**DRM RECEIVER**

***a project report Submitted by***

**A.BALAJI KAMALESH (UR11EC002)**

***in partial fulfillment for the award of the degree***

***of***

**BACHELOR OF TECHNOLOGY**

***in***

**ELECTRONICS AND COMMUNICATION ENGINEERING**

***under the supervision of***

**Dr.SHOBHA REKH**

****

**SCHOOL OF ELECTRICAL SCIENCES**

**KARUNYA UNIVERSITY**

(Karunya Institute of Technology and Sciences)

(Declared as Deemed-to-be-under Sec-3 of the UGC Act, 1956)

**Karunya Nagar, Coimbatore - 641 114. INDIA**

**JANUARY 2011**

**BONAFIDE CERTIFICATE**

Certified that this project report **“DRM RECEIVER”** is the bonafide work of “**A.BALAJI KAMALESH”** who carried out the project work under my supervision.

**SIGNATURE SIGNATURE**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Submitted for the Full Semester Viva Voce held on ……………………….

**Internal Examiner External Examiner**

LIST OF ABBREVIATIONS

AAC - Advanced Audio Coding

AM - Amplitude Modulation

BER - Bit Error Rate

CELP - Code Excited Linear Prediction

CRC - Cyclic Redundancy Check

DAB - Digital Audio Broadcasting

DRM - Digital Radio Mondiale

EEP - Equal Error Protection

ESC - Error Sensitivity Categories

FM - Frequency Modulation

HCR - Huffman Codeword Reordering

HF - High Frequency

HVXC - Harmonic Vector eXcitation Coding

ISO - International Organization for Standardization

LF - Low Frequency

MF - Medium Frequency

MPEG - Moving Picture Experts Group

MPS - MPEG Surround

MSC - Main Service Channel

OFDM - Orthogonal Frequency Division Multiplexing

OIRT - Organisation Internationale de Radiodiffusion et de Télévision

PNS - Perceptual Noise Substitution

PS - Parametric Stereo

QAM - Quadrature Amplitude Modulation

RF - Radio Frequency

rfa - reserved for future addition

rfu - reserved for future use

RVLC - Reversible Variable Length Coding

SAC - Spatial Audio Coding

SBR - Spectral Band Replication

SDC - Service Description Channel

SI - Side Information

TNS - Temporal Noise Shaping

UEP - Unequal Error Protection

VCB11 - Virtual Codebooks for Codebook 11

VXC - Vector eXcitation Coding

LIST OF SYMBOLS

*fR -* reference frequency of the emitted signal

*K -* number of active carriers in the OFDM symbol

*K*max *-* carrier index of the upper active carrier in the OFDM signal

*K*min *-* carrier index of the lower active carrier in the OFDM signal

*LMUX -* number of input bits per multiplex frame for the multilevel encoding

*NMUX  -* number of MSC cells (QAM symbols) per multiplex frame

*T -* elementary time period, equal to 831/3 μs (1/12 kHz)

*Tf  -* duration of the transmission frame

*Tg -* duration of the guard interval

*Ts  -* duration of an OFDM symbol

*Tsf  -* duration of the transmission super-frame built from the set of transmission frames

*Tu -* duration of the useful (orthogonal) part of an OFDM symbol, excluding the guard interval

*X\* -* complex conjugate of value X

*-* round towards plus infinity

*-* round towards minus infinity

CHAPTER 1

1.1 Introduction

The frequency bands used for broadcasting below 30 MHz are:

* Low Frequency (LF) band: from 148.5 kHz to 283.5 kHz, in ITU Region 1.
* Medium Frequency (MF) band: from 526.5 kHz to 1606.5 kHz, in ITU Regions 1 and 3 and from 525 kHz to 1705 kHz in ITU Region 2.
* High Frequency (HF) band: a set of individual broadcasting bands in the frequency range 2.3 MHz to 27 MHz, generally available on a Worldwide basis.

These bands offer unique propagation capabilities that permit the achievement of:

* Large coverage areas, whose size and location may be dependent upon the time of day, season of the year or period in the (approximately) 11 year sunspot cycle.
* Portable and mobile reception with relatively little impairment caused by the environment surrounding the receiver.

There is thus a desire to continue broadcasting in these bands, perhaps especially in the case of international broadcasting where the HF bands offer the only reception possibilities which do not also involve the use of local repeater stations. However, broadcasting services in these bands:

* Use analogue techniques.
* Are subject to limited quality.
* Are subject to considerable interference as a result of the long-distance propagation mechanisms which prevail in this part of the frequency spectrum and the large number of users.

As a direct result of the above considerations, there is a desire to affect a transfer to digital transmission and reception techniques in order to provide the increase in quality which is needed to retain listeners who, increasingly, have a wide variety of other programme reception media possibilities, usually already offering higher quality and reliability. In order to meet the need for a digital transmission system suitable for use in all of the bands below 30 MHz, the Digital Radio Mondiale (DRM) consortium was formed in early 1998. The DRM consortium is a non-profit making body which seeks to develop and promote the use of the DRM system worldwide. Its members include broadcasters, network providers, receiver and transmitter manufacturers and research institutes.

In March 2005, the DRM Consortium voted at its General Assembly to embark on extending the capability of the DRM system to provide digital radio services at higher transmission frequencies. This range includes:

* 47 MHz to 68 MHz (Band I) allocated to analogue television broadcasting;
* 65.8 MHz to 74 MHz (OIRT FM band);
* 76 MHz to 90 MHz (Japanese FM band);
* 87.5 MHz to 107.9 MHz (Band II) allocated to FM radio broadcasting.

This extension completes the family of digital standards for radio broadcasting.

1.2 General characteristic

1.2.1 System overview

The DRM system is designed to be used at any frequency below 174 MHz, with variable channelization constraints and propagation conditions throughout these bands. In order to satisfy these operating constraints, different transmission modes are available. A transmission mode is defined by transmission parameters classified in two types:

* signal bandwidth related parameters;
* transmission efficiency related parameters.

The first type of parameters defines the total amount of frequency bandwidth for one transmission. Efficiency related parameters allow a trade-off between capacity (useful bit rate) and ruggedness to noise.

1.2.2 System architecture

This gives a general presentation of the system architecture, based on the synoptic diagram of Figure 1 and describes the general flow of different classes of information (audio, data, etc.) and does not differentiate between different services that may be conveyed within one or more classes of information. A detailed description on the distribution of services onto those classes can be found in chapter 3.

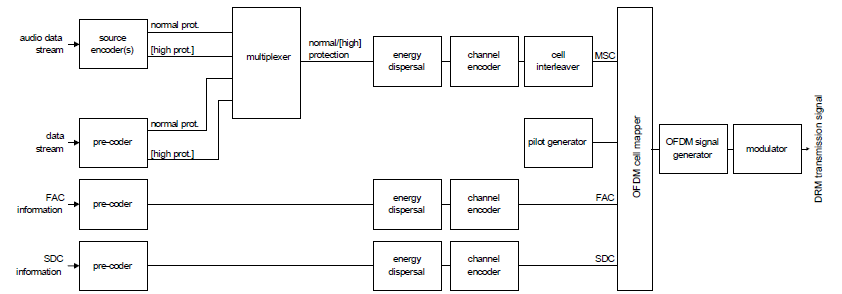


Figure 1: Conceptual DRM transmission block diagram

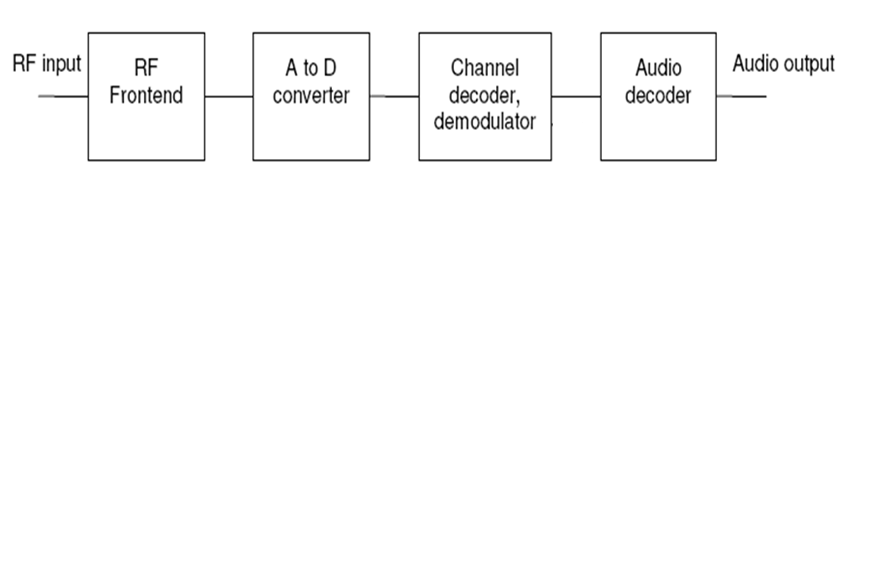


Figure 2: DRM Receiver block diagram

The source encoder and pre-coders ensure the adaptation of the input streams onto an appropriate digital transmission format. For the case of audio source encoding, this functionality includes audio compression techniques. The output of the source encoder(s) and the data stream pre-coder may comprise two parts requiring different levels of protection within the subsequent channel encoder. All services have to use the same two levels of protection. The multiplexer combines the protection levels of all data and audio services. The energy dispersal provides a deterministic selective complementing of bits in order to reduce the possibility that systematic patterns result in unwanted regularity in the transmitted signal.

The channel encoder adds redundant information as a means for quasi error-free transmission and defines the mapping of the digital encoded information onto QAM cells Cell interleaving spreads consecutive QAM cells onto a sequence of cells quasi-randomly separated in time and frequency, in order to provide robust transmission in time-frequency dispersive channels. The pilot generator provides means to derive channel state information in the receiver, allowing for a coherent demodulation of the signal. The OFDM cell mapper collects the different classes of cells and places them on the time-frequency grid. The OFDM signal generator transforms each ensemble of cells with same time index to a time domain representation of the signal. Consecutively, the OFDM symbol is obtained from this time domain representation by inserting a guard interval as a cyclic repetition of a portion of the signal. The modulator converts the digital representation of the OFDM signal into the analogue signal in the air. This operation involves digital-to-analogue conversion and filtering that have to comply with spectrum requirements as described in annex E.

CHAPTER 2

2 Transmission modes

2.1 Signal bandwidth related parameters

The current channel widths for radio broadcasting below 30 MHz are 9 kHz and 10 kHz. The DRM system is designed to be used:

* Within these nominal bandwidths, in order to satisfy the current planning situation;
* Within half these bandwidths (4.5 kHz or 5 kHz) in order to allow for simulcast with analogue AM signals;
* Within twice these bandwidths (18 kHz or 20 kHz) to provide for larger transmission capacity where and when the planning constraints allow for such facility.

The current channel raster (where defined) for radio broadcasting between 30 MHz and 174 MHz is 100 kHz. The DRM system is designed to be used with this raster

2.2 Transmission efficiency related parameters

For any value of the signal bandwidth parameter, transmission efficiency related parameters are defined to allow a tradeoff between capacity (useful bit rate) and ruggedness to noise, multipath and Doppler. These parameters are of two types:

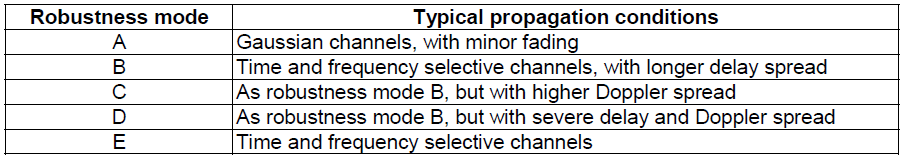
* Coding rate and constellation parameters, defining which code rate and constellations are used to convey data.
* OFDM symbol parameters, defining the structure of the OFDM symbols to be used as a function of the propagation conditions.

2.3 Coding rates and constellations

As a function of the desired protection associated within each service or part of a service, the system provides a range of options to achieve one or two levels of protection at a time. Depending on service requirements, these levels of protection may be determined by either the code rate of the channel encoder (e.g. 0.6, etc.), by the constellation order (e.g. 4-QAM, 16-QAM, 64-QAM), or by hierarchical modulation.

2.4 OFDM parameter set

The OFDM parameter set is presented in this paragraph. These values are defined for different propagation-related transmission conditions to provide various robustness modes for the signal.

Table 1: Robustness mode uses

In a given bandwidth, the different robustness modes provide different available data rates. Table 1 illustrates typical uses of the robustness modes. The transmitted signal comprises a succession of OFDM symbols, each symbol being made of a guard interval followed by the so-called useful part of the symbol. Each symbol is the sum of *K* sine wave portions equally spaced in frequency. Each sine wave portion, called a "cell", is transmitted with given amplitude and phase and corresponds to a carrier position. Each carrier is referenced by the index *k, k* belonging to the interval [ ] *k*min , *k*max ( *k* = 0 corresponds to the reference frequency of the transmitted signal).The time-related OFDM symbol parameters are expressed in multiples of the elementary time period *T* , which is equal to 831/3 μs. These parameters are:

* *Tg* : duration of the guard interval.
* *Ts* : duration of an OFDM symbol.
* *Tu* : duration of the useful (orthogonal) part of an OFDM symbol (i.e. excluding the guard interval).

The OFDM symbols are grouped to form transmission frames of duration*Tf* .

A certain number of cells in each OFDM symbol are transmitted with a predetermined amplitude and phase, in order to be used as references in the demodulation process. They are called "reference pilots" and represent a certain proportion of the total number of cells.



Table 2: OFDM symbol parameters

CHAPTER 3

3 Source coding

Within the constraints of broadcasting regulations in broadcasting channels below 30 MHz and the parameters of the coding and modulation scheme applied, the bit rate available for source coding is in the range from 8 kbit/s (half channels) to ≈20 kbit/s (standard channels) to up to ≈72 kbit/s (double channels). Within the constraints of broadcasting regulations in broadcasting channels between 30 MHz and 174 MHz and the parameters of the coding and modulation scheme applied, the bit rate available for source coding is in the range from 35 kbit/s to 185 kbit/s. In order to offer optimum quality at a given bit rate, the system offers different source coding schemes:

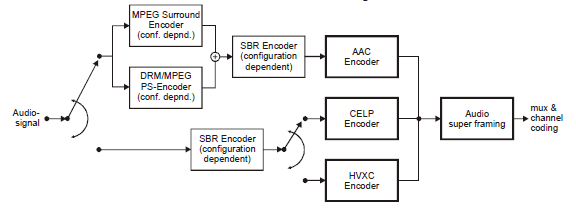
* A subset of MPEG-4 AAC (Advanced Audio Coding) including error robustness tools for generic mono and stereo audio broadcasting;
* A subset of MPEG-4 CELP speech coder for error robust speech broadcasting in mono, for cases when only a low bit rate is available or especially high error robustness is required;
* A subset of MPEG-4 HVXC speech coding for very low bit rate and error robust speech broadcasting in mono, especially well suited also for speech data base applications;
* Spectral Band Replication (SBR), an audio coding enhancement tool that allows the full audio bandwidth to be achieved at low bit rates. It can be applied to AAC, CELP and HVXC;
* Parametric Stereo (PS), an audio coding enhancement tool relevant to SBR that allows for stereo coding at low bit rates;
* MPEG Surround (MPS), an audio coding enhancement tool that allows for multichannel coding at low bit rates.

The bit-stream transport format of the source coding schemes has been modified to meet the requirements of the DRM system (Audio superframing). Unequal Error Protection (UEP) can be applied to improve the system behavior in error prone channels. Provision is made for further enhancement of the audio system by linking two DRM signals together.

3.1 Source coding modes

3.1.1 Overview

The source coding options in the DRM system are shown in figure 4. As described in chapter 3, the DRM system offers audio coding (AAC) and speech coding (CELP and HVXC). In addition, a high frequency reconstruction method (SBR) can be used to enhance the perceptual audio quality of the three different source coding schemes.



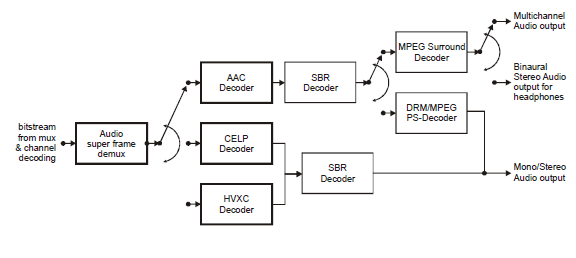
****

Figure 4: Audio source coding overview

In combination with the AAC codec, the MPEG PS tool can be used. Optionally, a multichannel reconstruction method (MPS) can be used on top of AAC/SBR to enable multichannel decoding.Special care is taken so that the encoded audio can be composed into audio super frames of constant length. Multiplexing and UEP of audio/speech services is done by means of the multiplex and channel coding units. Audio specific configuration information is transmitted in the SDC.

3.1.2 AAC audio coding

For generic audio coding, a subset of the MPEG-4 Advanced Audio Coding (AAC) toolbox chosen to best suit the DRM system environment is used. For example a standard configuration for use in one short wave channel could be 20 kbit/s mono AAC.

Specific features of the AAC stream within the DRM system are:

* Bit rate: AAC can be used at any bit rate. The granularity of the AAC bit rate is 20 bit/s for robustness modes A, B, C and D and 80 bit/s for robustness mode E.
* Sampling rates: permitted sampling rates are 12 kHz and 24 kHz for robustness modes A, B, C and D and 24 kHz and 48 kHz for robustness mode E. 48 kHz is only permitted if the SBR tool is not used.
* Transform length: the transform length is 960 to ensure that one audio frame corresponds to 80 ms or 40 ms (robustness modes A, B, C and D) or to 40 ms or 20 ms (robustness mode E) in time. This is required to harmonize CELP and AAC frame lengths and thus to allow the combination of an integer number of audio frames to build an audio super frame of 400 ms (robustness modes A, B, C and D) or 200 ms (robustness mode E) duration.
* Error robustness: a subset of MPEG-4 tools is used to improve the AAC bit stream error robustness in error prone channels (the MPEG-4 EP tool is not used).
* Audio super framing: 5 or 10 audio frames are composed into one audio super frame. For robustness modes A, B, C and D, the respective sampling rates are 12 kHz and 24 kHz producing an audio super frame of 400 ms duration; for robustness mode E, the respective sampling rates are 24 kHz and 48 kHz producing an audio super frame of 200 ms duration. The audio frames in the audio super frames are encoded together such that each audio super frame is of constant length, i.e. that bit exchange between audio frames is only possible within an audio super frame. One audio super frame is always placed in one logical frame in robustness modes A, B, C and D and in two logical frames in robustness mode E. In this way no additional synchronization is needed for the audio coding. Retrieval of frame boundaries and provisions for UEP are also taken care of within the audio super frame.
* UEP: better graceful degradation and better operation at higher BERs is achieved by applying UEP to the AAC bit stream. Unequal error protection is realized via the multiplex/coding units. For robustness mode E, the length of the higher protected part of an audio super frame must be a multiple of 2 bytes.

3.1.3 MPEG CELP coding

MPEG CELP speech coding is available in robustness modes A, B, C and D to allow for reasonable speech quality at bit rates significantly below the standard rate (for example "half rate" operation at 8 kbit/s). Possible scenarios for the use of the speech coder are:

* Dual/triple speech applications: instead of one audio programme at 20 kbit/s to 24 kbit/s, the channel contains two or three speech signals of 8 kbit/s to 10 kbit/s each, allowing simultaneous speech transmissions.
* Speech services in addition to an audio service.
* Simulcast transmissions: in case of analogue/digital simulcast only bit rates as low as 8 kbit/s may be available.
* Very robust speech applications: due to its nature a speech coder can be expected to offer higher robustness against channel errors. Therefore 8 kbit/s speech coding can be used to do ultra robust speech coding in one channel.
* Basic features of MPEG CELP coding are:
* 8 kHz or 16 kHz sampling rate.
* Bit rates between 4 kbit/s and 20 kbit/s.
* Error robustness.
* Composition of an integer number of CELP frames to build one audio super frame.

3.1.4 MPEG HVXC coding

MPEG-4 HVXC (Harmonic Vector eXcitation Coding) speech coding is available in robustness modes A, B, C and D to allow for reasonable speech quality at very low bit rates such as 2.0 kbit/s. The operating bit rates of HVXC open up new applications for DRM such as:

* Speech services in addition to an audio service.
* Multi-language application.
* Solid-state storage of multiple programmes such as news, data base in a card radio (e.g. total of about 4.5 hours of radio programmes can be stored in 4 MByte Flash memory).
* Time scale modification for fast playback/browsing of stored programme.
* Highly error robust transmission with or without hierarchical modulation scheme.

Basic features of HVXC coding are:

* 8 kHz sampling rate.
* Bit rates of 2.0 kbit/s and 4.0 kbit/s for fixed rate coding.
* Time scale and pitch modification of arbitrary amounts.
* Error robust syntax is supported, and a CRC tool can be used to improve the error resilience of the HVXC bitstream in error prone channels.
* Composition of a constant integer number of HVXC frames (20) to build one audio super frame.

3.1.5 SBR coding

To maintain a reasonable perceived audio quality at low bit rates, classical audio or speech source coding algorithms need to limit the audio bandwidth and to operate at low sampling rates. It is desirable to be able to offer high audio bandwidth also in very low bit rate environments. This can be realized by the use of Spectral Band Replication (SBR). The purpose of SBR is to recreate the missing high frequency band of the audio signal that could not be coded by the encoder. In order to do this in the best possible way, some side information needs to be transmitted in the audio bitstream, removing a small percentage of the available data rate from the audio coder. This side information is computed on the full bandwidth signal, prior to encoding and aids the reconstruction of the high frequencies after audio/speech decoding. SBR exists in two versions. The version difference is only reflected in the decoder design. High Quality SBR uses a complex filterbank whereas Low Power SBR uses a real-valued filterbank plus anti-aliasing modules. The Low Power version of SBR offers a significant reduction in complexity as compared to the High Quality version without compromthe present documentg too much on audio quality. AAC + SBR is defined in MPEG-4 Audio (High Efficiency AAC profile). SBR is also used in the configurations HVXC + SBR and CELP + SBR.

3.1.6 PS coding

For improved performance at low bitrate stereo coding, a Parametric Stereo (PS) coder is available. The PS tool can be used when running the configuration AAC + SBR (MPEG High Efficiency AAC profile). The general idea with PS coding is to send stereo image describing data as side information along with a downmixed mono signal. This stereo side information is very concise and only requires a small fraction of the total bitrate allowing the mono signal to have maximum quality for the total bitrate given. The stereo synthesis at the decoder reconstructs spatial properties but does not affect the total spectral energy. Hence, there is no colorization of the frequency spectrum compared to the mono compatible core signal. The target bitrates for applying parametric stereo coding on AAC + SBR are preferably any bitrate range where traditional stereo cannot be afforded. If the broadcast signal contains PS data, the PS tool as specified in MPEG-4 Audio [2] shall be used. In addition, the PS tool as specified in clause 5.7 may be used for robustness modes A, B, C and D.

3.2 Error concealment

For each audio coder and for the SBR and PS tools a description for the concealment of erroneous bit streams is given. The error concealment provided by a DRM decoder shall provide at least the same level of performance as the specified concealment tools, but may be enhanced by specific implementations.

3.3 MPEG Surround coding

An MPEG Surround (MPS) coder is available for mono/stereo compatible multichannel encoding. MPEG Surround is standardized in MPEG-D, Part-1 (ISO/IEC 23003-1 [10]). It describes:

* coding of multichannel signals based on a downmixed signal of the original multichannel signal, and associated spatial parameters. It offers lowest possible data rate for coding of multichannel signals, as well as an inherent mono or stereo downmix signal included in the data stream. Hence, a mono or stereo signal can be expanded to multi-channel by a very small additional data overhead;
* binaural decoding of the MPEG Surround stream, enabling a surround sound experience over stereo headphones;
* an Enhanced Matrix Mode that enables a multi-channel upmix from a stereo signal without any spatial parameters.

Receivers without multichannel decoding support can decode the unmodified mono or stereo core signal. Hence, MPEG Surround (Spatial Audio Coding, SAC) is capable of re-creating N channels based on M<N transmitted channels, and additional control data. In the preferred modes of operating the spatial audio coding system, the M channels can either be a single mono channel or a stereo channel pair. The control data represents a significantly lower data rate than required for transmitting all N channels, making the coding very efficient while at the same time ensuring compatibility with both M channel devices and N channel devices. The MPEG Surround standard incorporates a number of tools enabling a number of features that allow for broad application of the standard. A key feature is the ability to scale the spatial image quality gradually from very low spatial overhead towards transparency. Another key feature is that the compatible decoder input can be made compatible to existing matrix surround technologies. All tools are grouped to cover certain profiles. Receivers with a different number of output channels than the number of MPS target channels indicated by the SDC should still render the mutichannel audio signal according to the available number of output channels (possibly at a reduced quality compared to the case where the number of target channels matches the number of output channels).

3.4 UEP and audio super framing

Today's coding schemes are highly optimized in terms of coding efficiency and according to information theory this should lead to the fact, that the entropy of the bits is nearly equal. If this assumption is true, then the channel coding must be optimized, such that the total amount of residual errors usually referred to as Bit Error Rate (BER) is minimized. This criterion can be fulfilled by a channel coding method called Equal Error Protection (EEP), were all information bits are protected with the same amount of redundancy. However, the audible effects of errors are not independent of the part of the bitstream that was hit by the error. This behaviour of unequal error sensitivity is well known for source coding schemes that are used in broadcast and communication systems, like DAB (Eureka 147) or GSM. The optimized solution to cope with this unequal error sensitivity is called Unequal Error Protection (UEP). In such a system, higher protection is assigned to the more sensitive information, whereas lower protection is assigned to the less sensitive part of the bitstream. To accommodate for UEP channel coding, it is necessary to have frames with a constant length and a UEP profile that is constant as well for a given bit rate. Since AAC is a coding scheme with a variable length, several coded frames are grouped together to build one audio super frame. The bit rate of the audio super frame is constant. Since the channel coding is based on audio super frames, the audio super frames themselves consist of two parts: a higher protected part and a lower protected part. Therefore, the coded audio frames itself have to be split into these two parts. Further details on the audio super frame structure of AAC, CELP and HVXC are provided in the subsequent clauses. HVXC is intended for use with the EEP scheme only.

For robustness modes A, B, C and D, the audio super frame is mapped directly onto the logical frame, since both are of the same duration. For robustness mode E, the audio super frame is mapped onto two logical frames, since the audio super frame is of twice the duration of the logical frame. The mapping is performed such that the first half of the higher protected bytes followed by the first half of the lower protected bytes are mapped to logical frame *n* and the second half of the higher protected bytes followed by the second half of the lower protected bytes are mapped to logical frame *n+1*, as illustrated in figure 3.

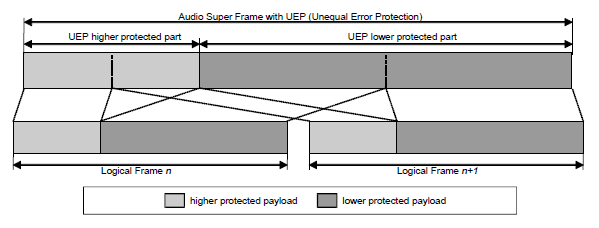


Figure 3: Mapping of audio super frame into two logical frames for robustness mode E

3.5 AAC coding

The following three clauses explain how the AAC, the AAC + SBR and the MPS enhanced frames fit into the audio super frame.

3.5.1 AAC

ISO/IEC 14496-3 defines the MPEG-4 Audio standard. The audio coding standard MPEG-4 AAC is part of the MPEG-4 Audio standard. From the possible audio object types, only the Error Robust (ER) AAC Scalable object type (Object Type ID = 20), which is part of the High Quality Audio Profile, is used in the DRM system. DRM specific usage of MPEG-4 AAC: Three error robustness tools may be used within an MPEG-4 ER AAC bitstream: HCR (Huffman Codeword Reordering), VCB11 (Virtual Codebooks for Codebook 11) and RVLC (Reversible Variable Length Coding). In the DRM system, all AAC bitstreams shall use the HCR tool, since this tool reduces the error sensitivity of the bitstream significantly with a minimum of overhead. The VCB11 tool shall be used, since for low bit rates, the VCB11 overhead is less than 1 %. The RVLC tool is not used, since it introduces a significant bit rate overhead that is a major drawback for the low bit rates used by DRM. The MPEG-4 AAC tool PNS (Perceptual Noise Substitution) is not used in DRM since SBR provides this functionality more appropriately. For DRM the 960 transform shall be used.

Robustness modes A, B, C and D:

* When 12 kHz sampling is used, 5 AAC frames shall be combined into one audio super frame.
* When 24 kHz sampling is used, 10 AAC frames shall be combined into one audio super frame.
* The AAC sampling rate shall be 24 kHz when the stereo mode is used.

Robustness mode E:

* When 24 kHz sampling is used, 5 AAC frames shall be combined into one audio super frame.
* When 48 kHz sampling is used, 10 AAC frames shall be combined into one audio super frame.

No standard extension\_payload() shall be used and the only allowed extensions are SBR (signalled via SDC) and MPS (signalled via SDC). The left and the right channel in one stereo audio frame are transmitted in an interleaved way to achieve a decreasing error sensitivity within the stereo frame. Any DRM AAC bitstream can easily be translated into an MPEG-4 ER compliant bitstream by applying the above rules. When the transmission is a base layer (the Base/Enhancement flag in the FAC is 0, see clause 6.3.3), the AAC frame corresponds to aac\_scalable\_main\_element() as defined in the MPEG-4 standard. The MPEG-4 standard defines how the bits for one raw error robust AAC audio frame are stored. Each element of the error robust AAC bitstream is assigned an error sensitivity category. In the DRM system there are two possible error robust AAC audio frames:

Mono audio frame

One mono audio frame consists of three consecutive parts, hereinafter called mono1, mono2 and mono3. Mono1 contains the Side Information (SI) bits, mono2 contains the Temporal Noise Shaping (TNS) bits and mono3 contains the spectral data bits. The error sensitivity decreases from mono1 to mono3.

Stereo audio frame

One stereo audio frame consists of seven consecutive parts, hereinafter called stereo1 (common side info), stereo2 (side info left channel), stereo3 (side info right channel), stereo4 (TNS left channel), stereo5 (TNS right channel), stereo6 (spectral data left channel), stereo7 (spectral data right channel). With this interleaving of left and right channel, the error sensitivity is decreasing from stereo1 to stereo7.

3.5.2 AAC audio super frame

Higher protected part

The higher protected part contains one header followed by num\_frames higher protected blocks. num\_frames is the number of audio frames in the audio super frame.

Header

The header contains information to recover the frame lengths of the num\_frames AAC frames stored in the audio superframe. All the frame lengths are derived from the absolute positions of the frame borders. These frame borders are stored consecutively in the header. Each frame border occupies 12 bits (unsigned integer, most significant bit first). The frame border is measured in bytes from the start of the AAC bitstream sequence. 4 padding bits are added in case num\_frames==10. num\_frames-1 frame borders are stored in the header.

Higher protected block

One higher protected block contains a certain amount of bytes from the start of each AAC frame, dependent upon the UEP profile. One 8-bit CRC check derived from the CRC-bits of the corresponding AAC frame follows (see annex D for CRC calculation). For a mono signal, the CRC-bits cover (mono1, mono2). For a stereo signal, the CRC-bits cover (stereo1, stereo2, stereo3, stereo4, stereo5).

Lower protected part

The lower protected bytes (the remaining bytes not stored in the higher protected part) of the AAC frames are stored consecutively in the lower protected part. Figure 4 illustrates an example audio super frame with 10 audio frames in the cases of equal and unequal error protection.

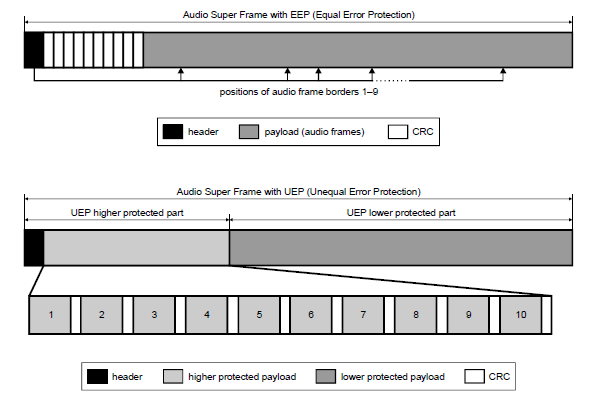
****

Figure 4: Example AAC audio super frame with 10 audio frames

3.5.3 AAC + SBR

The SBR sampling rate is twice the AAC sampling rate. One raw AAC + SBR frame contains an AAC part and a SBR part. The SBR part of the data is located at the end of the frame. The first bit in the SBR-bitstream is the last bit in the frame, and the SBR bits are thus written/read in reverse order. In this way, the starting points of respective part of the frame data are always easily found.

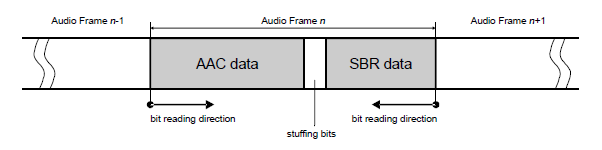


Figure 5: AAC + SBR frame

Both AAC and SBR data-sizes vary from frame to frame. The total size of the individual frames, now including the SBR data, can be derived from the aac\_super\_frame\_header() as described in clause 5.3.1. Thus no extra signalling due to the varying SBR bit rate is needed. The AAC + SBR frames are inserted into the audio super frame in the same manner as when SBR is not used. The details of the SBR-bitstream are described in clause 5.6.1.

3.5.4 AAC error concealment

The AAC core decoder includes a concealment function that increases the delay of the decoder by one frame. There are various tests inside the core decoder, starting with the CRC test and ending in a variety of plausibility checks. If such a check indicates an invalid bit stream, then concealment is applied. Concealment is also applied when the channel decoder indicates a distorted data frame. Concealment works on the spectral data just before the final frequency to time conversion. In case a single frame is corrupted, concealment interpolates between the preceding and the following valid frames to create the spectral data for the missing frame. If multiple frames are corrupted, concealment implements first a fade out based on slightly modified spectral values from the last valid frame. If the decoder recovers from the error condition, the concealment algorithm performs a fade-in on valid spectral values. Fade in might be delayed (suppressed) to deal with error conditions, where only a valid frame here and there is perceived.

4 TAG items specifying DRM multiplex

4.1 Audio status (rafs)

This TAG item is used to signal the status of the audio decoder with a time base of one audio unit as shown in figure 6.17. Because of the importance of this information this TAG item is mandatory for all RX\_STAT profiles except profile R. The given TAG value relates to the actual selected service (see clause 6.4.3.5). Whenever no information on audio unit decoding is available (e.g. if no audio service is selected or if no MSC data is available) then the TAG length of the TAG item "rafs" shall be zero.

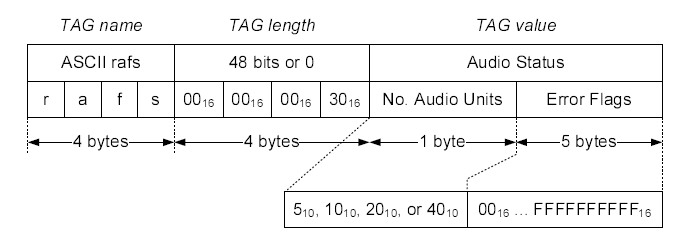
****

Figure 6.17: Audio status

No. audio units:number of audio units per audio super frame.

For robustness modes A to D, one audio super frame is of 400 ms duration and consists of:

* 5 audio units in case of MPEG-4 AAC with 12 kHz sampling rate, or
* 10 audio units in case of MPEG-4 AAC with 24 kHz sampling rate, or
* 20 audio units in case of MPEG-4 HVXC, or
* 10, 20 or 40 audio units in case of MPEG-4 CELP.

For robustness mode E, one audio super frame is of 200 ms duration and consists of:

* 5 audio units in case of MPEG-4 AAC with 24 kHz sampling rate, or
* 10 audio units in case of MPEG-4 AAC with 48 kHz sampling rate.

For further details of the audio coding used in the DRM system see ES 201 980 [1], clause 5.

Error flags:a 40-bit field carrying the binary-coded audio status with one bit representing the decoding status of one audio unit. The bit field is filled from the left; i.e. the MSb of the first byte describes the decoding status of the first audio unit of the audio super frame. When fewer than 40 audio units are present in the audio super frame, the corresponding unused bits are set to zero. The used values for the error flags have the following meaning:

* 02 audio units is ok
* 12 audio units is corrupted

4.2 Extended Audio Status (reas)

In addition to the TAG item "rafs" above (see clause 6.4.3.7) the status of the audio decoder can be described in more detail using the TAG item "reas" as shown in figure 6.18. This TAG item is mandatory for the higher datarate RX\_STAT profiles A, C and D only. The given TAG value relates to the actual selected service (see clause 6.4.3.5). Whenever no audio units are available (e.g. if no audio service is selected or if no MSC data is available) then the TAG length of the TAG item "reas" shall be zero. This more detailed differentiation of errors within an audio unit is done because errors which are located within the less sensitive part of the audio data can be concealed in many cases by the source decoding process in a way that they are possibly not recognized by a human listener. But errors which are located within the high sensitive data part will be audible in nearly all cases.

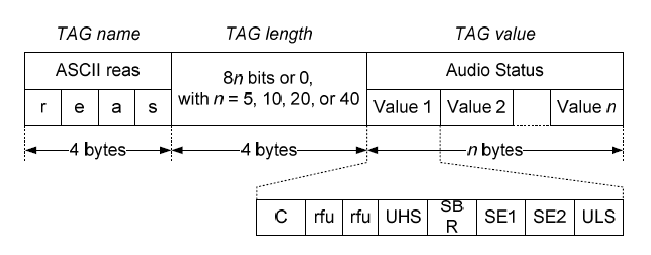


Figure 6.18: Extended audio status

Audio status values**:** 8-bit unsigned integer value specifying the status of one audio unit with eight separate error flags as explained below. A classification of errors may be done in the following way:

* 0016 audio unit is ok
* 0116 .. 0F16 audio unit is corrupted, but the errors are located within the less sensitive part
* 1016 .. FF16 audio unit is corrupted, errors are located within the high sensitive part and maybe also within the less sensitive part.

One byte of error status shall be provided by the audio decoder for each audio unit.

C (core CRC error flag):for MPEG-4 AAC and CELP coding this is the error status of the CRC check concerning the core data part; for MPEG-4 HVXC coding this is the error status of the CRC check concerning the ESC0 data part of the audio unit.

rfu**:** these two bits are reserved for future use and shall have the value zero.

UHS (Unspecified High Sensitive error flag):decoder dependent error flag for all other detected errors within the high sensitive data part of the audio unit.

SBR (SBR CRC error flag):error status of the CRC check concerning the SBR data part of the audio unit.

SE1 (Spectral data Error flag no. one):for MPEG-4 AAC coding this is the error status of the spectral data part (in case of stereo coding only of the left channel spectral data part); for MPEG-4 CELP coding this flag is not used and shall have the value zero; for MPEG-4 HVXC coding this is the combined error status of the CRC check concerning the ESC1, ESC2 and if present ESC3 data part of the audio unit.

SE2 (Spectral data Error flag no. two):for MPEG-4 AAC stereo coding this is the error status of the right channel spectral data part; for MPEG-4 AAC mono, CELP and HVXC coding this flag is not used and shall have the value zero.

ULS (Unspecified Less Sensitive error flag):decoder dependent error flag for all other detected errors within the less sensitive part of the audio unit.