ICS 35.040
CCS L 70

NATIONAL STANDARD OF THE PEOPLE'S REPUBLIC OF CHINA

GB/T 33475.7-2024

Information technology - High efficiency multimedia coding - Part 7: Image format

信息技术 高效多媒体编码 第7部分:图片文件格式

Issued on: May 28, 2024 **Implemented on: December 1, 2024**

Issued by: State Administration for Market Regulation;
Standardization Administration of the PRC

Table of Contents

Patent statement Patent statement in the introduction

The introduction states that a claim of conformity to the document may involve the use of patents relating to digital video coding and decoding technology in 8.1.2, and tabulates six of them by number, patent application number, title and holder; all six are held by Tencent Technology (Shenzhen) Company Limited and are titled as methods, systems, devices or intelligent terminals for processing image files. The issuing body takes no position on the authenticity, validity or extent of the patents. The holder has assured the issuing body that it is willing to negotiate licences with any applicant on reasonable and non-discriminatory terms and conditions, and its statement is on file with the issuing body, with a contact person, postal address, postcode, email address, telephone, fax and web address given for the digital audio and video coding technology standards working group. The introduction adds that content other than the patents listed may still involve patents that the issuing body is not responsible for identifying.

2 Normative references

The document cites GB/T 33475.1-2019 (Part 1 of the series, systems), GB/T 33475.2-2024 (Part 2, video) and ITU-R BT.601-2001 (studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios).

3 Terms and definitions

3 The terms defined in GB/T 33475.2-2024 apply, together with the terms defined here, most of which are also taken from that part. They are: reserved, certain values of syntax elements, noted as kept for future extension of the document; component, one of the three sample matrices of an image, luma and the two chroma, or a single sample within a matrix; parse process, the process of obtaining syntax elements from the bitstream; profile, a subset of the syntax, semantics and algorithms; luma, the signal representing the brightness of an image; chroma, the sample matrix or a single sample of either of the two colour difference signals Cr and Cb; RGB, the sample matrices or single samples of the three colour channels red, green and blue; alpha channel, used to record the transparency data of an image and to define transparent, opaque and semi-transparent areas, noted as also being called the grey scale channel; image sequence, the highest level syntax structure of the coded bitstream, noted as taking in one or more consecutive coded images; syntax element, the result of parsing a data unit in the bitstream; frame, the representation of the spatial information of a video signal made up of one luma sample matrix and two chroma sample matrices; picture frame, the representation of the spatial information of one image signal made up of the sample matrices of the three colour channels red, green and blue; byte, a string of eight bits; start code, a binary code word 32 bits long, noted as being unique in the whole bitstream, one of them marking the beginning of a syntax structure; field, made up of the same lines of the three sample matrices that make up a frame; picture, the set of texture or depth information made up of one frame or one field; YUV, a colour coding method in which Y stands for brightness and U and V are chroma; and image file, the set of related information made up of one frame or one field of image.

4 Abbreviated terms

4 AVS, audio and video coding standard; RGBA, the red green blue alpha colour space, that is the three primary colours plus a transparency channel; TPG, tiny portable graphics.

5 Conventions

5 The description of the bitstream syntax, of the parse process and of the decoding process is kept consistent with GB/T 33475.2-2024. The pseudocode used for the syntax and semantics of Clause 8 is written in the C language.

6 System framework

6 Figure 1 shows the framework of the image coding and decoding system, made up of an encoder and a decoder, following GB/T 33475.1-2019. The encoder takes source image data in and puts compressed image data out; the decoder takes compressed image data in and puts reconstructed image data out. The system is built on a video codec, the image format being packaged on top of the video coding so that the features of a still or moving image are obtained.

6 Figure 2 shows the coding and decoding flow. The encoder input is source RGB or RGBA data. RGB data is converted to YUV data, the YUV data is coded to give the colour frame data, and the colour frame data is packaged in the image file format to give a complete compressed image file. For RGBA data the encoder first separates the RGB data from the transparency data and codes them separately: the RGB data is converted to YUV and coded intra or inter to give the colour frame data, while the transparency data is set as the Y component of YUV data and coded intra or inter through the Y channel of the encoder to give the alpha channel frame data. The colour frame data and the alpha channel frame data are then merged and packaged, the packaging taking in the image header information segment, the alpha channel frame header and the colour frame header, to give a complete compressed image file.

6 The decoder input is the compressed image file and its output is RGB or RGBA data. The decoder first decodes the colour frame data to YUV data and converts it to RGB data. Where there is no alpha channel frame data the output is RGB data; where there is, the alpha channel frame data is decoded to YUV data, its Y component is taken as the transparency data and merged with the RGB data, and the output is RGBA data.

7 Structure of the coded bitstream

7.1 One image file is one image sequence, and the image sequence is the highest level syntax structure of the bitstream. Figure 3 shows the data structure of a complete image sequence. It is made up of the image header information segment, the colour frame data segment and the alpha channel frame data segment, the last being present only where the image has alpha channel data.

7.1 The image header information segment is the opening segment of the image sequence and has four parts: the image sequence header, the image feature information data segment, the image supplementary information data segment and the user defined information data segment. The image sequence header is mainly the identifier of the image file; the image feature information data segment describes key features such as the width, height and transparency of the image; and the image supplementary information data segment mainly describes EXIF information such as the time and place at which the image was taken. A string of coded image data follows the image header information segment.

7.1 The colour frame data segment is the bitstream produced by coding the RGB data of the image and has two parts, the colour frame header and the colour frame data. The colour frame header carries the feature information of each image in the sequence, every image frame having one. The colour frame data is the coded bitstream obtained by video coding the RGB information of each frame, that is, a bitstream conforming to GB/T 33475.2-2024. The alpha channel frame data segment is the bitstream produced by coding the transparency data and likewise has two parts, the alpha channel frame header, which describes the features of the alpha channel data frame and which every alpha channel

frame carries, and the alpha channel frame data, the coded bitstream obtained by video coding the transparency information of each frame.

7.2 An image file is made up of one image, for a still image, or of several, for a moving image. The document lays down that the images are frame images. One image is made up of one colour frame, or of one colour frame plus one alpha channel frame: an image without transparency has only the colour frame, an image with transparency has both. The coded data of one image runs from its colour frame header to the next colour frame header or to the end of the file. Decoding an image frame is made up of parsing the colour frame header and decoding the colour frame data; decoding an alpha channel frame is made up of parsing the alpha channel frame header and decoding the alpha channel frame data. The data format used for coding may be YUV data, in the 4:4:4, 4:2:0 or 4:0:0 format, or the RGB data may be coded directly.

8 Syntax and semantics of the bitstream

8.1.1 A start code is a particular string of characters. The document lays down the start codes of the image header information segment and of the image frame header data, the latter being of two kinds, the colour frame start code and the alpha channel frame start code; the start codes of the colour frame data and of the alpha channel frame data follow GB/T 33475.2-2024. A start code is made up of a start code prefix and a start code value; the prefix is the bit string 000000000000000000000001 and the start code is byte aligned. The start code value is an integer identifying the type. Table 1 gives the five types and their hexadecimal values: image feature information start code, B9; image supplementary information start code, BA; colour frame start code, BB; alpha channel frame start code, BC; and user defined information start code, BD.

8.1.2 The syntax is given in seven tables of C-style pseudocode, each row carrying the syntax element and its descriptor. Table 2 defines the image sequence: it calls the image sequence header, the sequence header and the extension and user data, then loops on the next 32 bits to call the colour frame header, the alpha channel frame header, the intra picture header or the inter picture header, followed by the extension and user data and the picture data, and closes on the video sequence end code or the video edit code, each written as a 32-bit fixed pattern. Table 3 defines the image sequence header: the image identifier and the codec identifier as 32-bit fixed patterns, the version identifier as an 8-bit unsigned value, then the image feature data, the image supplementary data and the image user data, each called where the next 32 bits match its start code. Table 4 defines the image feature information data segment: the start code, the data length as a 16-bit unsigned value, the alpha flag, the dynamic picture flag, the YUV format on two bits, the lossless flag and the YUV limited range flag as single bits, ten reserved bits, the image width and the image height as 24-bit unsigned values, and, where the dynamic picture flag is set, the number of frames as a 24-bit unsigned value. Table 5 defines the image