

# INTERNATIONAL STANDARD

# ISO/IEC 13818-7

Fourth edition  
2006-01-15

---

---

## Information technology — Generic coding of moving pictures and associated audio information —

### Part 7: Advanced Audio Coding (AAC)

*Technologies de l'information — Codage générique des images  
animées et du son associé —*

*Partie 7: Codage du son avancé (AAC)*

---

---

Reference number  
ISO/IEC 13818-7:2006(E)



**PDF disclaimer**

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

© ISO/IEC 2006

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office  
Case postale 56 • CH-1211 Geneva 20  
Tel. + 41 22 749 01 11  
Fax + 41 22 749 09 47  
E-mail [copyright@iso.org](mailto:copyright@iso.org)  
Web [www.iso.org](http://www.iso.org)

Published in Switzerland

# Contents

Page

|   |    |
|---|----|
| Foreword.....   | v  |
| Introduction .....  | vi |
| 1 Scope .....   | 1  |
| 1.1 General.....  | 1  |
| 1.2 MPEG-2 AAC Tools Overview.....  | 1  |
| 2 Normative References .....  | 7  |
| 3 Terms and Definitions .....   | 7  |
| 4 Symbols and Abbreviations .....   | 14 |
| 4.1 Arithmetic Operators .....  | 14 |
| 4.2 Logical Operators .....   | 15 |
| 4.3 Relational Operators .....  | 15 |
| 4.4 Bitwise Operators .....   | 16 |
| 4.5 Assignment .....  | 16 |
| 4.6 Mnemonics .....   | 16 |
| 4.7 Constants .....   | 16 |
| 5 Method of Describing Bitstream Syntax .....   | 16 |
| 6 Syntax .....  | 18 |
| 6.1 Audio Data Interchange Format, ADIF.....  | 18 |
| 6.2 Audio Data Transport Stream, ADTS .....   | 19 |
| 6.3 Raw Data.....   | 21 |
| 7 Profiles and Profile Interoperability .....   | 33 |
| 7.1 Profiles .....  | 33 |
| 7.2 Profile Interoperability.....   | 35 |
| 8 Overall Data Structure .....  | 36 |
| 8.1 AAC Interchange Formats .....   | 36 |
| 8.2 Raw Data.....   | 41 |
| 8.3 Single Channel Element (SCE), Channel Pair Element (CPE) and Individual Channel<br>Stream (ICS) ..... | 45 |
| 8.4 Low Frequency Enhancement Channel (LFE) .....   | 51 |
| 8.5 Program Config Element (PCE).....   | 51 |
| 8.6 Data Stream Element (DSE) .....   | 56 |
| 8.7 Fill Element (FIL) .....  | 56 |
| 8.8 Extension Payload .....   | 57 |
| 8.9 Tables.....   | 61 |
| 8.10 Figures .....  | 70 |
| 9 Noiseless Coding.....   | 70 |
| 9.1 Tool Description.....   | 70 |
| 9.2 Definitions .....   | 71 |
| 9.3 Decoding Process.....   | 73 |
| 9.4 Tables.....   | 76 |
| 10 Quantization .....   | 76 |
| 10.1 Tool Description.....  | 76 |
| 10.2 Definitions .....  | 76 |
| 10.3 Decoding Process.....  | 76 |
| 11 Scalefactors.....  | 77 |
| 11.1 Tool Description.....  | 77 |

|  |                                     |            |
|--|-------------------------------------|------------|
| 11.2   | Definitions.....                    | 77         |
| 11.3   | Decoding Process.....               | 78         |
| 12   | Joint Coding .....                  | 79         |
| 12.1   | M/S Stereo.....                     | 79         |
| 12.2   | Intensity Stereo .....              | 80         |
| 12.3   | Coupling Channel .....              | 82         |
| 13   | Prediction.....                     | 86         |
| 13.1   | Tool Description.....               | 86         |
| 13.2   | Definitions.....                    | 86         |
| 13.3   | Decoding Process.....               | 87         |
| 13.4   | Diagrams .....                      | 93         |
| 14   | Temporal Noise Shaping (TNS) .....  | 93         |
| 14.1   | Tool Description.....               | 93         |
| 14.2   | Definitions.....                    | 94         |
| 14.3   | Decoding Process.....               | 94         |
| 15   | Filterbank and Block Switching..... | 96         |
| 15.1   | Tool Description.....               | 96         |
| 15.2   | Definitions.....                    | 96         |
| 15.3   | Decoding Process.....               | 97         |
| 16   | Gain Control.....                   | 101        |
| 16.1   | Tool Description.....               | 101        |
| 16.2   | Definitions.....                    | 102        |
| 16.3   | Decoding Process.....               | 102        |
| 16.4   | Diagrams .....                      | 109        |
| 16.5   | Tables .....                        | 109        |
| <b>Annex A (normative) Huffman Codebook Tables.....</b>            |                                     | <b>111</b> |
| <b>Annex B (informative) Information on Unused Codebooks .....</b> |                                     | <b>130</b> |
| <b>Annex C (informative) Encoder .....</b>                         |                                     | <b>131</b> |
| <b>Annex D (informative) Patent Holders .....</b>                  |                                     | <b>189</b> |
| <b>Annex E (informative) Registration Procedure .....</b>          |                                     | <b>190</b> |
| <b>Annex F (informative) Registration Application Form .....</b>   |                                     | <b>192</b> |
| <b>Annex G (informative) Registration Authority .....</b>          |                                     | <b>193</b> |
| <b>Bibliography .....</b>  |                                     | <b>194</b> |

## Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC 13818-7 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This fourth edition cancels and replaces the third edition (ISO 13818-7:2004), which has been technically revised. It also incorporates the Technical Corrigendum ISO/IEC 13818-7:2004/Cor.1:2005.

ISO/IEC 13818 consists of the following parts, under the general title *Information technology — Generic coding of moving pictures and associated audio information*:

- *Part 1: Systems*
- *Part 2: Video*
- *Part 3: Audio*
- *Part 4: Conformance testing*
- *Part 5: Software simulation* [Technical Report]
- *Part 6: Extensions for DSM-CC*
- *Part 7: Advanced Audio Coding (AAC)*
- *Part 9: Extension for real time interface for systems decoders*
- *Part 10: Conformance extensions for Digital Storage Media Command and Control (DSM-CC)*
- *Part 11: IPMP on MPEG-2 systems*

## Introduction

The standardization body ISO/IEC JTC 1/SC 29/WG 11, also known as the Moving Pictures Experts Group (MPEG), was established in 1988 to specify digital video and audio coding schemes at low data rates. MPEG completed its first phase of audio specifications (MPEG-1) in November 1992, ISO/IEC 11172-3. In its second phase of development, the MPEG Audio subgroup defined a multichannel extension to MPEG-1 audio that is backwards compatible with existing MPEG-1 systems (MPEG-2 BC) and defined an audio coding standard at lower sampling frequencies than MPEG-1, ISO/IEC 13818-3.

The International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) draw attention to the fact that it is claimed that compliance with this document may involve the use of patents.

The ISO and IEC take no position concerning the evidence, validity and scope of this patent right.

The holder of this patent right has assured the ISO and IEC that he is willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statement of the holder of this patent right is registered with the ISO and IEC. Information may be obtained from the companies listed in Annex D.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights other than those identified in Annex D. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

# Information technology — Generic coding of moving pictures and associated audio information —

## Part 7: Advanced Audio Coding (AAC)

### 1 Scope

#### 1.1 General

This International Standard describes the MPEG-2 audio non-backwards compatible standard called MPEG-2 Advanced Audio Coding, AAC [1], a higher quality multichannel standard than achievable while requiring MPEG-1 backwards compatibility. This MPEG-2 AAC audio standard allows for ITU-R “indistinguishable” quality according to [2] at data rates of 320 kbit/s for five full-bandwidth channel audio signals.

The AAC decoding process makes use of a number of required tools and a number of optional tools. Table 1 lists the tools and their status as required or optional. Required tools are mandatory in any possible profile. Optional tools may not be required in some profiles.

**Table 1 — AAC decoder tools**

| Tool Name                       | Required / Optional |
|---------------------------------|---------------------|
| Bitstream Formatter             | Required            |
| Noiseless Decoding              | Required            |
| Inverse quantization            | Required            |
| Rescaling                       | Required            |
| M/S                             | Optional            |
| Prediction                      | Optional            |
| Intensity                       | Optional            |
| Dependently switched coupling   | Optional            |
| TNS                             | Optional            |
| Filterbank / block switching    | Required            |
| Gain control                    | Optional            |
| Independently switched coupling | Optional            |

#### 1.2 MPEG-2 AAC Tools Overview

The basic structure of the MPEG-2 AAC system is shown in Figure 1 and Figure 2. As is shown in Table 1, there are both required and optional tools in the decoder. The data flow in this diagram is from left to right, top to bottom. The functions of the decoder are to find the description of the quantized audio spectra in the bitstream, decode the quantized values and other reconstruction information, reconstruct the quantized spectra, process the reconstructed spectra through whatever tools are active in the bitstream in order to arrive at the actual signal spectra as described by the input bitstream, and finally convert the frequency domain spectra to the time domain, with or without an optional gain control tool. Following the initial reconstruction and scaling of the spectrum reconstruction, there are many optional tools that modify one or more of the spectra in order to provide more efficient coding. For each of the optional tools that operate in the spectral domain, the option to “pass through” is retained, and in all cases where a spectral operation is omitted, the spectra at its input are passed directly through the tool without modification.

The input to the bitstream demultiplexer tool is the MPEG-2 AAC bitstream. The demultiplexer separates the parts of the MPEG-AAC data stream into the parts for each tool, and provides each of the tools with the bitstream information related to that tool.

The outputs from the bitstream demultiplexer tool are:

- The sectioning information for the noiselessly coded spectra,
- The noiselessly coded spectra,
- The M/S decision information (optional),
- The predictor state information (optional),
- The intensity stereo control information and coupling channel control information (both optional),
- The temporal noise shaping (TNS) information (optional),
- The filterbank control information, and
- The gain control information (optional).

The noiseless decoding tool takes information from the bitstream demultiplexer, parses that information, decodes the Huffman coded data, and reconstructs the quantized spectra and the Huffman and DPCM coded scalefactors.

The inputs to the noiseless decoding tool are:

- The sectioning information for the noiselessly coded spectra, and
- The noiselessly coded spectra.

The outputs of the Noiseless Decoding tool are:

- The decoded integer representation of the scalefactors, and
- The quantized values for the spectra.

The inverse quantizer tool takes the quantized values for the spectra, and converts the integer values to the non-scaled, reconstructed spectra. This quantizer is a non-uniform quantizer.

The input to the Inverse Quantizer tool is:

- The quantized values for the spectra.

The output of the inverse quantizer tool is:

- The un-scaled, inversely quantized spectra.

The rescaling tool converts the integer representation of the scalefactors to the actual values, and multiplies the un-scaled inversely quantized spectra by the relevant scalefactors.

The inputs to the rescaling tool are:

- The decoded integer representation of the scalefactors, and
- The un-scaled, inversely quantized spectra.

The output from the scalefactors tool is:

- The scaled, inversely quantized spectra.

The M/S tool converts spectra pairs from Mid/Side to Left/Right under control of the M/S decision information in order to improve coding efficiency.

The inputs to the M/S tool are:

- The M/S decision information, and
- The scaled, inversely quantized spectra related to pairs of channels.

The output from the M/S tool is:

- The scaled, inversely quantized spectra related to pairs of channels, after M/S decoding.

Note The scaled, inversely quantized spectra of individually coded channels are not processed by the M/S block, rather they are passed directly through the block without modification. If the M/S block is not active, all spectra are passed through this block unmodified.

The prediction tool reverses the prediction process carried out at the encoder. This prediction process re-inserts the redundancy that was extracted by the prediction tool at the encoder, under the control of the predictor state information. This tool is implemented as a second order backward adaptive predictor. The inputs to the prediction tool are:

- The predictor state information, and
- The scaled, inversely quantized spectra.

The output from the prediction tool is:

- The scaled, inversely quantized spectra, after prediction is applied.

Note If the prediction is disabled, the scaled, inversely quantized spectra are passed directly through the block without modification.

The intensity stereo tool implements intensity stereo decoding on pairs of spectra.

The inputs to the intensity stereo tool are:

- The inversely quantized spectra, and
- The intensity stereo control information.

The output from the intensity stereo tool is:

- The inversely quantized spectra after intensity channel decoding.

Note The scaled, inversely quantized spectra of individually coded channels are passed directly through this tool without modification, if intensity stereo is not indicated. The intensity stereo tool and M/S tool are arranged so that the operation of M/S and intensity stereo are mutually exclusive on any given scalefactor band and group of one pair of spectra.

The coupling tool for dependently switched coupling channels adds the relevant data from dependently switched coupling channels to the spectra, as directed by the coupling control information.

The inputs to the coupling tool are:

- The inversely quantized spectra, and
- The coupling control information.

The output from the coupling tool is:

- The inversely quantized spectra coupled with the dependently switched coupling channels.

Note The scaled, inversely quantized spectra are passed directly through this tool without modification, if coupling is not indicated. Depending on the coupling control information, dependently switched coupling channels might either be coupled before or after the TNS processing.

The coupling tool for independently switched coupling channels adds the relevant data from independently switched coupling channels to the time signal, as directed by the coupling control information.

The inputs to the coupling tool are:

- The time signal as output by the filterbank, and
- The coupling control information.

The output from the coupling tool is:

- The time signal coupled with the independently switched coupling channels.

Note The time signal is passed directly through this tool without modification, if coupling is not indicated.

The temporal noise shaping (TNS) tool implements a control of the fine time structure of the coding noise. In the encoder, the TNS process has flattened the temporal envelope of the signal to which it has been applied. In the decoder, the inverse process is used to restore the actual temporal envelope(s), under control of the TNS information. This is done by applying a filtering process to parts of the spectral data.

The inputs to the TNS tool are:

- The inversely quantized spectra, and
- The TNS information.

The output from the TNS block is:

- The inversely quantized spectra.

Note If this block is disabled, the inversely quantized spectra are passed through without modification.

The filterbank / block switching tool applies the inverse of the frequency mapping that was carried out in the encoder. An inverse modified discrete cosine transform (IMDCT) is used for the filterbank tool. The IMDCT can be configured to support either one set of 128 or 1024, or four sets of 32 or 256 spectral coefficients.

The inputs to the filterbank tool are:

- The inversely quantized spectra, and
- The filterbank control information.

The output(s) from the filterbank tool is (are):

- The time domain reconstructed audio signal(s).

When present, the gain control tool applies a separate time domain gain control to each of four frequency bands that have been created by the gain control PQF filterbank in the encoder. Then, it assembles four frequency bands and reconstructs the time waveform through the gain control tool's filterbank.

The inputs to the gain control tool are:

- The time domain reconstructed audio signal(s), and
- The gain control information.

The output(s) from the gain control tool is (are):

- The time domain reconstructed audio signal(s).

If the gain control tool is not active, the time domain reconstructed audio signal(s) are passed directly from the filterbank tool to the output of the decoder. This tool is used for the scalable sampling rate (SSR) profile only.

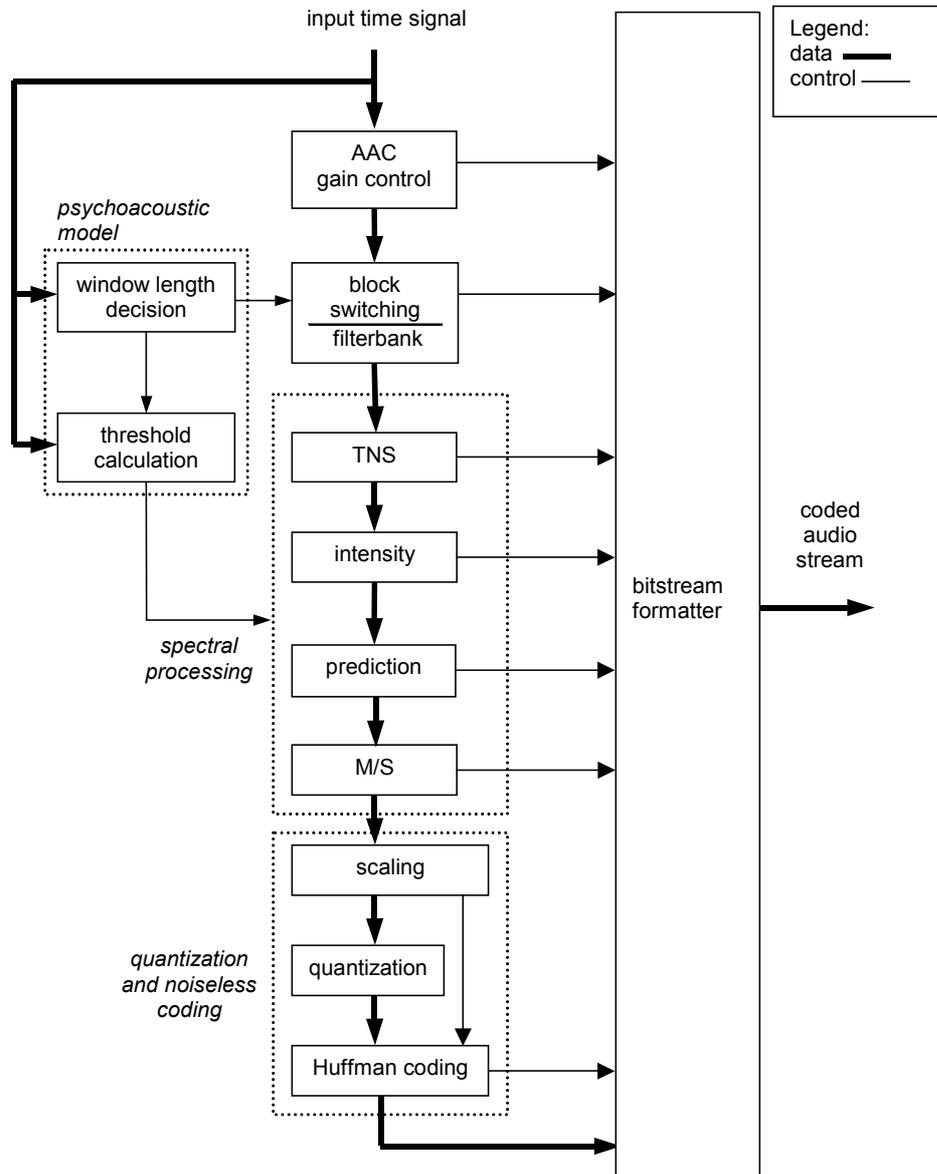


Figure 1 — MPEG-2 AAC Encoder Block Diagram

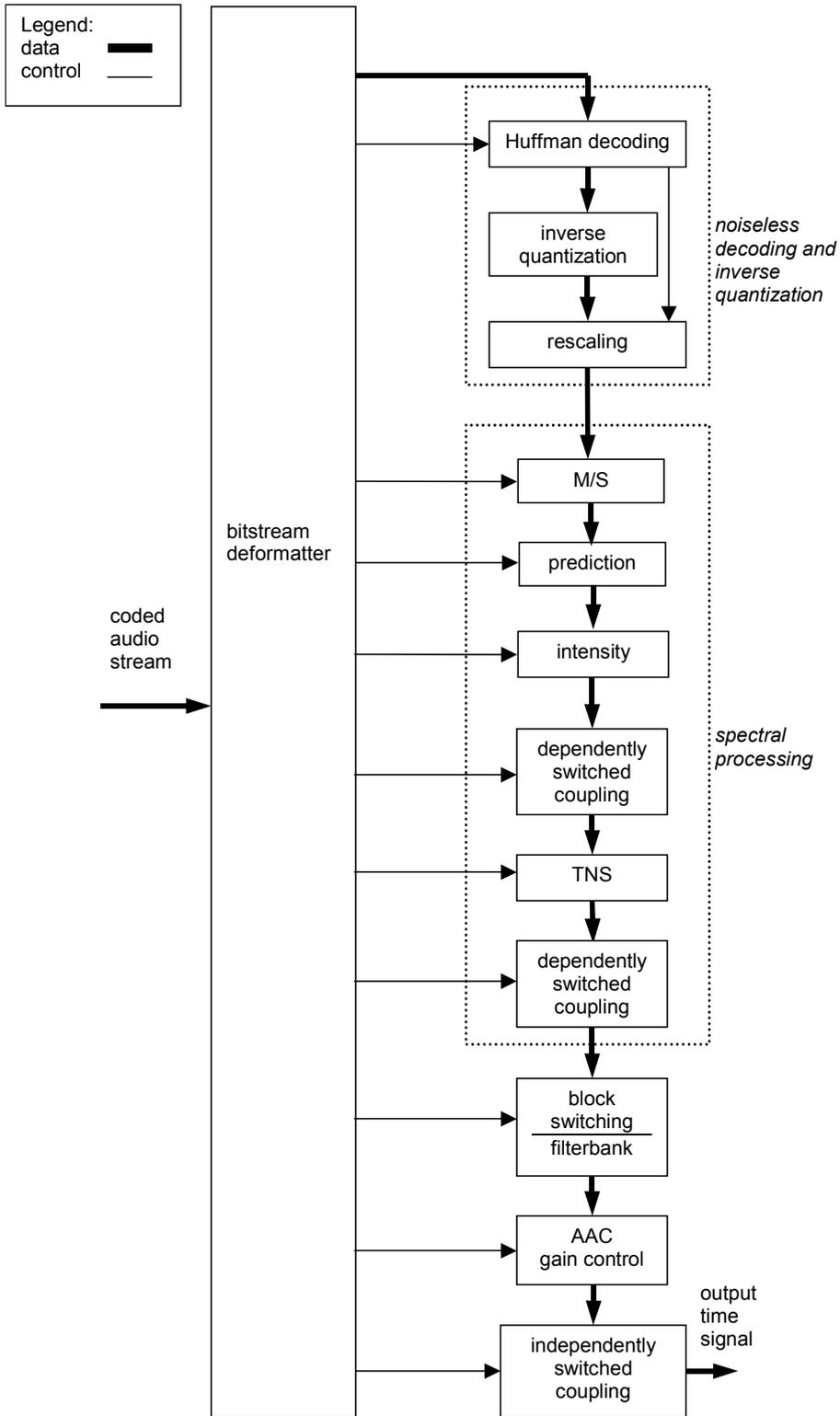


Figure 2 — MPEG-2 AAC Decoder Block Diagram

## 2 Normative References

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 11172-3: *Information technology — Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s — Part 3: Audio*

ISO/IEC 13818-1: *Information technology — Generic coding of moving pictures and associated audio information — Part 1: Systems*

ISO/IEC 13818-3: *Information technology — Generic coding of moving pictures and associated audio information — Part 3: Audio*

ISO/IEC 14496-3: *Information technology — Coding of audio-visual objects — Part 3: Audio*