INTERNATIONAL STANDARD

ISO/IEC 18477-7

First edition
2016-02-01

Information technology — Scalable compression and coding of continuous-tone still images —

Part 7:
HDR Floating-Point Coding

Technologies de l'information — Compression échelonnable et codage d'images plates en ton continu —

Partie 7: Codage de la virgule flottante en HDR

Reference number
ISO/IEC 18477-7:2016(E)

Contents

Page

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

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. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: Foreword - Supplementary information

The committee responsible for this document is ISO/IEC JTC 1, *Information technology*, SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

ISO/IEC 18477 contains the following parts under the general title *Information technology — Scalable compression and coding of continuous-tone still images*:

— *Part 1: Scalable compression and coding of continuous-tone still images*

— *Part 2: Extensions for high dynamic range images*

— *Part 3: Box file format*

— *Part 6: IDR Integer Coding*

— *Part 7: HDR Floating-Point Coding*

— *Part 8: Lossless and Near-lossless Coding*

— *Part 9: Alpha Channel Coding*

The following parts are to be published:

— *Part 4: Conformance testing*

— *Part 5: Reference software*

Introduction

This part of ISO/IEC 18477 specifies a coded codestream format for storage of continuous-tone high and low dynamic range photographic content. JPEG XT part 6 is a scalable image coding system supporting multiple component images consisting of floating point samples. It is by itself an extension of the coding tools defined in ISO/IEC 18477-1 and the box-based format defined in ISO/IEC 18477-3; the codestream is composed in such a way that legacy applications conforming to Rec. ITU-T T.81 | ISO/IEC 10918-1 are able to reconstruct a lower quality, low dynamic range, eight bits per sample version of the image. This standard low dynamic range image is typically constructed at encoder side by tonemapping from the high dynamic image; while the LDR image is always present, this part of ISO/IEC 18477 does not define a process that generates this image.

Today, the most widely used digital photography format, a minimal implementation of JPEG (specified in Rec. ITU-T T.81 | ISO/IEC 10918-1), uses a bit depth of 8; each of the three channels that together compose an image pixel is represented by 8 bits, providing 256 representable values per channel. If the dynamic range of the input scene is too large, however, an integer sample representation is no longer applicable and sample values need to be specified in floating point. These values typically are, or are proportional to physical radiance values of three primaries. These primaries may be device specific physical colours, or may be the basis of the CIE XYZ colourspace.

JPEG XT is primarily designed to provide coded data containing high dynamic range and wide colour gamut content while simultaneously providing eight bits per pixel low dynamic range images using tools defined in ISO 18477-1. The goal is to provide a backwards compatible coding specification that allows legacy applications and existing tool chains to continue to operate on codestreams conforming to this part of ISO/IEC 18477.

JPEG XT has been designed to be backwards compatible to legacy applications while at the same time having a small coding complexity; JPEG XT uses, whenever possible, functional blocks of Rec. ITU-T T.81 | ISO/IEC 10918-1 to extend the functionality of the legacy JPEG Coding System. It is optimized for storage and transmission of high dynamic range and wide colour gamut floating point images while also enabling low-complexity encoder and decoder implementations.

This part of ISO/IEC 18477 is an extension of ISO 18477-1, a compression system for continuous tone digital still images which is backwards compatible with Rec. ITU-T T.81 | ISO/IEC 10918-1. That is, legacy applications conforming to Rec. ITU-T T.81 | ISO/IEC 10918-1 will be able to reconstruct streams generated by an encoder conforming to this part of ISO/IEC 18477, though will possibly not be able to reconstruct such streams in full dynamic range, full quality or other features defined in this part of ISO/IEC 18477.

This part of ISO/IEC 18477 is itself based on ISO/IEC 18477-3 which defines a box-based file format similar to other JPEG standards. The aim of this part of ISO/IEC 18477 is to provide a migration path for legacy applications to support, potentially in a limited way, lossless coding and coding of high dynamic range images consisting of samples represented in floating point. Existing tools depending on the existing standards will continue to work, but will only be able to reconstruct a lossy and/or a low dynamic range version of the image contained in the codestream. This part of ISO/IEC 18477 specifies a coded file format, referred to as JPEG XT, which is designed primarily for storage and interchange of continuous-tone photographic content.

Information technology — Scalable compression and coding of continuous-tone still images —

Part 7: HDR Floating-Point Coding

1 Scope

This part of ISO/IEC 18477 specifies a coding format, referred to as JPEG XT, which is designed primarily for continuous-tone photographic content.

2 Normative references

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

ISO/IEC 18477-1:2015, *Information technology — Scalable compression and coding of continuous-tone still images — Part 1: Scalable compression and coding of continuous-tone still images*

ISO/IEC 18477-2, *Information technology — Scalable compression and coding of continuous-tone still images — Part 2: Extensions for high dynamic range images*

ISO/IEC 18477-3:2015, *Information technology — Scalable compression and coding of continuous-tone still images —Part 3: Box-Based File Format*

ISO/IEC 18477-6:2016, *Information technology — Scalable compression and coding of continuous-tone still images — Part 6: Coding of Intermediate Dynamic Range Images*

ISO/IEC/IEEE 60559, *Information technology — Microprocessor Systems — Floating-Point arithmetic*

Rec. ITU-T T.81 | ISO/IEC 10918–1:1994, *Information technology — Digital Compression and Coding of Continuous Tone Still Images — Requirements and Guidelines*

Rec. ITU-T BT.601, *Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios*

3 Terms, definitions, symbols and abbreviated terms

3.1 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

3.1.1
ASCII encoding
encoding of text characters and text strings according to ISO/IEC 10646

3.1.2
base decoding path
process of decoding legacy codestream and refinement data to the base image, jointly with all further steps until residual data is added to the values obtained from the residual codestream

3.1.3
base image
collection of sample values obtained by entropy decoding the DCT coefficients of the legacy codestream and the refinement codestream, and inversely DCT transforming them jointly

3.1.4
binary decision
choice between two alternatives

3.1.5
bitstream
partially encoded or decoded sequence of bits comprising an entropy-coded segment

3.1.6
block
8 × 8 array of samples or an 8 × 8 array of DCT coefficient values of one component

3.1.7
box
structured collection of data describing the image or the image decoding process embedded into one or multiple APP_{11} marker segments

Note 1 to entry: See ISO/IEC 18477-3:2015, Annex B for definition of boxes.

3.1.8
byte
group of 8 bits

3.1.9
coder
embodiment of a coding process

3.1.10
coding
encoding or decoding

3.1.11
coding model
procedure used to convert input data into symbols to be coded

3.1.12
(coding) process
general term for referring to an encoding process, a decoding process, or both

3.1.13
compression
reduction in the number of bits used to represent source image data

3.1.14
component
two-dimensional array of samples having the same designation in the output or display device

Note 1 to entry: An image typically consists of several components, e.g. red, green and blue.

3.1.15
continuous-tone image
image whose components have more than one bit per sample

3.1.16
data unit
8 × 8 block of samples of one component in DCT-based processes; a sample in lossless processes

3.1.17
decoder
embodiment of a decoding process

3.1.18
decoding process
process which takes as its input compressed image data and outputs a continuous-tone image

3.1.19
dequantization
inverse procedure to quantization by which the decoder recovers a representation of the DCT coefficients

3.1.20
discrete cosine transform
DCT
either the forward discrete cosine transform or the inverse discrete cosine transform

3.1.21
downsampling
procedure by which the spatial resolution of a component is reduced

3.1.22
encoder
embodiment of an encoding process

3.1.23
encoding process
process which takes as its input a continuous-tone image and outputs compressed image data

3.1.24
entropy-coded (data) segment
independently decodable sequence of entropy encoded bytes of compressed image data

3.1.25
entropy decoder
embodiment of an entropy decoding procedure

3.1.26
entropy decoding
lossless procedure which recovers the sequence of symbols from the sequence of bits produced by the entropy encoder

3.1.27
entropy encoder
embodiment of an entropy encoding procedure

3.1.28
entropy encoding
lossless procedure which converts a sequence of input symbols into a sequence of bits such that the average number of bits per symbol approaches the entropy of the input symbols

3.1.29
grayscale image
continuous-tone image that has only one component

3.1.30
high dynamic range
image or image data comprised of samples using a floating point representation

3.1.31
Huffman decoder
embodiment of a Huffman decoding procedure

3.1.32
Huffman decoding
entropy decoding procedure which recovers the symbol from each variable length code produced by the Huffman encoder

3.1.33
Huffman encoder
embodiment of a Huffman encoding procedure

3.1.34
Huffman encoding
entropy encoding procedure which assigns a variable length code to each input symbol

3.1.35
joint photographic experts group
JPEG
informal name of the committee which created this International Standard

Note 1 to entry: The "joint" comes from the ITU-T and ISO/IEC collaboration.

3.1.36
legacy codestream
collection of markers and syntax elements defined by Rec. ITU-T T.81 | ISO/IEC 10918-1 bare any syntax elements defined by the family ISO/IEC 18477 standards, i.e. the legacy codestream consists of the collection of all markers except those APP_{11} markers that describe JPEG XT boxes by the syntax defined in Annex A

3.1.37
legacy decoder
embodiment of a decoding process conforming to Rec. ITU-T T.81 | ISO/IEC 10918-1, confined to the lossy DCT process and the baseline, sequential or progressive modes, decoding at most four components to eight bits per component

3.1.38
lossless
descriptive term for encoding and decoding processes and procedures in which the output of the decoding procedure(s) is identical to the input to the encoding procedure(s)

3.1.39
lossless coding
mode of operation which refers to any one of the coding processes defined in this part of ISO/IEC 18477 in which all of the procedures are lossless

Note 1 to entry: See ISO/IEC 18477-8.

3.1.40
lossy
descriptive term for encoding and decoding processes which are not lossless

3.1.41
low dynamic range
image or image data comprised of data with no more than eight bits per sample

3.1.42
marker
two-byte code in which the first byte is hexadecimal FF and the second byte is a value between 1 and hexadecimal FE

3.1.43
marker segment
marker together with its associated set of parameters

3.1.44
pixel
collection of sample values in the spatial image domain having all the same sample coordinates, e.g. a pixel may consist of three samples describing its red, green and blue value

3.1.45
precision
number of bits allocated to a particular sample or DCT coefficient

3.1.46
procedure
set of steps which accomplishes one of the tasks which comprise an encoding or decoding process

3.1.47
quantization value
integer value used in the quantization procedure

3.1.48
quantize
act of performing the quantization procedure for a DCT coefficient

3.1.49
residual decoding path
collection of operations applied to the entropy coded data contained in the residual data box and residual refinement scan boxes up to the point where this data is merged with the base image to form the final output image

3.1.50
residual image
sample values as reconstructed by inverse quantization and inverse DCT transformation applied to the entropy-decoded coefficients described by the residual scan and residual refinement scans

3.1.51
residual scan
additional pass over the image data invisible to legacy decoders which provides additive and/or multiplicative correction data of the legacy scans to allow reproduction of high dynamic range or wide colour gamut data

3.1.52
refinement scan
additional pass over the image data invisible to legacy decoders which provides additional least significant bits to extend the precision of the DCT transformed coefficients

3.1.53
sample
one element in the two-dimensional image array which comprises a component

3.1.54
sample grid
common coordinate system for all samples of an image

Note 1 to entry: The samples at the top left edge of the image have the coordinates (0,0), the first coordinate increases towards the right, the second towards the bottom.

3.1.55
scan
single pass through the data for one or more of the components in an image

3.1.56
scan header
marker segment that contains a start-of-scan marker and associated scan parameters that are coded at the beginning of a scan

3.1.57
table specification data
coded representation from which the tables used in the encoder and decoder are generated and their destinations specified

3.1.58
(uniform) quantization
procedure by which DCT coefficients are linearly scaled in order to achieve compression

3.1.59
upsampling
procedure by which the spatial resolution of a component is increased

3.1.60
vertical sampling factor
relative number of vertical data units of a particular component with respect to the number of vertical data units in the other components in the frame

3.1.61
zero byte
0x00 byte

3.1.62
zig-zag sequence
specific sequential ordering of the DCT coefficients from (approximately) lowest spatial frequency to highest

3.2 Symbols

X width of the sample grid in positions

Y height of the sample grid in positions

Nf number of components in an image

$s_{i,x}$ subsampling factor of component i in horizontal direction

$s_{i,y}$ subsampling factor of component i in vertical direction

H_i subsampling indicator of component i in the frame header

V_i subsampling indicator of component i in the frame header

$v_{x,y}$ sample value at the sample grid position x,y

R_h additional number of dct coefficient bits represented by refinement scans in the legacy decoding path, $8 + R_h$ is the number of non-fractional bits (i.e. bits in front of the “binary dot”) of the output of the inverse dct process in the legacy decoding path

R_r additional number of dct coefficient bits represented by refinement scans in the residual decoding path. $p + R_r$ is the number of non-fractional bits of the output of the invers dct process in the residual decoding path, where p is the frame-precision of the residual image as recorded in the frame header of the residual codestream

R_b additional bits in the hdr image. $8 + R_b$ is the sample precision of the reconstructed hdr image

3.3 Abbreviations

ASCII American Standard Code for Information Interchange

LSB Least Significant Bit

MSB Most Significant Bit

HDR High Dynamic Range

LDR Low Dynamic Range

TMO Tone Mapping Operator

DCT Discrete Cosine Transformation

4 Conventions

4.1 Conformance language

This part of ISO/IEC 18477 consists of normative and informative text.

Normative text is that text which expresses mandatory requirements. The word “shall” is used to express mandatory requirements strictly to be followed in order to conform to this part of ISO/IEC 18477 and from which no deviation is permitted. A conforming implementation is one that fulfils all mandatory requirements.

Informative text is text that is potentially helpful to the user, but not indispensable and can be removed, changed or added editorially without affecting interoperability. All text in this part of ISO/IEC 18477 is normative, with the following exceptions: the Introduction, any parts of the text that are explicitly labelled as “informative”, and statements appearing with the preamble “NOTE” and behaviour described using the word “should”. The word “should” is used to describe behaviour that is encouraged but is not required for conformance to this part of ISO/IEC 18477.

The keywords “may” and “need not” indicate a course of action that is permissible in a conforming implementation.

The keyword “reserved” indicates a provision that is not specified at this time, shall not be used, and may be specified in the future. The keyword “forbidden” indicates “reserved” and in addition indicates that the provision will never be specified in the future.

4.2 Operators

NOTE Many of the operators used in this part of ISO/IEC 18477 are similar to those used in the C programming language.

4.2.1 Arithmetic operators

+ addition

- subtraction (as a binary operator) or negation (as a unary prefix operator)

* multiplication

/ division without truncation or rounding

4.2.2 Logical operators

||	Logical OR
&&	Logical AND
!	Logical NOT
∈	x ∈ {A, B} is defined as (x == A || x == B)
∉	x ∉ {A, B} is defined as (x != A && x != B)

4.2.3 Relational operators

>	greater than
>=	greater than or equal to
<	less than
<=	less than or equal to
==	equal to
!=	not equal to

4.2.4 Precedence order of operators

Operators are listed below in descending order of precedence. If several operators appear in the same line, they have equal precedence. When several operators of equal precedence appear at the same level in an expression, evaluation proceeds according to the associativity of the operator either from right to left or from left to right.

Operators	Type of operation	Associativity
(), [ ], .	Expression	Left to Right
–	Unary negation	
*, /	Multiplication	Left to Right
+, –	Addition and Subtraction	Left to Right
< , >, <=, >=	Relational	Left to Right

4.2.5 Mathematical functions

$\lceil x \rceil$	Ceil of x. Returns the smallest integer that is greater than or equal to x.
$\lfloor x \rfloor$	Floor of x. Returns the largest integer that is lesser than or equal to x.
|x|	Absolute value, is –x for x < 0; otherwise, x
sign(x)	Sign of x, 0 if x is zero, +1 if x is positive, -1 if x is negative.
clamp(x, min, max)	Clamps x to the range [min, max]: returns min if x < min, max if x > max or otherwise x.

x^a	Raises the value of x to the power of a. x is a non-negative real number, a is a real number. x^a is equal to exp(a × log(x)) where exp is the exponential function and log() the natural logarithm. If x is 0 and a is positive, x^a is defined to be 0.
log(x)	Natural logarithm to the base of e. log(x) is undefined for non-positive values of x.
exp(x)	The exponential function of x. exp(x) equals e^x where e is the Euler number. Furthermore, exp(log(x)) = x for all positive numbers x.

5 General

The purpose of this Clause is to give an informative overview of the elements specified in this part of ISO/IEC 18477. Another purpose is to introduce many of the terms which are defined in Clause 3. These terms are printed in *italics* upon first usage in this Clause.

The following are three elements specified in this part of ISO/IEC 18477.

a) An *encoder* is an embodiment of an *encoding process*. An encoder takes as input *digital source image data* and *encoder specifications* and by means of a specified set of *procedures*, generates as output *codestream*.

b) A *decoder* is an embodiment of a *decoding process*. A decoder takes as input a codestream and by means of a specified set of procedures, generates as output *digital reconstructed image data*.

c) The *codestream* is a compressed image data representation which includes all necessary data to allow a (full or approximate) reconstruction of the sample values of a digital image. Additional data might be required that define the interpretation of the sample data, such as colour space or the spatial dimensions of the samples.

5.1 High level overview on JPEG XT ISO 18477-7 (informative)

This part of ISO/IEC 18477 allows lossy coding of high dynamic range of photographic images in a way that is backwards compatible to Rec. ITU-T T.81 | ISO/IEC 10918-1. Decoders compliant to the latter standard will be able to parse codestreams conforming to this part of ISO/IEC 18477 correctly, albeit only in 8-bit sample precision with a limited dynamic range and potentially a limited colour gamut.

This part of ISO/IEC 18477 includes multiple tools to reach the above functionality, defined in Annex B and on. A short overview on these coding tools will be given in this subclause; an in-depth specification of the algorithm is given in Annex A.

The syntax of an ISO/IEC 18477-7 compliant codestream is specified in ISO/IEC 18477-3, that is, this part of ISO/IEC 18477 uses a syntax element denoted as “box” to annotate its syntactical elements. The definition of the box syntax element is not repeated here, and readers are referred to ISO/IEC 18477-3 for further details. Additional boxes besides those already specified in ISO/IEC 18477-3 are defined in ISO/IEC 18477-6:2016, Annex B and in Annex B. In addition, ISO/IEC 18477-6 defines coding tools such as **Refinement Coding**, specified in ISO/IEC 18477-6:2016, Annex D and the **residual codestream** carried by the Residual Data box.

The **Refinement Scan**, specified in ISO/IEC 18477-6:2016, Annex D, increases the bit precision of the DCT coefficients, i.e. it operates in the DCT domain. The mechanism used here is very similar to that of Subsequent Approximation Scans specified in Rec. ITU-T T.81 | ISO/IEC 10918-1:1994, Annex G. A legacy baseline, extended or progressive Huffman scan as defined in Annex G of the legacy standard defines the 12 most significant bits of the DCT coefficients. These initial scans are represented in the legacy codestream and are visible for any ISO/IEC 18477-1 compliant decoder. Refinement Scans decode now into up to four additional least significant bits in the same way subsequent approximation within Rec. ITU-T T.81 | ISO/IEC 10918-1 decode least significant bits of a progressive scan pattern. The difference between Refinement Scans and Subsequent Approximation Scans is only that in the latter case, the number of least significant bits is annotated in the scan header of the legacy codestream whereas Refinement Scans are hidden from legacy applications and do not alter the scan header of the legacy codestream. Their number is indicated in the Refinement Specification box within the Merging Specification box and not in the legacy codestream.

While Refinement Scans extend the bit precision within the DCT domain by up to four bits, **Residual Scans** extend the sample precision in the spatial (image) domain. While the entropy coded data of Residual Scans is hidden in the Residual Data box from legacy applications, its decoding process is **identical** to that of the legacy data: baseline, extended or progressive Huffman scan decodes the data in the Residual Data box to DCT coefficients, inverse quantization and inverse Discrete Cosine Transformation (DCT) compute from these coefficients the **residual image data**. An image merging process, defined in Annex A, computes from the **precursor image** reconstructed from the legacy codestream and the **residual image** a final HDR output image.

This merging process is depicted in Figure 1. It first performs chroma-upsampling to reconstruct a single sample on each point of the sample grid of the base image. Chroma upsampling is specified in ISO/IEC 18477-1:2015, Annex A. It then converts the colour space of the base image first from YCbCr into the Rec. ITU-T BT.601 colourspace. Optionally, the luma signal of the LDR image is computed from the BT.601 colourspace and used as **prescaling** of the chroma signals of the residual. The BT.601 image is optionally transformed again into a wider colourspace. For practical reasons, the transformations from YCbCr to BT.601 and from BT.601 into an output colourspace are combined into a single linear transformation matrix. This linear transformation is followed by a nonlinear point transformation acting separately on each of the output channels sample by sample. This point transformation can either be seen as an approximate inverse tone mapping operation, or as inverse gamma correction that transforms the nonlinear sample values of the LDR image into sample values proportional to physical radiance. It can either be specified by a parametric curve, or by an explicit lookup table. The output of this decoding path is transformed again by an optional colour transformation into the final HDR colourspace forming the **precursor image** which represents a rough imprecise approximation of the final HDR image, already in the correct HDR output colour space.

Processing continues with the decoding of the **Residual Image**. DCT coefficients of the residual image are decoded from the information in the Residual Data box, and their bit precision is extended by additional refinement scans decoded from the data in the **Residual Refinement box**. Processing proceeds with inverse quantization and inverse DCT transformation. The output undergoes chroma upsampling to generate a single sample per sample grid coordinate. The next processing step performs a nonlinear point transformation on each of the reconstructed channels, separately for each sample, resulting in an error image in a YCbCr type of colour space. Chroma samples are then optionally pre-scaled by a factor that is computed from the luminance value of the LDR image by a **nonlinear prescaling map**. Optionally, a **post-scaling factor** is computed from the luminance of the residual image through a **nonlinear post-scaling map**. The post-scaling factor scales a linear HDR image from limited dynamic range to full dynamic range in a later processing stage. Residual samples then undergo an inverse linear decorrelation transformation to map the sample values from the intermediate YCbCr colourspace into the target colourspace. This transformation is typically identical to the transformation matrix in the base decoding path, but does not need to be. After inverse decorrelation, the residual samples are mapped again by the **residual nonlinear point transformation**. The result of this operation is the **residual image.**

To form the final output image, sample values of the **precursor image** and the **residual image** are added together, plus an offset to make the residual image symmetric around zero. The result of this is scaled by the **post scaling factor** if the **nonlinear post-scaling** map is present. Results are then clamped to the range of the intermediate range output image and finally processed through the **output conversion.** The output conversion either convers samples from integer to floating point by a half-exponential map, or uses a table lookup or parametric curve for this conversion.

The detailed specification of the decoding and merging process is found in Annex A.

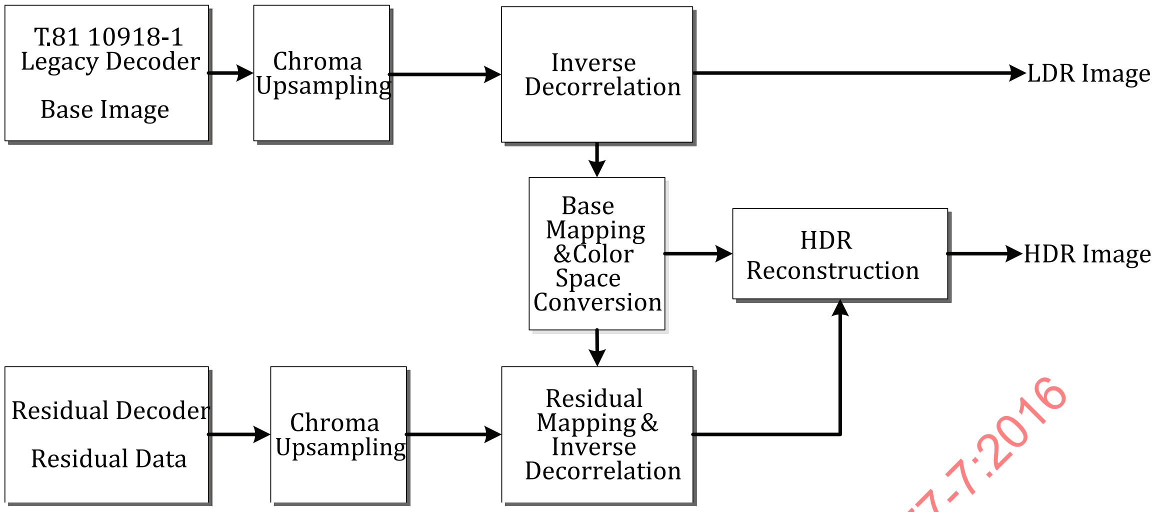


Figure 1 — Overview on the decoding process

5.2 Profiles

The Profiles define the implementation of a particular technology within the functional blocks of Figure 1. The profiles are described in Annex E.

5.3 Encoder requirements

There is **no requirement** in this part of ISO/IEC 18477 that any encoder shall support all profiles. An encoder is only required to meet the compliance tests and to generate the codestream according to the syntax and to limit the coding parameters to those valid within the profile it conforms to. Profiles are defined in Annex E.

5.4 Decoder requirements

A decoding process converts compressed image data to reconstructed image data. It **may** follow the decoding operation specified in the Recommendation | International Standard and ISO/IEC 18477-1 to generate an LDR image from the legacy codestream, and it **shall** follow the operations in this part of ISO/IEC 18477 to decode an HDR image from the data in the full file. The Decoder shall parse the codestream syntax to extract the parameters, the residual image and the base image. The parameters shall be used to merge the residual image with the base image into the reconstructed HDR Image.

In order to comply with this part of ISO/IEC 18477, a decoder

a) **may convert** a codestream conforming to this part of ISO/IEC 18477 **without considering the information in any box** into to a low dynamic range image, and

b) **shall** convert a conforming codestream within the profile it claims to be conforming to into a high dynamic range image.

Annex A
(normative)

Encoding and decoding process

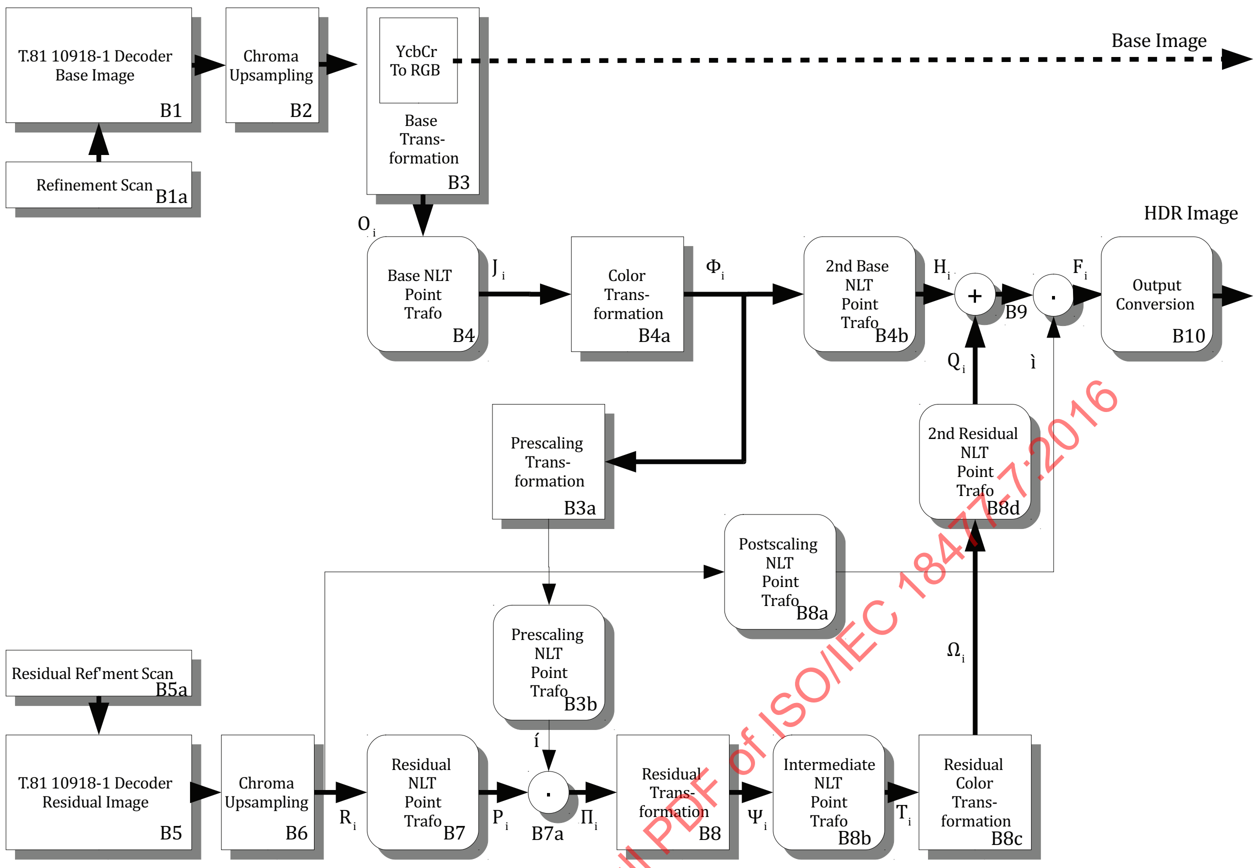
A.1 General

In this Annex and all of its subclauses, the flow charts and tables are normative only in the sense that they are defining an output that alternative implementations shall duplicate.

A.2 Decoding process (normative)

The decoding process relies on a layered approach to extend JPEG's capabilities. The encoder decomposes an HDR image into a base layer, which consists of a tone-mapped version of the HDR image, and an HDR residual layer. The combination of the data in the residual codestream and the legacy codestream allow the reconstruction of the fully HDR image. Both the metadata describing the merging process and the residual image are included in boxes invisible to legacy decoders. Such decoders will thus only see the tone mapped LDR image. While the base image complies to ISO/IEC 18477-1 and thus supports only the 8-bit extended or baseline, extended or progressive Huffman modes, the residual image may optionally be encoded in the 12-bit Huffman or progressive modes.

Figure A.1 illustrates the functionality of a compliant decoder.



NOTE Bold lines carry three (or one, for grey scale) components, thin lines scalar data. Round boxes implement point-transformations, square boxes (except B1, B1a, B5, B5a) multiplications by 3 × 3 matrices. Letters denote signal names.

Figure A.1 — High level overview of the decoding process of a compliant decoder

This subclause specifies the reconstruction process of an intermediate dynamic range image from a LDR image and residual image decoders shall follow. This process, which is an extension of the process defined in ISO/IEC 18477-6, consists of the following steps (see also Figure A.1).

— In steps B1 and B1a, reconstruct the base data from legacy codestream and the refinement codestream if a Refinement Data box is present. Refinement coding is specified in ISO/IEC 18477-6:2016, Annex D.

— In step B1 and B1a, apply the Inverse Quantization and Inverse Discrete Cosine Transformation as in Rec. ITU-T T.81 | ISO/IEC 10918-1.

— In step B2, the upsampling process specified in ISO/IEC 18477-1:2015, Annex A shall be followed to generate samples for all positions on the sample grid.

— In step B3, the linear transformation selected by the Base Transformation box defined in ISO/IEC 18477-6:2016, Annex B and Annex C shall be applied to inversely decorrelate the image components. ISO/IEC 18477-6:2016, Table C.1 defines which transformation to pick. The output of this block consists of either one or three samples per grid point O_i, depending on the number of components in the legacy codestream. The output of this transformation is rounded to integers and clipped to $[0, 2^{R_h+8}-1]$ where R_h is the number of refinement scans in the base image, see ISO/IEC 18477-6:2016, Annex D.

— In step B4, a point transformation shall be applied to each of the output components O_i. This process is selected according to the Base Nonlinear Point Transformation subbox of the Merging Specification box, implementing the Base Nonlinear Point Transformations and following the specifications of Annex C of ISO/IEC 18477-3:2015, Annex C. The outputs of this process are the predicted high dynamic range samples J_i. As above, i=1..Nf. The Base Nonlinear Point Transformation box is specified in ISO/IEC 18477-6:2016, Annex B.

— In step B4a, a colour transformation is applied to the input values J_i resulting in the output pixel values H_i. The transformation is selected by the Colour Transformation subbox of the Merging Specification box, which selects one of the transformations defined in ISO/IEC 18477-6:2016, Annex C. If Nf equals one, no transformation is performed. The Colour Transformation box is specified in ISO/IEC 18477-6:2016, Annex B. The output of this step is the linear precursor image samples Φ_i.

— In steps B3a and B3b, an optional prescaling transformation extracts the luminance from the linear precursor image Φ_i. A Nonlinear pre-scale map computes from the luminance a **pre-scaling factor** ν that is applied to the reconstructed chroma components P_2 and P_3 according to the procedure described in B.8. The luma value remains unaffected by the scaling process. If the nonlinear pre-scaling map is **not** present, the inferred value of the pre-scale factor shall be 1. The prescaling transformation is defined by the **Pre-scale Transformation box**, the nonlinear pre-scale map by the **Prescaling Nonlinear Point Transformation box**, both of which are specified in Annex B.

— In step B4b, the **Secondary Base Nonlinear Point Transformation box** applies a nonlinear transformation to the linear precursor image samples Φ_i creating the high dynamic range precursor image H_i. If this box is not present, the high dynamic range precursor image samples H_i are identical to the linear precursor image Φ_i.

— In steps B5 and B5a, the residual image shall be reconstructed from the data contained in the Residual Codestream box and the Residual Refinement box. The codestream contained in this box follows the specifications defined in Rec. ITU-T T.81 | ISO/IEC 10918-1. If a Residual Refinement box is present, the precision of the samples of the residual codestream shall be extended by refinement coding as specified in ISO/IEC 18477-6:2016, Annex D. The number of components of the residual image shall either be equal to the number of components signaled in the base image, or shall be one. If it is one, the inferred chrominance values are 2^{P-1+R_r} where P is the bit precision of the residual image specified by the frame header of the residual codestream and R_r the number of residual refinement scans. The R_r parameter is defined by the Refinement Specification subbox defined in ISO/IEC 18477-6:2016, Annex B.

— In steps B5 and B5a, apply Inverse Quantization and Inverse Discrete Cosine Transformation as in Rec. ITU-T T.81 | ISO/IEC 10918-1.

— In step B6, residual data is upsampled to the common sample grid following the specification of ISO/IEC 18477-1:2015, Annex A, resulting in the Residual Image R_i.

— In step B7, apply a point transformation to R_i that is defined by the Residual Point Transformation subbox of the Merging Specification box. The outputs of this operation are three residual sample values per pixel denoted by P_i.

— In step B7a, the chrominance values of the residual image P_i are multiplied by the prescaling factor ν, giving the pre-scaled values Π_i.

— In step B8, an inverse colour decorrelation shall be applied to the Π_i data. The transformation is defined by the Residual Transformation box, which is a subboxes of the Merging Specification box, and is specified in ISO/IEC 18477-6:2016, Annex B. The outputs of this process are the inversely decorrelated prediction errors Ψ_i.

— In step B8a, a postscaling factor μ is computed through the nonlinear post-scaling map from the luma component of the reconstructed residual image R. The post-scaling map is defined by the **Postscaling NonLinear Point Transformation box** defined in Annex B. If the Postscaling Nonlinear Point Transformation box is not present, the inferred value of the postscaling factor μ shall be 1.

— In step B8b, the **Intermediate Nonlinear Point Transformation box** defined in Annex B applies a gamma-correction on the decorrelated prediction errors Ψ_i generating the gamma-corrected decorrelated prediction errors T_i. This process is implemented by a Parametric Curve box. If the **Intermediate Nonlinear Point Transformation box** is not present, T_i equals Ψ_i.

— In step B8c, the **Residual Colour Transformation box** defined in Annex B applies an optional transformation of the gamma-corrected decorrelated prediction values T_i from the source colour space into the target colour space, defining the residual errors Ω_i. If the **Residual Colour Transformation box** is not present, Ω_i equals T_i.

— In step B8d, a secondary nonlinear point transformation selected by the **Secondary Residual Nonlinear Point Transformation box** is applied. This box selects for each component a nonlinearity, either a Floating Point Table Lookup box or a Parametric Curve box that maps the inversely decorrelated residual errors Ω_i into the final residual errors Q_i. If this box is not present, Q_i is set to Ω_i.

— In step B9, the intermediate dynamic range output F_i, i.e. the output of the decoding process, is reconstructed from H_i, the predicted high dynamic range signal, and the inversely decorrelation prediction errors Q_i. The output bit precision R_b is taken from the Output Conversion subbox of the Merging Specification box defined in Annex B. For this, compute

$$F_i := \mu \times \left(H_i + Q_i - 2^{Rb+8-1}\right)$$

— In step B10, the values F_i are further processed by the **output conversion** which is specified by the **Output Conversion box.** The output transformation box is defined in Annex B, and is an extension of the Output Transformation box defined in ISO/IEC 18477-6. Output conversion either performs a nonlinear point transformation for each component, or rounds the output to integers, then clamps to the range $[0,2^{Rb-1}]$ and then applies a half-exponential integer to floating point conversion specified in Annex D.

A.3 Encoding process (informative)

This part of ISO/IEC 18477 does not define a normative encoding process. Any encoding process that generates a file format that is compliant to this part of ISO/IEC 18477 is acceptable as long as this file format reconstructs to the desired image by the normative decoding algorithm.

A.4 Application notes (informative)

This part of ISO/IEC 18477 specifies a decoding process that is only normative in so far as it defines a decoding process whose output compliant implementations shall replicate. There is no need to follow the decoding process specified in A.2 precisely as long as its output is equivalent to the above. In fact, the above decoding process offers for many profiles (see Annex E for their definition) optimization opportunities implementations are encouraged to use.

This subclause is a non-exhaustive list of such opportunities that may lower the complexity of implementations or may lower coding errors due to imperfect implementations of mathematical functions such as x^a , log() or exp().

— The nonlinear point transformation processes defined by the Parametric Curve box include an input scaling and an output scaling mechanism that ensures that the arguments to the pre-defined functions of the Parametric Curve box are always between 0 and 1. This scaling process is defined in ISO/IEC 18477-3:2015, Annex C. Under some situations, this scaling process can be omitted because the output scaling of a nonlinear transformation cancels with the input scaling of the nonlinear transformation following it in the decoding chain, and the transformation between them is linear, allowing exchange it with linear scaling. This optimization can be applied to decoding steps B4, B4a, B4b and B8, B8b, B8c, B8d.

— The Output Conversion process together with the addition and the Secondary Base/Residual nonlinear Point Transformations, i.e. steps B4b, B8d, B9 and B10, can sometimes be accumulated into a single simpler composition process, giving decoders the freedom to merge these four steps into a single step. This happens, for example, for Profile B, see E.3, where the Secondary Base/Residual Nonlinear Point Transformations are logarithmic and the Output Conversion is an exponential map.

Implementations may want to optimize these steps into a single division operation. For details, see the informative note at the end of E.3.

— Pre-scaling and post-scaling operations, i.e. the multiplications by ν and μ in steps B7a and B9 can be omitted if the decoder can ensure that the scale factors are constant one. This happens, for example, in Profiles B, C and D where the Pre-scaling Nonlinear Transformation box and the Post-scaling Nonlinear Transformation box are absent.

Annex B
(normative)

Boxes

B.1 General

This Annex selects and refines a subset of the boxes defined in ISO/IEC 18477-3 for the purpose of high dynamic range image coding. It lists those boxes of ISO/IEC 18477-3 that are required for this part of ISO/IEC 18477. All other boxes are optional and its interpretation is outside the scope of this part of ISO/IEC 18477. Other parts of ISO/IEC 18477 or other standards may define their meaning, and decoders conforming to this part of ISO/IEC 18477 may ignore them.

Table B.1 lists the boxes required in this part of ISO/IEC 18477 that are paired with ISO/IEC 18477-3. Some of the boxes require additional specifications that are listed in subsequent clauses of this Annex.

Table B.1 — Boxes within this part of ISO/IEC 18477

Box Name	Box Type	Further definitions in which subclause of this part of ISO/IEC 18477
File Type box	0x66747970 ("ftyp").	
Legacy Data Checksum box	0x4C43484B ("LCHK")	
Residual Data box	0x52455349 ("RESI")	
Residual Refinement box	0x5246494e ("RFIN")	
Refinement Data box	0x46494e45 ("FINE")	
Merging Specification box	0x53504543 ("SPEC")	
Parametric Curve box	0x43555256 ("CURV")	
Integer Table Lookup box	0x544f4e45 ("TONE")	
Floating Point Table Lookup box	0x46544f4e ("FTON")	
Fix-point Linear Transformation box	0x4D545258 ("MTRX")	
Floating-point Linear Transformation box	0x46545258 ("FTRX")	
Output Conversion box	0x4F434F4E ("OCON")	B.2
Refinement Specification box	0x52535043 ("RSPC")	
Base Nonlinear Point Transformation Specification box	0x4C505453 ("LPTS")	B.3
Secondary Base Nonlinear Point Transformation Specification box	0x52505453 ("CPTS")	B.4
Residual Nonlinear Point Transformation Specification box	0x5152505453 ("QPTS")	B.5
Intermediate Residual Nonlinear Point Transformation Specification box	0x44505453 ("DPTS")	B.6
Secondary Residual Nonlinear Point Transformation Specification box	0x52505453 ("RPTS")	B.7

Table B.1 *(continued)*

Box Name	Box Type	Further definitions in which subclause of this part of ISO/IEC 18477
Prescaling Nonlinear Point Transformation Specification box	0x53505453 (“SPTS”)	B.8
Postscaling Nonlinear Point Transformation box	0x50505453 (“PPTS”)	B.9
Base Transformation box	0x4C545246 (“LTRF”)	B.10
Residual Transformation box	0x52545246 (“RTRF”)	B.11
Colour Transformation box	0x43545246 (“CTRF”).	B.12
Residual Colour Transformation box	0x44545246 (“DTRF”)	B.13
Prescaling Transformation box	0x53545246 (“STRF”)	B.14

B.2 Output Conversion box

This mandatory box defines the conversion process from the result of the base image/residual image merging process to the final floating point samples. It describes the final merging process and by that step B10 of the algorithm described in A.2. This box is already defined in ISO/IEC 18477-3:2015, Annex B, though its application to this part of ISO/IEC 18477 further constraints the value of its fields.

This box shall never appear top level in the file, but it shall be a subbox of the Merging Specification box defined in ISO/IEC 18477-3:2015, Annex B. Exactly one Output Conversion box shall appear in the Merging Specification box if a Merging Specification box exists.

Table B.2 constraints the parameters of the Output Conversion box as applied in this part of ISO/IEC 18477.

The detailed list of operations to be performed by the Output Conversion is specified in ISO/IEC 18477-3:2015, Annex B.

Table B.2 — Parameter constraints for the Output Conversion box

Parameter	Constraints within this Recommendation | International Standard	Meaning
R_b	8	The value of this field shall be eight. Number of additional bits available for high dynamic range data. The bit precision of the reconstructed high dynamic range image shall be computed as $8 + R_b = 16$.
Lf	0	The value of this field shall be 0. This field indicates whether the compression is lossy or lossless. ISO/IEC 18477-7 only specifies lossy coding. Decoders may use this field to decide whether to follow the DCT specification of ISO/IEC 18477-8 precisely or whether a customized DCT implementation is acceptable.

Table B.2 *(continued)*

Parameter	Constraints within this Recommendation | International Standard	Meaning
Oc	0..1	Half-exponential output-enable flag. If this flag is set, the output of the merging step shall be rounded to the nearest integer, then clipped, and the half-exponential map defined in Annex D shall be applied. The value of this flag shall be identical to the value of the Ce flag which controls clipping of output values. If the Oc flag is 1, the Ol flag shall be 0. See Table B.3 for a detailed definition of Output Conversion.
Ce	0..1	This field indicates whether the output shall be clipped to range $[0,2^{8+Rb}-1]$ if Oc is 0, or clipped to [-0x7bff,0x7bff] if Oc is 1 before processing the data further. The value of this flag shall be identical to the value of the Oc flag.
Ol	0..1	This field indicates whether an output lookup or point transformation is required. If so, output shall be processed further by a nonlinear point transformation defined by the to_i parameters. This flag shall be 0 whenever the Oc flag is 1.
to_0	0..15	This field selects the output point transformation for component 0 if Ol is set. It selects the Floating Point Table Lookup box or a Parametric Curve box for output mapping by finding the box whose M parameter matches to_0. If Ol is 0, this parameter is ignored.
to_1	0..15	Selects a nonlinear point output map for component 1 if Ol is set. Ignored if Ol is reset.
to_2	0..15	Selects a nonlinear point output map for component 2 if Ol is set. Ignored if Ol is reset.
to_3	0	Reserved for ITU | ISO/IEC purposes.

Table B.3 — Selection of the output conversion from the parameters

Oc	Ce	Ol	to_i	Output conversion process
1	1	0	ignored	Half exponential map for all components specified in Annex D
0	0	0	ignored	Direct output of F_i, the output the merging process are the reconstructed samples
0	0	1	Selects nonlinear point transformation for component i	The outputs of the merging process F_i are the input to a nonlinear point transformation selected by to_i. The value to_i selects a Floating Point Table Lookup box or a Parametric Curve box by finding the box whose M value equals the value of to_i
All other combinations				Reserved for ITU | ISO/IEC purposes.

The nonlinear point transformations selected by this box shall only be Parametric Curve boxes. The corresponding boxes referenced by this box appear at top level of the ISO/IEC 18477-3 compliant file or as subboxes of the Merging Specification box. The nonlinear point transformation itself is given by the process specified in ISO/IEC 18477-3:2015, Annex C. It requires four additional parameters, the input range R_w, R_e and the output range R_t, R_f. The two value pairs shall be given as follows:

$$R_w = 1 \qquad R_e = 0$$

$$R_t = 1 \qquad R_f = 0$$

The value of the rounding mode e shall be 0. Note that the value of e is ignored for this specific setting.

B.3 Base Nonlinear Point Transformation Specification box

This box defines the nonlinear point transformation between the output of the Base Transformation and the Colour Transformation. It thus defines step B4 in the decoder description in Annex A. At most one Base Nonlinear Point Transformation Specification box shall exist as a subbox of the Merging Specification box. It shall not appear at top-level of the file. The box layout of this box is that of the Nonlinear Transformation Specification box, defined in ISO/IEC 18477-3:2015, Annex B

The nonlinear point transformation selected by this box is an Integer Table Lookup box, a Floating Point Table Lookup box or a Parametric Curve box. The corresponding boxes referenced by this box appear at top level of the ISO/IEC 18477-3 compliant file or as subbox of the Merging Specification Box. The nonlinear point transformation itself is given by the process specified in ISO/IEC 18477-3:2015, Annex C. It requires four additional parameters, the input range R_w, R_e and the output range R_t, R_f. The two value pairs shall be given as follows:

$$R_w = 8 + R_h \qquad R_e = 0$$

$$R_t = 8 + R_b \qquad R_f = 0$$

The value R_h is the number of refinement scans in the base decoding path and is found in the Refinement Specification box. If the Refinement specification box is absent, the inferred value of R_h is 0.

The value R_b is found in the Output Conversion box, where R_b + 8 is the output precision of the image.

If this box does not exist, the implied nonlinear point transformation in the base decoding path is defined as if a Parametric Curve box with the identity function, i.e. with parameters t = 2 and e = 1, had been selected. That is, the input range is linearly scaled to the output range by the parameters R_w and R_t.

The type of this box shall be 0x4C505453, ASCII encoding of “LPTS”. The box structure and layout does not deviate from that in ISO/IEC 18477-3:2015, Annex B, neither apply any restrictions to parameters of the box.

B.4 Secondary Base Nonlinear Point Transformation Specification box

This box, if present, selects a secondary nonlinear point-transformation that is applied in the precursor image domain **after** the Colour Transformation. This box implements processing step B4a in the description of Annex A, i.e. it maps the linear precursor image sample values Φ_i into the final precursor image H_i to which the prediction residuals Q_i are added. If this box is not present, H_i is identical to Φ_i. The box layout of this box is that of the Nonlinear Transformation Specification box, defined in ISO/IEC 18477-3:2015, Annex B.

The nonlinear point transformation selected by this box is either a Floating Point Table Lookup Box or a Parametric Curve box. The corresponding boxes referenced by this box appear at top level of the ISO/IEC 18477-3 compliant file or as subboxes of the Merging Specification box. The nonlinear point transformation itself is given by the process specified in ISO/IEC 18477-3:2015, Annex C. It requires four additional parameters, the input range R_w, R_e and the output range R_t, R_f. The two value pairs shall be given as follows:

$$R_w = 8 + R_b \qquad R_e = 0$$

$$R_t = 8 + R_b \qquad R_f = 0$$

The value R_b is found in the Output Conversion box, where R_b + 8 is the output precision of the image.

If this box does not exist, the implied nonlinear point transformation in the base decoding path is defined as if a Parametric Curve box with the identity function had been selected, i.e. with parameters t = 2 and e = 1. That is, the input is identical to the output.

The type of this box shall be 0x43505453, ASCII encoding of “CPTS”. The box structure and layout does not deviate from that in ISO/IEC 18477-3:2015, Annex B, neither apply any restrictions to parameters of the box.

B.5 Residual Nonlinear Point Transformation Specification box

This box defines the nonlinear point transformation between the output of the residual DCT transformation and the Residual Transformation. It thus defines step B7 in the decoder description in Annex A. At most, one Residual nonlinear Point Transformation Specification box shall exist as a sub-box of the Merging Specification box. It shall not appear at top-level of the file. The box layout of this box is that of the Nonlinear Transformation Specification box, defined in ISO/IEC 18477-3:2015, Annex B. This box shall only be present if the Residual Data box is present.

The nonlinear point transformation selected by this box, if it is present, shall be a Parametric Curve box. References to Integer or Floating Point Table Lookup boxes are not permitted. The corresponding boxes referenced by this box appear at top level of the ISO/IEC 18477-3 compliant file or as subboxes of the Merging Specification box. The nonlinear point transformation itself is given by the process specified in ISO/IEC 18477-3:2015, Annex C. It requires four additional parameters, the input range R_w, R_e and the output range R_t, R_f. The two value pairs shall be given as follows:

$$R_w = P + R_r \qquad R_e = 0$$

$$R_t = 8 + R_b \qquad R_f = 0$$

The value R_r is the number of residual refinement scans in the base decoding path and is found in the Refinement Specification box. If the Refinement Specification box is absent, the inferred value of R_r is 0. The value of P is the bitdepth of the residual image; it is found in the frame header of the residual codestream, see Rec. ITU-T T.81 | ISO/IEC 10918-1:1994, Table B.2.

The value R_b is found in the Output Conversion box, where R_b + 8 is the output precision of the image.

If this box does not exist, the implied nonlinear point transformation in the base decoding path is defined as if a Parametric Curve box with the identity function had been selected, i.e. with parameters t = 2 and e = 0. That is, the input range is linearly scaled to the output range by the parameters R_w and R_t.

The type of this box shall be 0x5152505453, ASCII encoding of “QPTS”. The box structure and layout does not deviate from that in ISO/IEC 18477-3:2015, Annex B, neither apply any restrictions to parameters of the box.

NOTE The default rounding mode in the residual domain in the absence of this box is e = 0, unlike the default rounding mode in the base domain, which is e = 1. This deviation is intentional.

B.6 Intermediate Residual Nonlinear Point Transformation Specification box

This box, if present, selects an intermediate nonlinear point-transformation that is applied in the residual domain **after** the Residual Transformation but **before** the Residual Colour Transformation. This box implements processing step B8b in the description of Annex A, i.e. it maps the inversely decorrelated residual error values Ψ_i into gamma-corrected residual values T_i. If this box is not present, T_i is identical to Ψ_i. The box layout of this box is that of the Nonlinear Transformation Specification box, defined in ISO/IEC 18477-3:2015, Annex B.

The nonlinear point transformation selected by this box is either a Floating Point Table Lookup box or a Parametric Curve box. The corresponding boxes referenced by this box appear at top level of the ISO/IEC 18477-3 compliant file or as subboxes of the Merging Specification box. The nonlinear point transformation itself is given by the process specified in ISO/IEC 18477-3:2015, Annex C. It requires four additional parameters, the input range R_w, R_e and the output range R_t, R_f. The two value pairs shall be given as follows:

$$R_w = 8 + R_b \qquad R_e = 0$$

$$R_t = 8 + R_b \qquad R_f = 0$$

The value R_b is found in the Output Conversion box, where R_b+8 is the output precision of the image.

If this box does not exist, the implied nonlinear point transformation in the residual decoding path is defined as if a Parametric Curve box with the identity function had been selected, i.e. a Parametric Curve box with parameters t = 2 and e = 0. That is, the input is identical to the output.

The type of this box shall be 0x44505453, ASCII encoding of “DPTS”. The box structure and layout does not deviate from that in ISO/IEC 18477-3:2015, Annex B, neither apply any restrictions to parameters of the box.

B.7 Secondary Residual Nonlinear Point Transformation Specification box

This box, if present, selects a secondary nonlinear point transformation that is applied in the residual domain **after** the Residual Transformation. This box implements processing step B8d in the description of Annex A, i.e. it maps the inversely decorrelated residual error values Ω_i into the final residual error values Q_i that are added to the precursor image. If this box is not present, Q_i is identical to Ω_i. The box layout of this box is that of the Nonlinear Transformation Specification box, defined in ISO/IEC 18477-3:2015, Annex B.

The nonlinear point transformation selected by this box is either a Floating Point Table Lookup box or a Parametric Curve box. The corresponding boxes referenced by this box appear at top level of the ISO/IEC 18477-3 compliant file or as subboxes of the Merging Specification box. The nonlinear point transformation itself is given by the process specified in ISO/IEC 18477-3:2015, Annex C. It requires four additional parameters, the input range R_w, R_e and the output range R_t, R_f. The two values shall be given as follows:

$$R_w = 8 + R_b \qquad R_e = 0$$

$$R_t = 8 + R_b \qquad R_f = 0$$

The value R_b is found in the Output Conversion box, where R_b + 8 is the output precision of the image.

If this box does not exist, the implied nonlinear point transformation in the residual decoding path is defined as if a Parametric Curve box with the identity function had been selected, i.e with parameters t = 2 and e = 0. That is, the input is identical to the output.

The type of this box shall be 0x52505453, ASCII encoding of “RPTS”. The box structure and layout does not deviate from that in ISO/IEC 18477-3:2015, Annex B, neither apply any restrictions to parameters of the box.

B.8 Prescaling Nonlinear Point Transformation box

This box, if present, computes a point-transformation that computes a prescaling factor ν from the luminance of the LDR image, implementing step B3b of Figure A.1. If this box is present, the chroma components P_i of the residual image are scaled by ν, giving the scaled chroma residuals Π_i. This scaling operation, to be implemented as specified below, implements step B7a of Figure A.1.

If the Prescaling Nonlinear Point Transformation box is not present or the residual image consists only of a single component, the inferred value of ν shall be 1, or equivalently, the multiplication of step B7a is skipped.

Scaling of the chroma components, i.e. step B7a of Figure A.1 shall be implemented as follows. It computes the scaled residual components Π_i from the unscaled residual components P_i according to

$$\Pi_0 = P_0$$

$$\Pi_1 = \upsilon \times (P_1 - 2^{Rr + P - 1}) + 2^{Rr + P - 1}$$

$$\Pi_2 = \upsilon \times (P_2 - 2^{Rr + P - 1}) + 2^{Rr + P - 1}$$

where R_r is the number of refinement scans in the residual image, signaled by the Refinement Specification box and P is the bit precision signaled in the residual refinement frame header. The luminance component remains unscaled.

If the residual image consists only of a single component, step B7a is skipped and no scaling takes place.

NOTE The above procedure scales the chrominance values around the neutral value by υ.

This box shall never appear top level in the file, but it shall be a subbox of the Merging Specification box defined in ISO/IEC 18477-3:2015, Annex B. At most, one Prescaling Nonlinear Point Transformation box shall appear in the Merging Specification box.

The nonlinear point transformation selected by this box is either a Floating Point Table Lookup box or a Parametric Curve box. The corresponding boxes referenced by this box appear at top level of the ISO/IEC 18477-3 compliant file or as subboxes of the Merging Specification box. The nonlinear point transformation itself is given by the process specified in ISO/IEC 18477-3:2015, Annex C. It requires four additional parameters, the input range R_w, R_e and the output range R_t, R_f. The two value pairs shall be given as follows:

$$R_w = 8 + R_b \qquad R_e = 0$$

$$R_t = 1 \qquad R_f = 0$$

The type of this box shall be 0x53505453, ASCII encoding of “SPTS”. The box structure and layout does not deviate from that in ISO/IEC 18477-3:2015, Annex B, neither apply any restrictions to parameters of the box, however only parameter td_0 of the box is used.

B.9 Postscaling Nonlinear Point Transformation Specification box

This box, if present, selects a point-transformation that computes a postscaling factor μ from the luminance component (component 0) of the residual image. If this box is present, all components o are multiplied by μ after adding the residual image to the precursor image. It hence implements steps B8a and B9 of Figure A.1. If the Postscaling Nonlinear Point Transformation box is not present, the inferred value of μ shall be 1 and scaling of the output components shall be skipped. The box layout of this box is that of the Nonlinear Transformation Specification box, defined in ISO/IEC 18477-3:2015, Annex B.

This box shall never appear top level in the file, but it shall be a subbox of the Merging Specification box defined in ISO/IEC 18477-3:2015, Annex B. At most, one Postscaling Nonlinear Point Transformation box shall appear in the Merging Specification box.

The nonlinear point transformation selected by this box is either a Floating Point Lookup box or a Parametric Curve box. The corresponding boxes referenced by this box appear at top level of the ISO/IEC 18477-3 compliant file or as subboxes of the Merging Specification box. The nonlinear point transformation itself is given by the process specified in ISO/IEC 18477-3:2015, Annex C. It requires four additional parameters, the input range R_w, R_e and the output range R_t, R_f. The two values shall be given as follows:

$$R_w = P + R_r \qquad R_e = 0$$

$$R_t = 1 \qquad R_f = 0$$

where P is the precision of the residual image frame and R_r the number of residual refinement scans, taken from the Refinement Specification box.

The type of this box shall be 0x50505453, ASCII encoding of “PPTS”. The box structure and layout does not deviate from that in ISO/IEC 18477-3:2015, Annex B, neither apply any restrictions to parameters of the box.

B.10 Base Transformation box

This box defines the linear transformation between the output of the base entropy decoding process and the input of the base image nonlinear point transformation. It defines the transformation in step B3 of the decoding process specified in Annex A. The box structure and layout is already defined in ISO/IEC 18477-3:2015, Annex B, though its purpose is refined here. Its definition is intentionally identical to that given in ISO/IEC 18477-6:2016, Annex B.

There shall be at most, one Base Transformation box as subbox of the Merging Specification box. This box shall only exist if the number of components Nf is equal to 3, and it shall exist if the output bitdepth is larger than 8, i.e. the R_b parameter of the Output Conversion box is nonzero.

If the box does not exist, the inverse decorrelation transformation is the identity process if Nf is equal to 1. If the box does not exist and the number of components is equal to 3, the base transformation in step B3 of Annex A is defined by the Component Decorrelation Control Marker specified in ISO/IEC 18477-1:2015, Annex B, that is, it is the ICT if the Decorrelation Control Marker is absent or its cc value is equal to 1, it is the identity if the Decorrelation Control Marker exists and its cc value is 0. Details on the selection of the Base Transformation are specified in ISO/IEC 18477-6:2016, Table C.1.

The linear transformations specified in ISO/IEC 18477-6:2016, Annex C requires an additional level shift parameter R_s. The value of R_s for the Base transformation shall be given as follows

$$R_s = 8 + R_h - 1$$

where R_h is the number of refinement scans contained in the base decoding path. The value of R_h is found in the Refinement Specification box of the Merging Parameter box. It shall be 0 if no Refinement Specification box is present.

The type of the Base Transformation box shall be 0x4C545246, ASCII encoding of “LTRF”.

Table B.3 constraints the parameters of the Base Transformation box based on the parameters listed in ISO/IEC 18477-3:2015, Annex B.

Table B.4 — Parameter constraints for the Base Transformation box

Parameter	Constraints within this part of ISO/IEC 18477	Meaning
Xt	1, 2 or 5..15	Defines the linear transformation to be used as Base Transformation, see Table B.6 for the encoding of this field.
Re	0	Shall be 0.

B.11 Residual Transformation box

This box defines the linear transformation between the output of the nonlinear point transformation in the residual decoding path and the addition of the residual to the output of the Colour Transformation in the base decoding path. It defines the linear transformation in step B8 of the decoding process specified in Annex A. The box structure and layout is already defined in ISO/IEC 18477-3:2015, Annex B, though its purpose is refined here. Its definition is intentionally identical to that given in ISO/IEC 18477-6:2016, Annex B.

This box shall only exist as a subbox of the Merging Specification box specified in ISO/IEC 18477-3:2015, Annex B. It shall not appear top level. This box shall exist if and only if a Residual Data Box is present at the top level of the file and the number of components Nf equals 3.

The linear transformations specified in Annex C require an additional level shift parameter R_s. The value of R_s for the residual transformation shall be given as follows

$$R_s = 8 + R_b - 1$$

where R_b + 8 is the sample precision of the reconstructed HDR output image. The value of R_b can be found in the Output Conversion box, see B.2.

The type of this box shall be 0x52545246, ASCII encoding of “RTRF”.

Table B.5 constraints the parameters of the Residual Transformation box from the generic Linear Transformation Specification box defined in ISO/IEC 18477-3.

Table B.5 — Parameter Constraints for the Residual Transformation box

Parameter	Constraints within this part of ISO/IEC 18477	Meaning
Xt	1, 2 or 5..15	Defines the linear transformation to be used as Residual Transformation, see Table B.6 for the encoding.
Re	0	Shall be 0.

Table B.6 — Encoding of the Xt parameters of the Base and Residual Transformation box

Value	Transformation to be used
0	Reserved for ITU | ISO/IEC purposes.
1	The identity transformation shall be used.
2	The ICT Transformation as specified in ISO/IEC 18477-6:2016, Annex C shall be used.
3	Reserved for ITU | ISO/IEC purposes.
4	Reserved for ITU | ISO/IEC purposes.
5...15	The free form transformation (with offset shift) defined by the Integer or Floating Point Linear Transformation box whose M value matches the value of Xt shall be used. The Integer or Floating Point Linear Transformation boxes are specified in ISO/IEC 18477-3:2015, B.2 and B.3 and their application and implementation are specified in ISO/IEC 18477-6:2016, Annex C.

B.12 Colour Transformation box

This box defines the linear transformation between the output of the nonlinear point transformation in the base domain and the addition of the inversely decorrelated transformed residual. It defines the linear transformation in step B4a of the decoding process specified in Annex A. The box structure and layout is already defined in ISO/IEC 18477-3:2015, Annex B, though its purpose is refined here. The box is intentionally identical to the Colour Transformation box of ISO/IEC 18477-6, its definition is repeated here.

This box shall only exist as a subbox of the Merging Specification box specified in ISO/IEC 18477-3:2015, Annex B and it may only exist if the number of components in the image Nf equals 3. It shall not appear top level. If this box does not exist, the Colour Transformation shall be the identity transformation; otherwise the Xt parameter of the Colour Transformation box specifies the transformation matrix to pick.

The linear transformations specified in ISO/IEC 18477-6:2016, Annex C require an additional level shift parameter R_S. The value of R_S for the colour transformation shall be given as

$$R_S = -\infty \qquad \text{(i.e. no level shift)}$$

The type of this box shall be 0x43545246, ASCII encoding of "CTRF".

Table B.7 constraints the parameters from that defined in ISO/IEC 18477-3.

Table B.7 — Parameter Constraints for the Colour Transformation box

Parameter	Constraints within this part of ISO/IEC 18477	Meaning
Xt	1 or 5...15	Defines the linear transformation to be used as Base Transformation, see Table B.8 for the encoding.
Re	0	Shall be 0.

Table B.8 — Encoding of the Xt parameter of the Colour Transformation box

Value	Transformation to be used
0	Reserved for ITU | ISO/IEC purposes.
1	The identity transformation shall be used.
2..4	Reserved for ITU | ISO/IEC purposes.
5..15	The free form transformation (without offset shift) defined by the Integer or Floating Point Linear Transformation box whose M value matches the value of Xt shall be used. The Integer or Floating Point Linear Transformation boxes are specified in ISO/IEC 18477-3:2015, B.2 and B.3 and their application and implementation are specified in ISO/IEC 18477-6:2016, Annex C.

B.13 Residual Colour Transformation box

This box defines an additional linear transformation between the output of the nonlinear point transformation in the base domain and the addition of the inversely decorrelated transformed residual. It defines the linear transformation in step B8c of the decoding process specified in Annex A and computes the values Ω_i from the T_i. If this box is not present, the Ω_i shall be identical to T_i. The box structure and layout is already defined in ISO/IEC 18477-3:2015, Annex B, though its purpose is refined here.

This box shall only exist as a subbox of the Merging Specification box specified in ISO/IEC 18477-3:2015, Annex B and it may only exist if the number of components in the image Nf equals 3. It shall not appear top level. If this box does not exist, the Residual Colour Transformation shall be the identity transformation; otherwise, the Xt parameter of the Residual Colour Transformation box specifies the transformation matrix to pick.

The linear transformations specified in Annex C require an additional level shift parameter R_S. The value of R_S for the residual colour transformation shall be given as

$$R_S = -\infty \qquad \text{(i.e no level shift)}$$

The type of this box shall be 0x44545246, ASCII encoding of "DTRF".

Table B.9 constraints the parameters of this box from the generic box definition in ISO/IEC 18477-3.

Table B.9 — Parameter constraints of the Residual Colour Transformation box

Parameter	Constraints within this Recommendation | International Standard	Meaning
Xt	1, 5..15	Defines the linear transformation to be used as Base Transformation, see Table B.10 for the encoding.
DRe	0	Shall be 0.

Table B.10 — Encoding of the Xt parameter of the Residual Colour Transformation box

Value	Transformation to be used
0	Reserved for ITU | ISO/IEC purposes.
1	The identity transformation shall be used.
2..4	Reserved for ITU | ISO/IEC purposes.
5..15	The free form transformation (without offset shift) defined by the Integer or Floating Point Linear Transformation box whose M value matches the value of Xt shall be used. The Integer or Floating Point Linear Transformation boxes are specified in ISO/IEC 18477-3:2015, B.5 and B.6 and their application and implementation are specified in ISO/IEC 18477-6:2016, Annex C.

B.14 Prescaling Transformation box

This box defines the linear transformation that computes from the linear precursor image Φ_i the luminance. This luminance is mapped by the Prescaling Nonlinear Point Transformation box into a prescaling factor ν applied to the chroma components of the residual image. It defines the transformation in step B3a of the decoding process specified in Annex A. There shall be, at most, one Base Transformation box as subbox of the Merging Specification box. This box shall exist if and only if the number of components Nf is equal to 3 and the Prescaling Nonlinear Point Transformation Specification box exists. The box structure and layout is already defined in ISO/IEC 18477-3:2015, Annex B, though its purpose is refined here.

If the prescaling Nonlinear Point-Transformation Specification box exists and the number of components Nf is equal to 1, the input to the prescaling nonlinear point transformation is taken from the first (and only) component of the base image. The details of the merging process of residual/refinement data with LDR data is defined in Annex A.

The linear transformations specified in ISO/IEC 18477-6:2016, Annex C require an additional level shift parameter R_S. The value of R_S for the prescaling transformation shall be given as

$$R_s = 8 + R_h - 1$$

where R_h is the number of refinement scans contained in the base decoding path. The value of R_h is found in the Refinement Specification box of the Merging Parameter box. It shall be 0 if no Refinement Specification box is present.

The type of the Prescaling Transformation box shall be 0x53545246, ASCII encoding of “STRF”.

Table B.11 constraints the parameters of the Residual Transformation box from the generic Linear Transformation Specification box defined in ISO/IEC 18477-3:2015, Table B specifies the encoding of the Xt parameter of the box.

Table B.11 — Parameter Constraints for the Prescaling Transformation box

Parameter	Constraints within this part of ISO/IEC 18477	Meaning
Xt	2, 5..15	Defines the linear transformation to be used as Prescaling Transformation, see Table B.12 for the encoding.
Re	0	Shall be 0.

Table B.12 — Encoding of the Xt parameter of the Prescaling Transformation

Value	Transformation to be used
0..1	Reserved for ITU ISO/IEC purposes.
2	The forwards ICT Transformation as specified in Annex C shall be used.
3..4	Reserved for ITU ISO/IEC purposes.
5..15	The free form transformation with offset shift defined by the Integer or Floating Point Linear Transformation box whose M value matches the value of Xt shall be used. The Integer or Floating Point Linear Transformation boxes are specified in ISO/IEC 18477-3:2015, B.2 and B.3 and their application and implementation are specified in ISO/IEC 18477-6:2016, Annex C.

Annex C
(normative)

Multi-component decorrelation

C.1 General

In this Annex and all of its subclauses, the flow charts and tables are normative only in the sense that they are defining an output that alternative implementations shall duplicate.

This Annex defines the Multiple Component Decorrelation Transformations available as base, residual and colour transformations of the decoding process. This Annex is by itself an extension of ISO/IEC 18477-6:2016, Annex C as it defines one additional transformation that may be selected by the Prescaling Transformation box. Additional transforms also applicable in this part of ISO/IEC 18477 are found in ISO/IEC 18477-6:2016, Annex C.

A Multiple Component Decorrelation Transformation takes one or three input components, and generates from them one or three output components. A task dependent level shift by a value of 2^{R_S} ensures that chroma components are correctly centered and representable by unsigned values. The values of R_S that are applicable to the transformations are specified in B.10 to B.13.

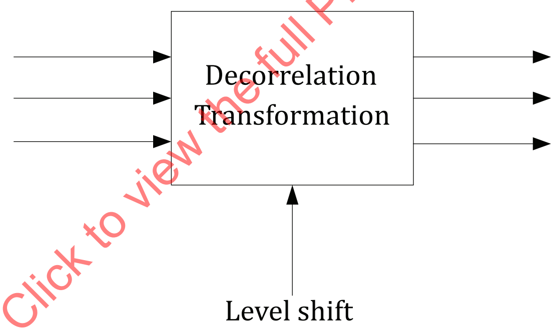


Figure C.1 — Input and output of a decorrelation transformation: Input components, output components and the DC level shift.

C.2 Irreversible forward multiple component transformation (forward ICT) (normative)

This transformation is used to decorrelate the low dynamic range versions of the image before encoding the components by a Rec. ITU-T T.81 | ISO/IEC 10918-1 conforming encoder. It is identical to the forward ICT defined in ISO/IEC 18477-1. Let L_0 to L_2 be the input pixel values, and R_S the number of level shift scale bits, then the output I_0 to I_2 shall be computed as

$$I_0 = 2{,}990\,0 \times L_0 + 0{,}587\,00 \times L_1 + 0{,}114\,00 \times L_2$$

$$I_1 = -0{,}168\,735\,891\,6 \times L_1 + 0{,}331\,264\,108\,4 \times L_1 + 0{,}5 \times L_2 + 2^{Rs}$$

$$I_2 = 0{,}5L_2 - 0{,}418\,687\,689\,2 \times L_1 - 0{,}081\,312\,410\,85 \times L_2 + 2^{Rs}$$

NOTE This transformation is intentionally identical to the forward ICT in ISO/IEC 18477-6. However, while only informative in ISO/IEC 18477-6, it is normative here. Note further that outputs I_1 and I_2 are not required for processing the Prescaling Transformation and their computation may be skipped.

Annex D
(normative)

Half-exponential output transformation

D.1 General

In this Annex and all of its subclauses, the flow charts and tables are normative only in the sense that they are defining an output that alternative implementations shall duplicate.

This Annex defines an output conversion that transforms integer sample values to floating point values. It is selected as final step for generating output sample values from the sum of base and residual image. This type of output conversion is applied whenever the Oc flag of the Output Conversion box specified in B.2 is set to 1.

If this output conversion is set, the (potential) floating point output of the merging step is rounded to the nearest integer, and then clamped to the range [-0x7bff,0x7bff]. This integer is then converted to floating point by the procedure specified in D.2.

D.2 Half-exponential output conversion (normative)

This Annex defines an optional output transformation denoted by Ψexp that may help to improve the compression performance for floating point sample values. This output transformation is effective whenever the Oc parameter of the Ouput Conversion box is 1, see B.2 for details on this box. The Ψexp map implements a half-exponential map that converts its integer input to a floating point output.

The Ψexp map is defined for non-negative inputs ($x \geq 0$) by the following steps.

— The integer input x shall be represented in a 16-bit binary representation.

— The resulting bit-pattern shall be re-interpreted as an ISO/IEC/IEEE 60559 (IEEE 754) floating point number. This floating point number is the final output of the output transformation.

For negative inputs ($x < 0$), the output of Ψexp is defined as Ψexp(x) := - Ψexp(-x-1).

NOTE It can be seen that this re-interpretation of bit-patterns on non-negative integers is equivalent to a piecewise linear approximation of the exponential function $f(x) \approx (2^{y-15})$ with $y = x \times 2^{-10}$. The definition for negative numbers ensures that Ψexp(-1) maps to the binary representation of the floating point number -0.

Annex E
(normative)

Profiles

E.1 General

This Annex defines four profiles that limit the choices of coding parameters offered by this part of ISO/IEC 18477. Profile A implements a decoding algorithm that is identical to the decoding algorithm specified in ISO/IEC 18477-2. Profile B specifies a decoding algorithm that computes the HDR output image as the quotient of the base image and the residual image. Profile C defines a simple additive profile that is compatible to the lossless HDR image coding specified in ISO/IEC 18477-8 and allows scalable lossy to lossless coding. Profile D implements a minimal HDR decoder that does not require a residual codestream. Coding performance of the profiles may differ, and will also depend on the input image.

Profiles are indicated by an additional entry in the Compatibility List of the File Type box, specified in ISO/IEC 18477-3:2015, Annex B. A decoder claiming performance to one of the profiles does not need to implement any mechanism that is not included in the profile, and it may use an optimized decoding procedure that bypasses some of the intermediate results specified in Annex A provided that its final output is equivalent to that of the hypothetical reference decoder.

E.2 Profile A

Figure E.1 gives a high-level overview of this profile. It is, with appropriate choices of parameters, a specialization of the decoding procedure described in Annex A. The profile indicator of this profile, i.e. the entry CL_i in the compatibility list of the File Type box shall be 0x78726464, ASCII encoding for “xrdd”

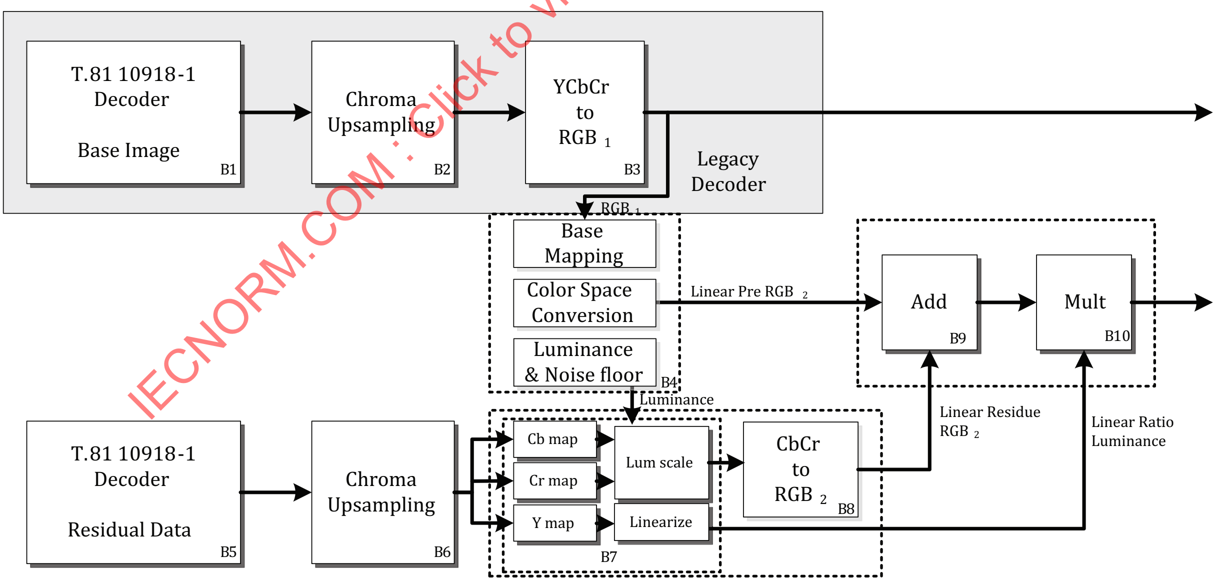


Figure E.1 — High level overview of the decoding process of a Profile A compliant decoder

The following constraints apply to a profile A compliant decoder.

— The residual image bit precision P shall be constrained to 8.

— The Refinement Specification box (see ISO/IEC 18477-6:2016, Annex D) shall be absent. This implies that no refinement scans are present and no Refinement Data or Residual Refinement boxes are present.

— The Xt value of the Base Transformation box shall be 2. That is, the base transformation shall be the ICT.

— The Base Nonlinear Point Transformation box shall select identical transformations for all components. Specifically, the transformation shall be defined by either a Parametric Curve box with a curve type t = 4 (gamma mapping) or a Floating Point Table Lookup box. If the Parametric Curve box is selected, it is constrained to the following parameters:

  e = 1 P_1 = 0,040 45 P_2 = 2,4 P_3 = 0,055

  This selects the in IEC 61966-2 (sRGB) nonlinearity and maps sRGB gamma corrected samples into a linear gamma space.

  NOTE The entries of the Floating Point Lookup Table box have to be rescaled compared to the entries found in the bga tag of ISO/IEC 18477-2. The former have a nominal range of [0,653 55] while the latter have a range of [0,1]. Thus, entries of the bga tag have to be multiplied by 655 35 before using them as entries in the Floating Point Table Lookup Table box referenced by the Base Nonlinear Point Transformation box.

  No secondary base transformation box shall be present.

— The Colour Transformation box shall be either absent, or shall have a Xt value of 1 or larger than 4, i.e. select either the identity transformation or a free-form transformation. The free form transformation shall point to a Floating Point Linear Transformation box.

— The Prescaling Transformation box shall be present if and only if the number of components is 3.

— The Prescaling Nonlinear Point Transformation boxes shall be present and shall select a Parametric Curve box of curve type t = 5 (linear ramp). Parameters shall be selected as follows

  e = 1 P_1 = noise floor value P_2 = noise floor value + 1

  This choice of parameters ensures that the prescaling value ν is computed as ν = L + Nf, where L is the luminance of the base image and Nf is the noise floor value specified in ISO/IEC 18477-2. R_w is here always 16.

— The Residual Nonlinear Point Transformation Specification box shall refer to Parametric Curve boxes with box types t = 5 (linear ramp). The parameters shall be selected as follows:

  e = 0 $P_1 = P_2 = 1/2$ for component 0 (residual luminance)

  e = 1 $P_1 = cb0 + 1/2$ $P_2 = cb1 + 1/2$ for component 1 (blue-yellow)

  e = 1 $P_1 = cr0 + 1/2$ $P_2 = cr1 + 1/2$ for component 2 (red-green)

  This choice of parameters implements the same scaling as in ISO/IEC 18477-2 and adds a relative offset of 2^{Rt-1}. This offset is then again subtracted by the Residual Transformation chroma offset. The choice $P_1 = P_2 = 1/2$ for component 0 creates an offset shift of 2^{Rb+8-1} in the output that is compensated by offset included when merging the residual and base image as indicated in Annex A. According to Annex B, R_w = 8 and R_t = 16.

— The Residual Transformation box shall exist if and only if Nf = 3. It shall signal an Xt value of 2 or larger than 4, selecting either the ICT or a free form decorrelation transformation.

— The Intermediate Residual Nonlinear Point Transformation box and the Residual Colour Transformation box shall not be present.

— The Postscaling Nonlinear Transformation box shall be present and shall either point to a Parametric Curve box of curve type t = 6 (exponential) or to a Floating Point Lookup box. In the former case, parameters shall be selected as follows:

$e = 1 \qquad P_1 = \ln 0 \qquad P_2 = \ln 1 \qquad P_3 = 1 \quad P_4 = 0$

This choice of parameters scales the residual luma value according to the exponential map as in ISO/IEC 18477-2. According to Annex B, $R_W = 8$ in the above formula.

— The Secondary Base Nonlinear Point Transformation box shall not be present and inversely decorrelated prediction errors Ω_i are identical to the final error residuals Q_i.

— The Secondary Residual Nonlinear Point Transformation box shall not be present and inversely decorrelated prediction errors Ω_i are identical to the final error residuals Q_i.

— The Output Conversion box shall have the flag values Ol = 1, Oc = 0 and Ce = 0 indicating that an output conversion is active. The to_i values of the Output Conversion box shall point to three identical or one Parametric Curve box of curve type t = 5. This Parametric Curve box shall contain the following parameters:

$e = 1 \qquad P_1 = 0 \qquad P_2 = 1/(2^{8+Rb}-1)$

This choice of the output conversion scales the output from the nominal range of 2^{Rb+8} to a unit range.

Suggestions on how to compute the parameters from an LDR, HDR image pair instead of the parameters from ISO/IEC 18477-2 can be found in Annex F.

E.3 Profile B

Profile B implements a simpler HDR decoding process than profile A requiring less scaling steps. Figure E.2 gives a high-level overview on this profile. The profile indicator of this profile, i.e. the entry CL_i in the compatibility list of the File Type box, shall be 0x78727864, ASCII encoding for "xrxd".

Profile B decomposes an HDR image into a base layer and an HDR residual layer. The LDR input image has to be represented in a linear space (no gamma correction) or should be linearized at encoding level. That is, sample values have to have sample values proportional to physical luminance. The base is either an exposed and clamped or a tone mapped version of the original floating point HDR image. This codestream will be the backwards compatible part and will be accessible by all legacy decoders. The residual image contains the fractional part of the tone mapped LDR image divided by the original HDR image carried out channel by channel in RGB space, and converted to YCrCb. The residual data is also encoded with an Rec. ITU-T T.81 | ISO/IEC 10918-1 encoder using only the baseline or extended Huffman or progressive 8 bit modes, and the Number of Components signaled in the residual frame header shall be identical to the number of components in the base image.

JPEG XT compliant decoders will combine the two as described below to recover the HDR image. The details of this merging process are outlined below.

Figure E.2 illustrates the functionality of a profile B decoder.

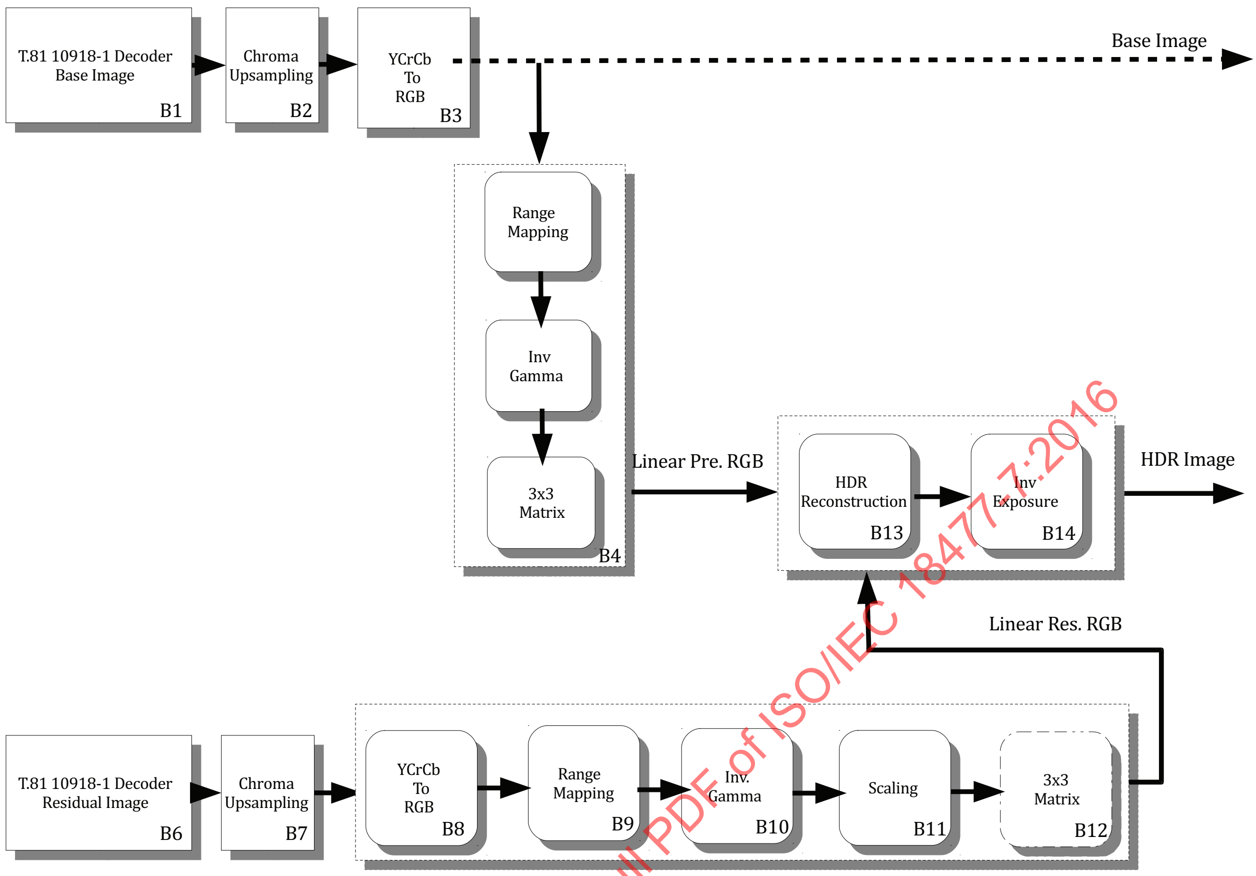


Figure E.2 — High level overview of the decoding process of a profile B compliant decoder

The following constraints apply to a profile B compliant decoder.

— The residual image bit precision P shall be constrained to 8.

— The number of components of the residual image in the Residual Code Stream box shall be identical to the number of components in the legacy codestream.

— The Refinement Specification box (see ISO/IEC 18477-3:2015, Annex D) shall be absent. This implies that no refinement scans are present and no Refinement Data or Residual Refinement boxes are present.

— All Parametric Curve boxes and Floating Point Transformation boxes shall appear as subboxes of the Merging Specification boxes. Parametric Curve boxes and Floating Point Transformation boxes shall not appear at top-level of the file.

— The Prescaling Nonlinear Point Transformation box and the Postscaling Nonlinear Point Transformation box shall be absent. That is, the prescaling factor ν and the postscaling factor μ shall both be 1.

— The Xt value of the Base Transformation box shall be either 2 or larger than 4. That is, the base transformation shall either be the ICT or a free-form transformation.

— The Base Nonlinear Point transformation box shall point to a single Floating Point Lookup box or three identical or a single Parametric curve box selecting a curve type of t = 4. Parameters for the Parametric curve box are constrained as follows:

$e = 1$ $P_1 = P_3 = 0$ $P_2 = ldr\_gamma$, or

$e = 1$ $P_1 = 0{,}040\ 45$, $P_2 = 2.4$, $P_3 = 0{,}055$

This selects an inverse gamma correction without knee value or sRGB nonlinearity

— The Colour Transformation box shall either be absent or specify a free-form transformation. The free form transformation shall be defined by a Floating Point Linear Transformation box.

— The Secondary Base Nonlinear Point Transformation box shall point to three identical or one Parametric Curve box selecting a parametric curve of curve type t = 7. The parameters of this map shall be as follows.

e = 1 $P_1 = 1$ $P_2 = 1$ $P_3 = \varepsilon$ $P_4 = 0$

This selects a logarithmic map. Parameter P_3 adds a small offset that avoids taking the logarithm of zero. A suggested value of ε is 10^{-7}.

— The Residual Nonlinear Point transformation box shall be absent.

— The Residual Transformation box shall either select a free-form transformation or have an Xt value of 1 indicating the ICT.

— The Intermediate Residual Transformation box shall point to one or three Parametric Curve boxes of type t = 8. The parameter shall be constraint as follows.

e = 1 P_1 = min P_2 = max P_3 = hdr_gamma

This selection of parameters first applies an inverse gamma correction of the residual image with the exponent *hdr_gamma* and then scales the output, including the output scaling of ISO/IEC 18477-3:2015, Annex C, to the range [min × $(2^{Rb+8} – 1)$, max × $(2^{Rb+8} – 1)$]. If the parameters differ between components, three boxes need to be written; otherwise, a single box is sufficient which is used by all components consistently.

— The Residual Colour Transformation box shall be present if and only if the Colour Transformation box is present. If it is present, its DXt parameter shall be identical to the DXt parameter of the Colour Transformation box, i.e. both boxes shall refer to the same Floating-point Linear Transformation box.

— The Secondary Residual Nonlinear Point Transformation box shall point to three identical or one Parametric curve box selecting a parametric curve of curve type t = 7. The parameters of this map shall be as follows:

e = 1 P_1 = -1,0 P_2 = 1,0 $P_3 = \varepsilon$ P_4 = 327 68/655 35

The negative sign of P_1 computes the negative logarithm instead of the positive logarithm. P_1 then scales the output of the Residual Transformation box to [0,1]. P_3 avoids a pole by normalizing the output, a suggested value for ε is 10^{-7}. P_4 corrects for the subtraction of the neutral value 2^{8+Rb-1} present in step B9 of the decoding procedure described in Annex A.

— The Output Conversion box shall signal the flag values Oc = 0 Ce = 0 Ol = 1. This enables the selection of an output nonlinearity. The to_i values of the Output Conversion box shall point to three identical or one Parametric Curve box of curve type t = 6. This Parametric Curve box shall contain the following parameters:

e = 1 $P_1 = 0$ $P_2 = 1/(2^{Rb+8}–1)$ P_3 = exposure-value $P_4 = 0$

The exposure value denotes here the choice of an exposure for the reconstructed image.

NOTE The parameter constraints of the Secondary Base, Intermediate and Residual Nonlinear Point Transformation box and the selection of the Output Conversion enforces that the merging process between the LDR base image and the residual image is given by the following formula:

$$\text{expval} \times \exp\left(\mu \times \left(\mathrm{H_i} + \mathrm{Q_i} - 2^{\mathrm{Rb+8-1}}\right)\right) = \text{expval} \times \exp\left(\log\left(\Phi_\mathrm{i} \Big/ \left(2^{\mathrm{Rb+8}} - 1\right)\right) - \log \Omega \Big/ \left(2^{\mathrm{Rb+8-1}}\right) + \varepsilon\right) + 2^{\mathrm{Rb+8-1}} - 2^{\mathrm{Rb+8-1}} = \text{expval} \times \varphi_\mathrm{i} \Big/ \left(\omega_\mathrm{i} + \varepsilon\right)$$

with $\omega_i = \Omega_i/(2^{Rb+8}-1)$ the residual and $\varphi_i = \Phi_i/(2^{Rb+8}-1)$ the base image sample values scaled to [0,1]. The value of μ is here always 1 due to the absence of the postscaling transformation.

A decoder conforming to profile B does not need to implement the logarithmic and exponential map of the Secondary Base and Residual Nonlinear Point Transformation and the Output Transformation, but may instead apply the right hand side of the above formula directly without going through all the intermediate steps. This will not only lower the algorithmic complexity, but also improve the accuracy of the algorithm. Furthermore, note that the output of the Secondary Base/Residual Nonlinear Point transformation box is scaled by the algorithm specified in ISO/IEC 18477-3:2015, Annex C by a factor of $2^{Rb + 8} - 1$ which is again removed by the input scaling of the Output Conversion box. That is, practical implementations may want to skip the output scaling of the Secondary Nonlinear Transformations and the input scaling of the Output Conversion box. Output scaling by the algorithm of Annex C in the Output Conversion box is also un-done by the choice of P_3 of the Parametric Curve box referenced by the Output Conversion box.

With these simplifications in place, decoding consists of the following steps.

— Scale the output of the LDR reconstruction process O_i from step B3 to the range [0,1] according to ISO/IEC 18477-3:2015, Annex C. In Figure E.2, this step is performed in step B4.

— Perform the pointwise nonlinearity transformation in step B4. This is part of the nonlinear point transformation; the outputs are the predicted high dynamic range sample values J_i. In profile B, the nonlinearity is either given by a lookup table, a sRGB-type of nonlinearity or a power law with the exponent taken from parameter P_2 of the Parametric Curve box pointed to by Base Nonlinear Point Transformation box.

— Since the chroma scaling factor ν is always 1 in profile B and the Residual Nonlinear Point Transformation is always the identity, the values Π_i are identical to the reconstructed residual sample values P_i scaled to the range $[0,2^{Rb}\text{-}1]$.

— In step B8, a YCbCr to RGB transformation (ICT) is applied to the reconstructed residual data P_i by applying a 3 × 3 matrix transformation and a level shift, generating the output Ψ_i. Due to the absence of a pre-scaling, the factor ν is 1 and Π_i equals P_i.

— Step B9 scales the values Ψ_i input to [0,1]. This step is implicit in the input and output scaling of the **Secondary Residual Nonlinear Point Transformation** and **Intermediate Residual Nonlinear Point Transformation**, and is specified in ISO/IEC 18477-3:2015, Annex C.

— In step B10, an inverse gamma correction is applied to the residual sample values Ψ_i giving the gamma-corrected decorrelated samples T_i. This operation is implemented by the **Intermediate Nonlinear Point Transformation**. The exponent comes from the P_3 value of the parametric Curve box pointed to by the Intermediate Nonlinear Point Transformation Specification box, the output scaling from the parameter P_1 and P_2. The exponent is signaled as value P_3 in the Parametric Curve box referenced by the Intermediate Nonlinear Point Transformation box.

— In step B11, the output of the inverse gamma mapping is scaled to the range [min, max] by multiplying the gamma-corrected values with (*min-max*) and adding an offset shift *min*. The operation is performed along with the inverse gamma correction in step B10 by the Parametric Curve box. The parameters *min* and *max* are signaled as parameters P_1 and P_2 of the Parametric Curve box signaled by the Intermediate Nonlinear Point Transformation box. Note that each component may refer to a different Parametric Curve box, thus exponent values and min/max values may differ between components.

— In step B12, a 3 × 3 matrix implementing a colour transformation is applied to the sample values T_i generating the values Ω_i. This transformation is defined by the Residual Colour Transformation box if it is present. If this box is not present, Ω_i is identical to T_i.

— In step B4, a 3 × 3 matrix implementing a colour transformation is applied to the base image sample values J_i generating the values Φ_i. This transformation is defined by the Colour Transformation box if it is present. If it is not present, Φ_i is identical to J_i.

— In steps B13 and B14, the residual and base image are merged by a division operation, that is, the output values are computed as $\text{expval} \times \varphi_i / (\omega_i + \varepsilon)$ where φ_i and ω_i are scaled versions of Φ_i and Ω_i. Note that this scaling and division process is implicitly signaled in the Merging Specification box by defining logarithmic nonlinear point transformations as Secondary Base Nonlinear Transformation and Secondary Residual Nonlinear Transformation and an output transformation that defines an exponential map. The value of expval comes from parameter P_3 of the Parametric Curve box referenced by the Output Conversion box.

E.4 Profile C

Profile C implements a HDR decoding process that is only a minor modification of the intermediate range compression specified in ISO/IEC 18477-6. Profile C decoders can be entirely implemented in integer logic and do not require floating point operations. Figure E.3 gives a high-level overview on this profile. The profile indicator of this profile, i.e. the entry CL_i in the compatibility list of the File Type box shall be 0x78726164, ASCII encoding for “xrad”.

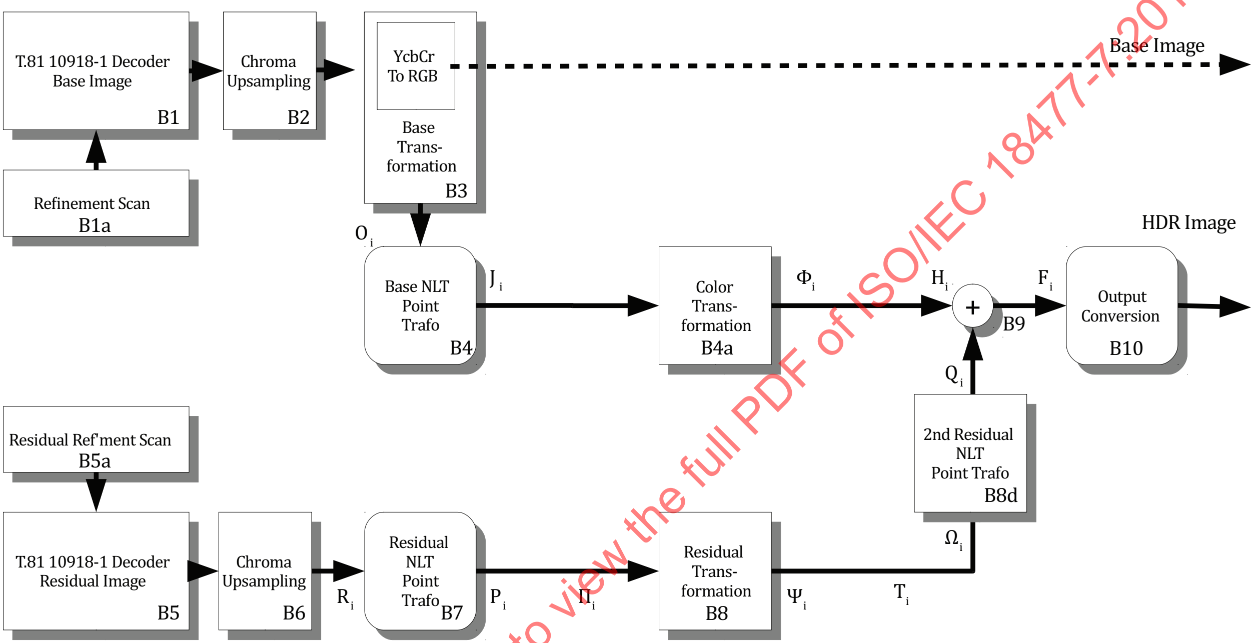


Figure E.3 — High level overview of the decoding process of a profile C compliant decoder

The following constraints apply to a profile C compliant decoder.

— The number of components of the residual image in the Residual Code Stream box shall be identical to the number of components in the legacy codestream.

— The Prescaling Nonlinear Point Transformation box and the Postscaling Nonlinear Point Transformation box shall be absent. That is, the prescaling factor ν and the postscaling factor μ shall both be 1.

— The Base Transformation box may be present and if so shall either select the Identity, the ICT or the free form transformation. In the latter case, Xt shall select an Integer Linear Transformation box.

— The Base Nonlinear Point Transformation box may be present and if so, shall select only Integer Lookup Table boxes.

— The Colour Transformation box may be present and if so, shall select an Integer Linear Transformation box.

— The Residual Nonlinear Point Transformation boxes may be present, but if so, may only reference a Parametric Curve box with t = 2 and e = 0. That is, simple linear scaling is applied to the reconstructed residual sample values.