is mainly used for large coverage areas by shortwave broadcasters such as BBC, Voice of America, Radio France, Voice of Russia, or Deutsche Welle. Interest from car radio manufacturers is increasing though DRM still lacks cheap mass market receivers and widespread use. The use of DRM in markets such as Russia and China will help to solve the receiver availability and acceptance problem.

DRM mainly uses HE-AAC v2 at rather low bitrates of around 24 kbit/s with an AAC core frame length of 960 samples for audio coding but also allows HVXC and CELP for pure speech content. Support for HVXC and CELP has been dropped in DRM+, which also offers higher possible data rates due to the wider channel

spacing in FM bands. Notable is the existence of a special version of the Parametric Stereo (PS) coding tool in the HE-AAC v2 codec of DRM. Recently, also the MPEG-4 version of PS has been allowed for DRM which makes it necessary to implement both PS versions in a DRM audio decoder as long as not all DRM broadcast audio encoders have moved to the MPEG-4 version of PS.

The next version of the DRM and DRM+ standard is expected to be finalized in 2009 and states MPEG Surround as the multi-channel audio codec of choice for this system with up to 7.1 output channels.

2.4 DMB (Digital Multimedia Broadcasting)

Developed as an extension to DAB for mobile TV applications, Digital Multimedia Broadcasting (DMB) uses MPEG-4 AVC (H.264) video coding and MPEG-4 HE-AAC v2 with a core frame length of 1024 samples for audio. MPEG-4 BSAC (Bit-Sliced Arithmetic Coding) is only required for audio in the South Korean market but nonetheless needs to be part of the decoder implementation if world wide usability of a receiver design is desired. Audio-only DMB radio services are planned in France. In 2009, an updated version of the DMB standard ETSI TS 102 428 will specify MPEG Surround for multi-channel audio transmissions and also the 'Technical Specifications for DMB Digital Radio Deployment in France' state MPEG Surround as the codec for multi-channel audio transmission.

2.5 HD Radio

HD Radio has been designed for the U.S. market for the digitization of the FM (88-108 MHz) and AM (0.5-1.6 MHz) broadcast bands by the U.S. company iBiquity. It offers a hybrid mode that makes it possible to transmit analog and digital signals simultaneously on the same frequency.

Although HD Radio is currently transmitted by about 1,800 stations in the U.S., receiver sales have been slow to develop, and HD Radio is not yet a standard feature of factory car head units. The simultaneous analog transmission of HD Radio allows for a gradual transition to digital broadcasts, which has the unfortunate consequence of inspiring no real urgency for consumers to replace existing receivers or for automotive manufacturers to include HD reception capability.

While the National Radio Systems Committee (a joint NAB/CEA committee that prepares digital broadcast standards for consideration by the U. S. Federal Communications Commission) has considered the adoption of surround for HD Radio, they chose not to conduct any technical evaluation of proposed systems, and instead have left the choice of surround system up to each broadcast station. MPEG Surround was

proposed for this use, along with several older matrix-based systems.

Some stations have chosen to offer surround programming using these matrix systems since they operate on analog signals and thus require no changes to the stereo transmission path, as long as the core codec is operated at a bitrate sufficient to preserve the stereo image.

MPEG Surround, being a digital system, requires the surround bitstream to be inserted into the transmitted bitstream. In return for the minor additional bitrate of this stream, MPEG Surround offers superior surround fidelity compared to matrix systems, and has been judged similar in quality at 96 kbit/s to discrete 6-channel transmission using the AC-3 codec at 384 kbit/s.

The technical details of the stereo core audio codec of HD Radio cannot be discussed here since it is proprietary. However, it can be noted MPEG Surround has been successfully integrated with the codec in both transmission and prototype receiving equipment, and successful over-the-air transmission trials of a system that combines the HD Radio stereo codec and MPEG Surround have been conducted in the Boston area.

At present, most broadcasters are focused on securing consumer interest in HD radio and developing an installed base of receivers. There is also commercial interest in dividing the 96 kbit/s payload of the HD Radio signal into two or perhaps three programs carried at a lower bitrate and thus lower audio quality.

Interestingly, the 5-1-5 mode of MPEG Surround means it is able to offer greater audio quality at lower bitrates, by transmitting a mono downmix signal. MPEG Surround has been demonstrated to offer good surround images at total bitrates as low as 24 kbit/s in this mode, and at 32 kbit/s in the 5-2-5 mode with a stereo downmix.

Despite these technical advantages, and the enhanced listening experience MPEG Surround offers, broadcasters are deferring surround broadcasts until they achieve success in developing an audience for their stereo or multicast services.

2.6 Satellite Radio

The merging XM and Sirius systems for American satellite radio have experimented with offering one or two matrix surround program channels in the past, but have not announced their plans for surround in the future combined service. Similarly, the Ondas system being deployed in Europe has not announced any decisions on surround programming. Both of these systems could make use of MPEG Surround for radio as well as future video programming.

2.7 DVB-H

An alternative to DMB in Europe is DVB-H [14], which extends the DVB-T technology in a system suitable for mobile reception. Fraunhofer IIS has operated a DVB-H broadcast station since 2005 and demonstrated DVB-H reception on mobile devices with MPEG Surround binaural decoding. MPEG Surround is currently being standardized for use as the surround codec in DVB-H.

2.8 ATSC Mobile

Similar in many conceptual respects to DVB-H, ATSC Mobile is a television service using a portion of existing U.S. terrestrial broadcast channels to deliver a robust signal to mobile devices. The Advanced Television Systems Committee, similar to the NRSC, prepares American digital TV standards for the FCC.

The ATSC-Mobile system uses H.264/AVC coding for the video signal and HE-AAC v2 coding for the audio program. A wide range of bitrates are considered for both, with the resolution and quality levels chosen by each broadcast station.

Although surround audio could offer an audio experience comparable to surround home viewing in the car, and a good rendering of surround to binaural listeners while mobile, the ATSC has chosen to defer consideration of surround until a potential future revision of the standard. Broadcasters are still uncertain of the formats and genres of content they will broadcast, and development of the standard has been focused and accelerated in order to compete with similar existing or expected television services offered by mobile wireless carriers.

3 AUTOMOTIVE DOCKING STATIONS FOR MPEG SURROUND PLAYBACK FROM PORTABLE MEDIA PLAYERS

In addition to broadcast signals, MPEG Surround content may also enter the car on portable media players. MPEG Surround is currently being proposed to the music industry as a practical means of distributing surround music to consumers, with the car being a natural surround listening environment.

3.1 Market Failure of Previous Surround Formats

Previous attempts to offer surround content have been based on physical media formats such as DVD-Audio or Super Audio CD. Although some automotive head unit manufacturers, notably Panasonic (with its Acura ELS Surround system) and Bose, have offered high-quality discrete DVD-Audio playback, others have not offered surround or have elected to feed all playback signals, both stereo and surround, through an upmixing engine that reduces the surround image quality. This lack of broad support for proper discrete surround playback is perhaps due to the limited amount of content available in the DVD-Audio and SACD formats.

These formats have required the music industry to provide duplicate physical media through their distribution channels to consumers, which is an expensive and inefficient process, resulting in a limited number (perhaps 10,000 titles overall) of surround titles being produced. While it could be argued that consumers may have migrated over time to one new format offering higher quality, the competition and resulting confusion between SACD and DVD-Audio made this impossible.

3.2 MPEG Surround for Electronic Music Distribution

MPEG Surround is being proposed to music labels and artists as a way to offer surround music to consumers through electronic distribution such as iTunes or other web music stores. This fits with the new web purchase model for music and allows labels or aggregators to offer one file for stereo and surround playback if they wish, eliminating even the trivial costs of duplicate releases in electronic media.

The trend to web music stores has largely been driven by the success of the iPod and similar personal music players. As these devices have grown in capacity, they have become the consumer's primary music library instead of a device to be downloaded from other PCs or storage devices. Consumers now expect to "take all their music with them" and to be able to dock these devices in their cars for playback.

As MPEG Surround offers artistic downmix capability and the surround bitstream can be hidden in a standard stereo file, legacy or portable devices see no difference between an MPEG Surround encoded file and a stereo file, playing both in stereo as the artist intended. When played through a MPEG Surround decoder in a docking station, home AVR, or car head unit, the full discrete-quality surround mix is reproduced.

Given the recent trend in the market towards higher quality and higher bitrates for music downloads, with many services offering stereo files at 256 or 320 kbit/s, it has been suggested that MPEG Surround content be offered at a similar quality level, with perhaps 200-250 kbit/s used for a stereo AAC downmix and 50-100 kbit/s for the surround bitstream.

3.3 Fraunhofer Prototype Docking Station

As a proof of the viability of this proposal, Fraunhofer constructed a prototype docking station for living-room use. MPEG Surround files on an iPod are played in stereo when listening through the iPod's earphones. When the iPod is docked, the files are read and decoded to surround by an ARM-based processor in the docking station.

To also illustrate how the docking station concept could work with existing receivers in the consumer's home, this prototype then encodes the surround signals with a

640 kbit/s AC-3 encoder for transmission over a Toslink optical cable to the existing AC-3 decoder in the receiver.

Thus, consumers could simply purchase such docking stations and plug them into their existing AVRs with a single cable to enjoy music they may have already purchased in surround. Of course, such stations could also include HDMI outputs with PCM surround signals for improved reproduction with the growing number of AVRs that support HDMI 1.3 inputs.



Figure 4: Prototype iPod Docking Station with MPEG Surround Decoder

3.4 Docking Stations for the Automotive Environment

The concept of this docking station can be easily extended to the car environment. The MPEG Surround decoder is available in both fixed-point and floating-point versions for a variety of embedded processors or DSP chips, with the prototype ARM-based design consuming a total 3 watts of power.

The ideal approach would be to combine the MPEG Surround decoder with an automotive head unit's existing stereo decoding software. Alternately, for systems (primarily aftermarket products) that have AC-3 Toslink inputs, a design similar to the living-room prototype could be connected.

Such capability would offer consumers the ease of listening to their "personal music library" in stereo over earphones and in surround while docked in their cars. Of course, as consumers upgrade or replace their music players, new ones could be equipped with decoders using MPEG Surround's binaural mode to allow surround playback over earphones. Alternatively, rear-seat entertainment docking stations could also offer the binaural reproduction mode of MPEG Surround, not only for pure audio but also for movies and other multimedia entertainment.

4 IMPLEMENTATION ASPECTS

After having shown the widespread use of the MPEG Surround technology in combination with different mono or stereo core audio codecs for a number of digital broadcast systems and other applications, the following sections of this paper will provide information on implementation aspects of these technologies with a special focus on embedded implementations of the decoder side as they will be required for example for car radio receiver designs.

4.1 Interfacing to the MPEG Surround Module

As explained previously the MPEG Surround module can be combined with various core audio codecs such as MPEG Layer-2, AAC-LC and HE-AAC, but also be used with plain PCM or the HD Radio codec. In the following the various existing interfacing options to the MPEG Surround decoder module will be explained.

4.1.1 QMF domain interface

The MPEG Surround decoder module employs a QMF analysis (Quadrature Mirror Filter bank) to obtain a uniformly-distributed, oversampled frequency representation of the audio signal before applying subsequent spatial processing steps. Since the SBR (Spectral Band Replication) tool of an HE-AAC decoder also works in the QMF domain and performs a QMF synthesis at its output stage, the two 64 band QMF synthesis and analysis banks can simply be omitted in the case where the MPEG Surround decoder module is to be combined with an HE-AAC core codec. The possibility to avoid these QMF processing steps, an important factor in the MPEG Surround system design, saves a significant amount of processing power which is crucial for low resource embedded implementations as they will be required for mass-market products using MPEG Surround such as in the automotive sector. Whether this is also applicable for a combination of MPEG Surround with the HD Radio codec depends on the exact internal design of the HD codec but it is publicly known that it does use SBR technology.

4.1.2 PCM time domain interface

For other combinations, such as MPEG Surround with Layer-2 for DAB Surround or with AAC-LC the core audio decoder has to fully process the data to PCM time domain representation and the MPEG Surround decoder module has to perform the QMF analysis.

Consequently, a multi-purpose MPEG Surround decoder module, usable for example for DAB and also DAB+ Surround, needs to provide two interface options to the core codec, one in the PCM time domain and one in the QMF domain, where for the latter it also has to be considered whether both, HE-AAC core codec and MPEG Surround module shall operate in the high

quality mode with a fully complex-valued QMF or in the low power mode where the QMF bank is only real-valued [2].

4.1.3 Frame length considerations

For the cases where an HE-AAC core codec interfaces to the MPEG Surround module in the QMF domain the implementation has to take into account the different number of QMF time slots used for the AAC core frame length 960, resulting in an SBR output of 1920 samples, with 30 QMF time slots and for the AAC core frame length 1024, resulting in an SBR output of 2048 samples, with 32 QMF time slots, respectively. Consequently, any standard compliant MPS decoder module needs to be parameterizable in that respect so that it can be combined e.g. with both, a HE-AAC core codec in DAB+/DRM flavor and in DMB/DVB-H flavor.

For cases where the interface of the core codec to the MPEG Surround decoder needs to be in the time domain it does make sense to align core codec and MPEG Surround framing in a way so that one MPEG Surround frame corresponds to an integer number of core codec frames, for example two. In the case of Layer-2 with a frame length of 1152 samples this would lead to 36 QMF time slots and for AAC-LC it is again 32 or 30 QMF time slots, just as for one HE-AAC core frame.

While the aforementioned number of QMF time slots and the alignment of core codec frames to MPEG Surround frames have proven to be reasonable it should be noted that these are not standardized but rather tuning parameters of the MPEG Surround encoder, which every MPEG Surround decoder module has to be able to cope with.

4.2 System Design Considerations

From a high level system design point of view for the audio decoder part of a digital radio car receiver the integration of the MPEG Surround decoder module looks comparably simple. Figure 5 shows the building blocks for the audio decoder subsystem of a combined DAB, DAB+, DMB, DRM, and DRM+ radio receiver as it is required for upcoming automotive products. The MPEG Surround decoder module can be found in the upper right of the diagram and is fed with mono or stereo downmix data in the PCM time domain from the Layer-2 or AAC-LC decoder or with QMF domain downmix data from HE-AAC's SBR module. It will output multi-channel audio, usually in 5.1 format, or a 2-channel binaural signal for rear-seat headphone entertainment purposes.

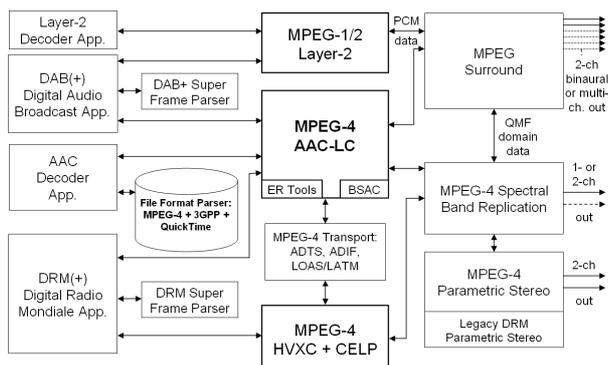


Figure 5: DAB/DAB+/DMB/DRM/DRM+ audio decoder modules

The building blocks in Figure 5 contribute in the following way to surround capable radio receivers:

For DAB Surround, the MPEG Layer-2 and the MPEG Surround modules are required besides the DAB decoder framework application.

DAB+ Surround requires the AAC-LC and SBR modules for mono or stereo HE-AAC operation and additionally the PS module for stereo HE-AAC v2. PS is replaced by MPEG Surround for multi-channel or binaural virtual surround operation. Besides that the DAB+ super frame parser including Reed-Solomon, Fire Code and CRC processing is necessary.

A Digital Radio Mondiale decoder consists of the AAC-LC core decoder with support for MPEG-4 Error Resilience (ER) [15], the SBR module and two flavors of PS as mentioned in chapter 2.3. The MPEG Surround module is used in multi-channel mode, in which PS is never active. The HVXC and CELP voice codecs are only required for DRM but not for DRM+ while both operation modes require the DRM super frame parser that also contains CRC and unequal error protection routines.

For DMB audio decoding only the AAC-LC, SBR, and PS modules are used. Again, PS is dropped in favor of MPEG Surround for multi-channel operation. DMB specifies an MPEG-2 transport stream layer that multiplexes audio, video, and auxiliary data into one single program but this shall not be discussed in more detail in this paper.

A multi-purpose audio decoder tool box for digital broadcast car radio receivers that is based on MPEG technology needs to offer the aforementioned modules in all required flavors regarding AAC core frame lengths, QMF time slots, and so on and needs to be highly configurable at compile and reconfigurable at run time so that an efficient re-use of the software modules for the various coding modes can be achieved.

4.3 Availability of MPEG Surround software

All MPEG Audio algorithms including the rather new MPEG Surround codec are available as reference

implementations from MPEG in the programming language C/C++. These reference software implementations are fully functional and standard compliant but mostly not very optimized in terms of processing power and memory consumption. Since MPEG Audio usually only standardizes bitstream syntax and decoder functionality, encoder implementations may vary considerably in the audio quality that they offer at a given bitrate.

Due to its foreseeable widespread use more highly optimized MPEG Surround encoder and decoder implementations are available already for PC platforms and also for many common embedded and digital signal processors (DSPs) as they are used in automotive digital radio receiver designs. If carefully implemented, DSP implementations of MPEG Audio algorithms can well achieve full standard compliance and high audio quality even on pure fixed-point arithmetic logic units as they are present in many of today's high volume embedded processor families used in the automotive sector.

5 RELATED TECHNOLOGIES

Much of today's available audio content unfortunately is not readily available in a multi-channel mix, although many songs or albums have had surround mixes prepared for potential SACD or DVD-Audio release. Multichannel content is required by an MPEG Surround codec to make full use of this technology. While it is absolutely possible to play back plain stereo content on MPEG Surround enabled receiver devices, there is an obvious desire to exploit the possibilities provided by MPEG Surround and thus offer an added value to consumers.

The use of upmixing algorithms that generate multi-channel audio from stereo source material is one way to solve this problem [9]. The upmix procedure can happen without using MPEG Surround on the receiver side, taking stereo audio from a digital radio receiver's audio codec, from an analog radio receiver, or storage media. On the other hand, upmixing on the broadcaster side in the studio and subsequent use of MPEG Surround for transmission of the newly generated multi-channel content has the advantage that the upmix process only has to happen once on the sending end, with highest possible quality due to basically unlimited processing power resources, and does not have to be performed in each receiver, thus saving resources and costs in automotive receiver devices.

One upmix algorithm that is in constant use in combination with MPEG Surround by German broadcasters for their DAB radio services is Fraunhofer's SX Pro.

6 CONCLUSIONS

Despite being a comparably young member of the MPEG Audio codec family, MPEG Surround has

gained widespread acceptance already in a number of digital broadcast system standards due to its ability to transmit high-quality and backwards compatible multi-channel audio at extremely low bitrates such as 48 or 32 kbit/s if combined with MPEG-4 HE-AAC. The adoption of MPEG Surround is an ongoing process not only for audio and multimedia broadcast systems but also for other applications that are relevant for the automotive sector such as the car docking station scenario, where the highly scalable MPEG Surround algorithm will be used with higher data rates for improved audio quality, or for rear-seat entertainment with the binaural rendering mode of MPEG Surround.

REFERENCES

- [1] ISO/IEC 23003-1:2007, “Information technology - MPEG audio technologies – Part 1: MPEG Surround”, 2007.
- [2] J. Herre, K. Kjörning, J. Breebaart, C. Faller, S. Disch, H. Purnhagen, J. Koppens, J. Hilpert, J. Röden, W. Oomen, K. Linzmeier, and K. S. Chong, “MPEG Surround – The ISO/MPEG Standard for Efficient and Compatible Multi-Channel Audio Coding”, *Proc. 122nd AES convention*, Vienna, Austria (2007).
- [3] J. Breebaart, J. Herre, L. Villemoes, C. Jin, K. Kjörning, J. Plogsties, and J. Koppens, “Multi-channel goes Mobile: MPEG Surround Binaural Rendering” *Proc. 29th AES conference*, Seoul, South Korea (2006).
- [4] L. Villemoes, J. Herre, J. Breebaart, G. Hotho, S. Disch, H. Purnhagen, and K. Kjörning, “MPEG Surround: The Forthcoming ISO Standard for Spatial Audio Coding” *Proc. 28th AES conference*, Pitea, Sweden (2006).
- [5] J. Breebaart, J. Herre, C. Faller, J. Röden, F. Myburg, S. Disch, H. Purnhagen, G. Hotho, M. Neusinger, K. Kjörning, and W. Oomen, “MPEG Spatial Audio Coding / MPEG Surround: Overview and Current Status” *Proc. 119th AES convention*, New York, USA (2005).
- [6] J. Herre, H. Purnhagen, J. Breebaart, C. Faller, S. Disch, K. Kjörning, E. Schuijers, J. Hilpert, and F. Myburg, “The Reference Model Architecture for MPEG Spatial Audio Coding” *Proc. 118th AES convention*, Barcelona, Spain (2005).
- [7] Audio Subgroup, “Report on MPEG Surround Verification Test”, ISO/IEC JTC 1/SC 29/WG 11 N8851, 2007
- [8] Olaf Korte, “New Standards and Codecs for European Digital Broadcasting” *Proceedings of 62nd NAB Broadcast Engineering Conference 2008*, Las Vegas, USA
- [9] Oliver Hellmuth, Steffen Bergweiler, Stefan Holzhäuser, Manfred Neumann, and Andreas Walther, “Challenges in Reproduction and Evaluation of Upmixed Audio in an Automotive Environment” *Proc. 124th AES convention*, Amsterdam, The Netherlands (2008).
- [10] ETSI ETS 300 401, “Radio broadcasting systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers”.
- [11] ETSI TS 102 563, “Digital Audio Broadcasting (DAB); Transport of Advanced Audio Coding (AAC) audio”.
- [12] ETSI TS 102 428, “Digital Audio Broadcasting (DAB); DMB video service; User Application Specification”.
- [13] ETSI ES 201 980, “Digital Radio Mondiale (DRM); System Specification”.
- [14] ETSI TS 102 005, “Digital Video Broadcasting (DVB); Specification for the use of Video and Audio Coding in DVB services delivered directly over IP protocols”.
- [15] ISO/IEC JTC1/SC29/WG11 (MPEG), International Standard ISO/IEC 14496-3: “Coding of audiovisual objects: Audio”.
- [16] ISO/IEC JTC1/SC29/WG11 (MPEG), International Standard ISO/IEC 11172-3: “Information technology -- Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s -- Part 3: Audio”.
- http://www.chiariglione.org/mpeg/working_documents/mpeg-d/sac/VT-report.zip