

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

T.807

Amendment 1
(03/2008)

SERIES T: TERMINALS FOR TELEMATIC SERVICES
Still-image compression – JPEG 2000

Information technology – JPEG 2000 image coding
system: Secure JPEG 2000

Amendment 1: File format security

Recommendation ITU-T T.807 (2006) – Amendment 1

ITU-T T-SERIES RECOMMENDATIONS
TERMINALS FOR TELEMATIC SERVICES

Facsimile – Framework	T.0–T.19
Still-image compression – Test charts	T.20–T.29
Facsimile – Group 3 protocols	T.30–T.39
Colour representation	T.40–T.49
Character coding	T.50–T.59
Facsimile – Group 4 protocols	T.60–T.69
Telematic services – Framework	T.70–T.79
Still-image compression – JPEG-1, Bi-level and JBIG	T.80–T.89
Telematic services – ISDN Terminals and protocols	T.90–T.99
Videotext – Framework	T.100–T.109
Data protocols for multimedia conferencing	T.120–T.149
Telewriting	T.150–T.159
Multimedia and hypermedia framework	T.170–T.189
Cooperative document handling	T.190–T.199
Telematic services – Interworking	T.300–T.399
Open document architecture	T.400–T.429
Document transfer and manipulation	T.430–T.449
Document application profile	T.500–T.509
Communication application profile	T.510–T.559
Telematic services – Equipment characteristics	T.560–T.649
Still-image compression – JPEG 2000	T.800–T.849
Still-image compression – JPEG-1 extensions	T.850–T.899

For further details, please refer to the list of ITU-T Recommendations.

Information technology – JPEG 2000 image coding system:
Secure JPEG 2000

Amendment 1

File format security

Summary

Amendment 1 to Rec. ITU-T T.807 | ISO/IEC 15444-8 specifies JPSEC file format derived from the ISO base file format and modifications to JPEG family file format (including JP2, JPX and JPM) for protection and secure adaptation of scalable pictures, which is possibly encrypted and/or authenticated by the owner. The pictures could be either static pictures or time-sequenced pictures. In particular, the amendment provides functionality to do the following:

- To store coded media data corresponding to different scalability levels. Elementary stream (ES) is used for this purpose.
- To define tracks describing the characteristics of the coded media data stored in ES. For example, the track should be able to indicate scalability level (resolution, layer, region, etc.) and the rate-distortion hints of the coded media data in order to facilitate easy and secure adaptation.
- To define new file format boxes to signal protection tools and parameters applied to coded media data or metadata. The protection tools can be applied to either static JPEG 2000 pictures or time-sequenced JPEG 2000 pictures.
- The protection tools defined in this amendment can be applied to JPEG family file formats including JP2, JPX and JPM and ISO-derived file formats such as MJ2 for motion JPEG.

Source

Amendment 1 to Recommendation ITU-T T.807 (2006) was approved on 15 March 2008 by ITU-T Study Group 16 (2005-2008) under Recommendation ITU-T A.8 procedure. An identical text is also published as Amendment 1 to ISO/IEC 15444-8.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at <http://www.itu.int/ITU-T/ipr/>.

© ITU 2009

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

CONTENTS

	<i>Page</i>
1) Clause 2: Normative references	1
2) Clause 3: Terms and definitions.....	1
3) Annex E: File format security	3
Annex E – File Format Security	3
E.1 Scope.....	3
E.2 Introduction	3
E.3 Extension to ISO base media file format	5
E.4 Elementary stream and sample definitions.....	14
E.5 Protection at file format level	16
E.6 Examples (Informative).....	18
E.7 Boxes defined in ISO/IEC 15444-12 (informative)	28

Document ID: 123456789

**INTERNATIONAL STANDARD ISO/IEC 15444-8
RECOMMENDATION ITU-T T.807**

**Information technology – JPEG 2000 image coding system:
Secure JPEG 2000**

Amendment 1

File format security

1) Clause 2: Normative references

Add the following references:

- Recommendation ITU-T T.803 (2002) | ISO/IEC 15444-4:2004, *Information technology – JPEG 2000 image coding system: Conformance testing*.
- ISO/IEC 13818-11:2004, *Information technology – Generic coding of moving pictures and associated audio information – Part 11: IPMP on MPEG-2 systems*.
- ISO/IEC 15444-6:2003, *Information technology – JPEG 2000 image coding system – Part 6: Compound image file format*.
- ISO/IEC 15444-12:2005, *Information technology – JPEG 2000 image coding system – Part 12: ISO base media file format (technically identical to ISO/IEC 14496-12)*.

2) Clause 3: Terms and definitions

a) Rewrite the first paragraph as follows (with the changes underlined):

For the purposes of this Recommendation | International Standard, the following definitions apply. The definitions defined in ITU-T Rec. T.800 | ISO/IEC 15444-1 clause 3 and ISO/IEC 15444-12:2005 clause 3 apply to this Recommendation | International Standard.

b) Add the following terms and definitions:

Normal decoder

Standard decoder is a process to decode a codestream that is fully compliant with the normative part of coding standard. Its behaviour is not defined if it tries to decode a non-compliant codestream.

Adaptive-format decoder

Adaptive-format decoder is a process to decode a codestream which is not fully compliant with the normative part of the coding standard. It shall reconstruct the media (possibly with low quality or resolution) even if the codestream has missing packets or inconsistent packet headers. For example, an adaptive-format decoder is able to understand a simply-transcoded codestream, such as the one that has its highest resolution packets removed.

Elementary Stream (ES)

Elementary streaming contains a sequence of samples, where each sample could be a video frame or a contiguous section of audio data. A sample in ES contains media data, ByteData structure, pointer structure, container structure, or any mixture of the above.

Self-Contained ES

Self-contained ES contains only media data, whose format is not defined in this amendment. The self-contained ES could be stored in MDAT box co-located with the file format specified in this amendment, or be stored in a separate file whose format is not specified by this amendment.

Composed ES

Composed ES may contain a mixture of ByteData, pointer and container structures, that is, its samples are composed with data from other elementary streams. A composed ES can either copy (using ByteData structure) or reference (using pointer) data from other ESes.

Scalable Composed ES

Scalable composed ES is made up of samples that may not be decodable by themselves. It may need to be combined with other scalable composed ESes to form a fully decodable codestream. Scalable composed ES is designed to support scalability, i.e., to make media data "thinable". For example, for a motion JPEG 2000 codestream where each picture has three layers, it can be divided into 3 scalable composed ESes: the first one consists of all layer 0 data, the second one consists of all layer 1 data and the third one consists of all layer 2 data.

Decodable Composed ES

Decodable composed ES is made up of samples that are decodable by themselves. It is designed for simple adaptation where the adaptor just needs to retrieve data pointed by pointer structure and remove the wrapper to form a fully scalable codestream. For example, for a motion JPEG 2000 codestream where each picture has three layers, it can form 3 decodable composed ESes: the first one consists of layer 0 data, the second one consists of layer 0 and layer 1 data and the third one consists of layer 0, 1 and 2 data.

Adaptor/transcoder

Adaptor/transcoder is a process to transform media data to lower scalability level, like lower resolution or lower quality or bit-rate, by removing portions of the file. The adaptor/transcoder can transform media data based on the information specified in this amendment. An adaptor/transcoder shall update byte offset values in file format parameters that are impacted by the process.

Secure adaptor/transcoder

Secure adaptor/transcoder is a process to transform encrypted or authenticated media data without necessity to decrypt or regenerate the MAC or signature. Thus, end-to-end security remains for the transcoded media data.

JPEG 2000-aware adaptor/transcoder

JPEG 2000-aware adaptor/transcoder combines one or more scalable composed ESes to form a fully decodable media codestream. It should have the capability to generate the headers and markers of media codestream and modify the packet index, such that the adapted codestream can be decoded by a normal decoder. It may also add empty packets to replace the removed ones, or it may insert POC marker.

Simple adaptor/transcoder

Simple adaptor/transcoder is able to transform data based on information specified by this amendment. It may not be capable of generating media headers or modifying packet indices. It simply retrieves data pointed by pointer structure and removes the wrappers, and the resulting codestream can be decoded by adaptive-format decoder, which can cope with missing packets and inconsistent headers.

Authentication adaptor/transcoder

An authentication adaptor/transcoder removes data that is not verifiable with the available media data and authentication data. For example, in a streaming system, some media packets may be lost during transmission. A file format receiver may reconstruct the received data to the best of its ability based on the available data. Then, an authentication adaptor/transcoder can determine which data can be verified, and then remove the packets that are not verified. The resulting file only contains the decodable, verified data.

Container

Container structure is used to wrap a sample in a composed ES. It might contain any number of ByteData or pointer structures, but is not allowed to contain another container structure.

Pointer

Pointer structure is used to reference a data segment in another ES. It must be contained inside a container structure.

ByteData

ByteData structure is used to wrap a data segment which is physically located in a composed ES. It must be contained inside a container structure.

4CC Code

4CC code is a 32-bit identifier, normally 4 printable characters. A 4CC code can be used to indicate the file type, the type of file format box, type of a file format track, type of a file format sample description and type of file format track reference. A 4CC code must be registered with a registration authority.

3) Annex E: File format security

Create a new annex and add the following text:

Annex E**File Format Security**

(This annex forms an integral part of this Recommendation | International Standard)

E.1 Scope

This annex specifies JPSEC file format derived from the ISO base file format and modifications to JPEG family file format (including JP2, JPX and JPM) for protection and secure adaptation of scalable pictures, which is possibly encrypted and/or authenticated by the owner. The pictures could be either static pictures or time-sequenced pictures. In particular, this annex provides functionality to do the following:

- To store coded media data corresponding to different scalability levels. Elementary stream (ES) is used for this purpose. There are three types of ESes, self-contained ES, scalable composed ES and decodable composed ES.
- To define tracks describing the characteristics of the coded media data stored in ES. For example, the track should be able to indicate scalability level (resolution, layer, region, etc.) and the rate-distortion hints of the coded media data in order to facilitate easy and secure adaptation.
- To define new file format boxes to signal protection tools and parameters applied to coded media data or metadata. The protection tools can be applied to either static JPEG 2000 pictures or time-sequenced JPEG 2000 pictures.
- The protection tools defined in this amendment can be applied to JPEG family file formats including JP2, JPX and JPM and ISO-derived file formats such as MJ2 for motion JPEG.

E.2 Introduction**E.2.1 Security protection at file format level**

This annex describes a JPSEC file format derived from the ISO base file format and modifications to JPEG family file format, to add security protection to JPEG 2000 pictures at the file format level. The protection applied at the file format level can be classified into two types: item-based protection and sample-based protection, both structures are defined by the ISO base file format. The item-based protection is designed to protect any byte ranges (including coded media data and metadata) while the sample-based protection is designed to protect time-sequenced media including JPEG 2000 pictures.

When the security tools applied change the data length, it shall update all pointers and length fields in all boxes, to ensure correct parsing by the reader.

E.2.2 Item-based protection

This annex describes two item-based protection schemes in the ISO base file format, by leveraging the syntax and structures specified by the JPSEC standard. Specifically, it describes schemes for decryption and authentication. Each item in the ItemLocationBox is protected by one or more protection schemes in the ItemProtectionBox. When multiple schemes are used (or chained together), the order in which they are applied may be significant and thus must be specified. This annex also specifies how such operations should be chained together. In addition, the

ItemDescriptionBox and ItemCorrespondingBox are added into the ISO base file format to allow the flexible processing properties that are provided by JPSEC. Specifically, the ItemDescriptionBox allows media-dependent metadata (such as resolution, quality layer, spatial region, and color space component) to be associated with different portions of the file. These descriptions can be provided regardless of whether protection is applied. When used with scalable coded pictures, this allows the file to be scaled down or transcoded without parsing or decoding the media data. In cases where protection is applied, this provides the benefit of enabling transcoding without requiring decryption.

E.2.3 Sample-based protection of scalable media

For time-sequenced pictures, this annex adds syntaxes to facilitate scalability at the file format level, including scalable composed elementary stream (ES), decodable composed ES, pointer structure, container structure and ByteData structure. The scalable coded pictures can be divided (either physically or virtually) into elementary streams at different scalability level, such that the adaptor/transcoder can "thin" media data with low complexity.

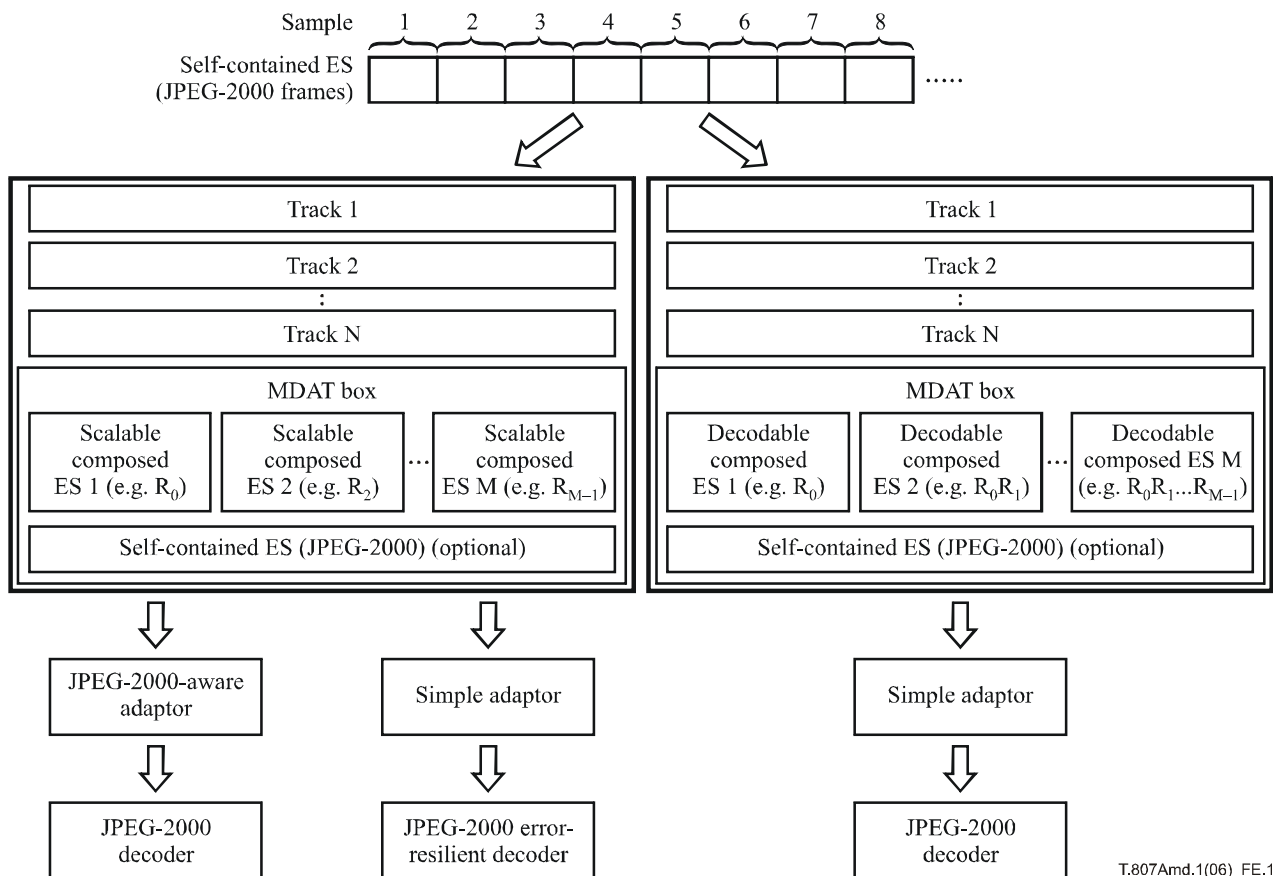
Figure E.1 gives an overview of the file format specified by this annex and also shows how the specified FF is used to adapt the media data.

Given a sequence of JPEG 2000 pictures (also referred to as *Self-contained ES*), there are two approaches to construct the file format. In the first approach, the MDAT box contains one or more *Scalable Composed ESes*, each of which corresponds to one scalability level of the media data, e.g., a resolution or a layer. The scalable composed ES must be stored in MDAT box that is co-located with the file format. The self-contained ES can be located in either MDAT box in the same file, or a different file whose format is not specified in this amendment. The scalable composed ES may not be decodable by itself, it may need to be combined with other scalable composed ESes to generate fully decodable JPEG 2000 pictures. In the second approach, the MDAT box contains one or more *Decodable Composed ESes*, and each ES constitutes fully decodable JPEG 2000 pictures by itself. Similarly, decodable composed ESes must be stored in MDAT box co-located with the file format, and the self-contained ES can be stored in either MDAT box in the same file, or a different file whose format is not specified by this annex.

Each scalable composed ES or decodable composed ES must be described by at least one track. The characteristics of the ES (like resolution, layer, and region) are indicated in SampleEntryBox inside each track.

To generate a fully decodable JPEG 2000 codestream from scalable composed ESes, a JPEG 2000-aware adaptor should have the capability to dynamically generate the image headers (based on the number of resolutions, layers and region in the adapted codestream), to insert empty packets or to insert POC markers as needed to make the resulting codestream decodable by any standard decoder. However, if a simple adaptor is used, the resulting codestream may have an inconsistent image header and there may be a missing packet, which require a JPEG 2000 adaptive-format decoder.

As a decodable composed ES is decodable by itself, a simple adaptor is sufficient to generate fully compliant JPEG 2000 pictures.



T.807Amd.1(06)_FE.1

Figure E.1 – System diagram for time-sequenced scalable media

Each elementary stream is described by at least one media track, and its characteristics are described in SampleDescriptionEntry or SampleGroupEntryBox within the track. It is possible that a single elementary stream is described by multiple tracks, each of which may describe different aspects of the elementary stream.

The sample-based protection can be applied to all samples or a group of samples in a scalable composed ES or decodable composed ES. If protection is applied to all samples, a ProtectionSchemeInfoBox signalling the parameters of the protection tool is added to the SampleDescriptionBox, which is then encapsulated as described in E.5.2. In addition, if protection is applied to a group of samples, a ProtectionSchemeInfoBox is added to their SampleGroupEntryBox, which is then encapsulated as described in E.5.4.

E.3 Extension to ISO base media file format

E.3.1 Overview

This subclause documents technical extensions (additional box types) to the ISO based media file format, which could be used for protection, adaptation, or secure adaptation of scalable coded pictures. However, the added box types could be used for other purposes as well. In particular, this subclause defines ProtectionSchemeInfoBox for the decryption tool and authentication tool, ItemDescriptionBox, ScalableSampleDescriptionEntry, ScalableSampleGroupEntry, and Generic Protected Box. All other boxes defined in ISO/IEC 15444-12 are still used as is.

E.3.2 Incorporate JPSEC codestream into ISO-driven file format

A JPSEC codestream can be placed as a payload in the 'mdat' box of the ISO base file format. In the Sample Description Box ('std'), the 'codingname' of the corresponding Sample Entry is defined to be 'jpsec', which is a registered identifier for JPSEC decoder. In this case, the security service is provided by JPSEC at codestream.

E.3.3 Protected file format brand

Files conforming to this Recommendation | International Standard may use 'ffsc' as the major brand in the File Type Compatibility Box.

ISO/IEC 15444-8:2007/Amd.1:2009 (E)

Files conforming to this Recommendation | International Standard, i.e., containing protection or authentication information may use 'ffsc' as a compatible brand in the File Type Compatibility Box.

There are uses of this Recommendation | International Standard which are compatible with JP2, JPX, MJ2, and JPM files. A typical use of this Recommendation | International Standard will leave the major brand of a file unchanged, but add boxes and thus add 'ffsc' as a compatible brand.

Thus brands including 'isom', 'iso2', 'jp2\040', 'jpx\040' and 'jpm\040' should be compatible.

The 'ffsc' compatible brand indicates the use of new boxes and new tools corresponding to the protection methods in JPSEC.

A file that has been protected, to the extent that an application intending to process the JP2, JPX, JPM, or other file type content will be unable to do so without using protection tools, may use the 'ffsc' major brand as the file type; such a protected file must not use a major brand for which it is no longer conformant.

E.3.4 Summary of boxes used

The ISO base media file format defines two structures to describe a presentation: the logical structure and media sequence structure. The logical structure uses the ItemLocationBox ('iloc') to describe an item which is the byte range or a series of byte ranges for a particular file, either a local file or a remote file. The media sequence structure uses the SampleGroupDescriptionBox ('spgd') or SampleDescriptionBox ('std') to describe the samples, which could be a frame of video, a time-contiguous series of video frames, or a time-contiguous compressed section audio.

Accordingly, the protection in the ISO base media file format level is classified into item-based protection and sample-based protection, as described in E.5.2 and E.5.4, respectively.

Several boxes are used from ISO/IEC 15444-12, these are marked as "Existing" in Table E.1. Boxes defined in this Recommendation | International Standard are listed as "New" in Table E.1. The definitions for these boxes depend on the definitions of Box and FullBox from ISO/IEC 15444-12, which are repeated for convenience in E.7.

Table E.1 – List of existing and new boxes

Box names						Status	Remarks
meta						Existing	Metadata
	iloc					Existing	Item location
	iproc					Existing	Item protection
		sinf				Existing	Protection scheme information box
			frma			Existing	Original format box
			schm			Existing	Scheme type box
			schl			Existing	Scheme information box
				gran		New	Granularity box
				vall		New	Value List box
				bcip		New	Block cipher box
					keyt	New	Key template box
				scip		New	Stream cipher box
					keyt	New	Key template box
				auth		New	Authentication box
					keyt	New	Key template box
	iinf					Existing	Item information box
	ides					New	Item description box
		dest				New	Description type box
		desd				New	Description data box
		vide				New	Visual item description entry
		j2ke				New	JPEG 2000 item description entry
	icor					New	Item correspondence box
...	...	...	...	...	...	...	...
stbl						Existing	Sample table box
	std					Existing	Sample description box