ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

T.800
Amendment 4
(05/2011)

SERIES T: TERMINALS FOR TELEMATIC SERVICES

Still-image compression – JPEG 2000

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

Amendment 4: Guidelines for digital cinema applications

Recommendation ITU-T T.800 (2002) – Amendment 4

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.829**
Still-image compression – JPEG XR	T.830–T.849
Still-image compression – JPEG-1 extensions	T.850–T.899

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

INTERNATIONAL STANDARD ISO/IEC 15444-1
RECOMMENDATION ITU-T T.800

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

Amendment 4

Guidelines for digital cinema applications

Summary

Amendment 4 to Recommendation ITU-T T.800 | ISO/IEC 15444-1 adds guidelines for the usage of JPEG 2000 in the framework of digital cinema applications. Three topics are addressed in these guidelines.

The first is dedicated to reliable multicast transmission of JPEG 2000 codestreams because digital cinema content may be delivered to the theatres by using several communication channels, and many of them can be wireless, such as DVB-T, DVB-S, and WiMAX.

The second topic is related to implementation guidelines for digital cinema distribution. One of the guidelines provides quantization steps for 2K visually lossless compression based on perceptual thresholds. Also included are visual weightings for digital cinema conditions, which improve the quality of the JPEG 2000 D-Cinema profiles. The third topic gives practical rules for using JPEG 2000 in motion picture archives.

NOTE – This amendment was developed as common text in collaboration with ISO/IEC JTC 1/SC 29/WG 1. For consistency with previous amendment approvals and publication decisions for Recommendation ITU-T T.800 Amd. 3, this amendment went through the AAP process as "Amendment 3" but it is published as Recommendation ITU-T T.800 (2002) Amendment 4 | ISO/IEC 15444-1:2004 Amendment 4.

History

Edition	Recommendation	Approval	Study Group
1.0	ITU-T T.800	2002-08-29	16
1.1	ITU-T T.800 (2002) Amd. 1	2005-09-13	16
1.2	ITU-T T.800 (2002) Cor. 1	2007-01-13	16
1.3	ITU-T T.800 (2002) Cor. 2	2007-08-29	16
1.4	ITU-T T.800 (2002) Amd. 2	2009-03-16	16
1.5	ITU-T T.800 (2002) Amd. 3	2010-06-22	16
1.6	ITU-T T.800 (2002) Amd. 4	2011-05-14	16

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 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/.

CONTENTS

INTERNATIONAL STANDARD
RECOMMENDATION ITU-T

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

Amendment 4

Guidelines for digital cinema applications

1)   Modification to Annex J

Add the following new clause J.16 at the end of Annex J (i.e., immediately following J.15.3):

J.16   Guidelines for digital cinema applications

This clause is intended as guidelines for the usage of JPEG 2000 in the framework of digital cinema applications. It is split in four parts: the first is dedicated to multicast transmission; the second presents the recommended visual frequency weightings for digital cinema environment; the third deals with the use of JPEG 2000 for film archive applications, and finally the fourth gives guidelines for digital cinema distribution.

J.16.1   Reliable multicast transmission of JPEG 2000 codestreams

J.16.1.1 Introduction

Digital Cinema (D-Cinema) content may be delivered to the theatres by using several communication channels, and many of them can be wireless, such as, for example, DVB-T [50], DVB-S [51], and WiMAX [52]. The destination of the video contents, in this case, can also (at reduced resolution/quality) be a mobile user, such as, for instance, an audience located on a train or bus.

In addition to the file transfer approach, already used for digital cinema transfer to the theatres, the use of streaming would allow for applications such as live events exhibition and (Near) movie on demand (in this case, however, additional methods for the reduction of the bit rate must be adopted).

The delivery of JPEG 2000 compressed video is a well-known challenge for wireless IP networks. In the case of TCP, the variable delay is unacceptable for many types of real-time applications. D-Cinema will introduce even more stringent conditions, in terms of both the required bandwidth and the high quality demand.

As concerns the distribution, the requirements can be specified in the following terms:

–   low required bandwidth occupancy, not only of the JPEG 2000 compressed content itself, but also considering some signalling overhead;

–   reliable transmission of the contents, by use of forward error correction techniques (FEC) and/or selective retransmission of lost data (selective ARQ).

A possible solution can be that of adopting a multicast protocol, which can be adapted to the high bit rate and low packet loss rate necessary for the distribution of D-Cinema content. If using multicast transmission, a reliable protocol could be used, which guarantees the sender that every receiver has a copy equal to the original. NORM protocol [56] can be a viable solution for such needs: it uses a NACK-oriented reliable multicast, achieving reliability by means of negative ACKs that are sent back from the receivers to the sender. In addition, it is possible to specify even a small FEC layer, which reduces the retransmission rate at the expense of some added overhead.

Moreover, in case of multicast transmission of JPEG 2000 codestreams, there are some constraints in limited network bandwidth or capabilities of mobile devices.

New mechanisms are required to distribute D-cinema contents to local theatres, high-end mobile users via wireless networks under the constrained conditions.

In forwarding packets to a mobile user, an intermediate distributor may have such functions such as adjusting the image quality or resolution for network bandwidth, network convergence, or capability of mobile device purposes.

J.16.1.2 Distribution model

Figure J.17 shows a possible scenario for content distribution. Different users, with different quality requirements/capabilities may be foreseen. In this scenario, an intermediate distributor is required with the role of adaptation of resolution/quality for forwarding capabilities.

In Figure J.17, cinema contents are transmitted from the production site (for example, studio post-production house) to regional theatres through satellite links. After that, some resolution or quality level is extracted from the original content and is distributed to local theatres, high-end mobile users, home and low-end mobile users by wireless network such as DVB-T, WiMAX and WLAN (IEEE 802.11n) [54].
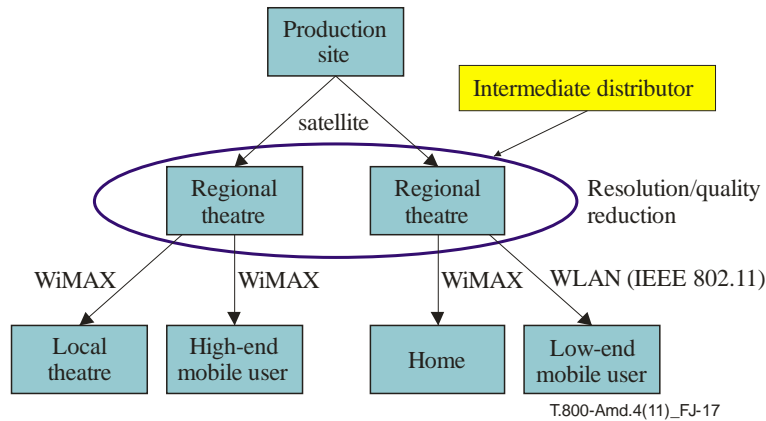
Figure J.17 – Distribution scenario for D-Cinema and live event

There are mainly three main problems in network delivery:

1) limited bandwidth;

2) convergence;

3) packet loss.

In forwarding a digital cinema stream to mobile users, an intermediate distributor such as a regional theatre might have functions of adjusting the image quality or resolution for network bandwidth, network convergence, or the capability of mobile device purposes.

J.16.1.3 Multicast protocol overview

NORM (NACK oriented reliable multicast) is a multicast protocol designed to provide end-to-end reliable multicast transfer. This protocol uses generic IP multicast capabilities, and on top of that tries to achieve reliable transfer using NACKs (negative acknowledgments). It can work both on reciprocal multicast network (i.e., wireless or wired LAN) or unidirectional link (such as a unidirectional satellite); in this way, it can satisfy all of the transmission needs for digital cinema distribution [53].

NORM repair capabilities are based on the use of negative acknowledgments (NACKs), sent from the receivers to the sender upon detection of an erroneous or missing packet. The sender transmits packets of data, segmented according to a precise strategy, each of which is identified by symbol numbers. As a receiver detects a missing packet, it initiates a repair request with a NACK message. Upon reception of NACKs, the sender prepares the appropriate repair messages, using FEC (forward error correction) blocks. Each receiver can reinitiate a repair procedure if it does not receive repair blocks.

Feedback suppression is applied, using a random back-off algorithm; each receiver, before sending a NACK for a certain packet, waits for a random-duration interval, during which it senses the medium and checks if other receivers requested a repair for the same packet: if so, it discards the NACK; otherwise, it sends it and waits for repair bits. When the sender receives the NACK, it enqueues repair bits (or the entire lost packet). Feedback suppression works efficiently in this protocol, and can achieve good results.

NACK packets can be sent both in multicast or unicast mode, using the sender's address. In the second case, feedback suppression can be achieved using multicast advertising messages, sent from each sender, which allow the receivers to know which packets have a pending repair request.

A NORM sender can even autonomously add FEC parity bits to each packet, thus enabling the receiver to correct errors and recover losses without starting a NACK procedure.
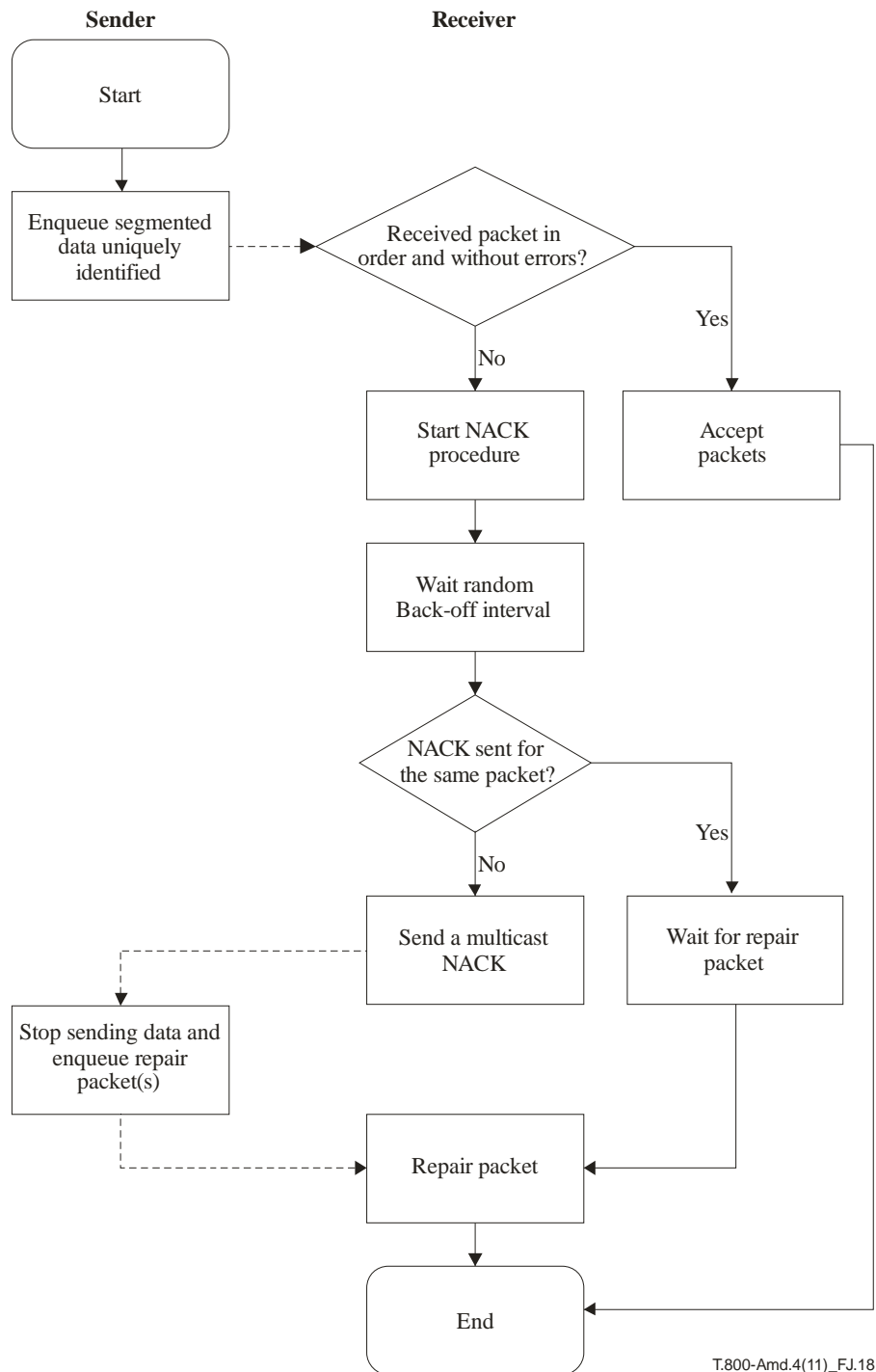
Figure J.18 – Sender and receiver sequence of operations

Figure J.18 shows a sequence of operations performed by a NORM sender and receiver during the normal transmission session. The sender node enqueues data packets, segmented according to parameters that can be changed by the user to accommodate for particular needs: in this case, JPEG 2000 codestreams are encapsulated and an additional header is included at the end of the NORM packet header. It also periodically enqueues control messages, such as round-trip collection and rate congestion control feedback.

Each receiver controls if the packet is in order and error-free: in this case, it accepts the packet and passes it to the destination application. Otherwise, it enters the NACK procedure: this consists in picking a random back-off delay, based on some parameters such as the greatest round-trip delay, usually supplied by the sender, and delaying NACK transmission until this interval is passed. In the meantime, if it receives a repair request for the same packet or the repair bits, the NACK is dropped; otherwise, it sends the NACK in multicast mode. Sending the NACK in multicast is useful for feedback suppression, as described previously.

When the sender receives the NACK, it stops usual data transmission and sends repair packets, in the form of a bit sequence derived from a Reed-Solomon (RS) code. With these repair bits, receivers are able to recover from transmission errors. If a receiver loses a repair packet too, it can resend a new NACK for the same packet after waiting for a new back-off interval.

J.16.1.4 Packetization strategy

To achieve a correct delivery of the DC content to all users, the JPEG 2000 video content must be carefully split into network packets that will be sent via the multicast protocol. To this purpose, three main procedures are needed:

1) preparation of the JPEG 2000 codestream (J2K);

2) splitting of codestreams into network packets;

3) adaptation of the multicast protocol packet header to J2K contest.

J.16.1.4.1 J2K codestream preparation

In the first stage, the J2K codestream is split into many pieces to be combined with the NORM header and UDP header as shown in Figure J.19. As a result, the network packet consists of the UDP header, the NORM header and one part of the J2K codestream.

T.800-Amd.4(11)_FJ-19
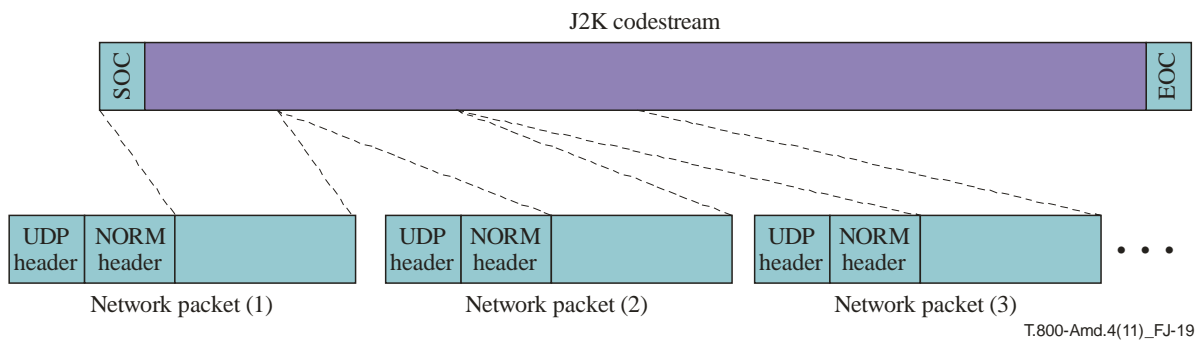
Figure J.19 – J2K codestream and network packets

In this case, the problem is that it is difficult to control the image quality because there is no information on what kinds of J2K packets are contained in network packets.

In order to solve the problem, a kind of J2K packet attribute information on the NORM J2K header is added to the following two cases.

– Case 1: An intermediate distributor that can easily control the image quality by discarding the network packets containing higher layers.

– Case 2: An intermediate distributor that can easily control the image resolution by discarding the network packets containing higher resolutions.

The "attribute" concept is provided to indicate the attribute of J2K packets, and it is similar to "Priority" in the JPEG 2000 RTP format. But "attribute" is simpler. "Attribute" is an incremental number classifying J2K packets by the same layer number and resolution level. Every J2K packet has an attribute number (> 0). The main header and tile-part header have a special attribute number ($= 0$).

There are two types of attribute numbering:

– "Layer-Resolution" attribute numbering

• Applied to LRCP progression order.

– "Resolution-Layer" attribute numbering

• Applied to RLCP, RPCL, PCRL or CPRL progression order.

J.16.1.4.1.1   \<Type-1\> "Layer-Resolution" attribute numbering

J2K packets are grouped by the layer and the resolution (Figure J.20).

T.800-Amd.4(11)_FJ-20
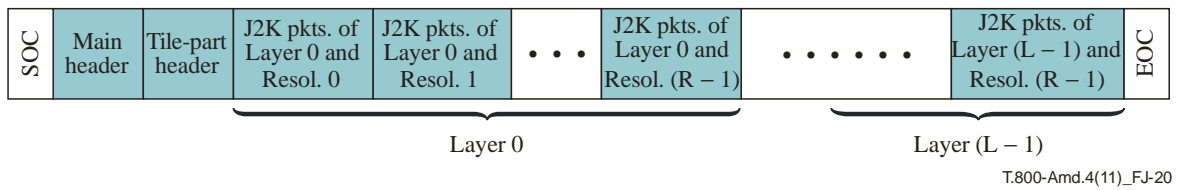
Figure J.20 – Attribute numbering (Layer-Resolution)

In "Layer-Resolution" attribute numbering, the attribute number is calculated as *(layer number) * (number of resolution levels) + (resolution level) + 1*.
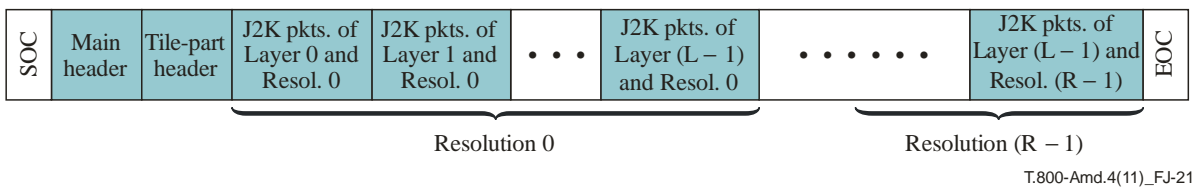
The attribute numbering can be applied for J2K codestream (LRCP).

For example:

- The attribute number of J2K packets belonging to "**Layer 0 and Resolution 0**" is **1**.
- The attribute number of J2K packets belonging to "**Layer 0 and Resolution 1**" is **2**.

   …
- The attribute number of J2K packets belonging to "**Layer l and Resolution r**" is "**l * R + r + 1**"

   …
- The attribute number of J2K packets belonging to "**Layer (L – 1) and Resolution (R – 1)**" is "**L * R**"

J.16.1.4.1.2   \<Type-2\> "Resolution-Layer" attribute numbering

J2K packets are grouped by the resolution and the layer (Figure J.21).

T.800-Amd.4(11)_FJ-21

Figure J.21 – Attribute numbering (Resolution-Layer)

In "Resolution-Layer" attribute numbering, the attribute number is calculated as *"(resolution number) * (number of layers) + (layer number) + 1"*.

The attribute numbering can be applied for J2K codestreams (RLCP, RPCL, PCRL or CPRL).

For example:

- The attribute number of J2K packets belonging to "**Layer 0 and Resolution 0**" is **1**.
- The attribute number of J2K packets belonging to "**Layer 1 and Resolution 0**" is **2**.

   …
- The attribute number of J2K packets belonging to "**Layer l and Resolution r**" is "**r * L + l + 1**"

   …
- The attribute number of J2K packets belonging to "**Layer (L – 1) and Resolution (R – 1)**" is "**L * R**"

### J.16.1.4.2 Network packet and attribute number

This clause explains the coordinating network packets (Figure J.19) to the J2K codestream. The J2K codestream in Figure J.22 is of the "Layer-Resolution" type, as shown in Figure J.20.

T.800-Amd.4(11)_FJ-22
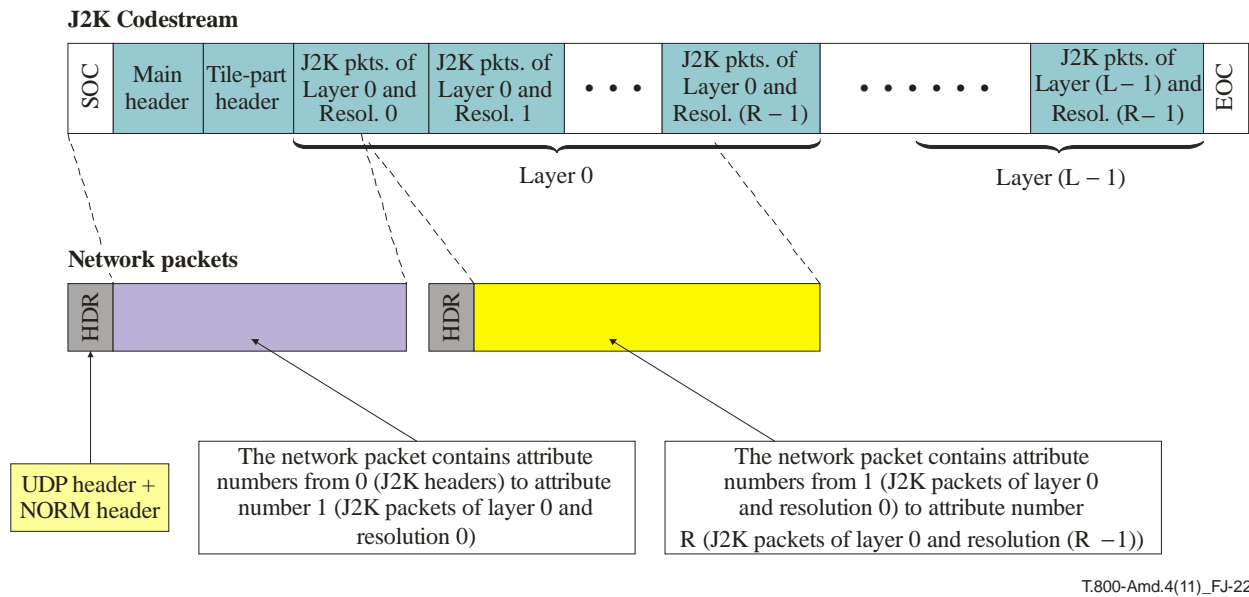
Figure J.22 – Network packet and attribute number

J.16.1.4.3 Multicast protocol header format

It is important that the sender and the receiver jointly minimize the probability of retransmissions; this can be done by making them aware of the underlying codestream-based file structure. A number of fields can be added to the base NORM protocol header, based on the fields utilized by the JPEG 2000 RTP streaming protocol.

T.800-Amd.4(11)_FJ-23
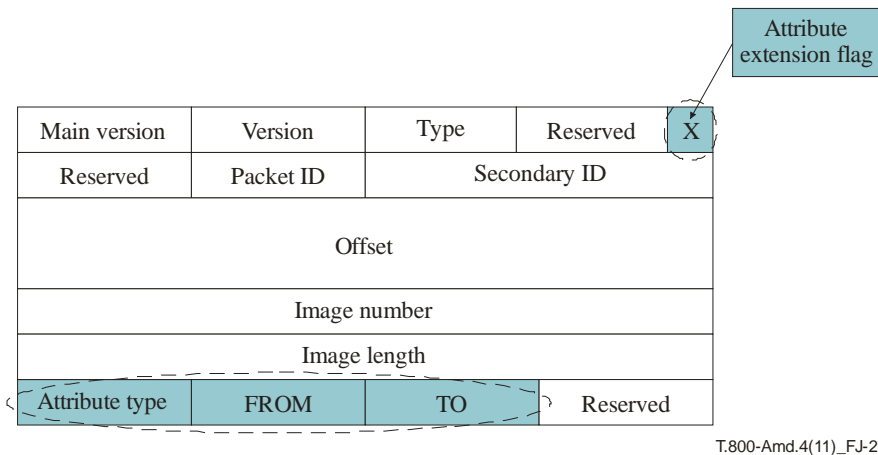
**Figure J.23 – Format of the header used for JPEG 2000 packetization**

A tabular view of the additional header fields is presented in Figure J.23; their meaning and aim are as follows:

–   *main version* (8 bits): it is always set to 0xCB. If the received packet has a different value for this field, it is discarded.

–   *version* (8 bits): minor version of the protocol, currently 1. Any packet with a different minor version is discarded from the receiver.

–   *type* (8 bits): it indicates the kind of transfer; in particular, it is related to the format of the codestream-containing file. Currently, it can be set to 0xF1 when a set of .j2c files (raw codestreams) is sent or to 0xF6 if a single .mj2 video file (wrapped file format) is being transferred. Unrecognized types would result in packet discarding.

–   *X* (1 bit): attribute extension flag, attribute extensions are valid only if X = 1.

– *packet ID* (8 bits): if a single codestream is fragmented at the sender, it indicates the progressive number of the transmission. This is especially useful when transmitting large codestreams, which can be divided into smaller portions at the data packet boundary. This is used for packet reordering upon receiving and checking missing packets.

– *secondary ID* (16 bits): it can be used to order different tracks inside a video file, or to differentiate among video, audio and synchronization data.

– *image number* (32 bits): progressive number of image (i.e., codestream) in a movie. This is used to reconstruct the video file in the correct order, and to check for the correct receiving of packets.

– *offset* (64 bits): it represents the offset of the first data in this packet, starting from the beginning of the file; it is directly expressed in bytes (since DC video content files can be as large as 250 GB, 8 bytes are necessary for seeking). The receiver uses the value carried by this file for correctly placing the received packet in the destination file.

– *image length* (32 bits): it denotes the total length of the codestream containing the image, expressed in bytes. It is used to check if the image has been completely received.

– *attribute type* (8 bits): 0 denotes "Layer-Resolution" attribute, 1 denotes "Resolution-Layer" attribute, other values are reserved.

– *FROM* (8 bits): attribute range, from.

– *TO* (8 bits): attribute range, to.

This header is added to each sent multicast packet for identification.

Figure J.24 is an example of the "attribute header value". In the first NORM J2K extension header, the attribute range is from "0" to "1", and in the second header, the attribute range is from "1" to "R".
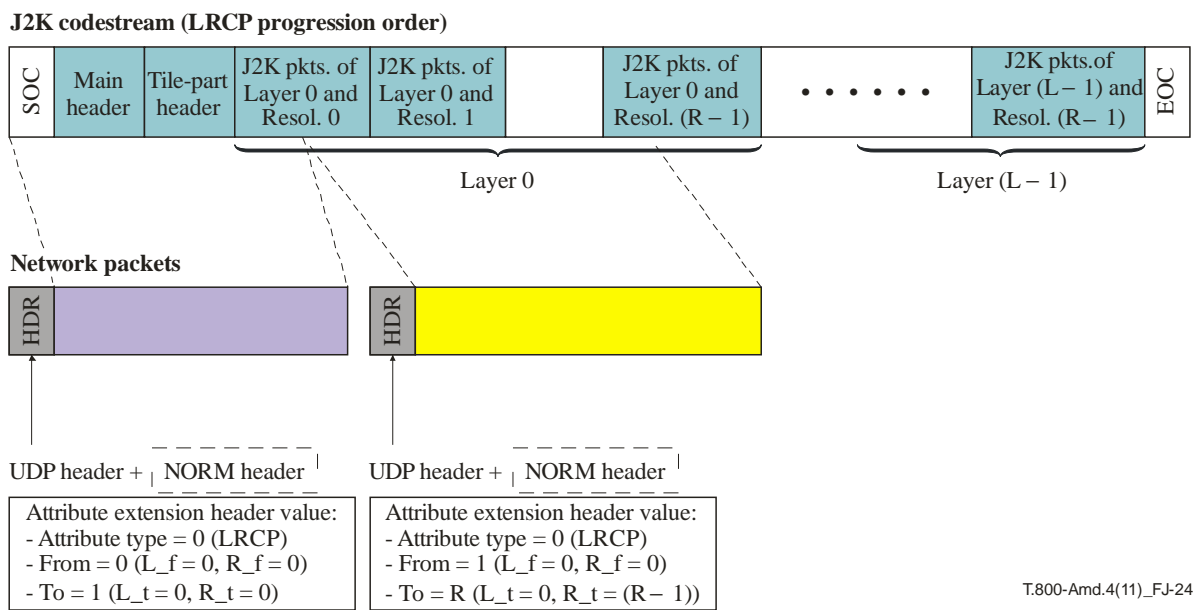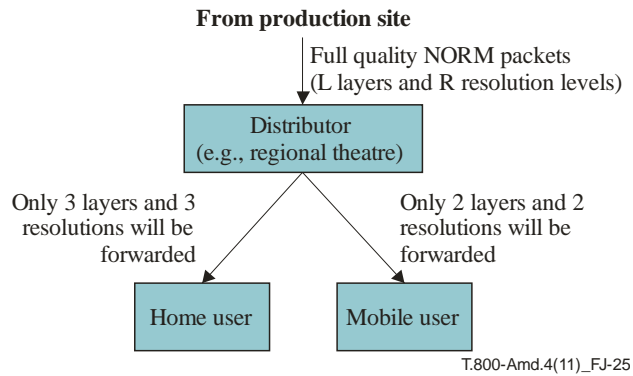

Figure J.24 – Attribute header value

J.16.1.5 Intermediate distributor function

This clause explains the function of the intermediate distributor. Full quality NORM packets ("L" layers and "R" resolutions) will be distributed from the production site to the intermediate distributor, as shown in Figure J.25.

The following workflow is an example of forwarding network packets to end users:

1) An intermediate distributor determines the number of layers and the number of resolution levels to be forwarded to the end user by judging from the network bandwidth or display.

2) After that, the distributor defines which network packets are to be forwarded, by using the attribute information of "FROM" and "TO", in the NORM J2K extension header.

3) Determination of the number of layers and resolution is out of scope in this Recommendation | International Standard. Instead, a method of how to forward network packets by checking the attribute extension is presented in the following examples.

From production site

Full quality NORM packets
(L layers and R resolution levels)

Distributor
(e.g., regional theatre)

Only 3 layers and 3
resolutions will be
forwarded

Only 2 layers and 2
resolutions will be
forwarded

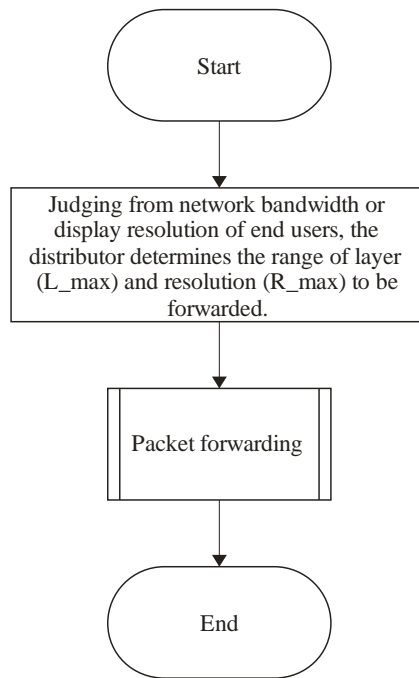Home user

Mobile user

T.800-Amd.4(11)_FJ-25

Figure J.25 – Example of forwarding network packets

Figure J.26 shows the procedure of forwarding packets by judging from the network bandwidth or display resolution of end users.

Figure J.27 shows the procedure of determining the attribute extension header value (FROM(R_f and L_f) and TO(R_t and L_t)) and forwarding those packets that satisfy the conditional sentence in the "Resolution-Layer" type.

Figure J.28 shows the procedure of determining the attribute extension header value (FROM(R_f and L_f) and TO(R_t and L_t)) and forwarding those packets that satisfy the conditional sentence in the "Layer-Resolution" type.

Start

Judging from network bandwidth or
display resolution of end users, the
distributor determines the range of layer
(L_{max}) and resolution (R_{max}) to be
forwarded.

Packet forwarding

End

T.800-Amd.4(11)_FJ-26

Figure J.26 – Packet forwarding in intermediate distributor

Start

Any more network packets?

No

Yes

Calculate layer number and resolution level of FROM attribute and TO attribute:
- $R\_f = INT\,((FROM - 1)/L)$
- $L\_f = (FROM - 1) - R\_f * L$
- $R\_t = INT\,((TO - 1)\,/L)$
- $L\_t = (TO - 1) - R\_t * L$

Skip the packet

End

No

If ($L\_f <= L\_max$ && $R\_f <= R\_max$ ‖ $L\_t <= L\_max$ && $R\_t <= R\_max$ ‖ $R\_f < R\_t$)

Yes

Forward the packet
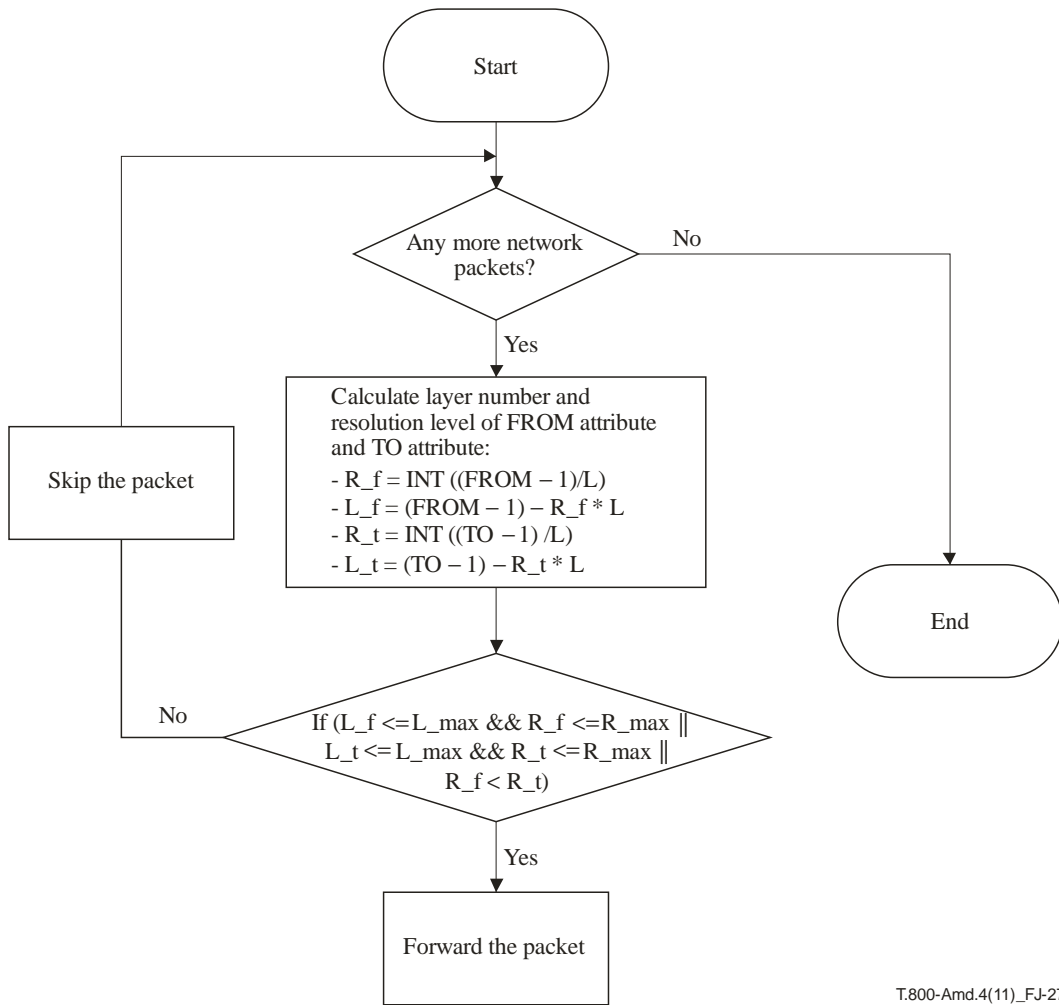
T.800-Amd.4(11)_FJ-27

Figure J.27 – Packet forwarding ("Resolution-Layer" type)
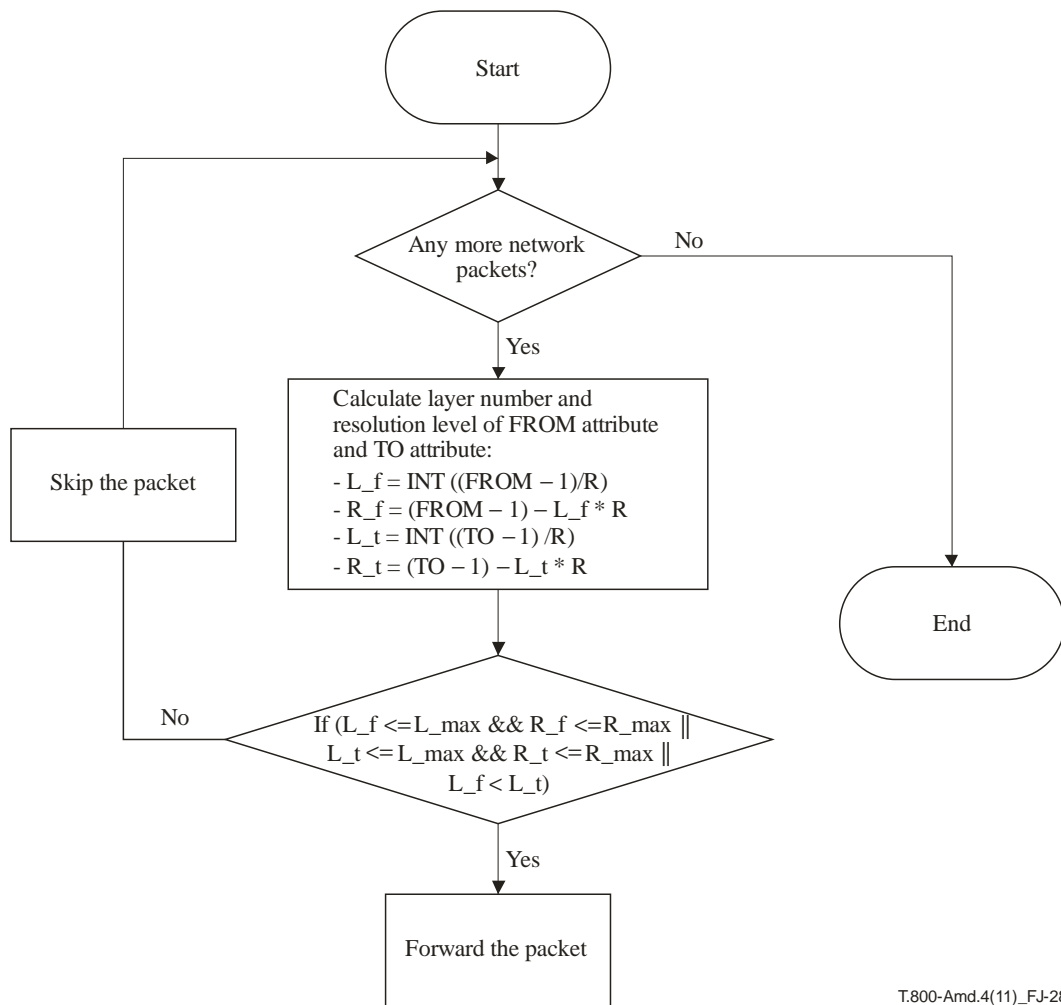
T.800-Amd.4(11)_FJ-28

Figure J.28 – Packet forwarding ("Layer-Resolution" type)

J.16.2 Implementation guidelines for digital cinema distribution

J.16.2.1 Experimental conditions

The following clauses explain frequency weightings and quantization steps for the 2K digital cinema obtained in specific experimental conditions:

    –   2K projector calibrated with respect to SMPTE recommendations;

    –   distance between the screen and the observer = 1.5 width of the displayed content.

J.16.2.2 Quantization steps for 2K visually lossless compression

Table J.27 provides quantization steps for 2K visually lossless compression. These latter are based on perceptual thresholds above which transparency (visually lossless) is achieved.

Table J.27 – Quantization steps for 2K visually lossless compression

Subband	C0	C1	C2
HL5	0.015625	0.03125	0.0078125
LH5	0.03125	0.03125	0.0078125
HH5	0.0625	0.0625	0.046875
HL4	0.00390625	0.0078125	0.00390625
LH4	0.00390625	0.0078125	0.00390625
HH4	0.015625	0.015625	0.0078125
HL3	0.00195313	0.00585938	0.00097656
LH3	0.00195313	0.00390625	0.00048828

**Table J.27 – Quantization steps for 2K visually lossless compression**

Subband	C0	C1	C2
HH3	0.00195313	0.00390625	0.00390625
HL2	0.00097656	0.00195313	0.00097656
LH2	0.00097656	0.00195313	0.00097656
HH2	0.00097656	0.00195313	0.00195313
HL1	0.00048828	0.00097656	0.00048828
LH1	0.00048828	0.00097656	0.00048828
HH1	0.00048828	0.00195313	0.00048828
LL	0.00048828	0.00097656	0.00048828

J.16.2.3   Visual frequency weighting in digital cinema environment

The different researches on the Human Visual System propose some models concerning the way that humans discern the surrounding world. For example, some models have been created to characterize the human sensitivity to brightness and to colours, with regards to the spatial frequencies. One of these models is the contrast sensitivity function (CSF).

To describe this function, it is necessary to introduce the notions of contrast threshold and contrast sensitivity. The contrast is a basic concept for the vision science. It is explained by the fact that the information represented in the visual system does not correspond to the absolute brightness level, but to the contrast. The latter is the ratio between the local intensity and the average intensity of the whole image. The most representative definitions for contrast are those proposed by Michelson and Weber [58]. These definitions have been expressed, for example, from simple experiences on the luminance, representing a sine-wave grating in gray level.

The neurons respond only to stimuli above a given contrast. The necessary contrast to provoke a neuron response is defined as the contrast threshold. The inverse of this threshold is called the contrast sensitivity. The latter varies with the spatial and temporal frequencies, which leads to the CSF concept.
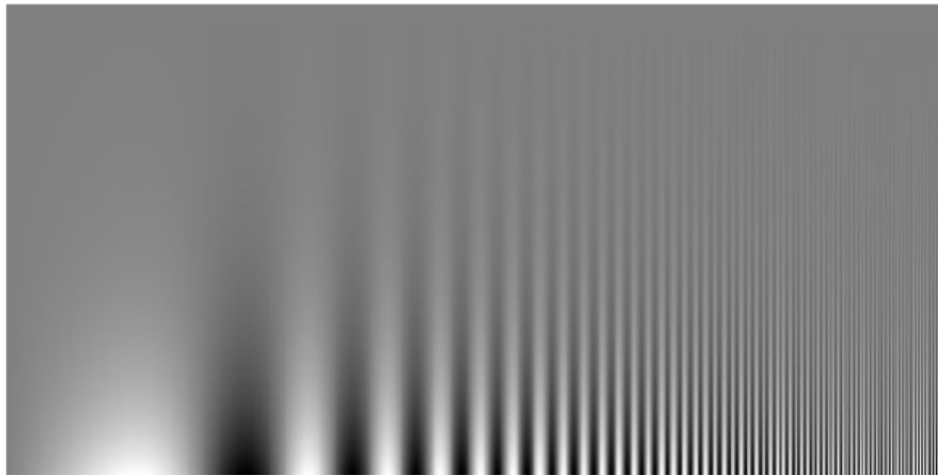
Figure J.29 – Experience of Campbell and Robson for the construction of the CSF:
modulation of a sine-wave grating for luminance

It is generally admitted that the human eye is more sensitive to spatial low frequencies than to high frequencies. Campbell and Robson [59] were the first to try to demonstrate the effect of the CSF with a representation as the one given by Figure J.29. The pixels luminance varies in a sine-wave way in the horizontal direction. The spatial frequency of this modulation increases exponentially from the left to the right, and the amplitude decreases exponentially from the bottom to the top of the image. The min and max values according to the horizontal axis remain constant. However, it is not visible when we observe this image. The observer has the impression that, on the left and right sides, the amplitude modulation was weaker. What the observer really perceives is his own contrast sensitivity function. If he could designate a line browsing the points where the stimulus confounds with the background of the image, he would finally have his own CSF curve of the luminance, as the one presented in Figure J.30.
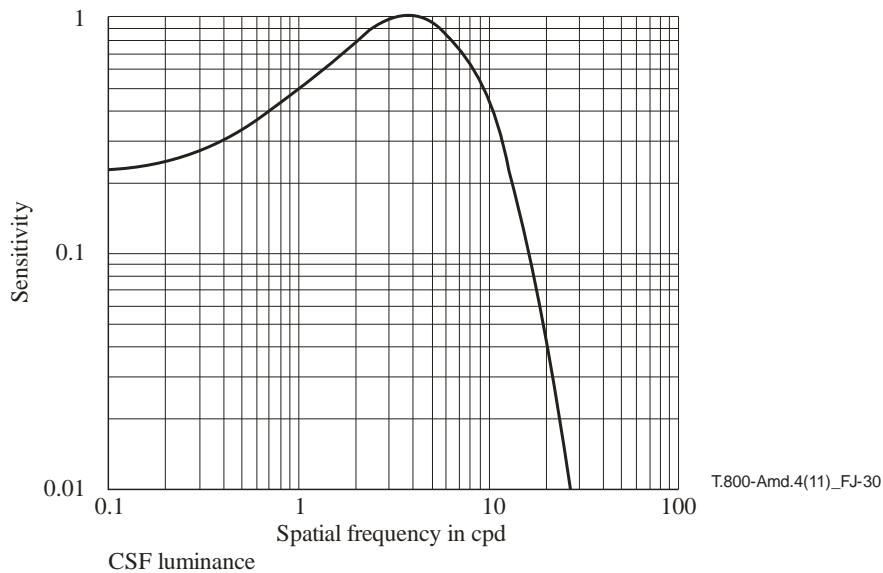
Figure J.30 – CSF of luminance

The physiological and psychophysical studies showed a low-pass behaviour of the eye and a high-pass behaviour of the retina. The general behaviour is therefore a band-pass with a sensitivity peak between 2 and 8 cpd (cycles per degree) and a cut-off frequency below 64 cpd. The most used model for this dependence has been proposed by Mannos and Sakrison [60] in 1974 for the case of the luminance, and joins the spatial sensitivity S to the spatial frequency f.

One of the most common visual optimization strategies in image compression is the use of the CSF described previously and which characterizes the variation of the visual system sensitivity in relation to the 2D spatial frequencies. In general, the human eye is less sensitive to high frequency errors (artefacts) than low frequency ones. The CSF could be used to determine the necessary relative precision for the different spatial frequencies, where the term of weighting is used to describe the asked proportional precision. To use the CSF, described in visual frequencies as the cycle per degree (cpd), it must be mapped on the domain of the digital frequencies of the compression, as the cycles per pixel (cpp). A correct weighting in frequency or in colour can lead to an important preservation of the details and textures, without the introduction of colour distortions.

In general, the visual weighting factors are more efficient for high observation distances or for large display or impression resolutions. Indeed, the weighting could also be used to reduce the artefacts of flickering in the Motion JPEG 2000 that results in a noise linked to the movement in a scene.

Digital cinema is quite a new application characterized by different viewing conditions that do not allow the use of already defined weighting factors without checking their validity. This is mainly due to the large aperture that introduces the peripheral vision in the process and low illumination that move our perception from the photopic domain to the photopic/mesopic one.

J.16.2.3.1 Recommended frequency weighting tables for digital cinema distribution

Table J.28 gives frequency weightings for digital cinema with respect to SMPTE recommendations 428 to 432.

Table J.28 – Recommended frequency weighting for digital cinema environment

a) Y weighting factors

Level	HL	LH	HH
5	0.260484	0.463463	0.651362
4	0.449002	0.754981	0.789797
3	0.912848	0.974562	0.976948
2	0.993774	0.997151	0.997898
1	1.000000	1.000000	1.000000

b) Cx weighting factors

Level	HL	LH	HH
5	0.088143	0.111436	0.109352
4	0.094904	0.149177	0.152355
3	0.190024	0.347601	0.362426
2	0.472931	0.636333	0.661777
1	0.741372	0.830503	0.865330

c) Cz weighting factors

Level	HL	LH	HH
5	0.204351	0.251607	0.262396
4	0.228638	0.420631	0.439335
3	0.567974	0.748305	0.754203
2	0.841546	0.913070	0.921399
1	0.951529	0.978377	0.976763

J.16.3   Implementation guidelines for the use of JPEG 2000 in film archive applications

J.16.3.1 Introduction

The digital film archive system architecture has to provide a solution for two major use cases with partly oppositional requirements. These are, on the one hand, long-term archival with the requirement to store the source imagery without any loss of information. This practice usually requires lossless compression and results, regarding the high spatial resolutions used in cinematic production, in large amounts of data. On the other hand, there is the requirement of frequent access to archived items. This practice usually does not require the original lossless quality of the source imagery. Instead the focus lies on easy and standardized access through local and remote channels. This results in the requirement for a significant reduction of the original amount of data and the restriction of certain coding parameters to ensure maximum compatibility with a wide range of decoding and playback equipment.

The above requirements lead to a two-tier system architecture with different file formats and JPEG 2000 profile specifications for long-term storage and frequent access (access or viewing copies), both specified in Annex A, Table A.46. Archived material may be stored in either one of the two formats or in both formats simultaneously depending on the type of usage scenario for the material. The long-term storage format (with profile indication Rsiz=5) should be used if cinematic material needs to be preserved digitally in all of its aspects. The access format (with profile indication Rsiz=6) should be used if access to the material is the main concern. The latter can be generated through transcoding from the former. The access format is close to the 2K digital cinema profile (with profile indication Rsiz=3), for simplified distribution of digital content to cinemas. It can also be used to transcode to other formats for delivery to end-users. For more information, see below.

Cinematic content usually consists not only of image sequences but also of additional data of various types. Typically, these are one or more sets of single or multi-channel audio data, timed text for subtitles and technical as well as descriptive metadata of different levels of details. All of this data should be stored together in one place, as described by the asset store approach standardized within the "open archival information system – Reference model" standardized as ISO 14721:2003.

J.16.3.2 Digital cinema archives

This clause provides practical guidelines for the use and application of the motion picture archive formats. The following paragraphs describe procedures that are meant to help in creating archive package files and how to use them to disseminate archived cinematic content to end users and to exchange archived materials between two or more archiving institutions. Implementations should be strict, when writing, and tolerant, when reading, the described formats.

As an alternative, the material exchange format, specified in SMPTE ST 377-1 [61], can also be considered to store JPEG 2000 motion picture content.

J.16.3.2.1 Ingesting into cinematic archives

Several ways exist to convert cinematic content to the motion picture archive formats. The content may come directly from various stages of an uncompressed digital production process or it can be a compressed digital distribution format, e.g., a digital cinema package (DCP).

It is also possible to ingest images originating from scanned film material as well as access copies that were created through transferring original film images to some other type of analogue or digital media, e.g., analogue or digital video tapes, DVD, etc.

J.16.3.2.1.1   Ingesting uncompressed digitally produced material

Digitally produced material includes images that were recorded with an electronic camera and did not undergo any further processing, as well as the results of different processing steps during post-production to the final cinematic product. This also includes scanned films that were subject of a digital restoration process. Common formats for this type of imagery are DPX, TIFF or TGA sequences. This type of material may also have its origin on a digital HDTV video tape format that has been ingested into a computer system. This type of material should be compressed using the JPEG 2000 profile for long-term storage of cinematic content. To be able to preserve the material in its original quality, lossless compression with the 5-3 reversible filter should be used. The original colour space and bit depth should not be changed.

J.16.3.2.1.2   Ingesting compressed digital cinema packages (DCPs)

The digital cinema package, as described in the different parts of SMPTE ST 429 [66], can be regarded as the digital equivalent to the release print on film material. Archival institutions in some cases may only be able to acquire a DCP of a certain production and not the uncompressed material originally used to create the DCP. The DCP should be converted without decoding and re-encoding of the contained JPEG 2000 images to the motion picture preservation storage format or the motion picture access format. All structural information of the DCP should be preserved in the metadata section to be able to exactly reconstruct the original DCP.

Direct archival of encrypted DCPs is discouraged because of the risks of loss of the encryption keys. Instead, the content of encrypted DCPs should be decrypted and stored, as described above, in order to ensure future accessibility. Security should be enforced by other means (e.g., physical security of storage media).

J.16.3.2.1.3   Ingesting scanned film material

The formats standardized in this Recommendation | International Standard can also be used to store cinematic content that is scanned from original film material. However, the choice of format and profile will usually depend on the preservation strategy for the original film. For digital long-term preservation, the original film material should be scanned at the highest reasonable spatial resolution and bit depth. These digital images should then be stored using the motion picture preservation format with the 5-3 reversible transform. Image data should be stored in the colour space that the film scanner had used. All available technical data of the film scanner should be stored as metadata. It is also possible to store additional image components (colour channels), e.g., infrared for scratch removal. This package can then be also used as a starting point for digital restoration procedures. For ingesting digitally restored images originating from film material, see clause J.16.3.2.1.1.

In many cases, it is desirable to keep the original film as long-term preservation element and to create digitized versions to facilitate access to the cinematic content. Depending on the requirements of the archival institutions, there are two main routes. For lower requirements, the film should be scanned and encoded according to the specifications for the motion picture access format. For higher requirements, the specifications for the motion picture preservation format should be used with parameters appropriate for the prospective use.

J.16.3.2.2 Disseminating cinematic archives

In the context of this Recommendation | International Standard dissemination means the delivery of archived cinematic content to end users. This does not necessarily mean that end users are required to have software and systems to decode and interpret the formats described in this Recommendation | International Standard. To create a format that can be

utilized by end users, a conversion process is necessary. The details and target formats of this conversion process largely depend on the requirements of each user group. Possible user groups may be digital cinemas, TV and movie production companies, schools and universities, scientists and others.

There may be a number of dissemination formats in consideration of the end user's many different requirements, such as:

- Digital cinema package (DCP) format as specified in the following documents:
  - SMPTE ST 429-3:2006 – D-Cinema Packaging – Sound and Picture Track File
  - SMPTE ST 429-4:2006 – D-Cinema Packaging – MXF JPEG 2000 Application
  - SMPTE ST 429-5:2009 – D-Cinema Packaging – Timed Text Track File
  - SMPTE ST 429-6:2006 – D-Cinema Packaging – MXF Track File Essence Encryption
  - SMPTE ST 429-7:2006 – D-Cinema Packaging – Composition Playlist
  - SMPTE ST 429-8:2007 – D-Cinema Packaging – Packing List
  - SMPTE ST 429-9:2007 – D-Cinema Packaging – Asset Mapping and File Segmentation
  - SMPTE ST 430-1:2006 – D-Cinema Operations – Key Delivery Message
  - SMPTE ST 430-2:2006 – D-Cinema Operations – Digital Certificate
- MPEG-4/Part-10 (Rec. ITU-T H.264/AVC) as specified in ISO/IEC 14496-10
- MPEG-2 as specified in ISO/IEC 13818 (including Rec. ITU-T H.222.0 | ISO/IEC 13818-1 and Rec. ITU-T H.262 | ISO/IEC 13818-2).

Usually, the motion picture access format is used as a starting point for the dissemination process because the compressed image data can be directly converted into a digital cinema package. It is also possible to create small-sized preview movies easily from this format by decoding only the lower resolution layers.

In special cases also, the motion picture preservation format may be used as a starting point for the creation of a dissemination package. The high quality contained in this format is, for example, needed to create uncompressed image sequences for digital restoration or re-mastering procedures.

J.16.3.2.3 Exchange of content between cinematic archives

To exchange cinematic content between different archival institutions both motion picture formats specified in this document can be used depending on the requirements of the institutions. Generally the motion picture access format should be preferred due to its stricter parameter set.

It is also possible to create new packages for exchange that, for example, contain only a subset of the image, audio, metadata or other content of the source format.

J.16.3.2.4 Use of the SMPTE material exchange format for motion picture preservation and access

For the access to and exchange of cinematic content, the MXF file format, as specified in SMPTE ST 377-1 [61], can be used as an alternative to the Motion JPEG 2000 format. The MXF operational pattern should be as simple as possible, but as complex as necessary. That is, elementary compositions should make use of the simplest operational pattern applicable, e.g., OP1a (SMPTE ST 378 [62]) for simple single-language files. However, complex compositions up to OP3c (SMPTE ST 408 [64]) can be used. The files should use the MXF generic container (GC) (SMPTE ST 379-1 [63]) with a frame-based essence mapping. Due to the asset store approach, all essence data should be contained within one single file. The JPEG 2000 codestreams should be mapped into the GC according to SMPTE ST 429-4 [66]. Audio data should be coded as described in Table 17 in SMPTE ST 377-1 [61]. It should be mapped into the MXF file format, according to the corresponding mapping document SMPTE ST 382 [65].

J.16.3.2.4.1 Application specification for the motion picture preservation format

There is no restriction on frame rates. All color spaces can be used, as long as they can be described with the picture essence descriptors, specified in SMPTE ST 377-1 [61].

J.16.3.2.4.2 Application specification for the motion picture access format

There is no restriction on frame rates. If distribution to digital cinema is intended, then frame rates of 24 fps or 48 fps are recommended. The X'Y'Z' color space as defined in SMPTE ST 428-1 [67] should be used and signalled in the RGBA picture essence descriptor defined in SMPTE ST 377-1 [61]. The values are defined in 6.4 of SMPTE ST 429-4 [66].

J.16.3.2.5 General procedures and archive management

The archival system used to handle and store the formats described in this Recommendation | International Standard should adhere to the OAIS reference model (ISO 14721:2003).

2)   Annex K

Add the following at the end of Annex K (i.e., immediately following K.8):

K.9   Guidelines for digital cinema applications

[50]   ETSI EN 300 744 V1.5.1 (2004-11), *Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television*. http://www.etsi.org

[51]   ETSI EN 300 421 V1.1.2 (1997-08), *Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for 11/12 GHz satellite services*. http://www.etsi.org

[52]   IEEE Std 802.16-2009, *IEEE Standard for Local and metropolitan area networks – Part 16: Air Interface for Broadband Wireless Access Systems*. http://standards.ieee.org/getieee802

[53]   IETF RFC 5371 (2008), *RTP Payload Format for JPEG 2000 Video Streams*. http://www.ietf.org/mail-archive/web/ietf-announce/current/msg05302.html

[54]   IEEE Std 802.11n-2009, *IEEE Standard for Information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks – Specific requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications – Amendment 5: Enhancements for Higher Throughput*. http://standards.ieee.org/getieee802

[55]   IETF RFC 5372 (2008), *Payload Format for JPEG 2000 Video: Extensions for Scalability and Main Header Recovery*. http://www.ietf.org/mail-archive/web/ietf-announce/current/msg05303.html

[56]   IETF RFC 3940 (2004), *Negative-acknowledgment (NACK)-Oriented Reliable Multicast (NORM) Protocol*.

[57]   Micanti, P., Baruffa, G., Frescura, F., (2008), *A Packetization Technique for D-Cinema Contents Multicasting over Metropolitan Wireless Networks*, Mobile WiMAX, February, Wiley, pp. 313-326.

[58]   Michelson, A.A., (1927), *Studies in Optics*. Chicago, IL: Univ. Chicago Press.

[59]   Campbell, F.W. and Robson, J. G. (1968), *Application of Fourier analysis to the visibility of gratings*, Journal of Physiology, London, 197: 551-566.

[60]   Mannos, J.L., and Sakrison, J.D., (1974), *The effects of a visual fidelity criterion on the encoding of images*, IEEE Trans. on Inform, Theory IT-20, pp. 525-536.

[61]   SMPTE ST 377-1:2011, *Material Exchange Format (MXF) – File Format specification*.

[62]   SMPTE ST 378:2004, *Television Material Exchange Format (MXF) – Operational pattern 1a (Single Item, Single Package)*.

[63]   SMPTE ST 379-1:2009, *Material Exchange Format (MXF) – MXF Generic Container*.

[64]   SMPTE ST 408:2006, *Television – MXF – Operational Patterns 1c, 2c and 3c*.

[65]   SMPTE ST 382:2007, *Material Exchange Format (MXF) – Mapping AES3 and Broadcast Wave Audio into the MXF Generic Container*.

[66]   SMPTE ST 429-4:2006, *D-Cinema Packaging – MXF JPEG 2000 Application*.

[67]   SMPTE ST 428-1:2006, *D-Cinema Distribution Master (DCDM) – Image Characteristics*.

[68]   SMPTE ST 429-3:2006, *D-Cinema Packaging – Sound and Picture Track File*.

[69]   SMPTE ST 429-5:2009, *D-Cinema Packaging – Timed Text Track File*.

[70]   SMPTE ST 429-6:2006, *D-Cinema Packaging – MXF Track File Essence Encryption*.

[71]   SMPTE ST 429-7:2006, *D-Cinema Packaging – Composition Playlist*.

[72]   SMPTE ST 429-8:2007, *D-Cinema Packaging – Packing List*.

[73]   SMPTE ST 429-9:2007, *D-Cinema Packaging – Asset Mapping and File Segmentation*.

[74]   SMPTE ST 430-1:2006, *D-Cinema Operations – Key Delivery Message.*

[75]   SMPTE ST 430-2:2006, *D-Cinema Operations – Digital Certificate.*

3)   Clause 2.2, Additional references

Add the following reference:

–   ISO 14721:2003, *Space data and information transfer systems – Open archival information system – Reference model.*

4)   Clause 4.1, Abbreviations

Add the following abbreviations:

CPRL   Component, Position, Resolution and Layer data packet progression order within a tile

LRCP   Layer, Resolution, Component and Position data packet progression order within a tile

PCRL   Position, Component, Resolution and Layer data packet progression order within a tile

RLCP   Resolution, Layer, Component and Position data packet progression order within a tile

RPCL   Resolution, Position, Component and Layer data packet progression order within a tile

SERIES OF ITU-T RECOMMENDATIONS

Series A   Organization of the work of ITU-T

Series D   General tariff principles

Series E   Overall network operation, telephone service, service operation and human factors

Series F   Non-telephone telecommunication services

Series G   Transmission systems and media, digital systems and networks

Series H   Audiovisual and multimedia systems

Series I   Integrated services digital network

Series J   Cable networks and transmission of television, sound programme and other multimedia signals

Series K   Protection against interference

Series L   Construction, installation and protection of cables and other elements of outside plant

Series M   Telecommunication management, including TMN and network maintenance

Series N   Maintenance: international sound programme and television transmission circuits

Series O   Specifications of measuring equipment

Series P   Terminals and subjective and objective assessment methods

Series Q   Switching and signalling

Series R   Telegraph transmission

Series S   Telegraph services terminal equipment

Series T   Terminals for telematic services

Series U   Telegraph switching

Series V   Data communication over the telephone network

Series X   Data networks, open system communications and security

Series Y   Global information infrastructure, Internet protocol aspects and next-generation networks

Series Z   Languages and general software aspects for telecommunication systems