INTERNATIONAL STANDARD ISO/IEC 23003-2:2010
TECHNICAL CORRIGENDUM 2

Published 2014-09-01

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

Information technology — MPEG audio technologies —

Part 2:
Spatial Audio Object Coding (SAOC)

TECHNICAL CORRIGENDUM 2

Technologies de l'information — Technologies audio MPEG —

Partie 2: Codage d'objet audio spatial (SAOC)

RECTIFICATIF TECHNIQUE 2

Technical Corrigendum 2 to ISO/IEC 23003-2:2010 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

ICS 35.040 **Ref. No. ISO/IEC 23003-2:2010/Cor.2:2014(E)**

Published in Switzerland

Replace the following text in the subclause "7.7.2.4 Stereo to binaural "x-2-b" processing mode" (inappropriate indexing):

$$v = \left(\mathbf{D}^1 + \mathbf{D}^2\right)\mathbf{E}\left(\mathbf{D}^1 + \mathbf{D}^2\right)^* + \varepsilon^2.$$

with:

$$v = \left(\mathbf{D}^0 + \mathbf{D}^1\right)\mathbf{E}\left(\mathbf{D}^0 + \mathbf{D}^1\right)^* + \varepsilon^2$$

Replace the following text in the subclause "7.7.2.4 Stereo to binaural "x-2-b" processing mode" (inappropriate and confusing indexing):

$$e^x_{i,j} = e_{i,j}\left(\frac{d^x_i}{d^1_i + d^2_i}\right)\left(\frac{d^x_j}{d^1_j + d^2_j}\right)$$

with:

$$e^x_{i,j} = e_{i,j}\left(\frac{d_{x,i}}{d_{0,i} + d_{1,i}}\right)\left(\frac{d_{x,j}}{d_{0,j} + d_{1,j}}\right)$$

Replace the following text in the subclause "7.7.2.4 Stereo to binaural "x-2-b" processing mode" (redundant definition of already introduced variables and inappropriate indexing):

The downmix matrix $\mathbf{D}^{l,x}$ of size $1\times N$ with elements $d^{l,x}_i$ can be found as

$$d^{l,1}_i = 10^{0.05\,DMG^l_i}\sqrt{\frac{10^{0.1DCLD^l_i}}{1+10^{0.1DCLD^l_i}}},\quad d^{l,2}_i = 10^{0.05\,DMG^l_i}\sqrt{\frac{1}{1+10^{0.1DCLD^l_i}}}.$$

The stereo downmix matrix $\mathbf{D}^l$ of size $2\times N$ with elements $d^l_{x,i}$ can be found as

$$d^l_{x,i} = d^{l,x}_i.$$

with:

where the downmix matrix $\mathbf{D}^x$ of size $1\times N$ and $\mathbf{D}$ of size $2\times N$ contain elements $d_{x,i}$, $(x = 0,1)$.

Replace the following text in the subclause "7.6.2.4 Derivation of the MPS parameters" (incorrect statement: "parameters are quantized, formatted as MPS bitstream"):

The obtained CLD, ICC and ADG parameters are quantized, formatted as MPS bitstream and fed into the MPS decoder.

with:

The obtained $\mathbf{D}_{\mathrm{ADG}}$, $\mathbf{D}_{\mathrm{CLD}}$ and $\mathbf{D}_{\mathrm{ICC}}$ parameters are fed into the MPS decoder.

Remove the following equation and its lead-in text in the subclause "7.6.3.3 Stereo Preprocessing" (this equation contradicts to the one given above, namely $\mathbf{P}_2 = \begin{cases} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & ,r_{1,2} > \varepsilon^2, \\ \mathbf{v}_{\mathrm{R}} diag\left(\mathbf{w}_{\mathrm{d}}\right) & ,otherwise. \end{cases}$ *):*

and finally the mix matrix,

$$\mathbf{P}_2 = \begin{pmatrix} \mathbf{v}_{\mathrm{R1}} & \mathbf{v}_{\mathrm{R2}} \end{pmatrix} \begin{pmatrix} \mathbf{w}_{\mathrm{d1}} & 0 \\ 0 & \mathbf{w}_{\mathrm{d2}} \end{pmatrix}$$

Replace the following text in the subclause "7.6.3.3 Stereo Preprocessing" (incomplete specification $\mathbf{P}_1$ *and inconsistent notation* $\mathbf{G}_{\mathrm{Mod}}$*):*

and the mix matrix $\mathbf{G}_{\mathrm{Mod}}$ is given as:

$$\mathbf{G}_{\mathrm{Mod}} = \begin{cases} diag\left(\mathbf{g}_{\mathrm{vec}}\right)\mathbf{G}, & r_{1,2} > 0, \\ \mathbf{G}, & otherwise. \end{cases}$$

Similarly, the mix matrix $\mathbf{P}_2$ is given as:

$$\mathbf{P}_2 = \begin{cases} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & ,r_{1,2} > \varepsilon^2, \\ \mathbf{v}_{\mathrm{R}} diag\left(\mathbf{w}_{\mathrm{d}}\right) & ,otherwise. \end{cases}$$

with:

The SAOC downmix preprocessor unit modifies the downmix according to the MPS parameters derived from SAOC data and rendering information. The matrices $\mathbf{G}$, $\mathbf{P}_1$ and $\mathbf{P}_2$ of size 2×2 are given as

$$\mathbf{G} = \begin{cases} diag\left(\mathbf{g}_{\mathrm{vec}}\right) \cdot \hat{\mathbf{G}}, & r_{1,2} > \varepsilon^2, \\ \hat{\mathbf{G}}, & otherwise. \end{cases}$$

$$\mathbf{P}_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \hat{\mathbf{G}},$$

$$\mathbf{P}_2 = \begin{cases} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & ,r_{1,2} > \varepsilon^2, \\ \mathbf{v}_R diag(\mathbf{w}_d) & ,otherwise. \end{cases}$$

Replace the term:

"section"

with:

"subclause"

in the text.

Replace the following text in the subclause "7.13.4 Low Power SAOC processing" (incorrect statement "transcoding (and stereo output) case"; incorrect equation: "$r_{1,2} > 0$"):

in the stereo downmix transcoding (and stereo output) case, always forcing $r_{1,2} > 0$ in section 7.6.3.3.

with:

in the stereo downmix case, always forcing $r_{1,2} > \varepsilon^2$ in subclause 7.5.5.1.1.

Remove the following text in the subclause "7.7.2 Downmix processor" (redundant text; incorrect statement for the SAOC transcoding mode "yielding the final output PCM signal"):

For the decoder mode of the SAOC system, the output signal of the downmix processor (represented in the hybrid QMF domain) is fed into the corresponding synthesis filterbank as described in ISO/IEC 23003-1:2007 yielding the final output PCM signal. The downmix processing incorporates the mono, stereo and, if required, subsequent binaural processing.

Note: The following text exists for the SAOC decoding mode:

The SAOC parameter processing unit calculates the upmix parameters ($\mathbf{G}$, $\mathbf{P}_1$ and $\mathbf{P}_2$) derived from the SAOC data (OLD, IOC, DMG, DCLD), rendering (and HRTF) information. The downmix processor applies these parameters and uses the corresponding synthesis filterbank yielding the final output PCM signal.

Replace the following text in the subclause "7.6.2.2 SAOC processing" (redundant and missing description):

Associated with each OTT element is the channel level difference (CLD) parameter that describes the relative level differences between the two output channels, and the inter-channel correlation (ICC) parameter for the desired cross-correlation between the output signals.

In contrast to the MPS the SAOC bitstream comprises the relative level of each audio object in the downmix signal, termed the object level difference (OLD) and the inter-object correlation (IOC).

The task of the SAOC parameter processing unit is to estimate all CLD and ICC parameters from the SAOC data and the rendering matrix. This process is performed for each OTT element independently.

with:

The Channel Level Difference (CLD) and Inter Channel Correlation (ICC) parameters associated with each OTT element describe the relative level differences and cross-correlation between two output channels, correspondently. The Channel Prediction Coefficient (CPC) parameters associated with TTT box represent Prediction Coefficients needed for reconstruction of the third output signal from two input ones. The Arbitrary Downmix Gains (ADG) parameters describe the modification of the downmix signal.

Replace the following text in the subclause "7.6.3.2.1 Stereo downmix ("x-2-5") processing mode" (redundant text; missing equation: term $\mathbf{J}$ *is undefended; incorrect definition: "*$\mathbf{J} = \left(\mathbf{DED}^*\right)^{-1}$*"; confusing notation):*

To avoid numerical problems when calculating the term $\mathbf{J} = \left(\mathbf{DED}^*\right)^{-1}$, $\mathbf{J}$ is modified. First the eigenvalues $\lambda_{1,2}$ of $\mathbf{J}$ are calculated, solving $\det(\mathbf{J} - \lambda_{1,2}\mathbf{I}) = 0$.

Eigenvalues are sorted in descending ($\lambda_1 \geq \lambda_2$) order and the eigenvector corresponding to the larger eigenvalue is calculated according to the equation above. It is assured to lie in the positive x-plane (first element has to be positive). The second eigenvector is obtained from the first by a – 90 degrees rotation:

$$\mathbf{J} = \left(\mathbf{v}_1 \mathbf{v}_2\right) \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \left(\mathbf{v}_1 \mathbf{v}_2\right)^* .$$

with:

where $\mathbf{J} \approx \left(\mathbf{DED}^*\right)^{-1}$.

The matrix $\mathbf{J}$ of size 2×2 is defined as

$$\mathbf{J} = \left(\mathbf{v}_1^J \quad \mathbf{v}_2^J\right) \begin{pmatrix} \hat{\lambda}_1^J & 0 \\ 0 & \hat{\lambda}_2^J \end{pmatrix} \left(\mathbf{v}_1^J \quad \mathbf{v}_2^J\right)^* ,$$

where both modified eigenvalues $\hat{\lambda}_i^J$ are determined as

$$\hat{\lambda}_i^J = \max\left(\lambda_i^J, c^J\right), \qquad c^J = \frac{1}{80}.$$

The eigenvalues $\lambda_{1,2}^J$ ($\lambda_1^J \geq \lambda_2^J$) of the matrix $\mathbf{J}$ and corresponding eigenvectors $\mathbf{v}_{1,2}^J$ are calculated solving the following characteristic equations

$$\det\left(\mathbf{DED}^* - \lambda_{1,2}^J \mathbf{I}\right) = 0 ,$$

$$\left(\mathbf{DED}^{*} - \lambda_{1,2}^{J}\mathbf{I}\right)\mathbf{v}_{1,2}^{J} = 0\,.$$

Replace the following equation:

$$\mathbf{T} = \begin{pmatrix} t_{1,1} & t_{1,2} \\ t_{2,1} & t_{2,2} \\ t_{3,1} & t_{3,2} \end{pmatrix} = \mathbf{A}_3\mathbf{D}^{*} + \varepsilon\mathbf{I}\,.$$

with:

$$\mathbf{T} = \begin{pmatrix} t_{1,1} & t_{1,2} \\ t_{2,1} & t_{2,2} \\ t_{3,1} & t_{3,2} \end{pmatrix} = \mathbf{A}_3\left(\mathbf{D}\right)^{-1} + \varepsilon\,.$$

Replace the following equation:

$$\mathbf{w}_{\mathrm{d1}} = \min\left(\sqrt{\frac{\lambda_1}{r_{\mathrm{d1}} + \varepsilon}}, 2\right), \quad \mathbf{w}_{\mathrm{d2}} = \min\left(\sqrt{\frac{\lambda_2}{r_{\mathrm{d2}} + \varepsilon}}, 2\right),$$

with:

$$\mathbf{w}_{\mathrm{d}} = \left(\min\left(\sqrt{\frac{\lambda_1^{R}}{\max\left(r_{1,1}^{d}, \varepsilon^2\right)}}, 2\right), \quad \min\left(\sqrt{\frac{\lambda_2^{R}}{\max\left(r_{2,2}^{d}, \varepsilon^2\right)}}, 2\right)\right).$$

Replace the following equation:

$$CLD_1 = 10\log_{10}\left(\frac{\max\left(f_{1,1}, \varepsilon^2\right)}{\max\left(f_{5,5}, \varepsilon^2\right)}\right), \qquad ICC_1 = \frac{\max\left(f_{1,5}, \varepsilon^2\right) f_{1,5}}{\sqrt{\max\left(f_{1,1}, \varepsilon^2\right)\max\left(f_{5,5}, \varepsilon^2\right)}},$$

with:

$$CLD_1 = 10\log_{10}\left(\frac{\max\left(f_{1,1}, \varepsilon^2\right)}{\max\left(f_{5,5}, \varepsilon^2\right)}\right), \qquad ICC_1 = \frac{\max\left(f_{1,5}, \varepsilon^2\right)}{\sqrt{\max\left(f_{1,1}, \varepsilon^2\right)\max\left(f_{5,5}, \varepsilon^2\right)}},$$

Replace the following text in "6.1 Payloads for SAOC":

```
for (i=0; i<bsNumDmxChannels + 1; i++) {
    idxPDG[i] = EcDataSaoc(PDG, i, numBands);                    Note 1
}
```

with: