ISO/IEC FDIS 21794-6~~:2024(en)~~

ISO/IEC JTC 1/SC 29

Secretariat: JISC

Date: 2025-~~01-17~~04-02

Information technology — Plenoptic image coding system (JPEG Pleno) —

Part 6:
Learning-based point cloud coding

FDIS stage

Formatted: Left

Style Definition

Formatted: zzCover large

Formatted: Right: 1.5 cm, Bottom: 1 cm, Gutter: 0 cm, Section start: New page, Header distance from edge: 1.27 cm, Footer distance from edge: 1.27 cm

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Regular, Font: Bold

Formatted: Cover Title_A2

Formatted: Regular, Font: Bold

Formatted: Left

Formatted: HeaderCentered

Formatted: Font: Bold

Formatted: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Formatted: Default Paragraph Font

Formatted: French (Switzerland)

Formatted: French (Switzerland)

Formatted: French (Switzerland)

Formatted: FooterPageRomanNumber

Contents

Formatted: Font: Bold, English (United Kingdom)
Formatted: HeaderCentered, Left
Formatted: Font: Bold
Formatted: Font: Bold, English (United Kingdom)
Formatted: Font: 10 pt
Formatted: FooterCentered, Left, Space Before: 0 pt, Tab stops: Not at 17.2 cm
Formatted: Font: 10 pt
Formatted: Font: 11 pt
Formatted: FooterPageRomanNumber, Left, Space After: 0 pt, Tab stops: Not at 17.2 cm

Formatted: HeaderCentered
Formatted: Font: Bold
Formatted: FooterPageRomanNumber

Formatted: Font: Bold, English (United Kingdom)
Formatted: HeaderCentered, Left
Formatted: Font: Bold
Formatted: Font: Bold, English (United Kingdom)
Formatted: Font: 10 pt
Formatted: FooterCentered, Left, Space Before: 0 pt, Tab stops: Not at 17.2 cm
Formatted: Font: 10 pt
Formatted: Font: 11 pt
Formatted: FooterPageRomanNumber, Left, Space After: 0 pt, Tab stops: Not at 17.2 cm

Foreword

ISO (the International Organization for Standardization) ~~is a~~ and IEC (the International Electrotechnical Commission) form the specialized system for worldwide ~~federation of national standards~~standardization. National bodies ~~(~~that are members of ISO ~~member bodies). The work~~or IEC participate in the development of ~~preparing~~ International Standards ~~is normally carried out~~ through ~~ISO~~ technical committees~~. Each member body interested in a subject for which a technical committee has been~~ established ~~has the right to be represented on that committee. International~~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. ~~ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.~~

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 ~~ISO documents~~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~~ 2 (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

~~Attention is drawn~~ISO and IEC draw attention to the possibility that ~~some of~~ the ~~elements~~implementation of this document may ~~be~~involve the ~~subject~~use of (a) patent~~ rights. ISO~~(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. 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 of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see ~~www.iso.org/iso/foreword.html~~www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

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

A list of all parts in the ISO/IEC 21794 series can be found on the ISO ~~website~~and IEC websites.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at ~~www.iso.org/members.html~~www.iso.org/members.html and www.iec.ch/national-committees.

Formatted: HeaderCentered
Formatted: Font: Bold
Formatted: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: FooterPageRomanNumber

Introduction

This ~~standard~~document is part of a series of standards for a system known as JPEG Pleno. This set of standards facilitates the capture, representation, exchange and visualization of plenoptic imaging modalities. A plenoptic image modality can be a light field, point cloud or hologram, which are sampled representations of the plenoptic function in the form of, respectively, a vector function that represents the radiance of a discretized set of light rays, a collection of points with position and attribute information, or a complex wavefront. The plenoptic function describes the radiance in time and in space obtained by positioning a pinhole camera at every viewpoint in 3D spatial coordinates, every viewing angle and every wavelength, resulting in a 7D function.

JPEG Pleno specifies tools for coding these modalities while providing advanced functionality at system level, such as support for data and metadata manipulation, editing, random access and interaction, protection of privacy and ownership rights.

The scope of ~~Part 6 of the standard~~this document is the specification of a learning-based coding standard for point clouds and associated attributes, offering a single-stream, compact compressed domain representation, supporting advanced flexible data access functionalities. In this context, learning-based refers to the use of machine learning technologies to learn an optimal compressed domain representation from supplied training data.

Information technology — Plenoptic image coding system (JPEG Pleno) ~~— Part 6: Learning-based point cloud coding —~~

Part 6: Learning-based point cloud coding

1 Scope

This ~~standard~~document defines the JPEG Pleno framework for learning-based point cloud coding.

This ~~standard~~document is applicable to interactive human visualization, with competitive compression efficiency compared to state of the art point cloud coding solutions in common use, and effective performance for 3D processing and machine-related computer vision tasks, and has the goal of supporting a royalty-free baseline.

This ~~standard~~document specifies a coded codestream format for storage of point clouds. ~~This standard also~~It provides information on the encoding tools. ~~The standard~~It also defines extensions to the JPEG Pleno File Format and associated metadata descriptors that are specific to point cloud modalities.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 6048-1[1]

~~ISO/IEC 21794-1, :—~~[1], Information technology — ~~Plenoptic~~*JPEG AI learning-based* image coding system ~~*(JPEG Pleno)*~~ — Part 1: ~~*Framework*~~*Core coding system*

~~REC. ITU-T T.800 |~~ ISO/IEC 15444-1[2], Information technology — JPEG 2000 image coding system — Part 1: Core coding system

~~REC. ITU-T T.801 |~~ ISO/IEC 15444-2[3], Information technology — JPEG 2000 image coding system — Part 2: Extensions

ISO/IEC 21794-1, *Information technology — Plenoptic image coding system (JPEG Pleno) — Part 1: Framework*

ISO/IEC 21794-2, Information technology — Plenoptic image coding system (JPEG Pleno) — Part 2: Light field coding

ISO/IEC 21794-3, Information technology — Plenoptic image coding system (JPEG Pleno) — Part 3: Conformance testing

~~[1] Under preparation. Stage at the time of publication: ISO/IEC DIS 6048-1:2024.~~
[1] Under preparation. Stage at the time of publication: ISO/IEC PRF 6048-1:2025.
[2] Similar to REC. ITU-T T.800 | ISO/IEC 15444-1
[3] Similar to REC. ITU-T T.801 | ISO/IEC 15444-2

Formatted: Left
Formatted: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers, Tab stops: Not at 0.7 cm + 1.4 cm + 2.1 cm + 2.8 cm + 3.5 cm + 4.2 cm + 4.9 cm + 5.6 cm + 6.3 cm + 7 cm
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers, Tab stops: Not at 0.7 cm + 1.4 cm + 2.1 cm + 2.8 cm + 3.5 cm + 4.2 cm + 4.9 cm + 5.6 cm + 6.3 cm + 7 cm
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Footer, Left, Space After: 0 pt, Tab stops: Not at 17.2 cm

ISO/IEC 21794-4, Information technology — Plenoptic image coding system (JPEG Pleno) — Part 4: Reference software

ISO/IEC 21794-5, Information technology — Plenoptic image coding system (JPEG Pleno) — Part 5: Holography

ISO/IEC 23090-5:2023, Information technology — Coded representation of immersive media — Part 5: Visual volumetric video-based coding (V3C) and video-based point cloud compression (V-PCC)

ISO/IEC 23090-9:2023, Information technology — Coded representation of immersive media — Part 9: Geometry-based point cloud compression

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

3 Terms, and definitions and abbreviations

For the purposes of this document, the terms and definitions given in ISO/IEC 21794-1, ISO/IEC 21794-2, ISO/IEC 21794-3, ISO/IEC 21794-4, ISO/IEC 21794-5, ISO/IEC 23090-5:2023, ISO/IEC 23090-9:2023, ISO/IEC 6048-1, REC. ITU-T T.800 | ISO/IEC 15444-1, REC. ITU-T T.801 | ISO/IEC 15444-2, ISO/IEC 60559 and the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obphttps://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/https://www.electropedia.org/

3.1 Terms and definitions

3.1 3.1.1
point
fundamental element of a point cloud comprising a position specified as 3D spatial coordinates and colour attributes

3.2 3.1.2
point cloud
unordered list of points

3.3 3.1.3
neural network layer
tensor operation that contains trainable parameters which receives and outputs a tensor

3.4 3.1.4
neural network module
set of neural network layers

3.5 3.1.5
neural network model
specified sequence of neural network modules (also called architecture) and corresponding trained parameters

3.6 3.1.6
trainable parameters
parameters of a neural network layer whose values require a training process based on input ground truth data to be set

3.7
~~**3.1.7**~~
trained parameters
parameters of a neural network model whose values have been set by a training process based on input ground truth data

3.8
~~**3.1.8**~~
dense tensor
representation of a 3D block as a regular array with four dimensions: horizontal, vertical, depth~~,~~ and channel dimension

3.9
~~**3.1.9**~~
sparse tensor
~~as opposed to a dense tensor, a sparse tensor is a~~ representation of a 3D block, where only non-zero elements are represented as a set of indices (or coordinates) C and associated values (or features) F

Note 1 to entry: The set of coordinates C is represented as a matrix $C \in Z^{P\times 3}$ and the associated features F are represented as a matrix $F \in R^{P\times N}$, where P is the number of non-zero elements and N is the number of channels. The remaining elements of a sparse tensor are zeros.

~~3.1.10~~ Note 2 to entry: Opposed to a dense tensor.

3.10
latent tensor
intermediate representation of point cloud data during encoding or decoding processes, as a sparse tensor

3.11
~~**3.1.11**~~
standard deviations tensor
tensor of unsigned 16 bits integers, used for entropy coding, denoted as σ

3.12
~~**3.1.12**~~
concatenation of sparse tensors
process where the features of two sparse tensors are concatenated

3.13
~~**3.1.13**~~
element-wise addition of sparse tensors
process where the features of two sparse tensors are added element-wise

3.14
~~**3.1.14**~~
sparse convolution layer
three-dimensional sparse convolution denoted as $SpConv\,(K_{ver} \times K_{hor} \times K_{dep}, N_{in}, N_{out}, s\downarrow)$

3.15
3.1.15
transposed sparse convolution layer
three-dimensional transposed sparse convolution denoted as $TSpConv(K_{ver} \times K_{hor} \times K_{dep}, N_{in}, N_{out}, s \uparrow)$ $TSpConv(K_{ver} \times K_{hor} \times K_{dep}, N_{in}, N_{out}, s \uparrow)$

3.16 3.1.16
generative transposed sparse convolution layer
three-dimensional generative transposed sparse convolution denoted as $GTSpConv(K_{ver} \times K_{hor} \times K_{dep}, N_{in}, N_{out}, s \uparrow)$ $GTSpConv(K_{ver} \times K_{hor} \times K_{dep}, N_{in}, N_{out}, s \uparrow)$

3.17
3.1.17
quantized sparse convolution layer
three-dimensional quantized sparse convolution is denoted as $qSpConv(K_{ver} \times K_{hor} \times K_{dep}, N_{in}, N_{out}, s \downarrow, d, p)$ $qSpConv(K_{ver} \times K_{hor} \times K_{dep}, N_{in}, N_{out}, s \downarrow, d, p)$

3.18 3.1.18
quantized generative transposed sparse convolution layer
three-dimensional quantized generative transposed sparse convolution is denoted as $qGTSpConv(K_{ver} \times K_{hor} \times K_{dep}, N_{in}, N_{out}, s \uparrow, d, p)$ $qGTSpConv(K_{ver} \times K_{hor} \times K_{dep}, N_{in}, N_{out}, s \uparrow, d, p)$

3.19
3.1.19
matrix multiplication
matrix multiplication (denoted as $\times$) $\times$) receives two two-dimensional arrays $input_1[h_{out}, L]$ $input_1[h_{out}, L]$ and $input_2[L, w_{out}]$. $input_2[L, w_{out}]$.

Note 1 to entry: This module produces a two-dimensional array $output$ $output$ of size $[h_{out}, w_{out}]$: $[h_{out}, w_{out}]$.

For $i = 0, \ldots, h_{out} - 1$ $= 0, \ldots, h_{out} - 1$ and $j = 0, \ldots, w_{out} - 1$: $j = 0, \ldots, w_{out} - 1$:

$$output[i,j] = \sum_{l=0}^{L-1} input_1[i,l] \cdot input_2[l,j].$$

3.1.20

$$output[i,j] = \sum_{l=0}^{L-1} input_1[i,l] \cdot input_2[l,j].$$

3.20
rectified linear unit
rectified linear unit is denoted as $ReLU$. $ReLU$

Note 1 to entry: This element-wise function is defined as:

$$ReLU(x) = \begin{cases} x, & if\ x \geq 0, \\ 0, & otherwise. \end{cases}$$

Formatted: Font: Bold
Formatted: HeaderCentered
Formatted: Term(s), Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: TermNum2, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: Term(s), Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: TermNum2, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: Term(s), Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: Note, Adjust space between Latin and Asian text, Adjust space between Asian text and
Formatted: Term(s), Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: Note, Adjust space between Latin and Asian text, Adjust space between Asian text and
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: FooterPageNumber, Space After: 0 pt, Line spacing: single
Formatted: Font: 11 pt