 ISO/IEC JTC 1/SC 29/WG 03 N0153

**ISO/IEC JTC 1/SC 29/WG 03  
MPEG Systems   
Convenorship: KATS (Korea, Republic of)**

**Document type:** Output Document

**Title:** Exploration of ISO/IEC 23001-17 Carriage of Uncompressed Video in ISOBMFF

**Status:** Approved

**Date of document:** 2021-01-22

**Source:** ISO/IEC JTC 1/SC 29/WG 03

**No. of pages:** 8

**Email of Convenor:** young.L @ samsung . com

**Committee URL:** <https://isotc.iso.org/livelink/livelink/open/jtc1sc29wg3>

**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 03 MPEG SYSTEMS**

**ISO/IEC JTC 1/SC 29/WG 03 N0153**

**January 2021, Virtual**

|  |  |
| --- | --- |
| **Title** | **Exploration of ISO/IEC 23001-17 Carriage of Uncompressed Video in ISOBMFF** |
| **Source** | **WG 03, MPEG Systems** |
| **Status** | **Approved** |
| **Serial Number** | **19985** |

Introduction

At MPEG#122, Telecom ParisTech presented an input contribution on uncompressed audio and video storage in ISOBMFF where several existing practices for storage were summarized and several missing features in ISOBMFF were identified [1]. While the audio aspects have been standardized as ISO/IEC 23003-5, the carriage of uncompressed video in ISOBMFF has not been resolved. It was pointed out that the current state of uncompressed video carriage in ISOBMFF has a lack of documentation and features. Furthermore, the carriage of uncompressed video should reuse as many existing tools as possible, but it is unclear what the state of the various specifications and formats is today.

At MPEG#132, Telecom ParisTech presented a second contribution on this topic, introducing a possible approach for carriage of uncompressed video [2], while Qualcomm presented a set of requirements for uncompressed video carriage based on the related activity in 3GPP [3]. The File Format group has agreed that there is a need for unified solution for carriage of uncompressed video in ISOBMFF in the industry and requested a subdivision of ISO/IEC 23001 into a next part 17.

At MPEG#133, Canon Research Centre France presented further use cases and requirements for the carriage of uncompressed video in ISOBMFF based on industrial cameras [4] and the group has agreed to start the exploration activity on 23001-17. This document can be considered as the starting point of this exploration, in which we try to collect best industry practices for carriage of uncompressed video, identify a set of features / requirements which should be addressed in the new MPEG specification, and identify any possible overlaps and gaps with current implementations.

Existing code points

In this section we investigate several existing methods for uncompressed video carriage. Please note that the term *uncompressed*, which is used throughout this document, is slightly inaccurate as chroma sub-sampling is also a form of data compression.

Technical Note TN2162

Technical Note [TN2162](https://developer.apple.com/library/archive/technotes/tn2162/_index.html) is originally known as Ice Floe Dispatch 19 is a part of the “Letters from the Ice Floe” from the QuickTime Engineering team which was originally published on December 14th, 1999 [5]. This document is no longer being updated and is stored in the Apple Documentation Archive, but it can be considered as being de facto standard for carriage of uncompressed video in QTFF [6]. Furthermore, several code points defined in TN2162 (‘2vuy’, ‘v210’) are now used in various implementations. For example, Final Cut Pro allows the users to extract the uncompressed version of the master file in two different formats packaged in a QTFF (.mov) file:

* *Uncompressed 8-bit 4:2:2*: which is using the ‘2vuy’ FOURCC
* *Uncompressed 10-bit 4:2:2*: which is using the ‘v210’ FOURCC

A list of Y’CbCr packing types as defined in TN2162 is listed below:

Table 1: Compression types for storing uncompressed Y’CbCr data

|  |  |  |
| --- | --- | --- |
| Compression Type | FourCC | Description |
| k422YpCbCr8CodecType | '2vuy' | 8-bit-per-component 4:2:2 |
| kComponentVideoCodecType | 'yuv2' | 8-bit-per-component 4:2:2 |
| k444YpCbCr8CodecType | 'v308' | 8-bit-per-component 4:4:4 |
| k4444YpCbCrA8CodecType | 'v408' | 8-bit-per-component 4:4:4:4 |
| k422YpCbCr16CodecType | 'v216' | 10,12,14,16-bit-per-component 4:2:2 |
| k444YpCbCr10CodecType | 'v410' | 10-bit-per-component 4:4:4 |
| k422YpCbCr10CodecType | 'v210' | 10-bit-per-component 4:2:2 |

Detailed information on the individual compression types from Table 1 can be found in [TN2162](https://developer.apple.com/library/archive/technotes/tn2162/_index.html). As we will show later, ‘2vuy’ and ‘v210’ are the most supported ones today.

SMPTE

Several SMPTE specifications also define the transport and storage of uncompressed video.

ST 2110

SMPTE ST 2110 is a suite of specifications that defines the RTP-based transport of digital media over IP networks. ST 2110 is mainly intended to be used by broadcasters in live production workflows where quality of the video is more important than the required bandwidth [7]. SDI to IP converters are typically used to provide a transition from SDI to all-IP networks [8].

Part 20 of the ST 2110 defines RTP-based transport of uncompressed video, which specifies the format in which samples of uncompressed video data are sent over IP [9]. In addition, it defines a set of required media type properties which are signaled using the Session Description Protocol including:

* Sampling (color difference sub-sampling structure)
  + YCbCr: 4:4:4 / 4:2:2 / 4:2:0 (BT.601, BT.709, BT.2020, BT.2100)
  + CLYCbCr: 4:4:4 / 4:2:2 / 4:2:0 (BT.2020)
  + ICtCp: 4:4:4 / 4:2:2 / 4:2:0 (BT. 2100)
  + RGB
  + XYZ (Defined in SMPTE ST 428-1)
* Depth
  + 8
  + 10
  + 12
  + 16
  + 16f (as defined in SMPTE ST 2065-1 and ITU-R BT.2100)
* Width and height
* Framerate
* Colorimetry
  + BT601, BT709, BT2020, BT2100
  + ST2065-1, ST2065-3
* Packing mode
* Interlacing
* Transfer characteristics
  + SDR, PQ, HLG, LINEAR, BT2100LINPQ, BT2100LINHLG, ST2065-1, ST428-1, DENSITY
* Range
* Pixel aspect ratio

Material Exchange Format

SMPTE also defines the ST 377-1 “Material Exchange Format (MXF)” [10], a File Format Specification which specifies the data structure for network transport and storage of multimedia data. Mapping of uncompressed pictures into the generic container of MXF are further specified in ST 384M [11]. Annex G of [10] provides further details on picture descriptor properties used in MXF including:

* Frame layout (progressive vs interlaced [4 types of interlacing])
* Sample rate
* Aspect Ratio
* Width and height
* Field dominance (Frame dominance in temporal order of interlaced frame)
* Alpha Transparency
* Transfer characteristic
* Image alignment offset (byte alignment of units)
* Image start/end offset (skip bytes)
* Picture Essence Coding (concept similar to FOURCC), other properties can be used in combination.
* Component bit depth
* Horizontal / Vertical Subsampling
* Color Siting (used to compute subsampled color difference values)
* Alpha sample depth
* Black / White reference level (luma value for black / white reference)
* Color range
* Pixel layout for RGBA only
* Palette for RGBA (reference colors by indexes)
* Scanning direction (8 modes lrtb, lrbt, lrbt, rlbt, tblr, tbrl, btlr, btrl)

MP4RA.org

There are also several ISOBMFF code points already registered at [MP4RA.org](http://mp4ra.org/) that are listed in Table 2 below. However, the linked specification [Digital Voodoo](http://www.digitalvoodoo.net/) is no longer available, so it is not really clear what exactly these 5 code points represent.

Table 2: Uncompressed video code points at mp4ra.org

|  |  |  |  |
| --- | --- | --- | --- |
| Code | Description | Handler | Specification |
| DVOO | Digital Voodoo 8 bit Uncompressed 4:2:2 codec | [Video](http://mp4ra.org/handler.html#vide) | [Digital Voodoo](http://mp4ra.org/specs.html#voodoo) |
| DVOR | Digital Voodoo intermediate raw | [Video](http://mp4ra.org/handler.html#vide) | [Digital Voodoo](http://mp4ra.org/specs.html#voodoo) |
| DVTV | Digital Voodoo intermediate 2vuy | [Video](http://mp4ra.org/handler.html#vide) | [Digital Voodoo](http://mp4ra.org/specs.html#voodoo) |
| DVVT | Digital Voodoo intermediate v210 | [Video](http://mp4ra.org/handler.html#vide) | [Digital Voodoo](http://mp4ra.org/specs.html#voodoo) |
| HD10 | Digital Voodoo 10 bit Uncompressed 4:2:2 HD codec | [Video](http://mp4ra.org/handler.html#vide) | [Digital Voodoo](http://mp4ra.org/specs.html#voodoo) |

Digital Voodoo appears to be an Australian video card manufacturer for professional workflows, but it seems that they are no longer in business as their website is down and their video boards are listed as deprecated on several resources such as [Viz Engine 4.0](https://documentation.vizrt.com/viz-engine-guide/4.0/Deprecated_Video_Boards.html). Nevertheless, some information on the above FOURCC’s could be found:

* ‘DVOO’: is implemented in FFmpeg as AV\_PIX\_FMT\_YUYV422 which is a packed YUV 4:2:2, 16bpp, Y0 Cb Y1 Cr, in addition DVOO is also mentioned [here](https://wiki.multimedia.cx/index.php?title=QuickTime_container).
* ‘DVOR’: according to several [websites](https://abcavi.kibi.ru/fourcc.php?fcc=DVOR&title=DVOR+-+BlueFish444+%28lossless+RGBA%2C+YUV+10-bit%29) is most likely a 16-to-10-bit RGBA
* ‘HD10’: is most likely a 10-bit YUVA (YUV + Alpha)

Formats

Many different format types are defined in a number of different places, here is a list of some of them:

* Microsoft
  + [Uncompressed RGB Formats](https://docs.microsoft.com/en-us/windows/win32/medfound/video-subtype-guids#uncompressed-rgb-formats)
  + [Recommended 8-bit YUV Formats](https://docs.microsoft.com/en-us/windows/win32/medfound/recommended-8-bit-yuv-formats-for-video-rendering)
  + [10-bit and 16-bit YUV Formats](https://docs.microsoft.com/en-us/windows/win32/medfound/10-bit-and-16-bit-yuv-video-formats)
* Apple
  + Core Media [Pixel Format Constants](https://developer.apple.com/documentation/coremedia/cmformatdescription/1564244-video_pixel_format_constants)
  + Core Video [Pixel Format Identifiers](https://developer.apple.com/documentation/corevideo/1563591-pixel_format_identifiers)
* GPAC
  + GF\_PixelFormat: [Supported pixel formats for video](https://doxygen.gpac.io/group__cst__grp.html#gaba58a85c71145a5a469688a6223d7f82)
* VideoLAN
  + [Documentation on YUV formats](https://wiki.videolan.org/YUV)
* FFmpeg
  + [AVPixelFormat](https://ffmpeg.org/doxygen/trunk/pixfmt_8h.html#a9a8e335cf3be472042bc9f0cf80cd4c5)
  + [AVCodecID](https://ffmpeg.org/doxygen/trunk/group__lavc__core.html#gaadca229ad2c20e060a14fec08a5cc7ce)
* GStreamer
  + [Raw Video Media Types](https://gstreamer.freedesktop.org/documentation/additional/design/mediatype-video-raw.html?gi-language=c): video/x-raw
* Fourcc.org
  + [YUV Pixel Formats](https://www.fourcc.org/yuv.php)
  + [RGB Pixel Formats](https://www.fourcc.org/rgb.php)
* Simple uncompressed video format for astronomical capturing [12]

From the above list it can be seen that the amount of different formats and associated FOURCCs currently implemented in various software is very large, and this fact should be taken into account when specifying the uncompressed carriage in ISOBMFF.

Requirements

Taking into account all four input contributions on this topic presented during last MPEG meetings, we construct a set of requirements [1][2][3][4]. When storing the uncompressed video format in ISOBMFF, signaling for the following attributes should be considered:

* Width and height
* Framerate and duration
* Chroma format
* Chroma Sub-sampling
* Picture aspect ratio
* Pixel aspect ratio
* Bit depth (chroma and luma for YUV, uniform for RGB)
* Color space format (YUV, RGB)
* Transfer characteristics
* YUV formats with side channel alpha / depth components
  + Note that the depth field of VisualSampleEntry in ISOBMFF is fixed to 0x0018 – images are in colour with no alpha.
* Field ordering signaling for interlaced content
* Support for several different pixel organizations, packing modes: planar, packed, semi-planar NV12/21 (hybrid planar Y + packed UV)
* Signaling for packed ordering (YUYV / YVYU / UYVY / VYUY, ARGB / ABGR / RGBA / BGRA, etc.)
* Define a sample format on how the data is stored in samples (interlacing, endianness)
* Take CICP metadata into account
* Embedded and using the data reference
* Composition of the images of the video in spatial zones with associated specific orders of groups of pixels (Figure 1)
* Support images consisting in “packed regions” (Figure 2)
* Access to spatial areas in a particular frame of the video (Figure 3)
* Other informative metadata e.g.: [3GPP raw schema](https://github.com/haudiobe/5G-Video-Content/blob/main/3gpp-raw-schema.json)

Graphical user interface

Description automatically generated

Figure 1: Zone and scan order in some industrial cameras where pixel data is available on a single tap. Pixel packets from different zones can be arranged in different orders [4].

|  |  |
| --- | --- |
| Application  Description automatically generated with medium confidence  Figure 2: Image as “packed regions” [4] | Figure 3: Image with regions of interest [4] |

Analysis of current implementations

In this section we investigate several methods on how to create ISOBMFF/QTFF files with uncompressed video data and what software is capable of playing them.

Technical tools

**FFmpeg**

FFmpeg is often referred to as a swiss army knife for multimedia processing. It can be used to generate a set of uncompressed video bitstreams as specified in [AVPixelFormat](https://ffmpeg.org/doxygen/trunk/pixfmt_8h.html#a9a8e335cf3be472042bc9f0cf80cd4c5). An example on how to generate a set of uncompressed video bitstreams can be found [here](https://github.com/podborski/isobmff/blob/raw2mov/createRaws.py). FFmpeg can even directly be used to package uncompressed video bitstreams into a .mov file. The example below packages ‘v210’ uncompressed bitstream into .mov using ‘v210’ sample entry:

ffmpeg -f v210 -s 176x144 -i gst\_v210\_176x144.raw -c copy v210\_gst.mov

**GPAC**

According to the release notes GPAC supports multiplexing of uncompressed video since version 1.0. Multiplexing uncompressed video data based on [TN2162](https://developer.apple.com/library/archive/technotes/tn2162/_index.html) seems to be considered by the developers.

**GStreamer**

A pipeline-based multimedia framework which can be used to generate uncompressed video formats as described in video/x-raw module: [Raw Video Media Types](https://gstreamer.freedesktop.org/documentation/additional/design/mediatype-video-raw.html?gi-language=c). An example on how to generate uncompressed video bitstreams can be found [here](https://github.com/podborski/isobmff/blob/raw2mov/createRaws.py).

**Final Cut Pro**

Video editing software capable of exporting and importing uncompressed ‘2vuy’ and ‘v210’ video to QTFF.

**HDRTools**

HDRTools is a software package which is used to support video standardization work in several groups e.g. JVET. It is highly configurable and already supports a wide range of uncompressed video formats. HDRTools does not currently support ISOBMFF multiplexing, but can be extended to provide such functionality. For the multiplexing part, ISOBMFF reference software ([libisomedia](https://github.com/MPEGGroup/isobmff)) can be used in a wrapper module of the HDRLib. This approach is currently implemented in a separate branch [feature\_mp4](https://gitlab.com/standards/HDRTools/-/tree/feature_mp4) and allows to create .mp4 files with the ‘v210’ format as defined in [TN2162](https://developer.apple.com/library/archive/technotes/tn2162/_index.html). The use of HDRTools in combination with libisomedia should be considered by the group while working on ISO/IEC 23001-17.

Playback of uncompressed files

To quickly test support for playback of various uncompressed sample entry types, we used a simple tool called [raw2mov](https://github.com/podborski/isobmff/tree/raw2mov/IsoLib/raw2mov). It packages the uncompressed video data of specified pixel format into samples of an ISOBMFF file and sets the specific sample entry type as a FOURCC from Table 3. FFmpeg and GStreamer [were used](https://github.com/podborski/isobmff/blob/raw2mov/createRaws.py) to generate the uncompressed video bitstreams which were packaged to ISOBMFF using raw2mov. In addition, Final Cut Pro was used to extract ‘2vuy’ and ‘v210’ QTFF files for verification. Also, the initial version of the [HDRTools](https://gitlab.com/standards/HDRTools/-/tree/feature_mp4) branch was used to verify ‘v210’ code point (note that horiz\_align\_pixels are currently not implemented in HDRTools and the width of the input image has to be a multiple of 48). Then we used 5 different players (QuickTime, MP4Client, ffplay, VLC and playbin) to test the coverage of each code point. Table 3 below summarizes the results.

It can be seen that ‘2vuy’, ‘v210’ and ‘raw ‘ (RGB24) are the most supported code points. Quite the opposite is the support of Digital Voodoo code points from Table 2, only ‘DVOO’ is supported by FFmpeg and all others didn’t play anywhere (not listed in the Table). GStreamer playbin was not able to play any of the files.

Table 3: On overview of player support for playing uncompressed video signaled with the FOURCC in a sample entry.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Code | Description | QT 10.5 (935.5) Mojave | QT 10.5 (1084.4.1) Big Sur | GPAC  1.1.0-DEV-rev114-g923e0e6e9-master | FFMPEG 4.3.1  ffplay | VLC 3.0.11.1 | GStreamer (1.18.0) playbin |
| ['2vuy'](https://mirror.informatimago.com/next/developer.apple.com/quicktime/icefloe/dispatch019.html#2vuy) | 8-bit-per-component 4:2:2 | OK | OK | OK | OK | OK | NO |
| ['yuv2'](https://mirror.informatimago.com/next/developer.apple.com/quicktime/icefloe/dispatch019.html#yuv2) | 8-bit-per-component 4:2:2 | wrong colors after convertion | NO | OK | wrong colors | OK | NO |
| ['v308'](https://mirror.informatimago.com/next/developer.apple.com/quicktime/icefloe/dispatch019.html#v308) | 8-bit-per-component 4:4:4 | NO | NO | wrong mapping | OK | NO | NO |
| ['v408'](https://mirror.informatimago.com/next/developer.apple.com/quicktime/icefloe/dispatch019.html#v408) | 8-bit-per-component 4:4:4:4 | NO | NO | NO | OK | NO | NO |
| ['v216'](https://mirror.informatimago.com/next/developer.apple.com/quicktime/icefloe/dispatch019.html#v216) | 10,12,14,16-bit-per-component 4:2:2 | wrong playback (white), thumbnail is fine | OK | NO | NO | NO | NO |
| ['v410'](https://mirror.informatimago.com/next/developer.apple.com/quicktime/icefloe/dispatch019.html#v410) | 10-bit-per-component 4:4:4 | playback is black, thumbnail wrong colors | playback is black, thumbnail wrong colors | wrong mapping | wrong colors | NO | NO |
| ['v210'](https://mirror.informatimago.com/next/developer.apple.com/quicktime/icefloe/dispatch019.html#v210) | 10-bit-per-component 4:2:2 | OK | OK | wrong mapping | OK | OK | NO |
| 'raw ' | RGB24 8-bit-per-component | OK | OK | OK | OK | OK | NO |
| ‘DVOO’ |  | --- | NO | NO | OK | NO | NO |
| 'j420' | planar YUV 4:2:0, 12bpp, full scale (JPEG) | OK (after convertion) | NO | wrong mapping | NO | NO | NO |
| 'I420' | planar YUV 4:2:0, 12bpp, full scale (JPEG) | NO | NO | NO | OK | OK | NO |
| 'IYUV' | planar YUV 4:2:0, 12bpp | NO | NO | NO | OK | OK | NO |
| 'yv12' | planar YUV 4:2:0, 12bpp | NO | NO | NO | OK | OK | NO |
| 'YVYU' | packed YUV 4:2:2, 16bpp, Y0 Cr Y1 Cb | NO | NO | NO | OK | OK | NO |
| 'RGBA' | packed RGBA 8:8:8:8, 32bpp, RGBARGBA... | OK (after convertion) | NO | NO | OK | OK | NO |
| 'ABGR' | packed ABGR 8:8:8:8, 32bpp, ABGRABGR... | OK (after convertion) | NO | NO | OK | NO | NO |

References

1. Jean Le Feuvre, "Raw AV in ISOBMF", Telecom ParisTech, MPEG#122 [m42594](https://dms.mpeg.expert/doc_end_user/documents/122_San%20Diego/wg11/m42594-v1-m42594.zip)
2. Jean Le Feuvre, "Storage of Raw Video in ISOBMFF", Telecom ParisTech, MPEG#132 [m55059](https://dms.mpeg.expert/doc_end_user/documents/132_OnLine/wg11/m55059-v2-m55059.zip)
3. Thomas Stockhammer, "Background and Requirements for Raw Video Format", Qualcomm Incorporated, MPEG#132 [m55286](https://dms.mpeg.expert/doc_end_user/documents/132_OnLine/wg11/m55286-v1-m55286_Raw.zip)
4. Eric Nassor, et al., "Inputs on uncompressed video File Format", Canon Research Centre France, MPEG#133 [m56032](https://dms.mpeg.expert/doc_end_user/documents/133_OnLine/wg11/m56032-v1-m56032_Onuncompressedvideostorage.zip)

1. [Technical Note TN2162](https://developer.apple.com/library/archive/technotes/tn2162/_index.html), "Uncompressed Y´CbCr Video in QuickTime Files", Apple Inc.

1. [QTFF](https://developer.apple.com/library/archive/documentation/QuickTime/QTFF/QTFFPreface/qtffPreface.html), "QuickTime File Format Specification", Apple Inc.
2. "CNN prepping new London facility for launch in March", [link](https://www.newscaststudio.com/2020/02/25/cnn-london-bureau/), Accessed 20 Jan 2021
3. "Inside the RSN’s New ST 2110 IP Production Facility in Wrigleyville", [link](https://www.sportsvideo.org/2020/08/11/marquee-sports-network-spotlight-part-1-inside-the-rsns-new-st-2110-ip-production-facility-in-wrigleyville/), Accessed 20 Jan 2021
4. SMPTE ST 2110-20:2017, “Professional Media Over Managed IP Networks: Uncompressed Active Video”
5. SMPTE ST 377-1:2011, "Material Exchange Format (MXF) — File Format Specification"
6. SMPTE 384M-2005: "Material Exchange Format (MXF) — Mapping of Uncompressed Pictures into the Generic Container"
7. H. Wilkens, G. Hahn, "Simple uncompressed video format for astronomical capturing", version 3, [link](https://free-astro.org/images/5/51/SER_Doc_V3b.pdf), Accessed 22 Jan 2021