Digital Imaging and Communications in Medicine (DICOM)

Supplement 202: Real-Time Video

Prepared by:

DICOM Standards Committee, Working Group 13
1300 N. 17th Street, Suite 900
Rosslyn, Virginia 22209 USA
VERSION: Final Text, September 16, 2019

Developed in accordance with: DICOM Workitem 2016-12-D

Table of Contents

Scope and Field of Application

This Supplement describes several new DICOM IODs and associated transfer syntaxes for the transport of real-time video, and/or audio, and associated medical data. These are referred to collectively as DICOM Real-Time Video (DICOM-RTV). The supplement defines an new IP-based DICOM Service for the broadcasting of real-time video to subscribers with a quality of service which is compatible with the communication inside the operating room (OR).

DICOM specified storage of medical video in endoscopy, microscopy or echography. But medical theaters such as the operating room (OR) are for the moment still using proprietary solutions to handle communication of real-time video and associated information like patient demographics, study description or 3D localization of imaging sources.

The new Real-Time Video Service supports interoperable devices inside the OR and beyond, enabling a better management of imaging information, impacting directly the quality of care.

Professional video (e.g., TV studios) equipment providers and users have defined in SMPTE (ST 2110 family of standards) a new standardized approach for conveying video and associated information (audio, ancillary data, metadata…). ST 2110-10 uses a multicast model rather than a peer-to-peer communication model.

SMPTE ST 2110 suite, elaborated on the basis of Technical Recommendation TR03 originated by the VSF (Video Services Forum), is used as a platform. DICOM is defining a mechanism to convey specific medical metadata along with the video in compliance with SMPTE ST 2110-10 which specifies the communication architecture.

DICOM-RTV restricts real-time communication to uncompressed video, since the underlying standards (SMPTE ST 2110 family) do not include any transport of compressed video yet. But the proposed mechanism for conveying the medical metadata along with the video (and audio) is fully compatible with transport of compressed video and it is anticipated that when underlying standards embrace compressed video, DICOM-RTV will be extended to support it. The Transfer Syntax UID mechanism enables the application to choose the nature of the compression (or not) of the associated video.

The Supplement does not define how the video will be stored or re-played. Only the method for feeding the recorder with the synchronized videos and associated metadata is specified by this supplement. Security aspects are out of scope of this supplement. Common security solutions (e.g., IPSEC, VLAN mechanisms) work with the proposed specification.

Changes to NEMA Standards Publication PS3.17-2019c

Digital Imaging and Communications in Medicine (DICOM)

Part 17: Explanatory Information

PPPP Real-Time Video Use Cases (Informative)

PPPP.1 INTRODUCTION

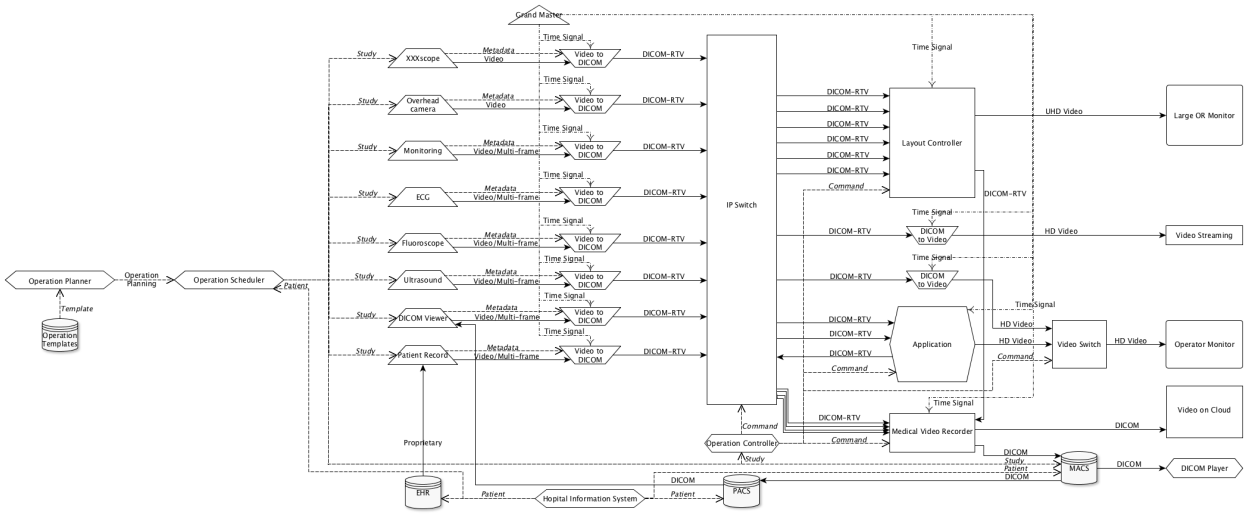


Figure PPPP.1-1: Overview diagram of operating room

As shown on Figure PPPP.1-1, the DICOM Real-Time Video (DICOM-RTV) communication is used to connect various video or multi-frame sources to various destinations, through a standard IP switch, instead of using a video switch. In the future, the equipment producing video will support DICOM-RTV natively but it is anticipated that the first implementations will rely on the use of converters to create a DICOM-RTV stream from the video stream (e.g., SDI) and associated metadata coming from information systems, through existing mechanisms (e.g., DICOM Worklist). Such converters have to be synchronized with the Grand Master which is delivering a very precise universal time. Similarly, the video receivers (e.g., monitors) will be connected to the central switch via a converter which has also to be synchronized via the Grand Master. The different DICOM-RTV streams can be displayed, recorded, converted or combined together for different use cases. The medical metadata in the DICOM-RTV streams can be used to improve the quality of the whole system, as explained in the following use cases.

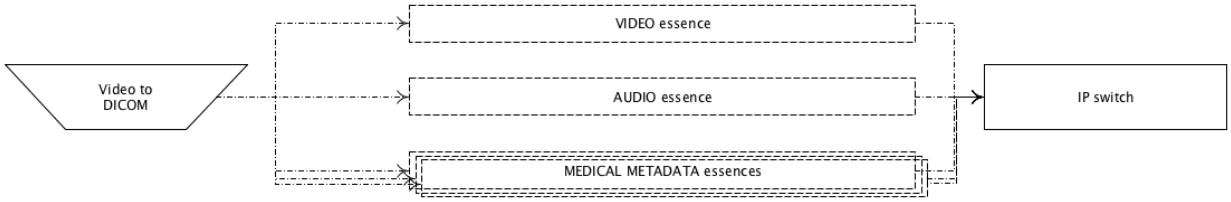


Figure PPPP.1-2: Real-Time Video stream content overview

As shown on Figure PPPP.1-2, the DICOM Real-Time Video stream is comprised of typically three different flows (“essences”) for respectively video, audio and medical metadata information, using the intrinsic capability of IP to convey different flows on the same medium, multiplexing three kinds of blocks. There will be thousands of blocks for each video frame, hundreds for each audio sample and one for the medical metadata associated to each video frame, respectively represented as “V” (video), “A” (audio) and “M” (metadata) on the Figure PPPP.1-3, which is the network view of the real-time streaming.

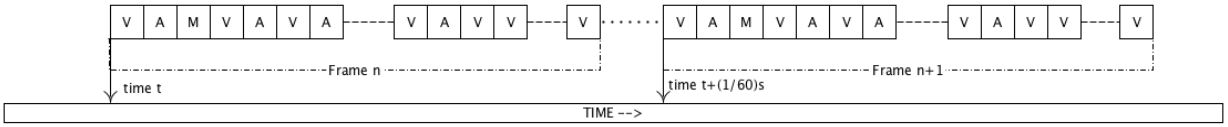


Figure PPPP.1-3: Real-Time Video transmission details

PPPP.2 USE CASE: DUPLICATING VIDEO ON ADDITIONAL MONITORS

In the context of image guided surgery, two operators are directly contributing to the procedure:

- a surgeon performing the operation itself, using relevant instruments;
- an assistant controlling the imaging system (e.g., laparoscope).

In some situations, both operators cannot stand on the same side of the patient. Because the control image has to be in front of each operator, two monitors are required, a primary one, directly connected to the imaging system, and the second one on the other side of the patient.

Additional operators (e.g., surgery nurse) might also have to see what is happening on additional monitors in order to anticipate actions (e.g., providing instrument).

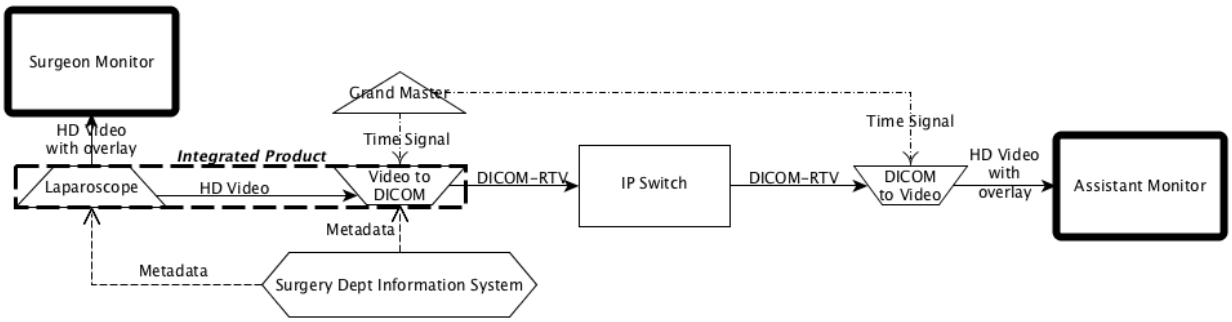


Figure PPPP.2-1: Duplicating on additional monitor

The live video image has to be transferred to additional monitors with a minimal latency, without modifying the image itself (resolution…). The latency between the two monitors (see Figure PPPP.2-1) should be compatible with collaborative activity for surgery where the surgeon is, for example, operating based on the primary monitor and the assistant is controlling the endoscope based on the second monitor. All equipment is synchronized with the Grand Master. The DICOM-RTV generation capability might be either an integrated part of the laparoscope product, or the laparoscope might send an HD video signal to the DICOM-RTV generator (Video-to-DICOM converter on the Figure PPPP.2-1). It is important that the converter be able to send video with or without a metadata overlay to the assistant monitor. This supplement addresses only the communication aspects, not the presentation.

PPPP.3 USE CASE: POST REVIEW BY SENIOR

A junior surgeon performs a procedure which apparently goes well. The next day, the patient experiences a complication requiring the surgeon to refer the patient to a senior surgeon.

In order to decide what to do, the senior surgeon:

- reviews and understands what happened;
- takes the decision to re-operate on the patient or not;
- accesses the videos of the first operation, if a new operation is performed.

Moreover, the junior surgeon has to review her/his own work in order to prevent against a new mistake.

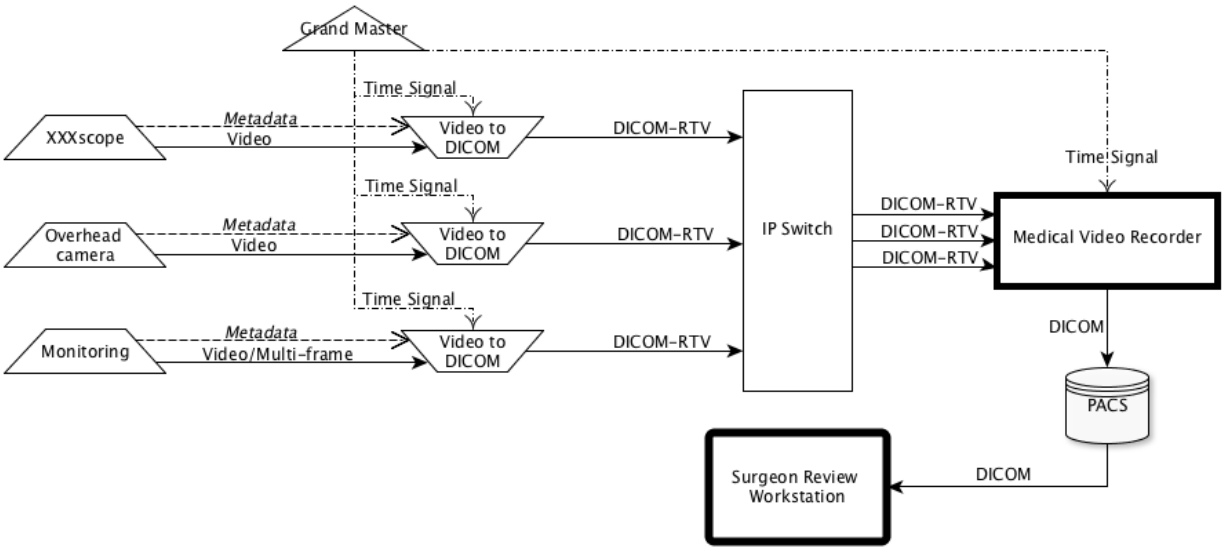


Figure PPPP.3-1: Recording multiple video sources

A good quality recording of video needs to be kept, at least for a certain duration, including all the video information (endoscopy, overhead, monitoring, …) and associated metadata from the surgery (see Figure PPPP.3-1). In this case, the metadata is coming directly from each device.. The recording has to maintain time consistency between the different video channels. Section PPPP.8.1 describes how the video could be captured and stored as a DICOM IOD using the present DICOM Store Service, as shown on the Figure PPPP.3-1, however the video could also be stored in another format. Such IODs could be retrieved and displayed using conventional DICOM workstation as shown on Figure PPPP.3-1. They could also be played back using DICOM-RTV as described in section PPPP.8.2.

PPPP.4 USE CASE: AUTOMATIC DISPLAY IN OPERATING ROOM (OR)

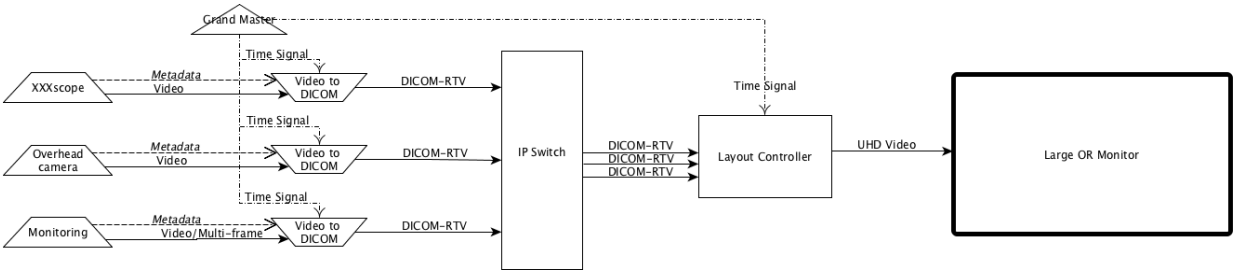


Figure PPPP.4-1: Displaying multiple source on one unique monitor

Some ORs have large monitors displaying a variety of necessary information. Depending on the stage of the procedure, the information to display changes. To improve the quality of the real-time information shared inside the OR, it is relevant to automate the changes of layout and content of such a display, based on the metadata conveyed along with the video (e.g., displaying the endoscope image only when the endoscope is inside the patient body).

All the video streams have to be transferred with the relevant metadata (patient, study, equipment…), as shown on the Figure PPPP.4-1. Mechanisms to select and execute the layout of images on the large monitor are not defined. Only the method for conveying the multiple synchronized videos along with the metadata, used as parameters for controlling the layout, is specified.

PPPP.5 USE CASE: AUGMENTED REALITY

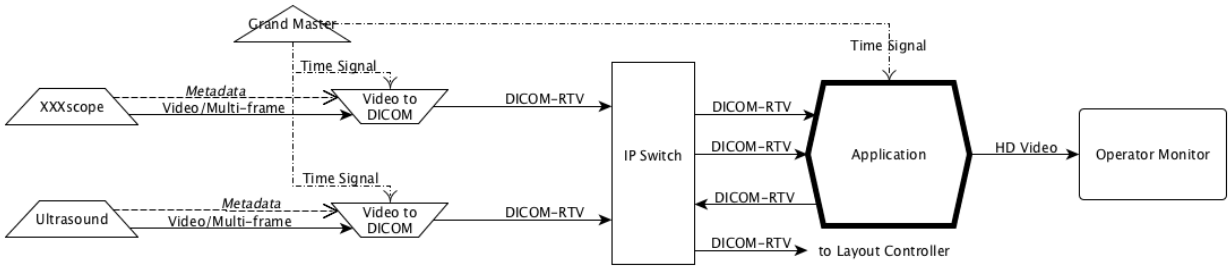


Figure PPPP.5-1: Application combining multiple real-time video sources

For image guided surgery, Augmented Reality (AR) applications enrich the live images by adding information as overlay, either 3D display of patient anatomy reconstructed from MR or CT scans, or 3D projections of other real-time medical imaging (3D ultrasound typically). In the second case, display devices (glasses, tablets…) show a real-time “combination” image merging the primary live imaging (endoscopy, overhead, microscopy…) and the real-time secondary live imaging (ultrasound, X-Ray…). The real-time “combination” image could also be exported as a new video source, through the DICOM Real-Time Video protocol.

All video streams have to be transferred with ultra-low latency and very strict synchronization between frames (see Figure PPPP.5-1). Metadata associated with the video has to be updated at the frame rate (e.g., 3D position of the US probe). The mechanisms used for generating augmented reality views or to detect and follow 3D position of devices are out of scope. Only the method for conveying the multiple synchronized video/multi-frame sources along with the parameters, that may change at every frame, is specified.

PPPP.6 USE CASE: ROBOTIC AIDED SURGERY

Robotic assisted surgery involves using image guided robots or “cobots” (collaborative robots) for different kinds of procedures. Different devices use the information provided by the robot (actual position, pressure feedback…) synchronized with the video produced by imaging sources. For effective haptic feedback, it may be necessary to convey such information at a frequency higher than the video frequency, i.e.; 400 Hz vs. 60 Hz for present HD video.

PPPP.7 EXAMPLE OF DICOM REAL-TIME VIDEO IMPLEMENTATION

The following example illustrates a specific implementation of the Generic Use Case 4: Augmented Reality described above.

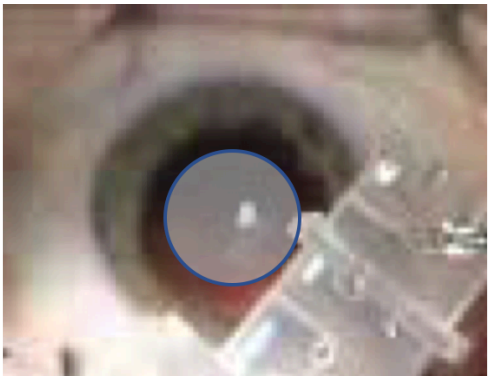


Figure PPPP.7-1: Example of implementation for Augmented reality based on optical image

The described use case is the replacement of the lens in cataract surgery (capsulorhexis). The lenses are manufactured individually, taking into account the patient's astigmatism. The best places for the incision, the position where the capsule bag should be torn and the optimal alignment for the new lens are calculated and a graphical plane is overlaid onto the optical path of the microscope to assist the surgeon, as shown in Figure PPPP.7-1.

Some solutions consist of a frame grabber in ophthalmology microscopes which grab video frames at 50 / 60 Hz. These frames are analyzed to identify the position and orientation of the eye and then a series of graphical objects are superimposed as a graphical plane onto the optical path to show the surgeon the best place to perform the incisions and how to orient the new lens to compensate the astigmatism.

Practically, the video frame grabbing takes 3 frames to be accessible to the image processor computing the series of graphical objects to be drawn as overlays on the optical image. It results in a delay between the frame used to create the objects and the one on which these objects are drawn. For safety reasons, it is important to record what the surgeon has seen. Due to the latency of the frame grabbing and the calculation of the positions of these graphical objects, the digital images are delayed in memory to also blend these objects onto the right digital image for the recording made in parallel.

DICOM Real-Time Video enables the storage of the recorded video and the frame by frame positions of these graphical objects separately. It might also be used to store other values associated with the streams such as the microscope's zoom, focus and light intensity values or the phaco's various settings, pressure, in the DICOM-RTV Metadata Flow. These separately stored flows could be later mixed together to aid in post-operative analysis or for teaching purposes. It would be possible to re-play the overlay either on the later image where the surgeon saw it, or on the image it was calculated from, to improve the algorithm. It would also reduce the workload of the machine during the operation because the blending of the video together with the display aids would be performed later during the post-operative analysis phase, and also maintain the original images.

The RTP Timestamp (RTS) of both video and DICOM-RTV Metadata Flows must match. Frame Origin Timestamp (FOTS) contained in DICOM-RTV Metadata must be consistent with RTP Timestamp,

enabling the proper synchronization between flows. As shown in the Figure PPPP.7-2, it is expected that the Frame Origin Timestamp relative of both the digital image and the overlays are set to T6 when the Image Datetime is T3 and the Referenced Image Datetime of the Mask is T0, represented as the T0 MASK.

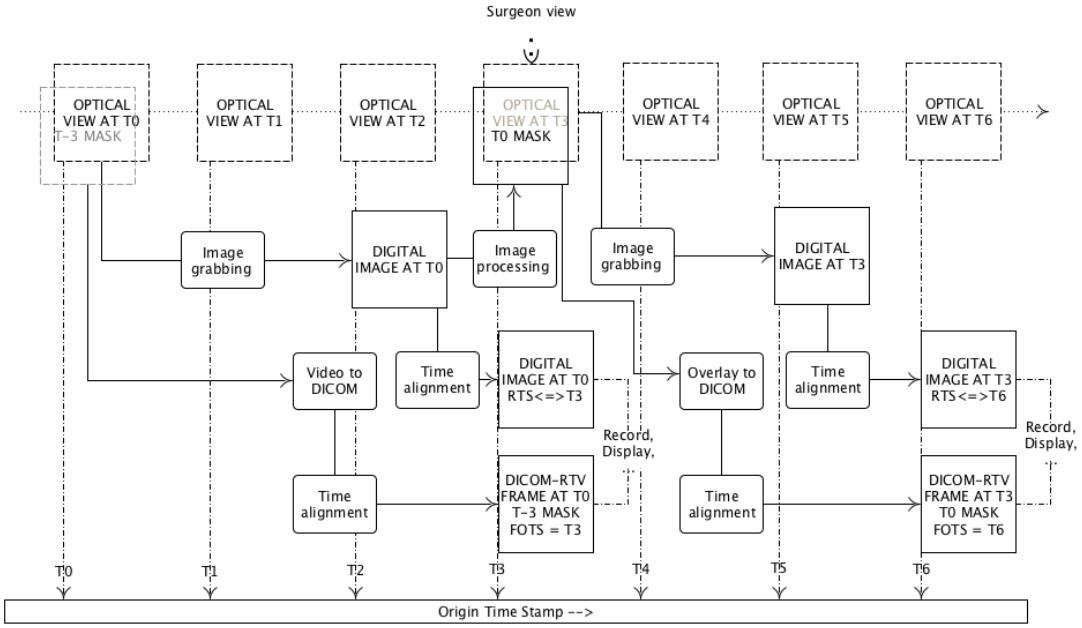


Figure PPPP.7-2: Example of implementation for Augmented reality based on optical image

Note

In the case the surgeon is viewing the digital image and not the optical image, the approach could be different, as shown in Figure PPPP.7-3.

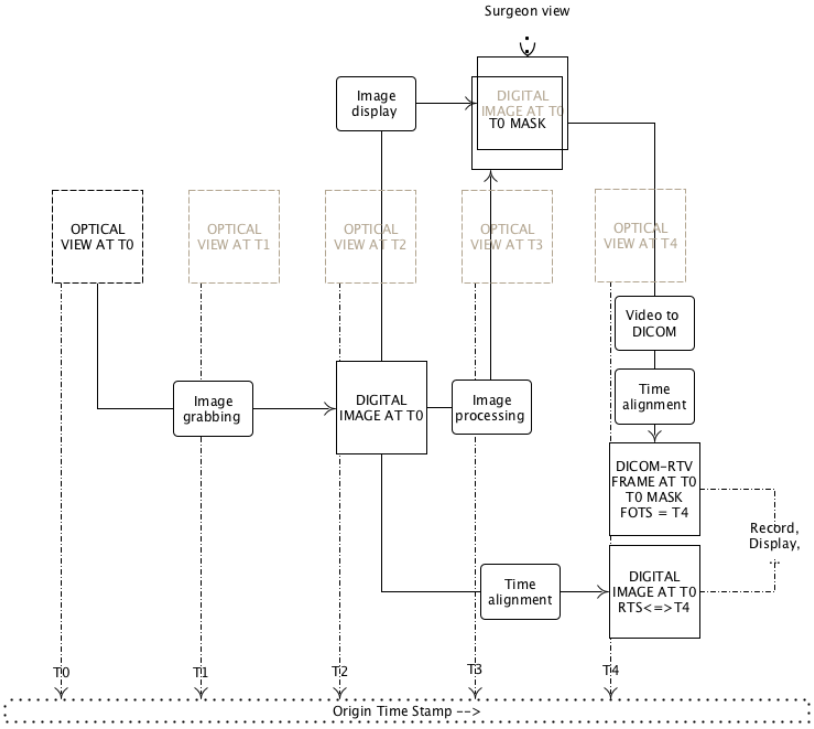


Figure PPPP.7-3: Example of implementation for Augmented reality based on digital image

PPPP.8 STORAGE CONSIDERATION

PPPP.8.1 Creating IOD from DICOM-RTV streams

It is reasonable to take some or all of an DICOM-RTV stream to create storage DICOM IOD. Transcoding the patient metadata and video content should be relatively straightforward. Some of the issues that have to be considered include how to get information describing origin equipment, etc.

Storage of video data, even received in real-time, is possible. However, how to initiate a DICOM-RTV stream based on a stored video is presently not described in the standard. Also, how to encode directly a received DICOM-RTV stream into a DICOM Video Instance is not fully described. An external decision (manual or automatic) is required to specify at least the start time and the end time of the portion of the stream to be stored. However, some principles can be established to ensure that receiving applications will actually find in the DICOM-RTV flow all the data items needed for the replay or storage of this data using DICOM Storage services. Regarding storage of this data using DICOM Storage services:

- “Pixel Data” and “Waveform Data” attributes of the DICOM (video) Composite Objects should be mapped from the corresponding payloads in media (e.g., video and audio) flows and associated SDP objects;

- The metadata attributes of the DICOM composite objects should be mapped from the DICOM-RTV metadata flows; attributes applicable to all frames (e.g., included in the Current Frame Functional Group Sequence) should be mapped from the static part of the DICOM-RTV

metadata; attributes applicable to a single frame (e.g., Per-frame Functional Group Sequence) should be mapped from the dynamic part of the DICOM-RTV Metadata;

- The “Cine” and “Multi-frame” modules, as well as the “Number of Waveform Samples” attribute, not present in the DICOM-RTV Metadata, are built from the values of the RTV Meta Information (e.g., Sample Rate), the dynamic payload of the relevant flows (e.g., Frame Numbers) and the external decisions (e.g., Start Time);

- Based on the choice of the application and on the possible presence of a DICOM-RTV Rendition flow, the DICOM composite object to be stored may gather or not the individual essences of the DICOM-RTV flows (e.g., video and audio contents in a single SOP instance using a MPEG2 Transfer syntax).

PPPP.8.2 Streaming DICOM-RTV from stored IOD

Regarding initiating a DICOM-RTV stream from a stored instance, the application should be able to regenerate the different DICOM-RTV flows, with the same synchronization characteristics, in compliance with SMPTE ST 2110-10.

- Subcase 1 is conventional video IODs e.g., ultrasound video/multi-frame or angio video/multi-frame.

- Subcase 2 is one or more video IODs that were previously DICOM-RTV, e.g., stored like PPPP.8.1.

- If the multiple stored IOD of the subcase 2 contain synchronization information extracted from DICOM, it should be possible to playback them with a good synchronization.

PPPP.9 EXAMPLE OF ENGINEERING IMPLEMENTATION

An example of implementation of the Video-to-DICOM converter presented in the use cases PPPP.2 above could respect the following approach:

- The metadata are sent from the Departmental System to the Video-to-DICOM converter through TCP/IP using classical protocols as DICOM Worklist or HL7 ORM.

- The video/multi-frame is sent through coaxial cable using classical video protocol (e.g., uncompressed HD video over Serial Digital Interface SDI).

- The time (“timestamp”) is sent through IP respecting PTP, for synchronizing all the senders and receivers, through “time alignment” mechanism described in SMPTE ST 2110-10.

- All this information is used to produce several RTP sessions over IP:

    - SMPTE ST 2110-20 compliant video flow.

- SMPTE ST 2110-10 compliant DICOM Metadata Flow, including payload header (RTV Meta Information) as well as dynamic payload part (DICOM Current Frame Functional Groups Module) for every frame, and including additionally the static payload part (DICOM Real-Time Video Endoscopic/Photographic Image IOD Modules) at least every second.

- If sound is provided:

  - SMPTE ST 2110-30 compliant audio flow.

  - SMPTE ST 2110-10 compliant DICOM Metadata Flow, including payload header (RTV Meta Information) as well as dynamic payload part (DICOM Current Frame Functional Groups Module) for every sample, and including additionally the static payload part (DICOM Real-Time Audio Waveform IOD Modules) at least every second.

  - SMPTE ST 2110-10 compliant DICOM Metadata Flow, including payload header and static payload part (DICOM Rendition Selection Document IOD Modules), at least every second, in order to associate the two flows above.

Note

Eventually, the laparoscope systems will embed the Video-to-DICOM converter, as shown on the "*Integrated Product*" box of the Figure PPPP.2-1.

PPPP.20 TRANSMITTING A STEREO VIDEO

The particular case of stereo vision, may either be solved by combining the contents into a single flow (Multiview video Coding) with inclusion of the C.7.6.28 Real-Time Acquisition Module in the metadata, or by separating contents into two flows (left content apart from right content) and then pairing them by using a (RTV Stereo Video) Rendition.

PS3.17: Add a new Annex Transport of Elementary Stream over IP as indicated.

QQQQ Transport of Elementary Stream over IP (Informative)

Carriage of audiovisual signals in their digital form across television plants has historically been achieved using coaxial cables that interconnect equipment through Serial Digital Interface (SDI) ports. The SDI technology provides a reliable transport method to carry a multiplex of video, audio and metadata with strict timing relationships.

The features and throughput of IP networking equipment having improved steadily, it has become practical to use IP switching and routing technology to convey and switch video, audio, and metadata essence within television facilities.

Existing standards such as SMPTE ST 2022-6:2012 have seen a significant adoption in this type of application where they have brought distinct advantages over SDI, albeit only performing Circuit Emulation of SDI (i.e.; Perfect bit-accurate transport of the SDI signal contents).

However, the essence multiplex proposed by the SDI technology may be considered as somewhat inefficient in many situations where a significant part of the signal is left unused if little or no audio and/or ancillary data has to be carried along with the video raster, as depicted in Figure QQQQ-1 below:

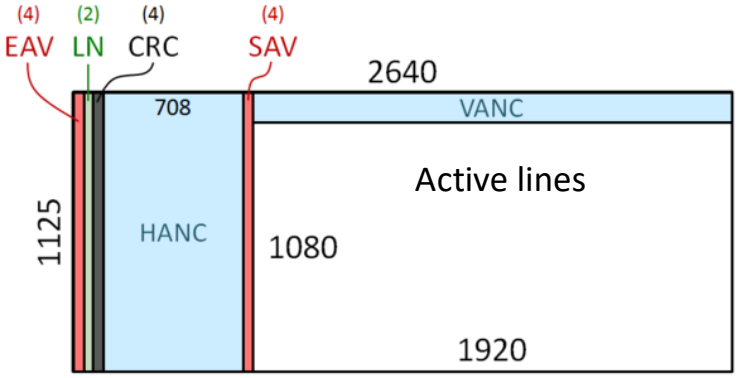


Figure QQQQ-1 Structure of a High Definition SDI signal

Note

> Acronyms on the Figure QQQQ-1 stand for: LN: line number; EAV: end of active video; SAV: start of active video; CRC: Cyclic Redundancy Code; HANC & VANC: horizontal & vertical ancillary data. The parentheses indicate the number of 8, 10 or 12 bits words used for each information.

As new image formats such as UHD get introduced, the corresponding SDI bit-rates increase, way beyond 10Gb/s and the cost of equipment at different points in a video system to embed, de-embed, process, condition, distribute, etc. the SDI signals becomes a major concern.

Consequently there has been a desire in the industry to switch and process different essence elements separately, leveraging the flexibility and cost-effectiveness of commodity networking gear and servers.

The Video Services Forum (VSF) has authored its Technical Recommendation #3 (a.k.a. VSF-TR03) describing the principles of a system where streams of different essences (namely video, audio, metadata to begin with) can be carried over an IP-based infrastructure whilst preserving their timing characteristics.

The TR03 work prepared by VSF has been handed off to the Society of Motion Picture & Television Engineers (SMPTE) for due standardization process, resulting in the SMPTE ST 2110 family of standards. SMPTE ST 2110-10, 20 and 30 were approved on September 18, 2017:

- ST 2110-10: System Timing and definitions;
- ST 2110-20: Uncompressed active video;
- ST 2110-21: Traffic Shaping Uncompressed Video;
- ST 2110-30: Uncompressed PCM audio;
- ST 2110-40: Ancillary data.

The ST 2110 family of standards expands over time and the corresponding DICOM components may consider adopting these extensions (e.g., compressed video, large metadata support…).

The system is intended to be extensible to a variety of essence types, its pivotal point being the use of the RTP protocol. In this system, essence streams are encapsulated separately into RTP before being individually forwarded through the IP network.

A system is built from devices that have senders and/or receivers. Streams of RTP packets flow from senders to receivers, however senders have no explicit awareness or coordination with the receivers. RTP streams can be either unicast or multicast, in which case multiple receivers can receive the stream over the network.

Devices may be adapters that convert from/to existing standard interfaces like HDMI or SDI, or they may be processors that receive one or more streams from the IP network, transform them in some way and transmit the resulting stream(s) to the IP network. Cameras and monitors may transmit and receive elementary RTP streams directly through an IP-connected interface, eliminating the need for legacy video interfaces.

Proper operation of the ST 2110 environment relies on a reliable timing infrastructure that has been largely inspired by the one used in AES67 for Audio over IP.

Inter-stream synchronization relies on timestamps in the RTP packets that are sourced by the senders from a common Reference Clock. The Reference Clock is distributed over the IP network to all participating senders and receivers via PTP (Precision Time Protocol version 2, IEEE 1588-2008).

Synchronization at the receiving device is achieved by the comparison of RTP timestamps with the common Reference Clock.

DICOM devices, which typically support NTP, will need to handle PTP to use this functionality, which may involve hardware changes. Each device maintains a Media Clock which is frequency locked to its internal time-base and advances at an exact rate specified for the specific media type. The media clock is used by senders to sample media and by receivers when recovering digital media streams. For video and ancillary data, the rate of the media clock is 90 kHz, whereas for audio it can be 44.1 kHz, 48 kHz, or 96 kHz.

For each specific media type RTP stream, the RTP Clock operates at the same rate as the Media Clock.

ST 2110-20 specifies a very generic mechanism for RTP encapsulation of a video raster. It supports arbitrary resolutions, frame rates, and introduces a clever pixel packing accommodating an extremely wide variety of bit depths and sampling modes. It is very heavily inspired from IETF RFC4175.

ST 2110-21 specifies traffic shaping and delivery timing of uncompressed video, in order to enable transport of multiple videos on the same physical network.

ST 2110-30 specifies a method to encapsulate PCM digital audio using AES67 to which it applies a number of constraints. AES67 is a technical standard for audio over IP and audio over Ethernet. The standard was developed by the Audio Engineering Society.

ST 2110-40 specifies a simple method to tunnel packets of SDI ancillary data present in a signal over the IP network and enables a receiver to reconstruct an SDI signal that will embed the ancillary data at the exact same places it occupied in the original stream.

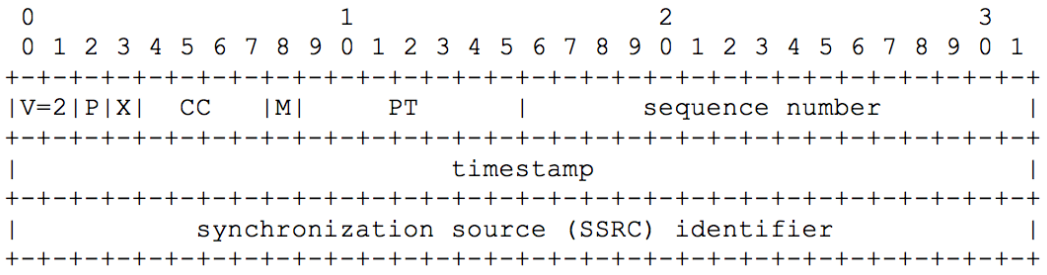
Sender devices construct one SDP (Session Description Protocol) object per RTP Stream. These SDP objects are made available through the management interface of the device, thereby publishing the characteristics of the stream they encapsulate, however no method is specified to convey the SDP object to the receiver. Implementations can rely on web URLs, files or documentation on media, or it can be configured on the receiver from product documentation since it can be relatively static. This SDP object provides the basic information a system needs in order to identify the available signal sources on the network.

It is worth noting that although ST 2110 currently describes the method for transporting video and audio, the same principles may be applied to other types of media by selecting the appropriate RTP payload encapsulation scheme, and complying to the general principles defined by ST 2110-10.

Some details of the ST 2110-10 are reproduced below for convenience. Refer to the original specifications for implementation.

The RTP header bits have the following format:

Figure QQQQ-2 RTP Header



With:

version (V): 2 bits	Version of RTP as specified in IETF RFC 3550.
padding (P): 1 bit	When set the packet contains padding octets at the end as specified in IETF RFC 3550.
extension (X): 1 bit	When set the fixed header is followed by an RTP header extension.
CSRC (CC): 4 bits	Number of CSRC identifiers as specified in IETF RFC 3550.
marker (M): 1 bit	For video it is set to 1 when the RTP packet is carrying the last video essence of a frame or the last part of a field as specified in SMPTE ST 2110-20.
payload type (PT)	Identifies the format of the payload. For a video or audio payload it is as specified in SMPTE ST 2110-10.
sequence number	Increments by one for each RTP data packet sent. It is as specified in IETF RFC 3550.
timestamp	Reflects the sampling instant of the first octet in the RTP data packet. It contains the timestamp as specified in SMPTE ST 2110-10.
SSRC	Identifies the synchronization source. It is as specified in IETF RFC 3550.

The RTP header extension bits have the following format:

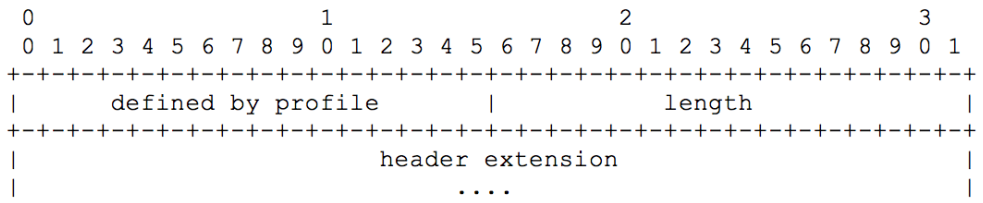


Figure QQQQ-3 RTP Header Extension

With:

defined by profile: 16 bits	It is defined by the type of header extension used.
length: 16 bits	Size of the header extension in 32-bits units. It does not include the 4 byte header extension (“defined by profile” + “length”).
header extension	The one-byte header extension form is described below. The total size of the header extension is a multiple of 4 bytes.

In complement to the SMPTE ST 2110 family of standards, AMWA (Advanced Media Workflow Association) has authored a recommendation called NMOS (Networked Media Open Specifications) which specifies the following header extensions:

- PTP Sync Timestamp:
  provides an absolute capture or playback timestamp for the Grain essence data, which consists of a 48-bit seconds field followed by a 32-bit nanosecond field. The length value in the extension header is 9.

- PTP Origin Timestamp:
  provides an absolute capture timestamp for the Grain essence data, which consists of a 48-bit seconds field followed by a 32-bit nanosecond field. The length value in the extension header is 9.

- Flow Identifier:
  a UUID which uniquely identifies the flow. The value is 16 bytes and therefore the length value in the extension header is 15.

- Source Identifier:
  a UUID which uniquely identifies the source. The value is 16 bytes and therefore the length value in the extension header is 15.

- Grain Duration:
  identifies the time period for which the video essence within the Grain should be displayed or the time period for which the audio essence should be played back, describing the length of a consistent video or audio sequence. It is a rational number consisting of a 4 byte numerator and 4 byte denominator. The value is 8 bytes and therefore the length value in the extension header is 7. Use of Grain Duration is optional.

- Grain Flags:
  The Grain Flags are a single byte with the following form:

```
0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|S|E| reserved |
+-+-+-+-+-+-+-+-+
```

Figure QQQQ-4 RTP Grain Flags

Start flag (S): 1 bit	This bit shall be set to 1 in the first packet of the Grain. Otherwise it shall be set to 0.
End flag (E): 1 bit	This bit shall be set to 1 in the last packet of the Grain. Otherwise it shall be set to 0.
Reserved: 6 bits	These bits are reserved for future use and should be set to 0. The length value of this extension header is 0.

Add a new NEMA Standards Publication PS3.22-2019d

Digital Imaging and Communications in Medicine (DICOM)

Part 22: Real-Time Communication

1 Scope

This standard specifies an SMPTE ST 2110-10 based service, relying on RTP, for the real-time transport of DICOM metadata. It provides a mechanism for the transport of DICOM metadata associated with a video or an audio flow based on the SMPTE ST 2110-20 and SMPTE ST 2110-30, respectively.

2 Normative References

[EBU-SMPTE-VSF], 2015. *Joint Task Force on Networked Media (JT-NM) Phase 2 Report- Reference Architecture v1.0 2015*

[RFC3550] IETF, July 2003. *RTP: A Transport Protocol for Real-Time Applications* https://tools.ietf.org/html/rfc3550

[RFC5285] IETF, July 2008. *A General Mechanism for RTP Header Extensions*. https://tools.ietf.org/html/rfc5285

[SMPTE ST 2110-10], 2017. *Professional Media over IP Networks: System Timing and Definitions*

[SMPTE ST 2110-20], 2017. *Professional Media over IP Networks: Uncompressed Active Video*

[SMPTE ST 2110-30], 2017. *Professional Media over IP Networks: PCM Digital Audio*

3 Definitions

DICOM Real-Time Video DICOM-RTV encompasses the DICOM-RTV Service, transport of related multimedia bulk data and the Real-Time IODs to which it may be applied.

DICOM-RTV Service Real-Time transport of metadata which characterize multimedia bulk data.

Essence Video, audio or data type of source, as defined in [EBU-SMPTE-VSF].

Flow A sequence of Grains from a Source; a concrete representation of content emanating from the Source, as defined in [EBU-SMPTE-VSF].

Grain Represents an element of Essence or other data associated with a specific time, such as a frame, or a group of consecutive audio samples, or captions, as defined in [EBU-SMPTE-VSF].

Rendition A collection of time-synchronized Flows intended for simultaneous presentation, providing a complete experience of a Source Group, as defined in [EBU-SMPTE-VSF].

Source An abstract concept that represents the primary origin of a Flow or set of Flows, as defined in [EBU-SMPTE-VSF].

4 Symbols and Abbreviated Terms

AVP Audio Video Profile

DICOM-RTV DICOM Real-Time Video

NMOS Networked Media Open Specifications

PTP Precision Time Protocol

RTP Real-Time Protocol

SDP Session Description Protocol

SMPTE Society of Motion Picture and Television Engineers

5 Conventions

6 Data Communication Requirements

DICOM Real-Time Video uses the RTP protocol as defined in SMPTE ST 2110-10.

6.1 INTERACTION

As shown on the Figure 6-1, a device can have multiple Sources, one for each Essence which corresponds of the type of bulk data (video, audio or medical metadata), each Source producing one or multiple Flows representing the same content in different formats (high definition, low definition, uncompressed, compress with or without loss…).

Several Sources may be grouped in a Source Group. A concrete experience of a Source Group is a Rendition, defined as a collection of time-synchronized Flows intended for simultaneous presentation (e.g., the audio channel of a surgical camera).

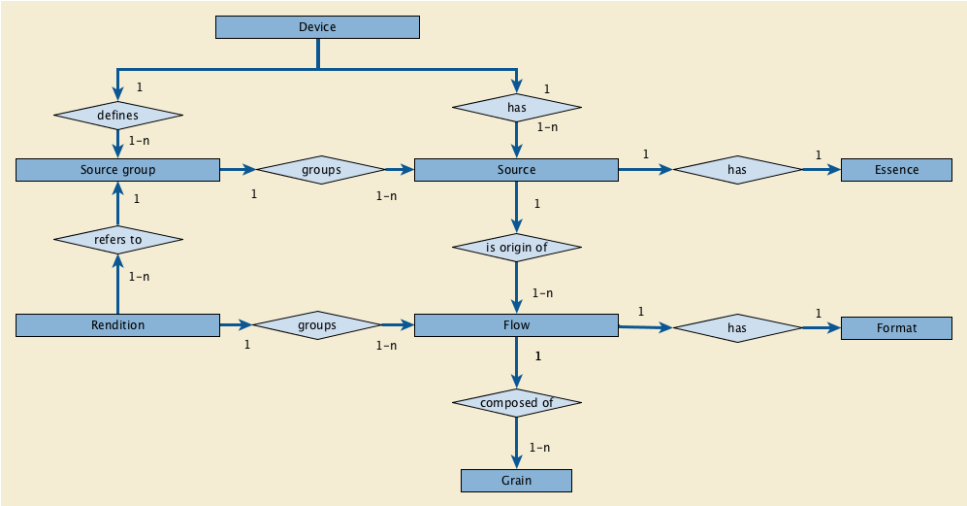


Figure 6-1. Real World diagram of DICOM-RTV

DICOM Real-Time Video standard specifies the communication mechanism for metadata, associated with real-time video and/or audio, originated from a medical imaging device. The mechanism involves one Source and one Flow of “DICOM Video Metadata Essence” for each video Flow and one Source and one Flow of “DICOM Audio Metadata Essence” for each audio Flow. Optionally, there is one Source and one Flow for the “DICOM Rendition Metadata” associating multiple Flows produced by the same device.

The interaction shall be as shown in Figure 6-2.

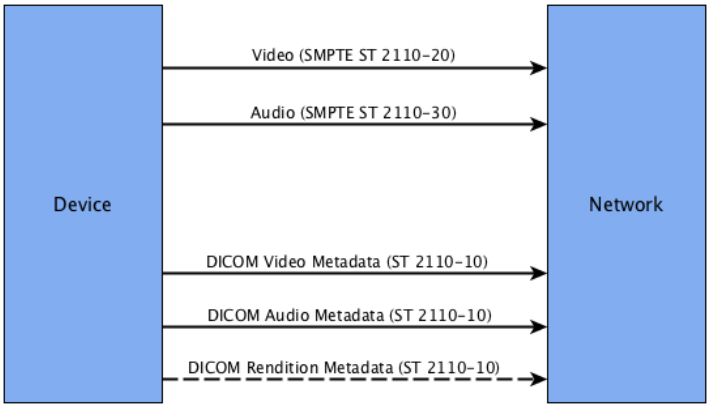


Figure 6-2. Interaction Diagram

SMPTE ST 2110-10 provides end-to-end network transport functions for applications transmitting real-time data. Content is transmitted in RTP sessions using RTP packets respecting SMPTE ST 2110-10.

A device can provide and/or consume content. A device that provides content has one or more Sources that can be of different Essences (e.g., Video and Audio). A Source is the origin of one or more Flows. Multiple Flows coming from the same Source are representations of the same content in different resolutions and/or codings. This is a broadcast/multicast protocol, so a device provides content whether or not a consuming device is present. A device that consumes content can subscribe/unsubscribe to available Flows.

The context and content of a video and/or audio Flow is described by a DICOM Metadata Flow, which is associated with each Flow. However the same DICOM Metadata Flow may be used to describe more than one Flow if their content is the same and their coding are close enough not to affect professional interpretation. A DICOM Rendition Metadata Flow may be used to associate multiple Flows provided by one device.

6.2 TRANSPORT

6.2.1 RTP Header

All Essences shall be transported with RTP according to SMPTE ST 2110-10 which requires that each Flow is described by an SDP object which specifies its content as well as connection details enabling the receiver to join the session. In addition to mandatory information specified in SMPTE ST 2110-10, for Audio and Video Essence, the SDP may also include the following information:

- PTP Sync Timestamp
- PTP Origin Timestamp
- Source Identifier
- Flow Identifier

  Note

  This information is the best way for associating multiple Flows originating from the same device. The presence of such information in the SDP implies that it is contained in the RTP Extended Header present in the first IP packet of a Grain (video frame, audio sample, metadata set…). It makes it possible to automatically associate and temporarily synchronize two Flows based on their content.

By definition, all the Flows according to SMPTE ST 2110-10 are synchronized by means of a common reference to the Universal Time, using PTP, with precision on the order of nanoseconds.

The RTP Header, for video and audio Flows, shall follow SMPTE ST 2110-20 and ST 2110-30, respectively.

The RTP Header, for DICOM Metadata Flows, shall follow SMPTE ST 2110-10. The clock rate shall be identical to the one defined in the referenced audio or video Flow. The following additional constraints apply:

extension (X): 1 bit	Shall be set to 1.
payload type (PT)	The value of payload type is selected from the range 96-127. It is recommended to avoid numbers frequently used for audio (97) and video

(96), and for example use 104 for DICOM Metadata Essence. The value shall be associated to the media type “application” and the subtype “dicom” in the SDP. E.g. (DICOM Metadata on port 12345):

```
m=application 12345 RTP/AVP 104
a=rtpmap:104 dicom/90000
```

For the DICOM Metadata Essence, the RTP Header Extension defined by NMOS shall be present, including the following information:

- PTP Sync Timestamp
- PTP Origin Timestamp
- Source Identifier
- Flow Identifier

The “defined by profile” part of the RTP Header Extension shall be set to 0xBEDE identifying that the one-byte header extension form is used, as specified in [RFC5285].

6.2.2 RTP Payload

The RTP Payload for audio and video Flows shall follow SMPTE ST 2110-20 and ST 2110-30, respectively.

The RTP Payload for DICOM Metadata Flows (audio, video and rendition) shall follow SMPTE ST 2110-10.

The RTP Payload for DICOM Metadata Flows consists of a DICOM dataset compliant with real-time communication.

The DICOM dataset is made of three parts:

- the RTV Meta Information part. This part shall be present in each Grain.
- the dynamic part containing information that varies over time (e.g., Origin Timestamp of the frame, Position of a probe, circle defining the eye. When it exists, this part shall be present in each Grain. The transmission rate of the dynamic part shall be identical to the rate of the associated Flow (e.g., one dataset per frame). This part is for the moment not applicable to DICOM Rendition Metadata.
- the static part containing information that doesn’t vary over time (e.g. Patient Name, Modality, …). This part will not be present in every Grain but shall be present at least in one Grain per second.

Note

The receiver cannot process information received from a sender until it receives DICOM Metadata including the static part, so it has to be sent at least every second in order to avoid a longer wait by the receiver when “connected” to a sender.

The transmission rate of DICOM audio flows will be typically of the range of 48kHz. The transmission rate of DICOM video flows will be typically of the range of 60Hz. The transmission rate of the DICOM Rendition Metadata Flow shall be at least 1Hz. It may be appropriate to use a higher frequency if there is a need for tight synchronization of associated Flows from a device (e.g., two videos of a stereo pair).

7 DICOM Real-Time Format

The DICOM Real-Time Format provides a means to encapsulate in an RTP session the Data Set representing a SOP Instance.

Figure 7-1 illustrates the encapsulation of a DICOM audio or video dataset in RTP. The byte stream of the Data Set is placed into the RTP Payload after the DICOM-RTV Meta Information. Each RTP session corresponds to a single SOP Instance.

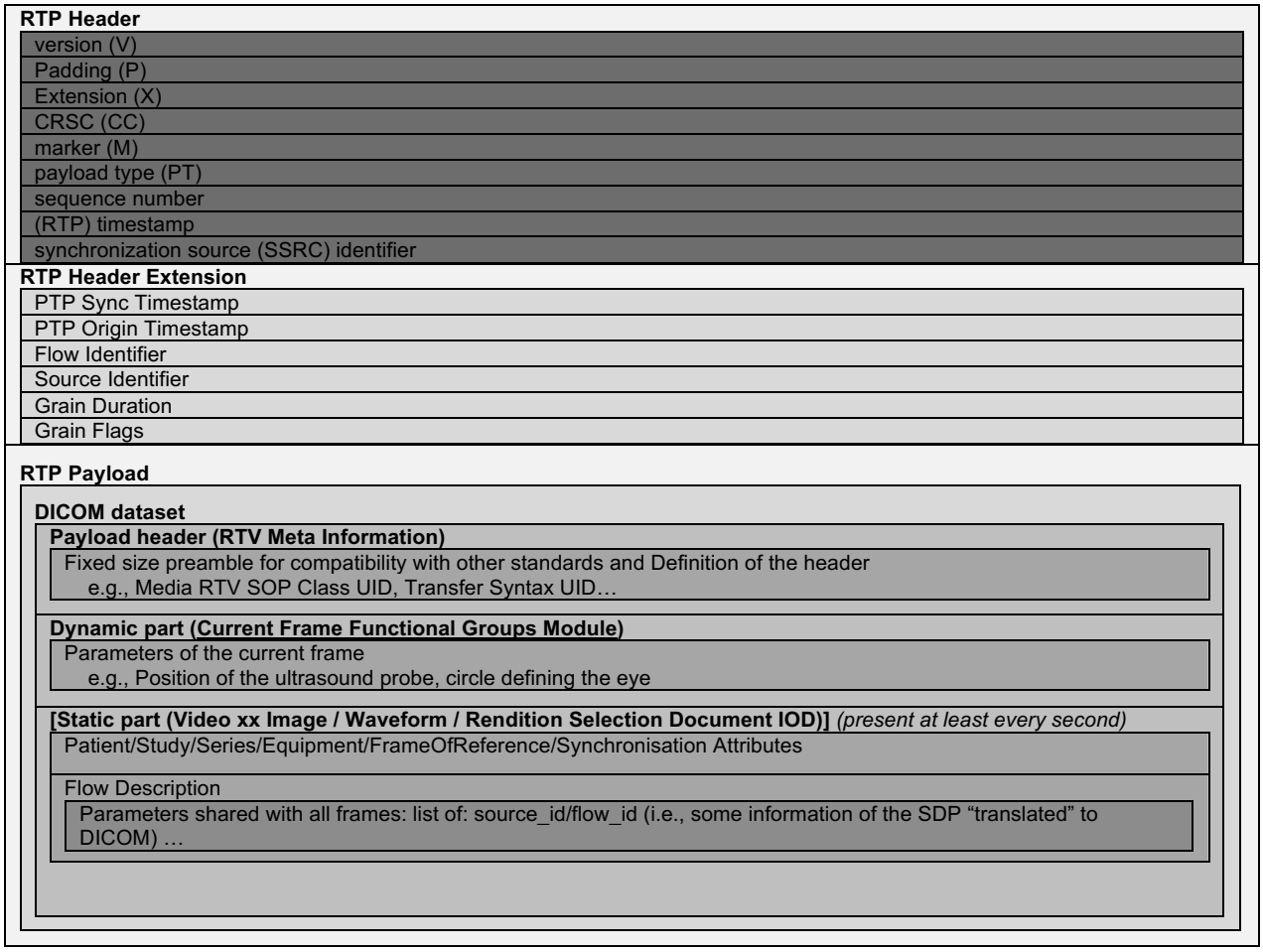


Figure 7-1. DICOM dataset encapsulation within RTP

7.1 RTV META INFORMATION

The RTV Meta Information includes identifying information on the encapsulated DICOM Data Set.

Note

The group number of the RTV Meta Information attributes (0002,xxxx) is lower than the one of other attributes in order to place the RTV Meta Information at the beginning of the payload, as is done in DICOM PS3.10.

Table 7.1-1. RTV Meta Information

Attribute Name	Tag	Type	Attribute Description
Header Preamble	*No Tag or Length Fields*	1	A fixed 128 byte field available for Application Profile or implementation specified use. If not used by an Application Profile or a specific implementation, all bytes shall be set to 00H. Receivers shall not rely on the content of this Preamble to determine that this payload is or is not a DICOM payload.
DICOM Prefix	*No Tag or Length Fields*	1	Four bytes containing the character string "DICM". This Prefix is intended to be used to recognize that this payload is or is not a DICOM payload.
File Meta Information Group Length	(0002,0000)	1	Number of bytes following this RTV Meta Element (end of the Value field) up to and including the last RTV Meta Element of the Group 2 RTV Meta Information
Transfer Syntax UID	(0002,0010)	1	Uniquely identifies the Transfer Syntax used to encode the referred bulk-data Flow. This Transfer Syntax does not apply to the RTV Metadata which is encoded using the Explicit VR Little Endian Transfer Syntax.
RTV Meta Information Version	(0002,0031)	1	This is a two byte field where each bit identifies a version of this RTV Meta Information header. In version 1 the first byte value is 00H and the second byte value is 01H.
RTV Communication SOP Class UID	(0002,0032)	1	Uniquely identifies the SOP Class associated with the Data Set. SOP Class UIDs allowed for RTV Communication are specified in section 7.2 STANDARD SOP CLASSES.
RTV Communication SOP Instance UID	(0002,0033)	1	Uniquely identifies the SOP Instance associated with the Data Set placed in the RTP Payload and following the RTV Meta Information.
RTV Source Identifier	(0002,0035)	1	The UUID of the RTP source that sends the RTV Metadata Flow.
RTV Flow Identifier	(0002,0036)	1	The UUID of the RTV Metadata Flow.
RTV Flow RTP Sampling Rate	(0002,0037)	1C	The rate of the dynamic part of the RTV Metadata Flow, the same as the bulk-data Flow rate. Required if RTV Metadata Flow includes a dynamic part.
RTV Flow Actual Frame Duration	(0002,0038)	3	Duration of image capture in msec.
Private Information Creator UID	(0002,0100)	3	The UID of the creator of the private information (0002,0102).
Private Information	(0002,0102)	1C	Contains Private Information placed in the RTV Meta Information. The creator shall be identified in (0002,0100). Required if Private Information Creator UID (0002,0100) is present.

7.2 STANDARD SOP CLASSES

The SOP Classes in the Real-Time Communication Class identify the Composite IODs to be sent. Table 7.2-1 identifies Standard SOP Classes.

Table 7.2-1. Standard SOP Classes

SOP Class Name	SOP Class UID	IOD Specification (defined in PS3.3)
Video Endoscopic Image Real-Time Communication	1.2.840.10008.10.1	Real-Time Video Endoscopic Image IOD
Video Photographic Image Real-Time Communication	1.2.840.10008.10.2	Real-Time Video Photographic Image IOD
Audio Waveform Real-Time Communication	1.2.840.10008.10.3	Real-Time Audio Waveform IOD
Rendition Selection Document Real-Time Communication	1.2.840.10008.10.4	Rendition Selection Document IOD

8 SECURITY CONSIDERATIONS

The metadata and ancillary streams usually contain Personally Identifiable Information (PII). The video and audio streams might contain protected information. The underlying SMPTE protocols do not specify any security protections to ensure confidentiality, integrity, or availability of the various data streams. DICOM does not specify any additions to the SMPTE protocols to provide such protection. Authorization and authentication of access to the DICOM-RTV Service is handled by configuration. Authentication is not re-confirmed at initiation of the underlying SMPTE protocols, and DICOM does not specify any additions to the SMPTE protocols for access control, authorization, or authentication.

The potential eavesdropping, replay, message insertion, deletion, modification, man-in-the-middle and denial of service attacks have not been analyzed. That analysis is up to the individual sites and installations.

Individual sites and installations will also need to perform their own assessments and selection of security mechanisms and add protections as necessary. The data rates and strict timing requirements for the data streams require careful analysis of any security mechanisms that are added. There do exist security mechanisms that operate at and below the IP level that can meet foreseen use cases, but there is insufficient experience or evidence to justify DICOM making a recommendation.

9 Conformance

An implementation claiming conformance to PS3.22 shall function in accordance with all its mandatory sections.

DICOM-RTV Services are used to transmit in real-time Composite SOP Instances. All Composite SOP Instances transmitted shall conform to the requirements specified in other Parts of the Standard.

An implementation may conform to the DICOM-RTV Services by supporting the role of origin device or receiving device, or both, for any of the Services defined in PS3.22.

The structure of Conformance Statements is specified in PS3.2.

An implementation shall describe in its Conformance Statement the Real-World Activity associated with its use of DICOM-RTV Services, including any proxy functionality between a DICOM-RTV and another service provided through DIMSE Service or RESTful (i.e.; storage of received video and audio with associated metadata).

In addition, the Conformance Statement document for a DICOM-RTV sending device shall specify how the receivers can get the content of the SDP objects describing the metadata and associated video and/or audio flows.

Changes to NEMA Standards Publication PS3.3-2019c

Digital Imaging and Communications in Medicine (DICOM)

Part 3: Information Object Definitions

Add new tables in the section A.1.4 Overview of the Composite IOD Module Content

Table A.1-11. Real-Time Object Modules Overview - Images

IODs Modules	RTV EN	RTV PH
Patient	M	M
Clinical Trial Subject	U	U
General Study	M	M
Patient Study	U	U
Clinical Trial Study	U	U
General Series	M	M
Clinical Trial Series	U	U
General Equipment	M	M
Enhanced General Equipment	M	M
Frame of Reference	C	C
Synchronization	M	M
General Image	M	M
General Reference	U	U
Real-Time Bulk Data Flow	M	M
Acquisition Context	M	M
Device	U	U
Specimen	C	C
VL Image	M	M
ICC Profile	M	U
SOP Common	M	M
Common Instance Reference	M	M

IODs Modules	RTV EN	RTV PH
Real-Time Acquisition	M	M
Current Frame Functional Groups	M	M

Table A.1-12. Real-Time Object Modules Overview - Waveforms

IODs Modules	RTV Basic Voice Audio
Patient	M
Clinical Trial Subject	U
General Study	M
Patient Study	U
Clinical Trial Study	U
General Series	M
Clinical Trial Series	U
General Equipment	M
Enhanced General Equipment	M
Synchronization	M
Waveform Identification	M
Real-Time Bulk Data Flow	M
Acquisition Context	M
SOP Common	M
Current Frame Functional Groups	M

Add a new section A.32.9 Real-Time Video Endoscopic Image IOD

A.32.9 Real-Time Video Endoscopic Image IOD

A.32.9.1 Real-Time Video Endoscopic Image IOD Description

The Real-Time Video Endoscopic Image IOD specifies the Attributes of Multi-frame Video Endoscopic Images transmitted in real-time.

A.32.9.2 Real-Time Video Endoscopic Image IOD Entity-Relationship Model

This IOD uses the E-R Model in Section A.1.2, with only the Image IE and Frame of Reference IE below the Series IE.

Table A.32.9-1. Real-Time Video Endoscopic Image IOD Modules

IE	Module	Reference	Usage
Patient	Patient	C.7.1.1	M
	Clinical Trial Subject	C.7.1.3	U
Study	General Study	C.7.2.1	M
	Patient Study	C.7.2.2	U
	Clinical Trial Study	C.7.2.3	U
Series	General Series	C.7.3.1	M
	Clinical Trial Series	C.7.3.2	U
Equipment	General Equipment	C.7.5.1	M
	Enhanced General Equipment	C.7.5.2	M
Frame of Reference	Frame of Reference	C.7.4.1	C – Required if multiple instances are spatially related or if the Camera Position Functional Group Macro is present in the Current Frame Functional Groups Module
	Synchronization	C.7.4.2	M
Image	General Image	C.7.6.1	M
	General Reference	C.12.4	U
	Real-Time Bulk Data Flow	C.7.6.26	M
	Acquisition Context	C.7.6.14	M
	Device	C.7.6.12	U
	Specimen	C.7.6.22	C - Required if the Imaging Subject is a Specimen
	VL Image	C.8.12.1	M
	ICC Profile	C.11.15	M

IE	Module	Reference	Usage
	SOP Common	C.12.1	M
	Common Instance Reference	C.12.2	M
	Real-Time Acquisition	C.7.6.28	M
	Current Frame Functional Groups	C.7.6.27	M

A.32.9.3 Real-Time Video Endoscopic Image IOD Content Constraints

The IOD shall not include audio. Captured patient voice or physiological sounds, healthcare professionals' commentary, or environmental sounds will be transported in a separate IOD.

A.32.9.3.1 Modality

The value of Modality (0008,0060) shall be ES.

Note

The use of a single value for Modality recognizes the fact that the same acquisition equipment is often used for different purposes (e.g., laparoscopy and colonoscopy). This means that Modality is not useful to distinguish one type of endoscopy from another when browsing a collection of studies. Therefore, the use of Procedure Code Sequence (0008,1032) and Anatomic Region Sequence (0008,2218) in the image instances and in the query response is recommended, though gathering sufficient information to populate these attributes in an unscheduled workflow environment (i.e.; in the absence of Modality Worklist) may require operator intervention.

A.32.9.3.2 Image Related Data Encoding

The Modality LUT Module, VOI LUT Module, Graphic Annotation Module and Overlay Plane Module shall not be present.

A.32.9.3.3 Anatomic Region Sequence

The Defined Context Group for Anatomic Region Sequence (0008,2218) shall be CID 4040 “Endoscopy Anatomic Regions”.

A.32.9.3.4 Current Frame Functional Groups

The Current Frame Functional Groups Module (see section C.7.6.27) shall be placed in the dynamic part of the RTP Payload.

Table A.32.9-2 specifies the Functional Group Macros included in the Current Frame Functional Groups Module for the Real-Time Video Endoscopic Image IOD.

Table A.32.9-2 Real-Time Video Endoscopic Image Functional Groups Macros

Functional Group Macro	Section	Usage
Time of Frame	C.7.6.16.2.29-1	M – shall not be used as a Shared Functional Group

Functional Group Macro	Section	Usage
Frame Content	C.7.6.16.2.2	M – shall not be used as a Shared Functional Group
Frame Usefulness	C.7.6.16.2.27	U – shall not be used as a Shared Functional Group
Camera Position	C.7.6.16.2.28	U

A.32.9.3.5 Time Distribution Protocol

The Time Distribution Protocol (0018,1802) of the Synchronization Module shall be set to “PTP”.

Add a new section A.32.10 Real-Time Video Photographic Image IOD

A.32.10 Real-Time Video Photographic Image IOD

A.32.10.1 Real-Time Video Photographic Image IOD Description

The Real-Time Video Photographic Image IOD specifies Multi-frame photographic Images, transmitted in real-time.

A.32.10.2 Real-Time Video Photographic Image IOD Entity-Relationship Model

This IOD uses the E-R Model in Section A.1.2, with only the Image IE below the Series IE.

Note

The Frame Pointers Module is included to managed the temporal synchronization with the video signal the IOD is referring, and optionally the spatial reference if each frame information contains a spatial position.

Table A.32.10-1. Real-Time Video Photographic Image IOD Modules

IE	Module	Reference	Usage
Patient	Patient	C.7.1.1	M
	Clinical Trial Subject	C.7.1.3	U
Study	General Study	C.7.2.1	M
	Patient Study	C.7.2.2	U
	Clinical Trial Study	C.7.2.3	U
Series	General Series	C.7.3.1	M
	Clinical Trial Series	C.7.3.2	U
Equipment	General Equipment	C.7.5.1	M
	Enhanced General Equipment	C.7.5.2	M

IE	Module	Reference	Usage
Frame of Reference	Frame of Reference	C.7.4.1	C – Required if multiple instances are spatially related or if the Camera Position Functional Group Module is present in the Current Frame Functional Groups Module
	Synchronization	C.7.4.2	M
Image	General Image	C.7.6.1	M
	General Reference	C.12.4	U
	Real-Time Bulk Data Flow	C.7.6.26	M
	Acquisition Context	C.7.6.14	M
	Device	C.7.6.12	U
	Specimen	C.7.6.22	C - Required if the Imaging Subject is a Specimen
	VL Image	C.8.12.1	M
	ICC Profile	C.11.15	U
	SOP Common	C.12.1	M
	Common Instance Reference	C.12.2	M
	Real-Time Acquisition	C.7.6.28	M
	Current Frame Functional Groups	C.7.6.27	M

A.32.10.3 Real-Time Video Photographic Image IOD Content Constraints

The IOD shall not include audio. Captured patient voice or physiological sounds, healthcare professionals' commentary, or environmental sounds will be transported in a separate IOD.

A.32.10.3.1 Modality

The value of Modality (0008,0060) shall be XC.

A.32.10.3.2 Image Related Data Encoding

The Modality LUT Module, VOI LUT Module, Graphic Annotation Module and Overlay Plane Module shall not be present.

A.32.10.3.3 Current Frame Functional Groups

The Current Frame Functional Groups Module (see section C.7.6.27) shall be placed in the dynamic part of the RTP Payload.

Table A.32.10-2 specifies the Functional Group Macros included in the Current Frame Functional Groups Module for the Real-Time Video Photographic Image IOD.

Table A.32.10-2 Real-Time Video Photographic Image Functional Group Macros

Functional Group Macro	Section	Usage
Time of Frame	C.7.6.16.2.29-1	M – shall not be used as a Shared Functional Group
Frame Content	C.7.6.16.2.2	M – shall not be used as a Shared Functional Group
Frame Usefulness	C.7.6.16.2.27	U – shall not be used as a Shared Functional Group
Camera Position	C.7.6.16.2.28	U

A.32.10.3.4 Time Distribution Protocol

The Time Distribution Protocol (0018,1802) of the Synchronization Module shall be set to “PTP”.

Add a new section A.34.11 Real-Time Audio Waveform IOD

A.34.11 Real-Time Audio Waveform IOD

A.34.11.1 Real-Time Audio Waveform IOD Description

The Real-Time Audio Waveform IOD specifies one-channel or two-channel digitized audio signals, transmitted in real-time.

A.34.11.2 Real-Time Audio Waveform IOD Entity-Relationship Model

This IOD uses the E-R Model in Section A.1.2, with only the Waveform IE and Frame of Reference IE below the Series IE.

A.34.11.3 Real-Time Audio Waveform IOD Module Table

Table A.34.11-1 specifies the Modules of the Real-Time Audio Waveform IOD.

Table A.34.11-1. Real-Time Audio Waveform IOD Modules

IE	Module	Reference	Usage
Patient	Patient	C.7.1.1	M
	Clinical Trial Subject	C.7.1.3	U
Study	General Study	C.7.2.1	M
	Patient Study	C.7.2.2	U

IE	Module	Reference	Usage
	Clinical Trial Study	C.7.2.3	U
Series	General Series	C.7.3.1	M
	Clinical Trial Series	C.7.3.2	U
Equipment	General Equipment	C.7.5.1	M
	Enhanced General Equipment	C.7.5.2	M
Frame of Reference	Synchronization	C.7.4.2	M
Waveform	Waveform Identification	C.10.8	M
	Real-Time Bulk Data Flow	C.7.6.26	M
	Acquisition Context	C.7.6.14	M
	SOP Common	C.12.1	M
	Current Frame Functional Groups	C.7.6.27	M

A.34.11.4 Real-Time Audio Waveform IOD Content Constraints

A.34.11.4.1 Modality

The value of Modality (0008,0060) shall be AU (audio).

A.34.11.4.2 Waveform Sequence

The number of Waveform Sequence (5400,0100) Items shall be 1.

A.34.11.4.3 Number of Waveform Channels

The value of Number of Waveform Channels (003A,0005) in the Waveform Sequence Item shall be 1 or 2.

A.34.11.4.4 Sampling Frequency

The value of Sampling Frequency (003A,001A) in each Waveform Sequence Item shall be 44.1 kHz, 48 kHz, or 96 kHz.

A.34.11.4.5 Channel Source

The Defined CID for the Channel Source Sequence (003A,0208) in each Channel Definition Sequence Item shall be CID 3000 “Audio Channel Source”.

A.34.11.4.6 Waveform Sample Interpretation

The value of Waveform Sample Interpretation (5400,1006) in each Waveform Sequence Item shall be SB or SS.

A.34.11.4.7 Current Frame Functional Groups Module

The Current Frame Functional Groups Module shall be placed in the dynamic part of the RTP Payload.

Table A.34.11-2 specifies the Functional Group Macros included in the Current Frame Functional Groups Module for the Real-Time Audio Waveform IOD.

Table A.34.11-2 Real-Time Audio Waveform Functional Group Macros

Functional Group Macro	Section	Usage
Time of Frame	C.7.6.16.2.29-1	M – shall not be used as a Shared Functional Group

A.34.11.4.8 Time Distribution Protocol

The Time Distribution Protocol (0018,1802) of the Synchronization Module shall be set to “PTP”.

Add an new section A.35.21 Rendition Selection Document IOD

A.35.21 Rendition Selection Document IOD

A.35.21.1 Rendition Selection Document IOD Description

The Rendition Selection Document IOD associates a group of time-synchronized Flows produced for a simultaneous presentation, transported using DICOM-RTV.

A.35.21.2 Rendition Selection Document IOD Entity-Relationship Model

This IOD uses the E-R Model in Section A.1.2, with only the SR Document IE below the Series IE.

A.35.21.3 Rendition Selection Document IOD Module Table

Table A.35.21-1 specifies the Modules of the Rendition Selection Document IOD.

Table A.35.21-1. Rendition Selection Document IOD Modules

IE	Module	Reference	Usage
Patient	Patient	C.7.1.1	M
	Clinical Trial Subject	C.7.1.3	U
Study	General Study	C.7.2.1	M
	Patient Study	C.7.2.2	U
	Clinical Trial Study	C.7.2.3	U
Series	Key Object Document Series	C.17.6.1	M
	Clinical Trial Series	C.7.3.2	U
Equipment	General Equipment	C.7.5.1	M
	Enhanced General Equipment	C.7.5.2	M

IE	Module	Reference	Usage
Frame Of Reference	Synchronization	C.7.4.2	M
Document	Key Object Document	C.17.6.2	M
	SR Document Content	C.17.3	M
	SOP Common	C.12.1	M

A.35.21.3.1 Rendition Selection Document IOD Content Constraints

A.35.21.3.1.1 Value Type

Value Type (0040,A040) in Content Sequence (0040,A730) of the SR Document Content Module is constrained to the following Enumerated Values (see Table C.17.3-7 for Value Type definitions):

Enumerated Values:

TEXT
CODE
UIDREF
PNAME
COMPOSITE
IMAGE
WAVEFORM
CONTAINER

The IMAGE and WAVEFORM Content Items shall only include reference to SOP Instance UID of DICOM-RTV Metadata Flows.

A.35.21.3.1.2 Relationship Constraints

Relationships between Content Items in the content of this IOD shall be conveyed in the by-value mode. See Table C.17.3-8 for Relationship Type definitions.

Note

Relationships by-reference are forbidden. Therefore, Referenced Content Item Identifier (0040,DB73) is not present in any of the Content Items within the SR Document Content Module.

Table A.35.21-2 specifies the relationship constraints of this IOD.

Table A.35.21-2. Relationship Content Constraints for Rendition Selection Document IOD

Source Value Type	Relationship Type (Enumerated Values)	Target Value Type
CONTAINER	CONTAINS	TEXT, IMAGE, WAVEFORM, COMPOSITE
CONTAINER	HAS OBS CONTEXT	TEXT, CODE, UIDREF, PNAME
CONTAINER	HAS CONCEPT MOD	CODE

A.35.21.3.1.3 Template Constraints

The document shall be constructed from TID 2010 “Key Object Selection” invoked at the root node.

Amend Section C.7.6.3.1.2 Photometric Interpretation

C.7.6.3.1.2 Photometric Interpretation

The value of Photometric Interpretation (0028,0004) specifies the intended interpretation of the image pixel data.

See PS3.5 Section 8.2.13 for constraints that apply when using DICOM Real-Time Video.

See PS3.5 for additional restrictions imposed by compressed Transfer Syntaxes.

The following values are defined. Other values are permitted if supported by the Transfer Syntax but the meaning is not defined by this Standard.

Defined Terms:

...

Add New Common Image Module: Real-Time Bulk Data Flow

C.7.6.26 REAL-TIME BULK DATA FLOW MODULE

This module references pixels/waveforms that are not contained within the DICOM dataset but conveyed in the associated Flows, as described in PS3.22 Section 6.1.

Table C.7.6.26-1 specifies the Attributes for the Real-Time Bulk Data Flow Module.

Table C.7.6.26-1 Real-Time Bulk Data Flow Module

Attribute Name	Tag	Type	Attribute Description
Real-Time Bulk Data Flow Sequence	(0034,000A)	1	Identifies the Source and Flows of a SMPTE ST 2110 communication providing Bulk Data. At least one item shall be included in this sequence.
>Source Identifier	(0034,0005)	1	UUID of the Source of the Flow(s). See C.7.6.26.1.1.
>Flow Identifier Sequence	(0034,0001)	1	Identifies the Flow(s) provided by the Source. One or more items shall be included in this sequence.
>>Flow Identifier	(0034,0002)	1	UUID of the Flow. See C.7.6.26.1.2.
>>Flow Transfer Syntax UID	(0034,0003)	1	UID of the encoding method of the referenced Flow See C.7.6.26.1.3.
>>Flow RTP Sampling Rate	(0034,0004)	1	Sampling rate in Hertz used by RTP for generating timestamp

			See C.7.6.26.1.4.

In case the Source is producing different Flows of the same Essence (e.g., video with two different sampling rates), the sequence must include at least the audio or video Flow that is associated with the present RTV Flow, and may in addition include the other Flow(s), enabling the receiver to subscribe to another Flow.

C.7.6.26.1 Real-Time Bulk Data Flow Module Attributes

C.7.6.26.1.1 Source Identifier

The Source Identifier (0034,0005) is a Universally Unique Identifier (UUID). The value is 128 bits long encoded in binary. It shall correspond to the value of the Source Identifier of the related bulk data Flow which may contain this Source Identifier in the RTP Extended Header. In case such Source Identifier is not present in the related bulk data flow, this Source Identifier shall be set to enable multiple IODs to refer the same Source.

C.7.6.26.1.2 Flow Identifier

The Flow Identifier (0034,0002) is a Universally Unique Identifier (UUID). The value is 128 bits long encoded in binary. It shall correspond to the value of the Flow Identifier of the bulk data Flow which may contain this Flow Identifier in the RTP Extended Header. In case such Flow Identifier is not present in the related bulk data flow, this Flow Identifier shall be set to enable multiple IODs to refer the same Flow.

C.7.6.26.1.3 Flow Transfer Syntax UID

The Flow Transfer Syntax UID (0034,0003) shall be the one relative to the corresponding Flow. The sequence shall have at least one item in which the Flow Transfer Syntax UID and the Flow RTP Sampling Rate correspond to the Transfer Syntax UID (0002,0010) and RTV Flow RTP Sampling Rate (0002,0037), respectively, of the DICOM-RTV Meta Information Header.

C.7.6.26.1.4 Flow RTP Sampling Rate

The Flow RTP Sampling Rate (0034,0004) shall be the one defined in the SDP of the corresponding Flow.

Add New Module: Current Frame Functional Groups Module

C.7.6.27 Current Frame Functional Groups Module

Table C.7.6.27-1 defines the Attributes related to the current frame when the IOD is transported using Real-Time Communication. It corresponds to the “per-frame” attribute for the not Real-Time IODs.

Note

The group number (0x0006) of the Sequence containing the Attributes of the Current Frame Functional Groups Module is used to contain Attributes that are specific to real-time transfer. If this information needs to be kept in a persistent object, its contents have to be copied into an Attribute that has a group number of 0x0008 or higher. The Current Frame Functional Groups Module is relative to the dynamic part of the RTP Payload. Its group number is lower than that of other Attributes in order to be placed before the static part of the RTP Payload, but higher than the Attributes of the RTV Metadata Information, in order to be placed after the header of the RTP Payload, as is done in DICOM PS3.10.

Table C.7.6.27-1 Current Frame Functional Groups Module Attributes

Attribute Name	Tag	Type	Attribute Description
Current Frame Functional Groups Sequence	(0006,0001)	1	Sequence that contains the Functional Groups Sequence Attributes corresponding to the current frame or audio sample. Only a single Item shall be included in this Sequence.
>Include one or more Functional Group Macros that are contained in every frame.			For each IOD that includes this module, a table is defined in which the permitted Functional Group Macros and their usage is specified.

C.7.6.27.1 Current Frame Functional Groups Module Attributes

C.7.6.27.1.1 Frame Origin Timestamp

Frame Origin Timestamp (0034,0007) contains 10 bytes conforming with the IEEE 1588:2008 (PTPv2) standard. IEEE 1588:2008 represents seconds and nanoseconds since Epoch, defined as 1 of January, 1970, at 00:00:00 TAI (International Atomic Time). The first 6 bytes contain the number of seconds, and the last 4 bytes contain the number of nanoseconds. The Time Source and Time Distribution Protocol values defined in C.7-7. Synchronization Module Attributes describe how the time was obtained.

If the RTP Header extension of a frame contains an PTP Origin Timestamp, that values shall be used for the Frame Origin Timestamp. Otherwise, the Frame Origin Timestamp value can be derived from Flow RTP sampling rate, the RTP Timestamp, and the reference time for the RTP Timestamp.

Frame Origin Timestamp is used to pair content from different Flows (payload flow with metadata flow), through a time alignment mechanism.

Add New Macro: Frame Usefulness Macro

C.7.6.16.2.27 Frame Usefulness Macro

Table C.7.6.16.2.27-1 specifies the attributes of the Frame Usefulness Functional Group Macro, related to the clinical usefulness of the current frame.

Table C.7.6.16.2.27-1 Frame Usefulness Functional Group Macro Attributes

Attribute Name	Tag	Type	Attribute Description
Frame Usefulness Group Sequence	(0034,0009)	1	Sequence that contains the Functional Groups Sequence Attributes corresponding to the current frame or audio sample. Only a single Item shall be included in this Sequence.
>Includes Information	(0034,000C)	3	Whether or not the current frame includes useful information, i.e. not noise nor void Enumerated value: **YES** **NO**
>Includes Imaging Subject	(0034,0008)	3	Whether or not the current frame includes the imaging subject. Enumerated value: **YES** **NO**

Add New Macro: Camera Position Macro

C.7.6.16.2.28 Camera Position Macro

Table C.7.6.16.2.28-1 specifies the attributes of the Camera Position Functional Group Macro, related to the position of the camera or the acquisition device for the current frame, with regards to the patient.

Table C.7.6.16.2.28-1 Camera Position Functional Group Macro Attributes

Attribute Name	Tag	Type	Attribute Description
Camera Position Group Sequence	(0034,000B)	1	Sequence that contains the Functional Groups Sequence Attributes corresponding to the current frame or audio sample. Only a single Item shall be included in this Sequence.
>Render Projection	(0070,1602)	1	Projection style. Enumerated Values: **PERSPECTIVE**
>Viewpoint Position	(0070,1603)	1	Position of the viewpoint in volume space. A point (x,y,z) in the Frame Reference Coordinate System referenced in the Frame of Reference Module.

Attribute Name	Tag	Type	Attribute Description
>Viewpoint LookAt Point	(0070,1604)	1	Point the viewpoint is looking at. A point (x,y,z) in the Frame Reference Coordinate System referenced in the Frame of Reference Module.
>Viewpoint Up Direction	(0070,1605)	1	Vertical orientation of the view. A vector (x,y,z) in the Frame Reference Coordinate System referenced in the Frame of Reference Module.
>Render Field of View	(0070,1606)	1	The field of view specified as a 6-tuple of values (X_{left}, X_{right}, Y_{top}, Y_{bottom}, $Distance_{near}$, $Distance_{far}$) in the Viewpoint Coordinate System, in mm. See Section C.11.30.1.

Add New Macro: Time of Frame Macro

C.7.6.16.2.29 Time of Frame Macro

Table C.7.6.16.2.29-1 specifies the attributes of the Time of Frame Functional Group Macro, related to the time of frame.

Table C.7.6.16.2.29-1 Time of Frame Functional Group Macro Attributes

Attribute Name	Tag	Type	Attribute Description
Time of Frame Group Sequence	(0034,000D)	1	Time characteristics of the current frame or audio sample. Only a single Item shall be included in this Sequence.
>Frame Origin Timestamp	(0034,0007)	1	This timestamp contains the capture time of the payload content for this frame or audio sample. It is in TAI, not UTC.

Add New Module: Real-Time Acquisition Module

C.7.6.28 Real-Time Acquisition Module

Table C.7.6.28-1 defines the Attributes related to all frames when the IOD is transported using Real-Time Communication.

Table C.7.6.28-1 Real-Time Acquisition Module Attributes

Attribute Name	Tag	Type	Attribute Description
Shared Functional Group Sequence	(5200,9229)	1	Sequence that contains the Functional Group Macros that are shared for all frames.

Attribute Name	Tag	Type	Attribute Description
			Only a single Item shall be included in this Sequence.
>Include one or more Functional Group Macros that are shared by all frames.			For each IOD that includes this module, a table is defined in which the permitted Functional Group Macros and their usage is specified. The Item may be empty if the requirements for inclusion of the Functional Groups are not satisfied.
Stereo Pairs Present	(0022,0028)	1	The multi-frame pixel data consists of left and right stereoscopic pairs. See Section C.7.6.28.1.1 for further explanation. Enumerated Values: **YES** **NO**

C.7.6.28.1 Real-Time Acquisition Module Attributes

C.7.6.28.1.1 Stereo Pairs Present

Stereo Pairs Present (0022,0028) shall have the value of YES when frame is encoded as left and right stereoscopic pair.

Changes to NEMA Standards Publication PS3.5-2019c

Digital Imaging and Communications in Medicine (DICOM)

Part 5: Data Structures and Encoding

Add New Section: Constraints for Pixel Data in DICOM-RTV

8.2.13 Constraints for SMPTE ST 2110-20 Uncompressed Active Video for DICOM-RTV

This section describes the constraints applying to pixel data carried in the DICOM-RTV Flow (separated from DICOM-RTV Metadata Flow) and fully described in SMPTE ST 2110-20.

The following table describes constraints on the SMPTE ST 2110-20 Video Flow in terms of the valid values for the corresponding DICOM attributes in the DICOM-RTV Metadata Flow:

- Samples per pixel
- Bits Allocated
- Bits Stored
- High Bit

Table 8.2.13-1: Constraints Applicable to Attributes describing Pixel Data

Samples per Pixel (0028,0002)	Bits Allocated (0028,0100)	Bits Stored (0028,0101)	High Bit (0028,0102)
3	8,16,16,16	8,10,12,16	7,9,11,15

DICOM Photometric Interpretation is based on CCIR 601 (aka ITU-R BT.601), therefore some restrictions apply to the possible combination of Sampling System and Colorimetry parameters as stated by SMPTE ST 2110-20.

Table 8.2.13-2: List of supported SMPTE ST 2110-20 Parameter Combinations

SMPTE ST 2110-20		DICOM Photometric Interpretation (0028,0004)
Sampling system	Colorimetry	
RGB	BT601	RGB
YCbCr-4:4:4	BT601	YBR_FULL
YCbCr-4:2:2	BT601	YBR_FULL_422
YCbCr-4:2:0	BT601	YBR_PARTIAL_420

Some other SMPTE ST 2110-20 parameter combinations do not correspond to existing DICOM photometric interpretations, so their use is currently not permitted. The table 8.2.x-3 lists the unsupported combinations.

Table 8.2.13-3: List of unsupported SMPTE ST 2110-20 Parameter Combinations

SMPTE ST 2110-20	
Sampling system	Colorimetry
RGB	BT2020, BT709, BT2100, ST2065-1, ST2065-3
YCbCr-4:4:4	BT2020, BT709, BT2100
YCbCr-4:2:2	BT2020, BT709, BT2100
YCbCr-4:2:0	BT2020, BT709, BT2100
CLYCbCr-4:4:4	BT2020
CLYCbCr-4:2:2	BT2020
CLYCbCr-4:2:0	BT2020
ICtCp-4:4:4	BT2100
ICtCp-4:2:2	BT2100
XYZ	XYZ
KEY	

Add New Section: Transfer Syntaxes for DICOM-RTV

10.15 TRANSFER SYNTAX FOR SMPTE ST 2110-20 UNCOMPRESSED PROGRESSIVE ACTIVE VIDEO

This Transfer Syntax is used for Uncompressed Video pixels carried in a DICOM-RTV Flow (separated from DICOM-RTV Metadata Flow) as described by the SMPTE ST 2110-20 standard, in the case the video is progressive (e.g., 1080p). The main parameters of the transfer syntax are described in the Annex A.8.

10.16 TRANSFER SYNTAX FOR SMPTE ST 2110-20 UNCOMPRESSED INTERLACED ACTIVE VIDEO

This Transfer Syntax is used for Uncompressed Video pixels carried in a DICOM-RTV Flow (separated from DICOM-RTV Metadata Flow) as described by the SMPTE ST 2110-20 standard, in the case the video is interlaced (e.g., 1080i). The main parameters of the transfer syntax are described in the Annex A.9.

10.16.1 Interlaced vs. Progressive video

Interlaced video supports transmitting video with a smaller bandwidth. One frame contains only odd lines and the next one contains only even lines. Interlaced video is acceptable for display but may cause problems in image processing. It is recommended to use progressive video. However, in case an original interlaced video signal is converted in the DICOM-RTV format, it is recommended to maintain the interlaced format and let the processing application deal with it.

10.17 TRANSFER SYNTAX FOR SMPTE ST 2110-30 PCM DIGITAL AUDIO

This Transfer Syntax is used for audio channel data carried in a DICOM-RTV Flow (separated from DICOM-RTV Metadata Flow) as described by the SMPTE ST 2110-30 standard. The main parameters of the transfer syntax are described in the Annex A.10.

Add New Section to Annex A : SMPTE ST 2110-20 Uncompressed Progressive Active Video

A.8 : SMPTE ST 2110-20 UNCOMPRESSED PROGRESSIVE ACTIVE VIDEO Transfer Syntax

This Transfer Syntax is used in a DICOM-RTV Metadata Flow in order to describe the accompanying SMPTE ST 2110-20 Video Flow.

DICOM attributes:

- Samples per Pixel (0028,0002)
- Photometric Interpretation (0028,0004)
- Bits Allocated (0028,0100)
- Bits Stored (0028,0101)
- High Bit (0028,0102)

are still applicable with some accommodations below.

As DICOM Photometric Interpretation (0028,0004) values {YBR_FULL, YBR_FULL_422, YBR_PARTIAL_420} are based on CCIR 601 (aka BT.601), DICOM-RTV supports only the following:

- SMPTE ST 2110-20 YCbCr-4:4:4 sampling system
- SMPTE ST 2110-20 RGB sampling system
- SMPTE ST 2110-20 YCbCr-4:2:2 sampling system

- SMPTE ST 2110-20 YCbCr-4:2:0 sampling system

The Table A.8-1 describes the different color resolution.

Table A.8-1: DICOM attributes for different color resolution

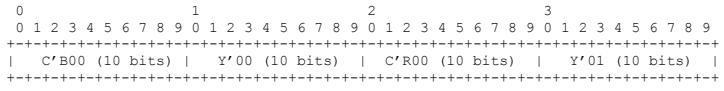
Bit Depth	Samples per Pixel (0028,0002)	Bits Allocated (0028,0100)	Bits Stored (0028,0101)	High Bit (0028,0102)
8	3	8	8	7
10	3	16	10	9
12	3	16	12	11
16	3	16	16	15

The way of encoding pixels shall respect SMPTE ST2110-20.

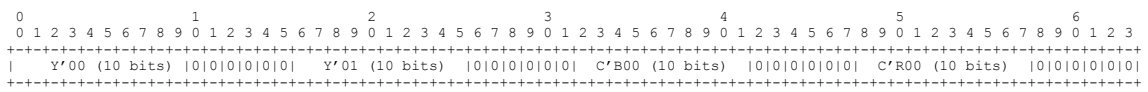
Note

This encoding is different than the encoding of Pixel Data (7FE0,0010). Example, for YBR_FULL_422 10bits:

- SMPTE ST 2110-20 Video Flow YCbCr 4:2:2 10 bits



- DICOM Pixel Data (7FE0,0010) YBR_FULL_422 10 bits



A DICOM Transfer Syntax for SMPTE ST 2110-20 Uncompressed Progressive Active Video shall be identified by a UID value of:
• 1.2.840.10008.1.2.7.1 corresponding to the SMPTE ST 2110-20 Professional Media over IP Networks: Uncompressed Active Video standard for the progressive video.

Add New Section to Annex A : SMPTE ST 2110-20 Uncompressed Interlaced Active Video

A.9 : SMPTE ST 2110-20 UNCOMPRESSED INTERLACED ACTIVE VIDEO Transfer Syntax

This Transfer Syntax is used in a DICOM-RTV Metadata Flow in order to describe the accompanying SMPTE ST 2110-20 Video Flow.

The parameters are similar to the ones described in the SMPTE ST 2110-20 Uncompressed Progressive Active Video (Annex A.8), but the frames are interlaced, one frame containing only odd lines and the next one containing only even lines.

A DICOM Transfer Syntax for SMPTE ST 2110-20 Uncompressed Interlaced Active Video shall be identified by a UID value of:

- 1.2.840.10008.1.2.7.2 corresponding to the SMPTE ST 2110-20 Professional Media over IP Networks: Uncompressed Active Video standard for the interlaced video.

Add New Section to Annex A: SMPTE ST 2110-30 PCM Audio

A.10 : SMPTE ST 2110-30 PCM AUDIO Transfer Syntax

This Transfer Syntax is used in a DICOM-RTV Metadata Flow in order to describe the accompanying SMPTE ST21110-30 Audio Flow.

DICOM attributes:

- Number of Waveform Channels (003A,0005) is limited to 15
- Number of Waveform Samples (003A,0010) is limited to 96 (1ms at 96 kHz)
- Sampling Frequency (003A,001A) shall either be 44100, 48000 or 96000
- Waveform Bits Stored (003A,021A) shall either be 16 or 24
- Waveform Bits Allocated (5400,1004) shall either be 16 or 24
- Waveform Sample Interpretation (5400,1006) shall either be US, SS or OB

Table A.10-1: ST 2110-30 and DICOM sampling frequency

ST 2110-30 Sampling Frequency	Sampling frequency (0003,001A)
44.1 kHz	44100
48 kHz*	48000
96 kHz	96000

* 48 kHz is recommended by SMPTE

Table A.10-2: Waveform Sample Interpretation

Bit Depth	Waveform Bits Stored (003A,021A)	Waveform Bits Allocated (5400,1006)	Waveform Sample Interpretation (5400,1006)	Wave Sample Interpretation meaning
16	16	16	SS	signed16-bit linear
16	16	16	US	unsigned16-bit linear
24	24	24	OB	24 bit linear

Table A.10-3: Example of Number of Waveform Samples for 48kHz for basic Audio (Mono or Stereo)

Bit Depth	Waveform Bits Stored (003A,021A)	Numbers of Waveform Channels (003A,0005)	Number of Waveform Sample (003A,0010)	Resulting packet Length (1ms)
16	16	1	48	96
24	24	1	48	144
16	16	2	48	192
24	24	2	48	288

SMPTE ST 2110-30 restricts the audio Flow:

- Sampling frequency is either 44.1 kHz, 48 kHz or 96 kHz, 48 kHz being the recommended value
- Coding scheme is either L16 (16-bit linear) or L24 (24-bit linear)
- Packet time should be 1ms (but could get down to 125 µs)
- Number of Waveform Channels is limited to 15

A DICOM Transfer Syntax for SMPTE ST 2110-30 PCM Digital Audio shall be identified by a UID value of:

- 1.2.840.10008.1.2.7.3 corresponding to the SMPTE ST 2110-30 Professional Media over IP Networks: PCM Digital Audio.

Changes to NEMA Standards Publication PS3.6-2019c

Digital Imaging and Communications in Medicine (DICOM)

Part 6: Data Dictionary

Amend Table 6-1. Registry of DICOM Data Elements

Table 6-1. Registry of DICOM Data Elements

Tag	Name	Keyword	VR	VM	
…					
(0034,0001)	Flow Identifier Sequence	FlowIdentifierSequence	SQ	1	
(0034,0002)	Flow Identifier	FlowIdentifier	OB	1	
(0034,0003)	Flow Transfer Syntax UID	FlowTransferSyntaxUID	UI	1	
(0034,0004)	Flow RTP Sampling Rate	FlowRTPSamplingRate	UL	1	
(0034,0005)	Source Identifier	SourceIdentifier	OB	1	
(0034,0007)	Frame Origin Timestamp	FrameOriginTimestamp	OB	1	
(0034,0008)	Includes Imaging Subject	IncludesImagingSubject	CS	1	
(0034,0009)	Frame Usefulness Group Sequence	FrameUsefulnessGroupSequence	SQ	1	
(0034,000A)	Real-Time Bulk Data Flow Sequence	RealTimeBulkDataFlowSequence	SQ	1	
(0034,000B)	Camera Position Group Sequence	CameraPositionGroupSequence	SQ	1	
(0034,000C)	Includes Information	IncludesInformation	CS	1	
(0034,000D)	Time of Frame Group Sequence	TimeOfFrameGroupSequence	SQ	1	

Amend Table 7-1. Registry of DICOM File Meta Elements

Table 7-1. Registry of DICOM File Meta Elements

Tag	Name	Keyword	VR	VM	
…					
(0002,0031)	RTV Meta Information Version	RTVMetaInformationVersion	OB	1	
(0002,0032)	RTV Communication SOP Class UID	RTVCommunicationSOPClassUID	UI	1	
(0002,0033)	RTV Communication SOP Instance UID	RTVCommunicationSOPInstance UID	UI	1	
(0002,0035)	RTV Source Identifier	RTVSourceIdentifier	OB	1	
(0002,0036)	RTV Flow Identifier	RTVFlowIdentifier	OB	1	
(0002,0037)	RTV Flow RTP Sampling Rate	RTVFlowRTPSamplingRate	UL	1	
(0002,0038)	RTV Flow Actual Frame Duration	RTVFlowActualFrameDuration	FD	1	

Add new Section 9 to Annex A.

9 REGISTRY OF DICOM DYNAMIC RTP PAYLOAD ELEMENTS

Table 9-1. Registry of DICOM Dynamic RTP Payload Elements

Tag	Name	Keyword	VR	VM	
(0006,0001)	Current Frame Functional Groups Sequence	CurrentFrameFunctionalGroupsSequence	SQ	1	

A REGISTRY OF DICOM UNIQUE IDENTIFIERS (UIDS) (NORMATIVE)

Add new UIDs to Annex A.

UID Value	UID Name	UID Type	Part
1.2.840.10008.1.2.7.1	SMPTE ST 2110-20 Uncompressed Progressive Active Video	Transfer Syntax	PS3.5
1.2.840.10008.1.2.7.2	SMPTE ST 2110-20 Uncompressed Interlaced Active Video	Transfer Syntax	PS3.5
1.2.840.10008.1.2.7.3	SMPTE ST 2110-30 PCM Digital Audio	Transfer Syntax	PS3.5
1.2.840.10008.10.1	Video Endoscopic Image Real-Time Communication	SOP Class	PS3.22
1.2.840.10008.10.2	Video Photographic Image Real-Time Communication	SOP Class	PS3.22
1.2.840.10008.10.3	Audio Waveform Real-Time Communication	SOP Class	PS3.22
1.2.840.10008.10.4	Rendition Selection Document Real-Time Communication	SOP Class	PS3.22

Add new UIDs to Annex A. Table A-3. Context Group UID Values:

Context UID	Context Identifier	Context Group Name	Comment
1.2.840.10008.6.1.1303	CID 7070	Real Time Video Rendition Titles	PS3.5

Changes to NEMA Standards Publication PS3.15-2019c

Digital Imaging and Communications in Medicine (DICOM)

Part 15: Security and System Management Profiles

Modify E.1.1 De-identifier:

An Application may claim conformance to an Application Level Confidentiality Profile and Options as a de-identifier if it protects and retains all Attributes as specified in the Profile and Options. Protection in this context is defined as the following process:

...

7. If the Dataset being de-identified is being stored within a DICOM File, then the File Meta Information including the 128 byte preamble, if present, shall be replaced with a description of the de-identifying application. Otherwise, there is a risk that identity information may leak through unmodified File Meta Information or preamble. See PS3.10.

8. If the Dataset being de-identified is being communicated by DICOM Real-Time Video, then the File Meta Information including the 128 byte preamble, if present, shall be replaced with a description of the de-identifying application. Otherwise, there is a risk that identity information may leak through unmodified File Meta Information or preamble. See PS3.22.

Add Table E.1-1. Application Level Confidentiality Profile Attributes:

Attribute Name	Tag	Retd. (from PS3.6)	In Std. Comp. IOD (from PS3.3)	Basic Prof.	Rtn. Safe Priv. Opt.	Rtn. UIDs Opt.	Rtn. Dev. Id. Opt.	Rtn. Inst. Id. Opt.	Rtn. Pat. Chars. Opt.	Rtn. Long. Full Dates Opt.	Rtn. Long. Modif. Dates Opt.	Clean Desc. Opt.	Clean Struct. Cont. Opt.	Clean Graph. Opt.
Flow Identifier Sequence	(0034,0001)	N	Y	D										
Flow Identifier	(0034,0002)	N	Y	D										
Source Identifier	(0034,0005)	N	Y	D										
Frame Origin Timestamp	(0034,0007)	N	Y	D						K	C			

Changes to NEMA Standards Publication PS3.16-2019c

Digital Imaging and Communications in Medicine (DICOM)

Part 16: Content Mapping Resource

Modify Context Group 7010 in Part 16 Annex B DCMR (Normative)

CID 7010 Key Object Selection Document Title

Resources: **HTML | FHIR JSON | FHIR XML | IHE SVS XML**
Type: **Extensible**
Version: **~~20170914~~20190915**
UID: **1.2.840.10008.6.1.490**

Table CID 7010. Key Object Selection Document Title

DCM	128221	Tumor Board Input Used
DCM	128208	For Tumor Registry
DCM	128222	Tumor Registry Input Used
DCM	128207	For Clinical Trial Submission
DCM	128223	Clinical Trial Submission Input Used
Include CID 7023 "RT Process Output"		
Include CID 7024 "RT Process Input"		
Include CID 7025 "RT Process Input Used"		
Include CID 7014 "Export Additional Information Document Titles"		
Include CID 7070 "Real-Time Video Rendition Titles"		

Add Context Group XXX to Part 16 Annex B DCMR (Normative)

CID 7070 Real Time Video Rendition Titles

Type: **Extensible**
Version: **20190915**

Table CID 7070. Real Time Video Rendition Titles

Coding Scheme Designator	Code Value	Code Meaning
DCM	130370	RTV Rendition
DCM	130371	RTV Audio and Video Rendition
DCM	130372	RTV Stereo Video Rendition
DCM	130373	RTV Audio and Stereo Video Rendition

(Add the following definitions to Part 16 Annex D DICOM Controlled Terminology Definitions (Normative)

Code Value	Code Meaning	Definition	Notes
130370	RTV Rendition	DICOM objects communicated in time-synchronized flows using DICOM-RTV, intended for a simultaneous presentation	
130371	RTV Audio and Video Rendition	Audio and Video DICOM objects communicated in time-synchronized flows using DICOM-RTV, intended for a simultaneous presentation	
130372	RTV Stereo Video Rendition	Two Video DICOM objects communicated in time-synchronized flows using DICOM-RTV, intended for a stereo video	
130373	RTV Audio and Stereo Video Rendition	Audio and two Video DICOM objects communicated in time-synchronized flows using DICOM-RTV, intended for a stereo video	

Changes to NEMA Standards Publication PS3.2-2019c

Digital Imaging and Communications in Medicine (DICOM)

Part 2: Conformance

Add New Annex L : Conformance Statement Sample DICOM-RTV Service Provider (Informative)

L Conformance Statement Sample DICOM-RTV Service Provider (Informative)

Disclaimer:

This document is an example DICOM Conformance Statement for a fictional application service called EXAMPLE-RTV-SERVICE produced by a fictional vendor called EXAMPLE-IMAGING-PRODUCTS.

As stated in the annex title, this document is truly informative, and not normative. A conformance statement of an actual product might implement additional services and options as appropriate for its specific purpose. In addition, an actual product might implement the services described in a different manner and, for example, with different characteristics and/or sequencing of activities. In other words, this conformance statement example does not intend to standardize a particular manner that a product might implement DICOM-RTV functionality.

L.0 Cover Page

Company Name: EXAMPLE-IMAGING-PRODUCTS

Product Name: EXAMPLE-RTV-SERVICE

Version: 1.0-rev. A.1

Internal document number: 1024-1960-xx-yy-zz rev 1

Date: YYYYMMDD

L.1 Conformance Statement Overview

This fictional product EXAMPLE-RTV-SERVICE implements the DICOM-RTV services for sending video and associated metadata, to be consumed in real-time by other compliant devices. The EXAMPLE-RTV-SERVICE is only available as a plug in option for the EXAMPLE-INTEGRATED-MODALITY. All of the networking, database, and other services are provided by the EXAMPLE-INTEGRATED-MODALITY. This conformance claim refers to the conformance claim for the EXAMPLE-INTEGRATED-MODALITY for all such services.

Table L.1-1 provides an overview of the network services supported by EXAMPLE-RTV-SERVICE.

Table L.1-1. Network Services

Network Service	User of Service (SCU)	Provider of Service (SCP)
DICOM Real-Time Video (DICOM-RTV)		
DICOM-RTV	No	Yes

L.2 Table of Contents

A table of contents shall be provided to assist readers in easily finding the needed information.

L.3 Introduction

L.3.1 Revision History

Table L.3.1-1. Revision History

Document Version	Date of Issue	Author	Description
1.1	March 8th, 2018	ECR	Initial version for PC

L.3.2 Audience, Remarks, Terms and Definitions, Basics of DICOM Communication, Abbreviations, References

See example text in Section A.3.

L.3.3 Additional Remarks for This Example

This document is a sample DICOM Conformance Statement created for DICOM PS3.2. It is to be used solely as an example to illustrate how to create a DICOM Conformance Statement for a DICOM-RTV Service Provider. The subject of the document, EXAMPLE-RTV-SERVICE, is a fictional product.

L.4 Networking

L.4.1 Implementation Model

L.4.1.1 Application Data Flow

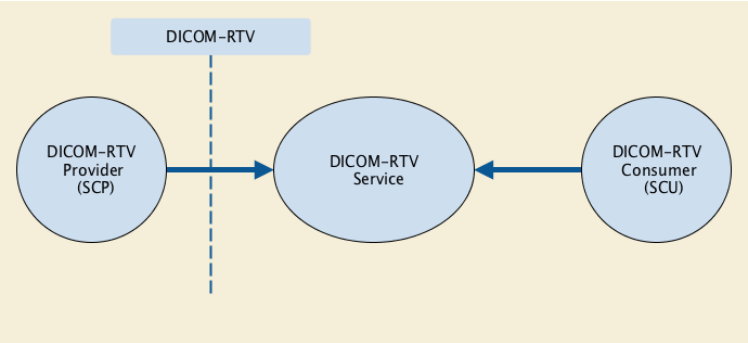


Figure X.4.1-1. Application Data Flow Diagram

The DICOM-RTV Service Application provides multiple DICOM-RTV compliant Flows, transported in RTP over IP, that can be consumed by one or multiple other DICOM-RTV Service Application(s).

L.4.1.2 Functional Definition of AEs

L.4.1.2.1 Functional Definition of RTV Service Application

The DICOM-RTV Service is Active when the equipment produces video content.

L.4.2 AE Specifications

This AE complies with Section 6.2 “TRANSPORT” in PS3.22, specification for DICOM-RTV.

L.4.2.1 DICOM-RTV Application Entity Specifications

L.4.2.1.1 SOP Classes

EXAMPLE-RTV-SERVICE provides Standard Conformance to the following SOP Classes:

Table L.4.2-1. SOP Classes for DICOM-RTV AE

SOP Class Name	SOP Class UID	SCU	SCP
Video Photographic Image Real-Time Communication	1.2.840.10008.10.2	No	Yes

Some restrictions applies on the Real-Time Communications:

Table L.4.2-2. DICOM-RTV Instances Specification

Category	Restrictions
Transfer Syntaxes Supported	SMPTE ST 2110-20 Uncompressed Progressive Active Video
Photometric interpretation	RGB
Bit depth	10

Table L.4.2-3. DICOM-RTV Screen Resolutions

Rows	Columns	Frame rate	Video Type	Progressive or Interlaced
1080	1920	25	25 Hz HD	P
1080	1920	29.97, 30	30 Hz HD	P
1080	1920	25	25 Hz HD	I
1080	1920	29.97, 30	30 Hz HD	I
720	1280	25	25 Hz HD	P
720	1280	29.97, 30	30 Hz HD	P
720	1280	50	50 Hz HD	P
720	1280	59.94, 60	60 Hz HD	P

The resolution is defined by the equipment configuration, and is reflected in the SDP object.

L.4.2.2.4 Connection Policies

L.4.2.2.4.1 General

The consumer shall get the SDP object on the following URL: http://<local-IP-address-of-the-device>/SDP.

L.4.2.2.4.2 Number of Connections

EXAMPLE-RTV-SERVICE is provided in multicast. The limit of simultaneous connection depends on the local network infrastructure.

L.4.3 Network Interfaces

L.4.3.1 Physical Network Interface

EXAMPLE-RTV-SERVICE uses the network interface from the hosting EXAMPLE-INTEGRATED-MODALITY. See its conformance claim for details.

L.4.3.2 Additional Protocols

EXAMPLE-RTV-SERVICE uses the network services from the hosting EXAMPLE-INTEGRATED-MODALITY. See its conformance claim for details.

L.4.3.3 IPv4 and IPv6 Support

This product supports both IPv4 and IPv6 connections.

L.4.4 Configuration

L.4.4.1 DICOM-RTV Interface

The EXAMPLE-RTV-SERVICE is configured to define the following parameters expressed in the SDP object:

- the RTP payload type (PT) used for the video is 96
- the RTP payload type (PT) used for DICOM-RTV Metadata is 104.

L.5 Media Interchange

Not applicable.

L.6 Support of Character Sets

All EXAMPLE-RTV-SERVICEs support Unicode UTF-8 for all communications.

L.7 Security

Has to be managed at the individual sites and installations.

L.8 Annexes

L.8.1 IOD Contents

See conformance claim for the EXAMPLE-INTEGRATED-MODALITY. The modules and fields contained in the DICOM-RTV metadata are reflecting the values of the corresponding ones in the EXAMPLE-INTEGRATED-MODALITY X-Ray Radiofluoroscopic Image Storage IOD.

L.8.2 Data Dictionary of Private Attributes

No private attributes are used.

L.8.3 Coded Terminology and Templates

See conformance claim for EXAMPLE-INTEGRATED-MODALITY.

L.8.4 Standard Extended / Specialized / Private SOP Classes

Not Applicable.

L.8.5 Private Transfer Syntaxes

Private transfer syntaxes are not supported.

Add New Annex M : Conformance Statement Sample DICOM-RTV Service Provider (Informative)

M Conformance Statement Sample DICOM-RTV Service Consumer (Informative)

Disclaimer:

This document is an example DICOM Conformance Statement for a fictional application service called EXAMPLE-RTV-DISPLAY produced by a fictional vendor called EXAMPLE-Viewing PRODUCTS.

As stated in the annex title, this document is truly informative, and not normative. A conformance statement of an actual product might implement additional services and options as appropriate for its specific purpose. In addition, an actual product might implement the services described in a different manner and, for example, with different characteristics and/or sequencing of activities. In other words, this conformance statement example does not intend to standardize a particular manner that a product might implement DICOM-RTV functionality.

M.0 Cover Page

Company Name: EXAMPLE-Viewing PRODUCTS

Product Name: EXAMPLE-RTV-DISPLAY

Version: 1.0-rev. A.1

Internal document number: 1024-1960-xx-yy-zz rev 1

Date: YYYYMMDD

M.1 Conformance Statement Overview

This fictional product EXAMPLE-RTV-DISPLAY implements the DICOM-RTV services for consuming video, audio and associated metadata, provided by another compliant device, and displaying the information in a window on the screen. The EXAMPLE-RTV-DISPLAY is only available as a plug in option for the EXAMPLE-INTEGRATED-MODALITY. All of the networking, database, and other services are provided by the “SAMPLE DICOM Image Viewer”. This conformance claim refers to the conformance claim for the “SAMPLE DICOM Image Viewer” for all such services.

Table M.1-1 provides an overview of the network services supported by EXAMPLE-RTV-DISPLAY.

Table M.1-1. Network Services

Network Service	User of Service (SCU)	Provider of Service (SCP)
DICOM Real-Time Video (DICOM-RTV)		
DICOM-RTV	Yes	No

M.2 Table of Contents

A table of contents shall be provided to assist readers in easily finding the needed information.

M.3 Introduction

M.3.1 Revision History

Table M.3.1-1. Revision History

Document Version	Date of Issue	Author	Description
1.1	March 8th, 2018	ECR	Initial version for PC

M.3.2 Audience, Remarks, Terms and Definitions, Basics of DICOM Communication, Abbreviations, References

See example text in Section A.3.

M.3.3 Additional Remarks for This Example

This document is a sample DICOM Conformance Statement created for DICOM PS3.2. It is to be used solely as an example to illustrate how to create a DICOM Conformance Statement for a DICOM-RTV Service Provider. The subject of the document, EXAMPLE-RTV-DISPLAY, is a fictional product.

M.4 Networking

M.4.1 Implementation Model

M.4.1.1 Application Data Flow

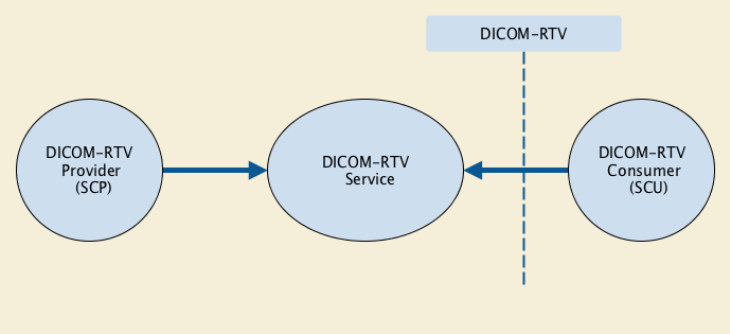


Figure M.4.1-1. Application Data Flow Diagram

The DICOM-RTV Service Application consumes one or multiple DICOM-RTV compliant Flows, transported in RTP over IP, that is/are provided by one other DICOM-RTV Service Application.

M.4.1.2 Functional Definition of AEs

M.4.1.2.1 Functional Definition of RTV Service Application

The DICOM-RTV Service is Active when the real-time display feature of the equipment is running and some video and/or audio content is provided.

M.4.2 AE Specifications

This AE complies with Section 6.2 “TRANSPORT” in PS3.22, specification for DICOM-RTV.

M.4.2.1 DICOM-RTV Application Entity Specifications

M.4.2.1.1 SOP Classes

EXAMPLE-RTV-SERVICE provides Standard Conformance to the following SOP Classes:

Table M.4.2-1. SOP Classes for DICOM-RTV AE

SOP Class Name	SOP Class UID	SCU	SCP
Video Photographic Image Real-Time Communication	1.2.840.10008.10.2	Yes	No
Audio Waveform Real-Time Communication	1.2.840.10008.10.3	Yes	No
Rendition Selection Document Real-Time Communication	1.2.840.10008.10.4	Yes	No

Some restrictions applies on the Real-Time Communications:

Table M.4.2-2. DICOM-RTV Instances Specification

Category	Restrictions
Transfer Syntaxes Supported	SMPTE ST 2110-20 Uncompressed Progressive Active Video, SMPTE ST 2110-30 PCM Digital Audio
Photometric interpretation	RGB
Bit depth (video)	10
Number of Waveform Channels	2
Bit depth (audio)	16 (signed 16-bits linear)
Sampling Frequency	48 kHz

Table M.4.2-3. DICOM-RTV Screen Resolutions

Rows	Columns	Frame rate	Video Type	Progressive or Interlaced
1080	1920	25	25 Hz HD	P
1080	1920	29.97, 30	30 Hz HD	P
1080	1920	25	25 Hz HD	I
1080	1920	29.97, 30	30 Hz HD	I
720	1280	25	25 Hz HD	P
720	1280	29.97, 30	30 Hz HD	P
720	1280	50	50 Hz HD	P
720	1280	59.94, 60	60 Hz HD	P

The resolution is automatically determined based on the one provided by the sent video.

M.4.2.2.4 Connection Policies

M.4.2.2.4.1 General

The URL to be accessed by the equipment to get the SDP object is set by configuration.

M.4.2.2.4.2 Number of Connections

EXAMPLE-RTV-DISPLAY is consuming multicast communication.

M.4.3 Network Interfaces

M.4.3.1 Physical Network Interface

EXAMPLE-RTV-DISPLAY uses the network interface from the hosting “SAMPLE DICOM Image Viewer”. See its conformance claim for details.

M.4.3.2 Additional Protocols

EXAMPLE-RTV-DISPLAY uses the network services from the hosting “SAMPLE DICOM Image Viewer”. See its conformance claim for details.

M.4.3.3 IPv4 and IPv6 Support

This product supports both IPv4 and IPv6 connections.

M.4.4 Configuration

M.4.4.1 DICOM-RTV Interface

The EXAMPLE-RTV-DISPLAY uses the network parameters (IP, port…) defined in the SDP.

M.5 Media Interchange

Not applicable.

M.6 Support of Character Sets

EXAMPLE-RTV-DISPLAY supports only Unicode UTF-8 for all communications.

M.7 Security

Has to be managed at the individual sites and installations.

M.8 Annexes

M.8.1 IOD Contents

Not Applicable.

M.8.2 Data Dictionary of Private Attributes

No private attributes are used.

M.8.3 Coded Terminology and Templates

Not Applicable.

M.8.4 Standard Extended / Specialized / Private SOP Classes

Not Applicable.

M.8.5 Private Transfer Syntaxes

Private transfer syntaxes are not supported.