INTERNATIONAL STANDARD
ISO/IEC 14496-3:2001/Amd.1:2003/Cor.1:2004(E)

TECHNICAL CORRIGENDUM 1

Published 2004-06-01

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION • МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ • ORGANISATION INTERNATIONALE DE NORMALISATION
INTERNATIONAL ELECTROTECHNICAL COMMISSION • МЕЖДУНАРОДНАЯ ЭЛЕКТРОТЕХНИЧЕСКАЯ КОМИССИЯ • COMMISSION ÉLECTROTECHNIQUE INTERNATIONALE

Information technology — Coding of audio-visual objects —

Part 3: **Audio**

AMENDMENT 1: Bandwidth extension

TECHNICAL CORRIGENDUM 1

Technologies de l'information — Codage des objets audiovisuels —

Partie 3: Codage audio

AMENDEMENT 1: Extension de largeur de bande

RECTIFICATIF TECHNIQUE 1

Technical Corrigendum 1 to ISO/IEC 14496-3:2001/Amd.1:2003 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

NOTE This document specifies the first corrigendum to ISO/IEC 14496-3:2001/Amd.1:2003. With the exception of the correction for Table 4.54 which is only relevant for multi-channel decoding of ER streams, all the corrections strive to correct errors in the standard text and thus aligning the standard text with the reference software.

ICS 35.040 **Ref. No. ISO/IEC 14496-3:2001/Amd.1:2003/Cor.1:2004(E)**

Published in Switzerland

In ISO/IEC 14496-3:2001/Amd.1, 4.4.2.8 Payloads for the audio object type SBR, replace Table 4.54 by the table below (modified parts are marked in gray):

Table 4.54A – Syntax of sbr_extension_data()

Syntax	No. of bits	Mnemonic

sbr_extension_data(id_aac, crc_flag)
{
num_sbr_bits = 0;
 	
if (crc_flag) {
bs_sbr_crc_bits;
	10	**uimsbf**

num_sbr_bits += 10;
}
 	
if (sbr_layer != SBR_STEREO_ENHANCE) {
		Note 1

num_sbr_bits += 1; if (**bs_header_flag**)	**1**	
num_sbr_bits += sbr_header();		Note 2

}
 	
num_sbr_bits += sbr_data(id_aac, bs_amp_res);
		Note 2

num_align_bits = (8*cnt - 4 - num_sbr_bits) % 8; **bs_fill_bits**;	num_align_bits	**uimsbf**
return ((num_sbr_bits + num_align_bits + 4) / 8) }		
Note 1: When the SBR tool is used with a non-scalable AAC core coder, the value of the helper variable sbr_layer is SBR_NOT_SCALABLE. When the SBR tool is used with a scalable AAC core coder, the value of the helper variable sbr_layer depends on the current layer and the scalability configuration of the AAC core coder as defined in Table 4.86 in subclause 4.5.2.8.2.4. Note 2: sbr_header() and sbr_data() return the number of bits read (cnt is a parameter in extension_payload()).		

In ISO/IEC 14496-3:2001/Amd.1:2003, subclause 4.4.2.8 Payloads for the Audio Object Type SBR, Table 4.55A, replace:

Note 3: If this bit is not set the default values for the underlying bitstream elements should be used.

by:

Note 3: If this bit is not set the default values for the underlying bitstream elements shall be used disregarded any previous value.

In ISO/IEC 14496-3:2001/Amd.1, 4.4.2.8 Payloads for the audio object type SBR, Table 4.58, Table 4.59, and Table 4.60, add the part marked in gray:

if (**bs_extended_data**) {	**1**	
cnt = **bs_extension_size**;	**4**	**uimsbf**
if (cnt == 15)		
cnt += **bs_esc_count**;	**8**	**uimsbf**
num_bits_left = 8 * cnt;		
while (num_bits_left > 7) {		
bs_extension_id;	**2**	**uimsbf**
num_bits_left -= 2;		
sbr_extension(bs_extension_id, num_bits_left);		Note 1
}		
bs_fill_bits	num_bits_left	
}		

In ISO/IEC 14496-3:2001/Amd.1, subclause 4.5.2.8.2.3 SBR Extension Payload for the Audio Object Types ER AAC LC and ER AAC LTP replace:

The extension payload shall not include both DRC extension elements and SBR extension elements simultaneously. If SBR extension elements are used, DRC extension elements are prohibited.

by:

The SBR extension elements shall be placed after any other extension elements.

In ISO/IEC 14496-3:2001/Amd.1, subclause 4.5.2.8.2.4 SBR Extension Payload for the Audio Object Types AAC Scalable and ER AAC Scalable, replace:

The scalable SBR data is embedded into the MPEG-4 stream in the same way as for non-scalable SBR data elements, by means of using the extension_payload().

by:

The scalable SBR data is embedded into the MPEG-4 stream in the same way as for non-scalable SBR data elements, by means of using the extension_payload(). The SBR extension elements shall be placed after any other extension elements.

In ISO/IEC 14496-3:2001/Amd.1, subclause 4.6.18.2.2 SBR specific notation, replace:

$$\mathbf{Q}_{Mapped}(m-lsb,l)=\mathbf{Q}_{Orig}(i,k(l)),\mathbf{f}_{TableNoise}(i)\le m<\mathbf{f}_{TableNoise}(i+1),0\le i<N_Q,0\le l<L_E$$

where $k(l)$ is defined by $\mathbf{t}(l)\ge\mathbf{t}_Q(k(l)),\mathbf{t}(l+1)\le\mathbf{t}_Q(k(l)+1)$

should be interpreted as follows. $\mathbf{Q}_{Mapped}(m-lsb,l)$ equals $\mathbf{Q}_{Orig}(i,k(l))$

by:

$$\mathbf{Q}_{Mapped}\left(m-k_x,l\right)=\mathbf{Q}_{Orig}\left(i,k\left(l\right)\right),\mathbf{f}_{TableNoise}(i)\le m<\mathbf{f}_{TableNoise}(i+1),0\le i<N_Q,0\le l<L_E$$

where $k\left(l\right)$ is defined by $\mathbf{t}\left(l\right)\ge\mathbf{t}_Q\left(k\left(l\right)\right),\mathbf{t}\left(l+1\right)\le\mathbf{t}_Q\left(k\left(l\right)+1\right)$

should be interpreted as follows. $\mathbf{Q}_{Mapped}\left(m-k_x,l\right)$ equals $\mathbf{Q}_{Orig}\left(i,k\left(l\right)\right)$

In ISO/IEC 14496-3:2001/Amd.1, subclause 4.6.18.2.5 Constants, replace:

ε	A constant to avoid division by zero, e.g. 96 dB below maximum signal input.

by:

$\varepsilon=1$	A constant to avoid division by zero, e.g. 96 dB below maximum signal input.

In ISO/IEC 14496-3:2001/Amd.1, subclause 4.6.18.2.6, add:

W	is the subband matrix where the QMF filtered subband samples are stored.

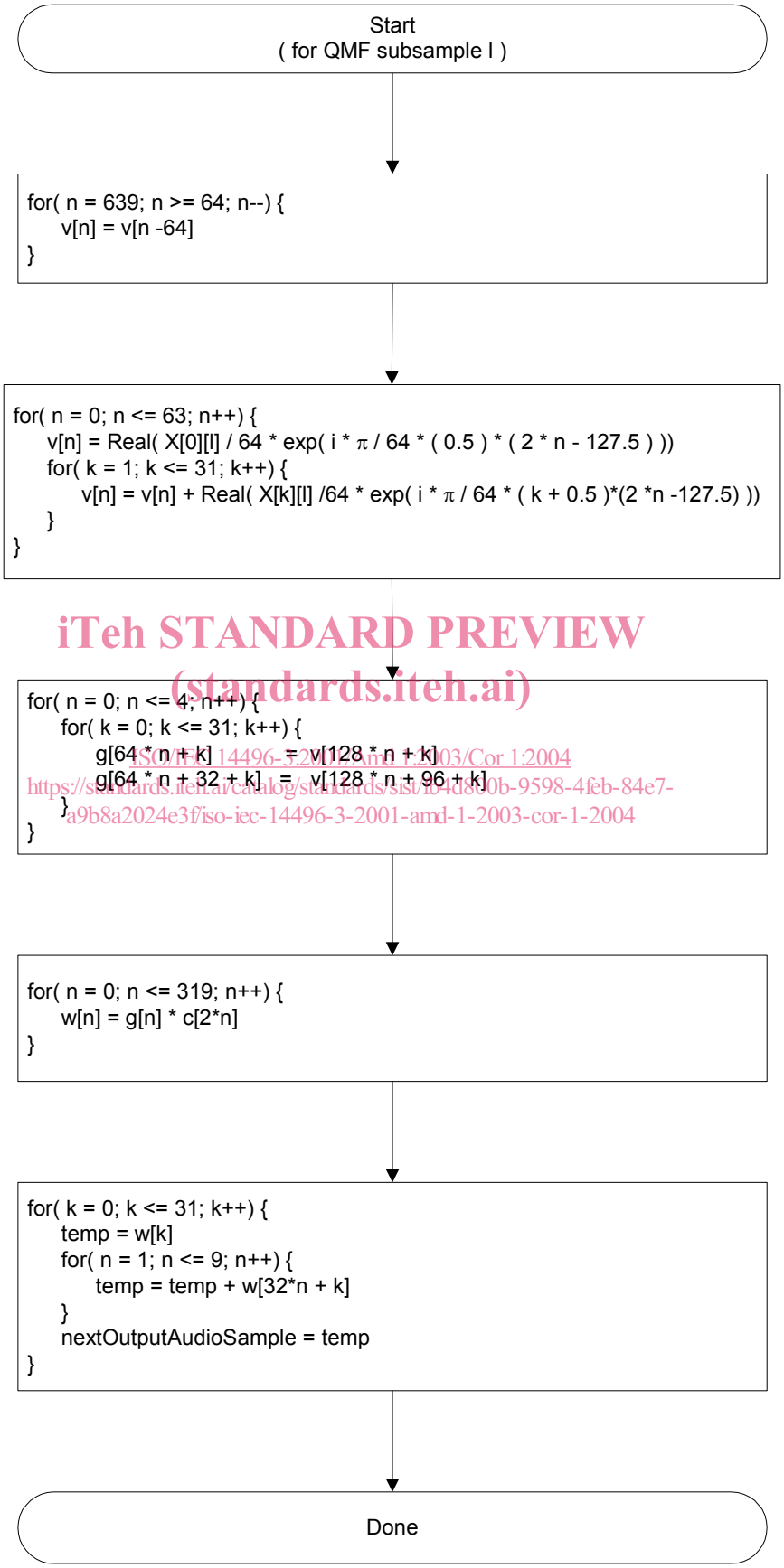
In ISO/IEC 14496-3:2001/Amd.1, subclause 4.6.18.4 and 4.6.18.8.2, replace all instances (in text flowcharts and pictures) of

$\mathbf{X}_{Low}$

by:

W

In ISO/IEC 14496-3:2001/Amd.1, subclause 4.6.18.4.3 Down sampled synthesis filterbank, in Figure 4.44, where it says "127", it should say "127.5", hence replace it by the following figure:



In ISO/IEC 14496-3:2001/Amd.1, subclause 4.6.18.5 SBR tool overview, replace:

start-up l_{Temp}, lsb' and $bsco'$ are set to zero. And where
$bsco = \max\left(INT\left(maxAACLine \cdot 32 / frameLength\right) - lsb, 0\right)$, and where

by:

start-up l_{Temp}, k_x' and $bsco'$ are set to zero. Where $bsco = 0$ unless a scalable core coder is used, for which
$bsco = \max\left(INT\left(maxAACLine \cdot 32 / frameLength\right) - k_x, 0\right)$, and where

In ISO/IEC 14496-3:2001/Amd.1, subclause 4.6.18.5 SBR Tool overview, replace:

The output from the filtering is stored in the matrix:

$$\mathbf{X}_{Low}\left(k,\ l\ +\ t_{HFGen}\right),\ 0\ \le\ k\ <\ 32,\ 0\ \le\ l\ <\ numTimeSlots \cdot RATE$$

by:

The subband filtered low band is defined by $\mathbf{X}_{Low}$ according to:

$$\mathbf{X}_{Low}\left(k,\ l\ \right) = \begin{cases} \mathbf{W}(k, l - t_{HFGen}) & , 0 \le k < k_x,\ t_{HFGen} \le l < l_f + t_{HFGen} \\ 0 & , k_x \le k < 32,\ t_{HFGen} \le l < l_f + t_{HFGen} \\ \mathbf{W}'(k, l + l_f - t_{HFGen}) & , 0 \le k < k'_x,\ 0 \le l < t_{HFGen} \\ 0 & , k'_x \le k < 32,\ 0 \le l < t_{HFGen} \end{cases}$$

where $\mathbf{W}'$ is the $\mathbf{W}$ matrix from the previous frame, and k'_x is the k_x value from the previous frame, and where $l_f = numTimeSlots \cdot RATE$. If scalable SBR is used the following apply instead of the equation above:

$$\mathbf{X}_{Low}\left(k,\ l\ \right) = \begin{cases} \mathbf{W}(k, l - t_{HFGen}) & , 0 \le k < 32,\ t_{HFGen} \le l < l_f + t_{HFGen} \\ \mathbf{W}'(k, l + l_f - t_{HFGen}) & , 0 \le k < 32,\ 0 \le l < t_{HFGen} \end{cases}$$

In ISO/IEC 14496-3:2001/Amd.1, subclause 4.6.18.5 SBR Tool overview, replace:

$$\mathbf{X}(k,l) = \begin{cases} \mathbf{X}_{Low}\left(k, l + t_{HFAdj}\right) & , 0 \le k < k_x' + bsco',\ 0 \le l < l_{Temp} \\ \mathbf{X}_{Low}\left(k, l + t_{HFAdj}\right) & , 0 \le k < k_x + bsco,\ l_{Temp} \le l < numTimeSlots \cdot RATE \\ \mathbf{Y}\left(k, l + t_{HFAdj}\right) & , k_x' + bsco' \le k < k_x + M,\ 0 \le l < l_{Temp} \\ \mathbf{Y}\left(k, l + t_{HFAdj}\right) & , k_x + bsco \le k < k_x + M,\ l_{Temp} \le l < numTimeSlots \cdot RATE \\ 0 & , \max\left(k_x + bsco, k_x + M\right) + M \le k < 64,\ 0 \le l < numTimeSlots \cdot RATE \end{cases}$$ "

by:

$$\mathbf{X}(k,l)=\begin{cases}\mathbf{X}_{Low}(k,l+t_{HFAdj}) & ,0 \le k < k_x'+bsco',\ 0 \le l < l_{Temp}\\ \mathbf{Y}'(k,l+t_{HFAdj}+l_f) & ,k_x'+bsco' \le k < k_x'+M',\ 0 \le l < l_{Temp}\\ 0 & ,\max(k'_x+bsco',k'_x+M') \le k < 64,\ 0 \le l < l_{Temp}\\ \mathbf{X}_{Low}(k,l+t_{HFAdj}) & ,0 \le k < k_x+bsco,\ l_{Temp} \le l < l_f\\ \mathbf{Y}(k,l+t_{HFAdj}) & ,k_x+bsco \le k < k_x+M,\ l_{Temp} \le l < l_f\\ 0 & ,\max(k_x+bsco,k_x+M) \le k < 64,\ l_{Temp} \le l < l_f\end{cases}$$

where

$l_f = numTimeSlots \cdot RATE$

and

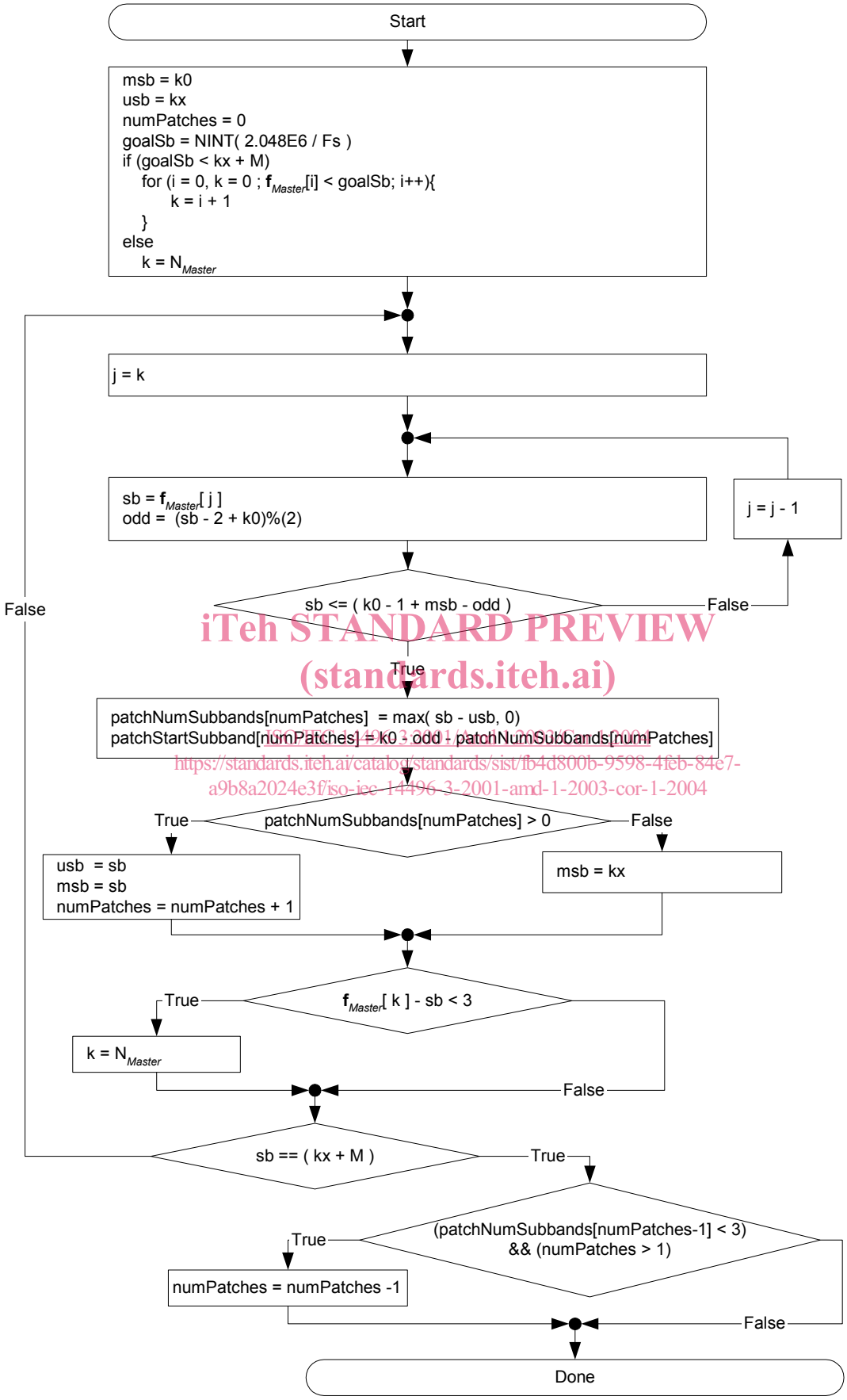
In ISO/IEC 14496-3:2001/Amd.1, subclause 4.6.18.6.2 Inverse filtering , replace:

where **bs_invf_mode´** are the **bs_invf_mode** values from the previous SBR frame.

by:

where **bs_invf_mode´** are the **bs_invf_mode** values from the previous SBR frame, and are assumed to be zero for the first frame.

In ISO/IEC 14496-3:2001/Amd.1, subclause 4.6.18.6.3 HF generator, replace Figure 4.46 by the following figure:



In ISO/IEC 14496-3:2001/Amd.1, subclause 4.6.18.7.2 Mapping, replace:

$$\mathbf{S}_{IndexMapped}(m-k_x,l)=\begin{cases}0 & if\ m\neq INT\left(\dfrac{\mathbf{f}_{TableHigh}(i+1)+\mathbf{f}_{TableHigh}(i)}{2}\right)\\ \mathbf{S}_{Index}(i)\cdot\delta_{Step}(i,l) & if\ m=INT\left(\dfrac{\mathbf{f}_{TableHigh}(i+1)+\mathbf{f}_{TableHigh}(i)}{2}\right)\end{cases}$$

for $\mathbf{f}_{TableHigh}(i)\le m<\mathbf{f}_{TableHigh}(i+1)$, $0\le i<N_{High}$, $0\le l<L_E$

where

$$\delta_{Step}(i,l)=\begin{cases}1 & if\ (l\ge l_A)\,OR\,(\mathbf{S}'_{Index}(i)=1)\\ 0 & otherwise\end{cases},$$

by:

$$\mathbf{S}_{IndexMapped}(m-k_x,l)=\begin{cases}0 & if\ m\neq INT\left(\dfrac{\mathbf{f}_{TableHigh}(i+1)+\mathbf{f}_{TableHigh}(i)}{2}\right)\\ \mathbf{S}_{Index}(i)\cdot\delta_{Step}(m-k_x,l) & if\ m=INT\left(\dfrac{\mathbf{f}_{TableHigh}(i+1)+\mathbf{f}_{TableHigh}(i)}{2}\right)\end{cases}$$

for $\mathbf{f}_{TableHigh}(i)\le m<\mathbf{f}_{TableHigh}(i+1)$, $0\le i<N_{High}$, $0\le l<L_E$

where

$$\delta_{Step}(m,l)=\begin{cases}1 & if\ (l\ge l_A)\,OR\,(\mathbf{S}'_{IndexMapped}(m,L'_E-1)=1)\\ 0 & otherwise\end{cases},$$

and replace:

and $\mathbf{S}'_{Index}(i)$ is $\mathbf{S}_{Index}(i)$ of the previous SBR frame.

by:

and $\mathbf{S}'_{IndexMapped}$ is $\mathbf{S}_{IndexMapped}$ of the previous SBR frame for the same frequency range. If the frequency range is larger for the current frame, the entries for the QMF subbands not covered by the previous $\mathbf{S}_{IndexMapped}$ are assumed to be zero.

In ISO/IEC 14496-3:2001/Amd.1, subclause 4.6.18.7.2 Mapping, replace:

$$\mathbf{S}_{Mapped}(m-k_x,l)=\delta_S(i,l),\ l_i\le m<u_i,\begin{cases}u_i=\mathbf{F}\left(k(i+1,l),\mathbf{r}(l)\right)\\ l_i=\mathbf{F}\left(k(i,l),\mathbf{r}(l)\right)\end{cases}$$

for $0 \le i < N_{High}$, $0 \le l < L_E$ where $k(i,l)$ is defined by

$$\begin{cases} \mathbf{F}(i,HI) \ge \mathbf{F}(k(i,l),LO), \mathbf{F}(i+1,HI) \le \mathbf{F}(k(i,l)+1,LO) & ,\mathbf{r}(l) = LO \\ k(i,l) = i & ,\mathbf{r}(l) = HI \end{cases}$$

and where

$$\delta_S(i,l) = \begin{cases} 1 & ,1 \in \{\mathbf{S}_{IndexMapped}(k,l): \quad \mathbf{F}(k(i,l),\mathbf{r}(l)) \le k < \mathbf{F}(k(i+1,l),\mathbf{r}(l))\} \\ 0 & ,otherwise \end{cases}.$$

In order to handle the varying frequency resolution of the envelope scalefactors, $k(i,l)$ is introduced. For a given high frequency resolution band, $k(i,l)$ gives the proper indices to the corresponding low frequency resolution band of which the former is a subset, if the current SBR envelope is of low frequency resolution. Finally, the $\delta_S(i,l)$ function returns one if any entry in the $\mathbf{S}_{IndexMapped}$ matrix is one within the given boundaries, i.e. if an additional sinusoid is present within the present frequency band.

by:

$$\mathbf{S}_{Mapped}(m-k_x,l) = \delta_S(i,l), l_i \le m < u_i, \begin{cases} u_i = \mathbf{F}(i+1,\mathbf{r}(l)) \\ l_i = \mathbf{F}(i,\mathbf{r}(l)) \end{cases}$$

for $0 \le i < \mathbf{n}(\mathbf{r}(l))$, $0 \le l < L_E$

where

$$\delta_S(i,l) = \begin{cases} 1 & ,1 \in \{\mathbf{S}_{IndexMapped}(j-k_x,l): \quad \mathbf{F}(i,\mathbf{r}(l)) \le j < \mathbf{F}(i+1,\mathbf{r}(l))\} \\ 0 & ,otherwise \end{cases}.$$

The $\delta_S(i,l)$ function returns one if any entry in the $\mathbf{S}_{IndexMapped}$ matrix is one within the given boundaries, i.e. if an additional sinusoid is present within the present frequency band. The $\mathbf{S}_{Mapped}$ matrix is hence one for all QMF subbands in the scalefactor bands where an additional sinusoid shall be added.

In ISO/IEC 14496-3:2001/Amd.1, subclause 4.6.18.7.6 Calculation of Additional HF Signal Components, replace the equation:

$$\mathbf{S}_M(m,l) = \sqrt{\mathbf{E}_{OrigMapped}(m,l) \cdot \frac{\mathbf{S}_{Mapped}(m,l)}{1+\mathbf{Q}_{Mapped}(m,l)}} \quad , \quad 0 \le m < M, 0 \le l < L_E$$

by:

$$\mathbf{S}_M(m,l) = \sqrt{\mathbf{E}_{OrigMapped}(m,l) \cdot \frac{\mathbf{S}_{IndexMapped}(m,l)}{1+\mathbf{Q}_{Mapped}(m,l)}}, \quad 0 \le m < M, 0 \le l < L_E$$

In ISO/IEC 14496-3:2001/Amd.1, subclause 4.6.18.7.5 Calculation of gain, replace the equation:

$$\mathbf{G}(m,l) = \begin{cases} \sqrt{\dfrac{\mathbf{E}_{OrigMapped}(m,l)}{(\varepsilon + \mathbf{E}_{Curr}(m,l)) \cdot (1+\delta(l) \cdot \mathbf{Q}_{Mapped}(m,l))}} & if \quad \mathbf{S}_M(m,l) = 0 \\ \sqrt{\dfrac{\mathbf{E}_{OrigMapped}(m,l)}{(\varepsilon + \mathbf{E}_{Curr}(m,l))} \cdot \dfrac{\mathbf{Q}_{Mapped}(m,l)}{(1+\mathbf{Q}_{Mapped}(m,l))}} & if \quad \mathbf{S}_M(m,l) \neq 0 \end{cases}, 0 \le m < M, 0 \le l < L_E$$

by:

$$\mathbf{G}(m,l) = \begin{cases} \sqrt{\dfrac{\mathbf{E}_{OrigMapped}(m,l)}{(\varepsilon + \mathbf{E}_{Curr}(m,l)) \cdot (1+\delta(l) \cdot \mathbf{Q}_{Mapped}(m,l))}} & if \quad \mathbf{S}_{Mapped}(m,l) = 0 \\ \sqrt{\dfrac{\mathbf{E}_{OrigMapped}(m,l)}{(\varepsilon + \mathbf{E}_{Curr}(m,l))} \cdot \dfrac{\mathbf{Q}_{Mapped}(m,l)}{(1+\mathbf{Q}_{Mapped}(m,l))}} & if \quad \mathbf{S}_{Mapped}(m,l) \neq 0 \end{cases}, 0 \le m < M, 0 \le l < L_E$$

In ISO/IEC 14496-3:2001/Amd.1, subclause 4.6.18.7.5 Calculation of gain, replace the equation:

$$\mathbf{G}_{Max_{Temp}}(k,l) = \sqrt{\frac{\varepsilon_0 + \sum_{i=\mathbf{f}_{TableLim}(k)}^{\mathbf{f}_{TableLim}(k+1)-1} \mathbf{E}_{OrigMapped}(i,l)}{\varepsilon_0 + \sum_{i=\mathbf{f}_{TableLim}(k)}^{\mathbf{f}_{TableLim}(k+1)-1} \mathbf{E}_{Curr}(i,l)}} \cdot \mathbf{limGain}(bs\_limiter\_gains), \quad 0 \le k < N_L, 0 \le l < L_E$$

$$\mathbf{G}_{Max}(m,l) = \min\left(\mathbf{G}_{Max_{Temp}}(k(m),l), 10^5\right), \quad 0 \le m < M, 0 \le l < L_E$$

where $k(m)$ is defined by $\mathbf{f}_{TableLim}(k(m)) \le m < \mathbf{f}_{TableLim}(k(m)+1)$,

by:

$$\mathbf{G}_{Max_{Temp}}(k,l) = \sqrt{\frac{\varepsilon_0 + \sum_{i=\mathbf{f}_{TableLim}(k)-k_x}^{\mathbf{f}_{TableLim}(k+1)-1-k_x} \mathbf{E}_{OrigMapped}(i,l)}{\varepsilon_0 + \sum_{i=\mathbf{f}_{TableLim}(k)-k_x}^{\mathbf{f}_{TableLim}(k+1)-1-k_x} \mathbf{E}_{Curr}(i,l)}} \cdot \mathbf{limGain}(bs\_limiter\_gains), \quad 0 \le k < N_L, 0 \le l < L_E$$

$$\mathbf{G}_{Max}(m,l) = \min\left(\mathbf{G}_{Max_{Temp}}(k(m),l), 10^5\right), \quad 0 \le m < M, 0 \le l < L_E$$

where $k(m)$ is defined by $\mathbf{f}_{TableLim}(k(m)) \le m + k_x < \mathbf{f}_{TableLim}(k(m)+1)$,