INTERNATIONAL STANDARD ISO/IEC 14496-14:2003
TECHNICAL CORRIGENDUM 1

Published 2006-04-01

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 14:
MP4 file format

TECHNICAL CORRIGENDUM 1

Technologies de l'information — Codage des objets audiovisuels —

Partie 14: Format de fichier MP4

RECTIFICATIF TECHNIQUE 1

Technical Corrigendum 1 to ISO/IEC 14496-14:2003 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 14496-14:2003/Cor.1:2006(E)**

© ISO/IEC 2006 – All rights reserved

--`,,```,,,,````-`-`,,`,,`,`,,`---

Published in Switzerland

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS

Not for Resale

Add the following Annex B after Annex A:

Annex B
(informative)

Handling of Audio Timestamps and Profile/Level Indication

As specified in ISO/IEC 14496-3 (Audio), the decoder produces a compositionUnit as output for every accessUnit it receives as input. An edit-list can be used to indicate the desired audio output (that is, the valid samples) from amongst the set of samples in the output compositionUnits. For example, an edit list might specify that the system using the decoder discard the first 1 024 audio samples (possibly the result from decoding a pre-roll accessUnit), and also discard some final samples of the decoded waveform (those resulting from rounding up the length of the wave-form to an audio frame boundary). This enables exact 'round trip' processing, whereby the output of the decoder has the same length as the input to the encoder, with the audio in the same temporal position.

Systems that see only the edit may feel that they are able to discard data not needed by the edits. When the analogous situation arises in video (when edits do not fall on random-access points) they are aware of the need to keep data back to the random access point preceding the start of the edit.

In this case the file can specify the need for "pre-roll" using a pre-roll sample group, for example a pre-roll value of −1 (minus one), to indicate to the system using the decoder that it must start the sequence of accessUnits presented to the decoder with the accessUnit immediately prior to the accessUnit whose corresponding compositionBuffer contains the start of the desired audio. This includes the cases of starting at the beginning of the audio (the start of the edit list), random access, or where the user has performed further editing in the encoded domain.

In addition, care should be taken when an audio signal can be decoded in either a backwards-compatible or enhanced fashion. As specified in ISO/IEC 14496-3, the timestamp (constructed from the time-to-sample table) applies to the backwards-compatible decoding. If a decoder applies additional 'delay' to the output waveform (i.e. the audio appears later in the audio output waveform than if just backwards-compatible decoding were performed), then the system using the decoder must be informed of this delay so that it can compensate for it, in order to maintain correct temporal behaviour (including synchronization).

If it is desired to label audio streams with their profile and level indications, an ExtensionProfileLevelDescriptor may be inserted in the ES_Descriptor, as stored in the 'esds' box of the audio stream. This is especially useful in cases where no IOD is present in the file.

--`,,```,,,,````-`-`,,`,,`,`,,`---

© ISO/IEC 2006 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale