**INTERNATIONAL STANDARD ISO/IEC 14496-3:2005**
TECHNICAL CORRIGENDUM 2

Published 2008-04-15

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

TECHNICAL CORRIGENDUM 2

*Technologies de l'information — Codage des objets audiovisuels —*

*Partie 3: Codage audio*

*RECTIFICATIF TECHNIQUE 2*

Technical Corrigendum 2 to ISO/IEC 14496-3:2005 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information.*

Note that there is no Technical Corrigendum 1 to ISO/IEC 14496-3:2005 since its content was incorporated in ISO/IEC 14496-3:2005.



**ICS 35.040** **Ref. No. ISO/IEC 14496-3:2005/Cor.2:2008(E)**



Published in Switzerland

*In subclause 8.5.1, replace Table 8.19 with:*

**Table 8.19 – Channel configuration in case channelConfiguration == 2**

| mode_ext | Full channel configuration | nrof_channels |
|---|---|---|
| 00 | Dual mono (ch0=left, ch1=right) | 2 |
| 01 | Parametric Stereo | 1 |
| 10 | Reserved | Na |
| 11 | Reserved | Na |

*In subclause 8.5.2, replace:*

**s_phi[sf][ch][n]** – For sub-frame sf and channel ch, this represents the phase parameter of the n-th sinusoid. This value is converted into a phase value in radians in the range [-π, π> conforming to

$$sp_q[n] = 2 \cdot sp_e \cdot \text{s\_phi[sf][ch][n]},$$

where $sp_e$ represents the absolute phase error ( $sp_e = \frac{\pi}{32}$ ) and $sp_q$ represents the dequantized absolute phase (in radians) The allowed range for s_phi is [-16, 15]; the representation level +16 is represented by –16 (because +π == -π).

*with:*

**s_phi[sf][ch][n]** – For sub-frame sf and channel ch, this represents the phase parameter of the n-th sinusoid. This value is converted into a phase value in radians in the range [-π, π> conforming to

$$sp_q[n] = 2 \cdot sp_e \cdot \text{s\_phi[sf][ch][n]},$$

where $sp_e$ represents the absolute phase error ( $sp_e = \frac{\pi}{32}$ ) and $sp_q$ represents the dequantized absolute phase (in radians). The allowed range for s_phi is [-16, 15].



*In subclause 8.5.2, replace:*

The default and the fine quantization grids for IID, iid_quant = %0 and iid_quant = %1 are as provided in 8.B.18 and 8.B.17 respectively.

*with:*

The default and the fine quantization grids for IID, iid_quant = %0 and iid_quant = %1 are as provided in Table 8.25 and Table 8.26 respectively.

 

*In subclause 8.5.2, replace:*

**enable_ext** – The PS extension layer is enabled using the enable_ext bit. If it is set to %1 the IPD and OPD parameters are sent. If it's disabled, i.e. %0, the extension layer is skipped.

*with:*

**enable_ext** – The PS extension layer is enabled using the enable_ext bit. If it is set to %1 the IPD and OPD parameters are sent. If it is disabled, i.e. %0, the extension layer is skipped and **enable_ipdopd** is set to %0.



*In subclause 8.5.2, replace:*

**iid_dt[e]** – This flag describes for envelope index e, whether the IID parameters are coded differentially over time (iid_dt==%1) or over frequency (iid_dt==%0). In the case iid_mode is different from the previous envelope (e-1), iid_dt[e] shall have the value 0% forcing frequency differential coding.

*with:*

**iid_dt[e]** – This flag describes for envelope index e, whether the IID parameters are coded differentially over time (iid_dt==%1) or over frequency (iid_dt==%0). In the case iid_mode of the current ps_data() element is different from iid_mode of the previous ps_data() element, iid_dt[0] shall have the value 0% forcing frequency differential coding.

*In subclause 8.5.2, replace:*

**icc_dt[e]** – This flag describes for envelope index e, whether the ICC parameters are coded differentially over time (icc_dt==%1) or over frequency (icc_dt==%0). In the case icc_mode is different from the previous envelope (e-1), icc_dt[e] shall have the value 0% forcing frequency differential coding.

*with:*

**icc_dt[e]** – This flag describes for envelope index e, whether the ICC parameters are coded differentially over time (icc_dt==%1) or over frequency (icc_dt==%0). In the case icc_mode of the current ps_data() element is different from icc_mode of the previous ps_data() element, icc_dt[0] shall have the value 0% forcing frequency differential coding.

*In subclause 8.5.2, replace:*

**enable_ipdopd** – The application of IPD and OPD parameters in the bit-stream is denoted by enable_ipdopd. If set (enable_ipdopd==%1) IPD and OPD parameters are sent, if disabled (enable_ipdopd==%0) no IPD and OPD parameters are sent for the current frame in the bit-stream. The quantization grid for both IPD and OPD is provided in Table 8.31. If no IPD or OPD data is sent in the bit-stream, all IPD and OPD parameters are set to 0 (i.e. index=0).

*with:*

**enable_ipdopd** – The application of IPD and OPD parameters in the bit-stream is denoted by enable_ipdopd. If set (enable_ipdopd==%1) IPD and OPD parameters are sent, if disabled (enable_ipdopd==%0) no IPD and OPD parameters are sent for the current frame in the bit-stream. In the case enable_iid==%0, enable_ipdopd shall not be set to %1. The quantization grid for both IPD and OPD is provided in Table 8.31. If no IPD or OPD data is sent in the bit-stream, all IPD and OPD parameters are set to 0 (i.e. index=0).

*In subclause 8.5.2, replace:*

**ipd_dt[e]** – This flag describes for envelope index e, whether the IPD parameters are coded differentially over time (ipd_dt==%1) or over frequency (ipd_dt==%0). In the case iid_mode is different from the previous envelope (e-1), ipd_dt[e] shall have the value 0% forcing frequency differential coding.

*with:*

**ipd_dt[e]** – This flag describes for envelope index e, whether the IPD parameters are coded differentially over time (ipd_dt==%1) or over frequency (ipd_dt==%0). In the case iid_mode of the current ps_data() element is different from iid_mode of the previous ps_data() element, ipd_dt[0] shall have the value 0% forcing frequency differential coding.

*In subclause 8.5.2, replace:*

**opd_dt[e]** – This flag describes for envelope index e, whether the OPD parameters are coded differentially over time (opd_dt==%1) or over frequency (opd_dt==%0). In the case iid_mode is different from the previous envelope (e-1), opd_dt[e] shall have the value 0% forcing frequency differential coding.

*with:*

**opd_dt[e]** – This flag describes for envelope index e, whether the OPD parameters are coded differentially over time (opd_dt==%1) or over frequency (opd_dt==%0). In the case iid_mode of the current ps_data() element is different from iid_mode of the previous ps_data() element, opd_dt[0] shall have the value 0% forcing frequency differential coding.

*In subclause 8.5.2, ipd_par_df[e][b], replace:*

where ipd_par[e][b-1] represents the IPD index of the previous IPD value for envelope e. The IPD value *ipd*[*b*] is obtained by using ipd_par[e][b] as an index to Table 8.28.

*with:*

where ipd_par[e][b-1] represents the IPD index of the previous IPD value for envelope e. The IPD value *ipd*[*b*] is obtained by using ipd_par[e][b] as an index to Table 8.31.

*In subclause 8.5.2, opd_par_df[e][b], replace:*

where opd_par[e][b-1] represents the OPD index of the previous OPD value for envelope e. The OPD value *opd*[*b*] is obtained by using opd_par[e][b] as an index to Table 8.28.

*with:*

where opd_par[e][b-1] represents the OPD index of the previous OPD value for envelope e. The OPD value *opd*[*b*] is obtained by using opd_par[e][b] as an index to Table 8.31.

*In subclause 8.6.2.2, replace:*

For a continuation ($sf = [K+1, K+\kappa-1]$), the representation levels s_delta_cont_freq_pha[sf][ch][n] are converted to a quantized prediction error $\Delta[sf][ch[n]$ using Table 8.35 with index == 2.

*with:*

For a continuation (*sf* = [*K*+1, *K*+κ-1]), the representation levels s_delta_cont_freq_pha[sf][ch][n] are converted to a quantized prediction error Δ[*sf*][*ch*][*n*] using Table 8.35 with index == 2.

*In subclause 8.6.2.2, replace:*

In order to avoid very small or very large entries in prediction error, the adaptation is only done if the absolute value of the inner level, 0.75*c*[*sf*+*1*][*ch*][*p*], is between π/128 and 3π/8. In the case were the inner level is less than or equal to π/128 or greater than or equal to 3π/8, the scale factor *c*[*sf*+1][*ch*][*p*] is set to *c*[*sf*][*ch*][*n*]. This is illustrated in Table 8.35.

*with:*

In order to avoid very small or very large entries in prediction error, the adaptation is only done if the absolute value of the inner level, 0.75*c*[*sf*+*1*][*ch*][*p*], is between π/128 and 3π/8.

*In subclause 8.6.3.3.4, replace:*

where $k = [\mathrm{n\_nrof\_den} - 1..0]$ and

*with:*

where $m = [\mathrm{n\_nrof\_den} - 1..0]$ and

*In subclause 8.6.4.6.1, substitute Table 8.44 by:*

**Table 8.44 - The number of stereo bands depends on the number of parameters for IID and ICC.**

| Number of IID parameters | number of ICC parameters | number of stereo bands |
|---|---|---|
| 10 | 10 | 20 (i.e., 10,20 stereo band configuration) |
| 10 | 20 | |
| 20 | 10 | |
| 20 | 20 | |
| 10,20 | 34 | 34 |
| 34 | 10,20 | |
| 34 | 34 | |

*In subclause 8.6.4.6.1, replace:*

The averaging process denoted by e.g. (2*$idx_0$ + $idx_1$)/2 in Table 8.45 and Table 8.46 is carried out for the integer index representation $idx_k$ of the IID or ICC parameters prior to dequantization, according to ANSI-C integer arithmetic.

*with:*

The averaging process denoted by e.g. ($idx_0$ + $idx_1$)/2 in Table 8.45 and (2*$idx_0$ + $idx_1$)/2 Table 8.46 is carried out for the integer index representation $idx_k$ of the IID or ICC parameters prior to dequantization, according to ANSI-C integer arithmetic.



 

*Replace in subclause 8.6.4.6.2:*

In order to generate the QMF subband signals for the subband samples $n = n_e + 1 \ldots n_{e+1}$ the parameters at position $n_e$ and $n_{e+1}$ are required as well as the subband domain signals $s_k(n)$ and $d_k(n)$ for $n = n_e + 1 \ldots n_{e+1}$ (see Figure 8.25). For IPD/OPD, the parameters at position $n_{e-1}$ are needed in addition. $n_e$ represents the start position for envelope e. In the case frameclass == %1 (VAR_BORDERS), the border positions $n_e$ are obtained by borderposition[e]. In the case frameclass == %0 (FIX_BORDERS), the border positions $n_e$ are obtained by means of

$$n_e = \left\lfloor \frac{numQMFSlots * (e+1)}{num\_env} \right\rfloor - 1, \qquad e = [0, \ldots, num\_env - 1].$$

*with:*

In order to generate the QMF subband signals for the subband samples $n = n_e + 1 \ldots n_{e+1}$ the parameters at position $n_e$ and $n_{e+1}$ are required as well as the subband domain signals $s_k(n)$ and $d_k(n)$ for $n = n_e + 1 \ldots n_{e+1}$ (see Figure 8.25). For IPD/OPD, the parameters at position $n_{e-1}$ are needed in addition. $n_e$ represents the start position for envelope e. In the case e=0, $n_{-1}$ represents the last parameter position of the previous stereo frame. For the first stereo frame the IPD/OPD parameters at position $n_{-1}$ are initialized to zero. In the case frameclass == %1 (VAR_BORDERS), the border positions $n_e$ are obtained by border_position[e]. In the case frameclass == %0 (FIX_BORDERS), the border positions $n_e$ are obtained by means of

$$n_e = \left\lfloor \frac{numQMFSlots * (e+1)}{num\_env} \right\rfloor - 1, \qquad e = [0, \ldots, num\_env - 1].$$

*In subclause 8.6.4.6.3.1. replace:*

If the number of stereo bands changes from 10,20 in the previous frame to 34 in the current frame, the coefficients $h_{11}(b)$, $h_{12}(b)$, $h_{21}(b)$, and $h_{22}(b)$ at the end of the previous frame are mapped from 20 to 34 stereo bands according to Table 8.45 (by substituting $idx_b$ by $h_{ij}(b)$, where $ij$ is 11, 12, 21, or 22) prior to further processing as defined in subclause 8.6.4.6.3 and the IPD/OPD smoothing state variables are reset, i.e., $opd(b, n_{e-1}) = 0$, $ipd(b, n_{e-1}) = 0$, $opd(b, n_e) = 0$, and $ipd(b, n_e) = 0$. The frequency resolution of the hybrid QMF analysis filterbank (see subclause 8.6.4.3) is changed instantaneously to the 34 stereo band configuration. The state variables of the decorrelation process are reset to zero (see Table 8.47).

If the number of stereo bands changes from 34 in the previous frame to 10,20 in the current frame, the coefficients $h_{11}(b)$, $h_{12}(b)$, $h_{21}(b)$, and $h_{22}(b)$ at the end of the previous frame are mapped from 34 to 20 stereo bands according to Table 8.46 (by substituting $idx_b$ by $h_{ij}(b)$, where $ij$ is 11, 12, 21, or 22) prior to further processing as defined in subclause 8.6.4.6.3 and the IPD/OPD smoothing state variables are reset, i.e., $opd(b, n_{e-1}) = 0$, $ipd(b, n_{e-1}) = 0$, $opd(b, n_e) = 0$, and $ipd(b, n_e) = 0$. The frequency resolution of the hybrid QMF analysis filterbank (see subclause 8.6.4.3) is changed instantaneously to the 20 stereo band configuration. The state variables of the decorrelation process are reset to zero (see Table 8.47).

*with:*

If the number of stereo bands changes from 10,20 in the previous frame to 34 in the current frame, the coefficients $h_{11}(b)$, $h_{12}(b)$, $h_{21}(b)$, and $h_{22}(b)$ at the end of the previous frame are mapped from 20 to 34 stereo bands according to Table 8.45 (by substituting idx$_b$ by $h_{ij}(b)$, where $ij$ is 11, 12, 21, or 22). Then the coefficients $H_{11}(k,n)$, $H_{12}(k,n)$, $H_{21}(k,n)$ and $H_{22}(k,n)$ for the end of the previous frame are derived according to the four equations given in subclause 8.6.4.6.3 prior to further processing as defined in subclause 8.6.4.6.4. The IPD/OPD smoothing state variables are reset, i.e., $opd(b,n_{e-1})=0$, $ipd(b,n_{e-1})=0$, $opd(b,n_e)=0$, and $ipd(b,n_e)=0$. The frequency resolution of the hybrid QMF analysis filterbank (see subclause 8.6.4.3) is changed instantaneously to the 34 stereo band configuration. The state variables of the decorrelation process are reset to zero (see Table 8.47).

If the number of stereo bands changes from 34 in the previous frame to 10,20 in the current frame, the coefficients $h_{11}(b)$, $h_{12}(b)$, $h_{21}(b)$, and $h_{22}(b)$ at the end of the previous frame are mapped from 34 to 20 stereo bands according to Table 8.46 (by substituting idx$_b$ by $h_{ij}(b)$, where $ij$ is 11, 12, 21, or 22). Then the coefficients $H_{11}(k,n)$, $H_{12}(k,n)$, $H_{21}(k,n)$ and $H_{22}(k,n)$ for the end of the previous frame are derived according to the four equations given in subclause 8.6.4.6.3 prior to further processing as defined in subclause 8.6.4.6.4. The IPD/OPD smoothing state variables are reset, i.e., $opd(b,n_{e-1})=0$, $ipd(b,n_{e-1})=0$, $opd(b,n_e)=0$, and $ipd(b,n_e)=0$. The frequency resolution of the hybrid QMF analysis filterbank (see subclause 8.6.4.3) is changed instantaneously to the 20 stereo band configuration. The state variables of the decorrelation process are reset to zero (see Table 8.47).



*In subclause 8.6.4.6.4. replace:*

The intermediate values for $H_{11}(k,n)$, $H_{12}(k,n)$, $H_{21}(k,n)$ and $H_{22}(k,n)$ at positions $n=n_e+1 \ldots n_{e+1}$ are obtained by means of linear interpolation conforming to.

$$\begin{aligned}
H_{11}(k,n) &= H_{11}(k,n_e)+(n-n_e)\frac{H_{11}(k,n_{e+1})-H_{11}(k,n_e)}{n_{e+1}-n_e} \\
H_{12}(k,n) &= H_{12}(k,n_e)+(n-n_e)\frac{H_{12}(k,n_{e+1})-H_{12}(k,n_e)}{n_{e+1}-n_e} \\
H_{21}(k,n) &= H_{21}(k,n_e)+(n-n_e)\frac{H_{21}(k,n_{e+1})-H_{21}(k,n_e)}{n_{e+1}-n_e} \\
H_{22}(k,n) &= H_{22}(k,n_e)+(n-n_e)\frac{H_{22}(k,n_{e+1})-H_{22}(k,n_e)}{n_{e+1}-n_e}
\end{aligned}.$$

*with:*

The intermediate values for $H_{11}(k,n)$, $H_{12}(k,n)$, $H_{21}(k,n)$ and $H_{22}(k,n)$ at positions $n=n_e+1 \ldots n_{e+1}$ are obtained by means of linear interpolation conforming to.

$$
\begin{aligned}
H_{11}(k,n) &= H_{11}(k,n_e) + (n-n_e)\frac{H_{11}(k,n_{e+1}) - H_{11}(k,n_e)}{n_{e+1}-n_e} \\
H_{12}(k,n) &= H_{12}(k,n_e) + (n-n_e)\frac{H_{12}(k,n_{e+1}) - H_{12}(k,n_e)}{n_{e+1}-n_e} \\
H_{21}(k,n) &= H_{21}(k,n_e) + (n-n_e)\frac{H_{21}(k,n_{e+1}) - H_{21}(k,n_e)}{n_{e+1}-n_e} \\
H_{22}(k,n) &= H_{22}(k,n_e) + (n-n_e)\frac{H_{22}(k,n_{e+1}) - H_{22}(k,n_e)}{n_{e+1}-n_e}
\end{aligned}.
$$

Special cases:

a) For the first region of a stereo frame (see subclause 8.6.4.4), with $n = 0 \ldots n_0 - 1$ the following applies:

$$
\begin{aligned}
H_{11}(k,n) &= H_{11}(k,n_{-1}) + n\frac{H_{11}(k,n_0) - H_{11}(k,n_{-1})}{n_0} \\
H_{12}(k,n) &= H_{12}(k,n_{-1}) + n\frac{H_{12}(k,n_0) - H_{12}(k,n_{-1})}{n_0} \\
H_{21}(k,n) &= H_{21}(k,n_{-1}) + n\frac{H_{21}(k,n_0) - H_{21}(k,n_{-1})}{n_0} \\
H_{22}(k,n) &= H_{22}(k,n_{-1}) + n\frac{H_{22}(k,n_0) - H_{22}(k,n_{-1})}{n_0}
\end{aligned},
$$

where $H_{xx}(k,n_{-1})$ represents the $H_{xx}(k,n_{num\_env-1})$ coefficients obtained from the previous stereo frame. Note that for the first stereo frame $H_{xx}(k,n_{-1})$ is initialized to zero.

b) For the last region of a stereo frame (see 8.6.4.4) with $n = n_{num\_env-1} \ldots numQMFSlots - 1$ the following applies:

$$
\begin{aligned}
H_{11}(k,n) &= H_{11}(k,n_{num\_env-1}) \\
H_{12}(k,n) &= H_{12}(k,n_{num\_env-1}) \\
H_{21}(k,n) &= H_{21}(k,n_{num\_env-1}) \\
H_{22}(k,n) &= H_{22}(k,n_{num\_env-1})
\end{aligned}.
$$



*In subclause 8.6.4.6.3.2. replace:*

In the case IPD and OPD are enabled as indicated by (enable_ipdopd==1) the following procedure is applied. First the IPD and OPD values are smoothed over time according to

$$
\begin{aligned}
\varphi_{opd}(b) &= \angle\left\{\frac{1}{4}\exp\left(j \cdot opd(b,n_{e-1})\right) + \frac{1}{2}\exp\left(j \cdot opd(b,n_e)\right) + \exp\left(j \cdot opd(b,n_{e+1})\right)\right\} \\
\varphi_{ipd}(b) &= \angle\left\{\frac{1}{4}\exp\left(j \cdot ipd(b,n_{e-1})\right) + \frac{1}{2}\exp\left(j \cdot ipd(b,n_e)\right) + \exp\left(j \cdot ipd(b,n_{e+1})\right)\right\}
\end{aligned}.
$$