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: Draft Letter Ballot, November 8, 2018
Developed in accordance with: DICOM Workitem 2016-12-D

Table of Contents

**** *Editorial content – to be removed before Final Text* ****

TODO:

Editor's Notes

External sources of information

Editorial Issues and Decisions

#	Issue	Status

Closed Issues

#	Issues
1	Name of the supplement ➔ ("Real-Time Video" proposed).
2	Do we specify use case(s) and which level of detail? ➔ Some typical use cases at high level and one example with more details.
3	Do we embrace also reconstructed medical imaging (e.g., live US, live RF) or only (visible light) video? ➔ Visible light only at this stage.
4	How shall we deal with proper understanding and proper referencing of SMPTE/VSF documents ➔ Reference and some high level summary, as well as some examples.
5	How shall we proceed with the medical metadata, either using a VSF/SMPTE defined mechanism or a pure RTP one, respecting the classical DICOM encoding? ➔ The solution consists in conveying DICOM Datasets using SMPTE ST 2110-10 mechanism.
7	Selection of metadata to be conveyed and why (justified based on the use cases). ➔ Be very selective. Limit for the moment the metadata of the existing Video IOD. Which frequency for sending the metadata (every frame?). ➔ Persistent metadata at least every second. Changing metadata every frame or sample.
8	Is there a mechanism to register (in SMPTE or others) for a domain specific options? ➔ No available static RTP Payload Type.
9	Shall we define a new Service-Object Pair for existing IODs, since the service is new? ➔ Yes we defined new IODs from existing ones (adding a "Real-time" in the name).
10	Shall we document more the difference between reference times (UTC vs. TAI) and how to deal with potential conversion between them? ➔ Offer the possibility to use either UTC or TAI but no explanation on how to perform the conversion from one to the other.
11	Should the supplement include an existing IOD example to understand which kind of extension is needed? ➔ Several video IODs and one audio IOD are documented.
12	Should the supplement include a new video IOD to understand how we would design such IODs if unconstrained? ➔ Yes (see above)
13	What are the storage semantics if any of this streaming service? For example, is the storage of described IOD implied by streaming? Or is it a separate subsequent step by the SCU/SCP? Or should it be an explicit parameter of the streaming service? ➔ The storage mechanism is not described in the Real-Time flow. However, all the information necessary for creating the storage IOD is contained within the flow but some additional "decisions" (start/end time, transfer syntax) will be required.
14	Do we address the archiving feature? ➔ NO, out of the scope of this supplement but adding one explanatory section.
15	Do we make Identifiers required in the SMPTE flows? ➔ Yes, they are mandatory in the DICOM-RTV Metadata flow and recommended only in the video/audio flows, since they are not even described in the present ST 2110-xx family of standards. It may be revisited if SMPTE proposes some mechanisms for pairing the flows through identifiers. For the moment, in case of absence of identifiers in the video/audio flow, the matching shall rely on the mandatory SDP object.

16	The reliability of the metadata on a real-time stream is not guaranteed (e.g. having a wrong value in a DICOM field due packet corruption). Do we propose a mechanism which manages integrity if there is an available one? ➔ NO, after verification, the level of errors on such network is 10^{-13} and SMPTE ST 2110-10 considers the network is reliable.
17	Include update of the Basic Voice Audio IOD Modules? ➔ No since a new audio IOD is created.
18	Draft a new IOD that contains only the Patient, Study, Equipment, Synchronization, Series modules that you actually need for the first frame (shared) and another IOD for what you actually need for each subsequent frame (of anything), which doesn't replicate what is already known from the stream's own metadata (like timing unless there is something from the module). ➔ No, the other option has been selected, i.e.; creating new IODs containing both persistent tags and changing tags (through a dedicated macro).
19	Have we to insert the RTP Timestamp in the metadata, in order to use it (instead of the Origin Time Stamp) for synchronizing two flows which were recorded and replayed? ➔ Origin Time Stamp will enable to re-synchronize two flows and in any case RTP Timestamp is thrown away when recorded.
20	Potential patent on storing medical metadata separately from the video. ➔ No, not identified at the moment.
21	Shall we differentiate “static” content vs. “dynamic” content in order to optimize size of transmitted data? ➔ Yes, transmitting “static” content only every second vs. “dynamic” content every frame/sample.
22	Shall we define a required or recommended mechanism to access the SDP object enabling to start the DICOM-RTV session? ➔ SDP object is linked to the SMPTE ST 2110-10 standard and DICOM shall continue to rely on it. Several mechanisms exist already. If SMPTE proposes a mechanism, DICOM could adopt it but DICOM has not to define it.
23	How do we manage the situation when the “static” metadata is changing? Shall we add another type of header containing the “static” attributes that may change (e.g. another side of the patient is concerned) along the time (e.g., series number/UID because the nature of the video content has changed)? Or do we start a new stream? ➔ When a change happens in the “static” metadata, the sender shall include the “static” metadata in the payload and create a new SOP Instance UID, while keeping the same Source and Flow Identifier.
24	Have we to document the real-time transmission of stereo video, either using a simple stereo flow or using two different flows with a synchronization mechanism. The particular case of stereo vision, may either be solved by combining the contents into a single flow (Multiview video Coding) or by separating contents into single flows (left content apart from right content) and then pairing them by using a Rendition. Synchronization is defined by SMPTE ST 2110-10 through PTP. Rendition IOD is defined in A.XX ➔ see the informative section on it
25	How to deal with the situation where a flow is created from the extraction of a previous flow? Should the Frame Extraction Module be applicable (or adapted, or replaced) ➔ out of the scope of this supplement. Could be addressed after the recording use case has been explored
26	How can we manage Transfer Syntax UID, to describe the different kinds of Bulk Data, one or multiple for the video and for the audio complying with SMPTE ST 2110-20 and 30, respectively, depending on the some options? ➔ three Transfer Syntaxes are documented, with corresponding UIDs, two for the progressive and interlaced video, respectively, and one for the audio.
27	Do we need to extend some existing constraints in PS 3-3 and PS 3-5 to support up 16 bits for color resolution ➔ should be addressed separately from this supplement since it concerns also the HEVC/H.265 Transfer Syntax.
28	How to convey new information brought by new HDR (High Dynamic Range) standards such as ITU BT.2100, BT.709, BT.2020 and how does it relate with ICC profiles in DICOM? ➔ to be solved with the color resolution topic (see #27).
29	How to describe interlaced vs progressive flows (may be related to #26) ➔ beyond Transfer Syntax UID (see #26), a conditional Functional Group Macro is defined for specifying the parity of the frame lines.
30	Do we have to customize the TID 2010 Key Object Selection to extend the values in order to reflect the new usages. Different propositions: 1°) Duplication of TID dedicated to Rendition and then creating new codes, 2°) Extension of the existing TID (with "RTV Rendition" as a possible Document title) and then use "Key Object Description" to document the Rendition ➔ The Option 2° has been retained
31	Shall we exclude the Waveform Annotation module? ➔ Yes, We propose to exclude the Waveform Annotation module from Real-Time Audio Waveform IOD, because it is too specific of waveforms and a more general framework for real-time annotation should be preferred. If we want to annotate the waveform, we will have a separate flow for that.
32	DICOM-RTV is intended to convey dynamic parameters along with the data flow (but Video Endoscopic, Microscopic and Photographic IODs do not seem to require such specific dynamic parameters). WG13 may further analyze whether optical parameters (such as zoom factor, focal position, aperture) should be

	considered. ➔ the supplement includes some Real-Time Video Image Macros.
33	Replay of DICOM-RTV flows shall be possible and is mentioned in the XX.7 "Storage Consideration". So far it is not described in this document. ➔ see #25.
34	Address the storage of audio independently of the video, creating a new audio IOD for the purpose? ➔ nothing forbids to record the sound alone.
35	How to manage proprietary tags (e.g., in the RTV Meta Information) ➔ the mechanism enabling private attributes and Private Functional Groups applies also for DICOM-RTV metadata.
36	Insert Frame Type in Stereo Pair in the Current Frame Functional Group Macro to document if the current frame is corresponding to Left or Right ➔ this information is included in the video flow
37	Insert a section in Part 5 to describe how to encode UUID in binary? ➔ out of the scope of the present supplement. Could be addressed elsewhere (CP…).
38	Reference standards for describing ICC Profiles if exist? ➔ see #27.
39	Instead of "duplicating" existing video IODs, create only one for all the Real-Time Video types ➔ approach that has been retained.
40	Remove the Frame Origin Time Stamp from the Frame Functional Group Macro since it is contained in the RTP Header Extension which is mandatory for DICOM Metadata flow. ➔ no, to be able to record it later for further replay.
41	How compact the frame related information shall be and so shall we use another mechanism more compact for conveying frame based information? ➔ the main purpose of the standard is to convey video and compare to the size of video information, size of metadata is negligible. If, in the future, DICOM-RTV is deployed also for conveying the signal only, the topic could be revisited.
42	Shall we make mandatory that the DICOM metadata is exactly synchronized with the video, e.g., if a frame is dropped from the video flow, the corresponding metadata will be dropped from the DICOM Metadata flow. ➔ no, the metadata can contain information that makes sense, even in the absence of the corresponding frame (calculation of interpolation of 3D position for example), so the receiver will take the decision on what to do with the metadata without corresponding video frame, including ignoring metadata.
43	To avoid to duplicate information that is contained with the SDP object, with risk on inconsistency, the parameters describing the image (rows, columns…), the Image Pixel and the Waveform Modules are not included in the IODs. Shall it include them however, since the information is not contained in the media flows themselves? ➔ this option minimizes the risk of inconsistency but increases the complexity the recording which will require to access the SDP object to obtain the information on Image and Waveform. The supplement documents the information which is normally contained in the DICOM dataset.
44	For the purpose of public comment, the PS3.5 Transfer Syntax definition is described in terms of Image Pixel Data Module Attributes, even though that Module is not sent in the stream; input is sought on how best to describe the constraints in terms of SDP terminology. Shall we maintain this approach? ➔ using the same kind of definition will help the developer familiar with DICOM to adopt the new standard. It will also help to convert DICOM-RTV flows in DICOM Video IODs for storage.
45	Do we need to narrow the content of the Rendition Document, e.g., restrict it to referring to IMAGE / WAVEFORM, or keep all the possibilities of the Key Object Document template (TID 2010), i.e., IMAGE / WAVEFORM / COMPOSITE ? ➔Two CID have been defined and they may be completed by additional one(s) if necessary in the future, through Change Proposals.
46	Which other standards for the Time Distribution to add (e.g., GPS)? ➔ out of the scope of the supplement. The DICOM-RTV enabled equipment will rely on a PTP server which may be updated through GPS but not the Time Distribution from GPS will not go directly to the DICOM-RTV enabled equipment.

Open Issues

#	Issues	Status

******** ***End of Editorial content – to be removed before Final Text*** ********

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.

135
136
137
138
139
140
141

142 **Changes to NEMA Standards Publication PS 3.17-20xx**

143 **Digital Imaging and Communications in Medicine (DICOM)**

144 **Part 17: Explanatory Information**

145
146

XX Real-Time Video Use Cases (Informative)

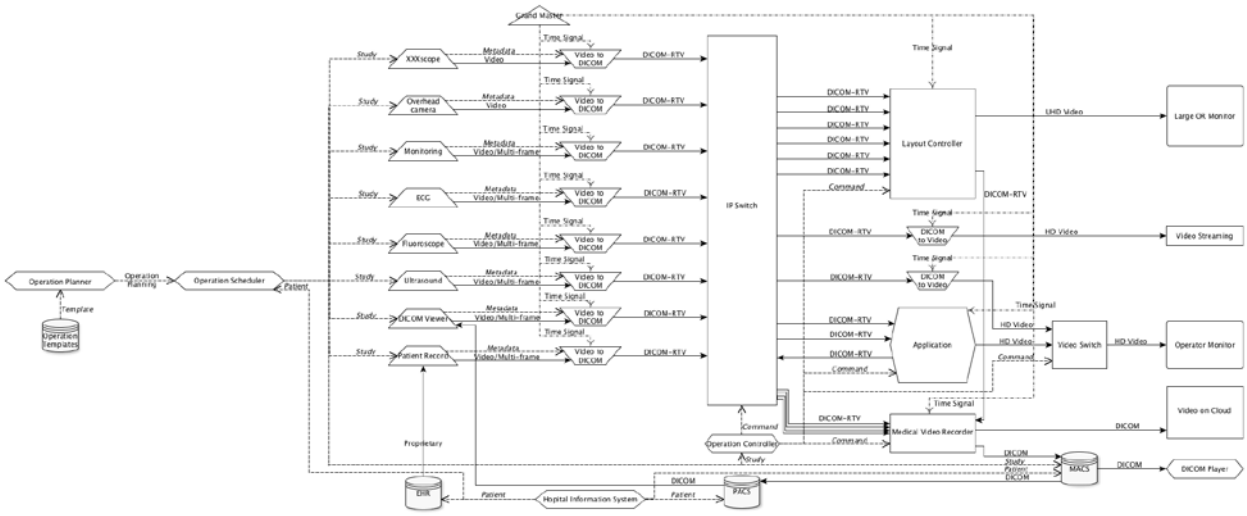


Figure XX.0-1: Overview diagram of operating room

As shown on Figure XX.0-1, 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.

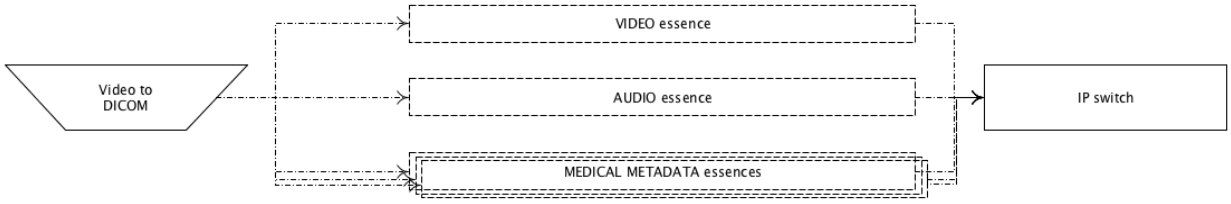


Figure XX.0-2: Real-Time Video stream content overview

As shown on Figure XX.0-2, 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 XX.0-3, which is the network view of the real-time streaming.

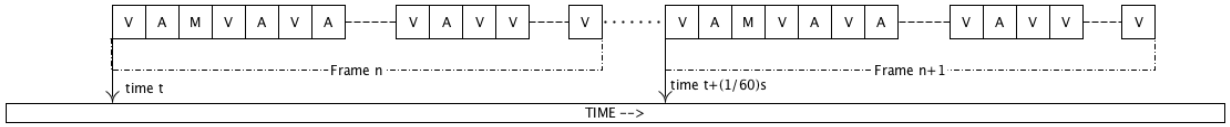


Figure XX.0-3: Real-Time Video transmission details

XX.1 USE CASE 1: 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).

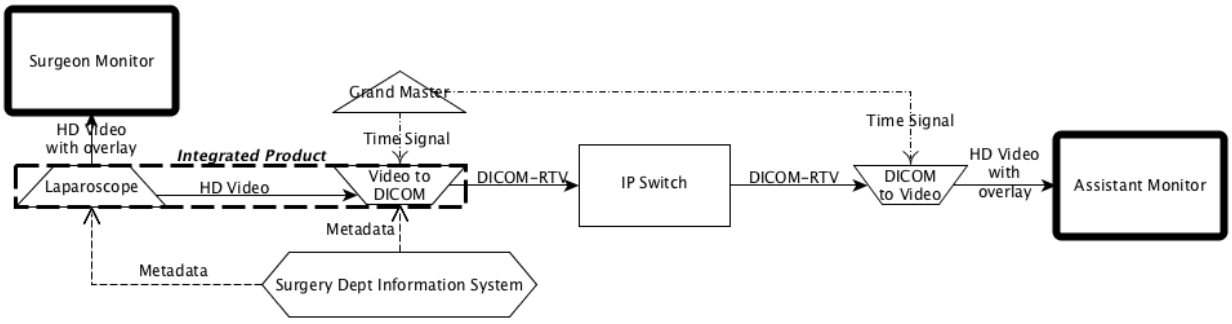


Figure XX.1-1: Duplicating on additional monitor

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 XX.1-1) 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. 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.

XX.2 USE CASE 2: 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.

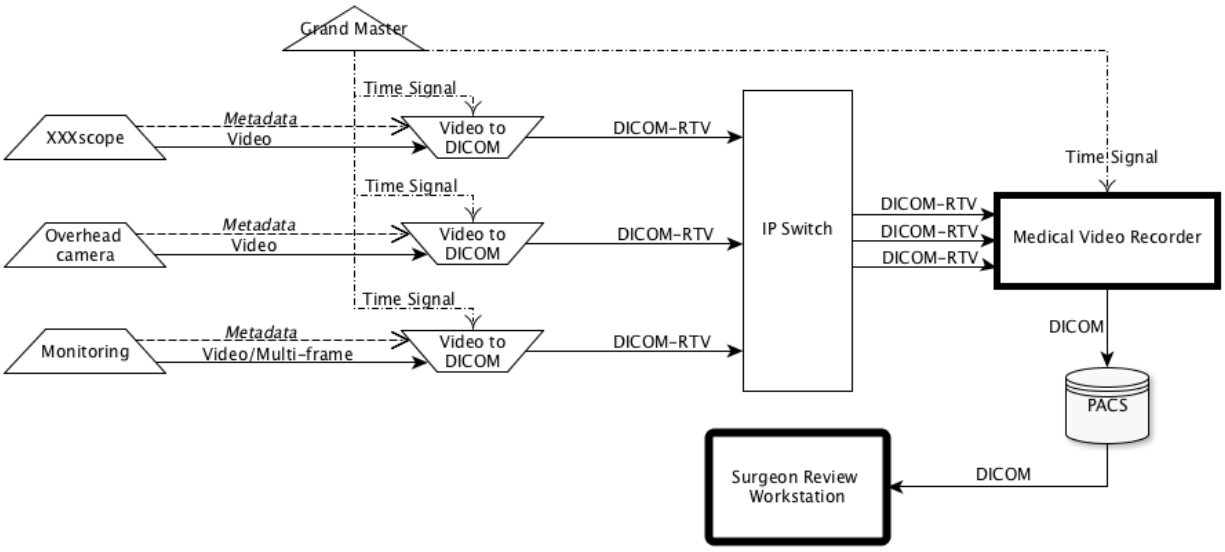


Figure XX.2-1: Recording multiple video sources

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 XX.2-1). In this case, the metadata is coming directly from each device.. The recording has to maintain time consistency between the different video channels. Section XX.7.1 describes how the video could be captured and stored as a DICOM IOD using the present DICOM Store Service, as shown on the Figure XX.2-1, 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 XX.2-1. They could also be played back using DICOM-RTV as described in section XX.7.2.

XX.3 USE CASE 3: AUTOMATIC DISPLAY IN OPERATING ROOM (OR)

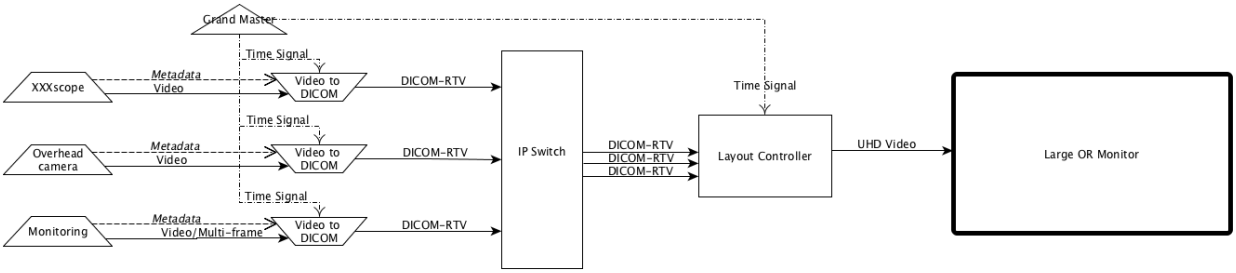


Figure XX.3-1: Displaying multiple source on one unique monitor

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 XX.3-1. 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.

XX.4 USE CASE 4: AUGMENTED REALITY

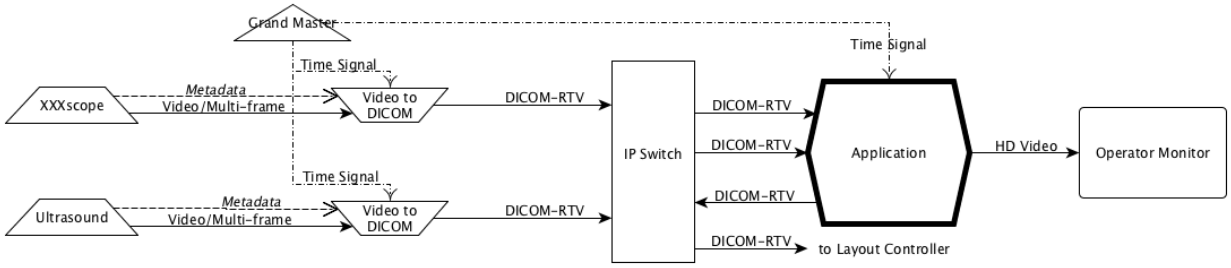


Figure XX.4-1: Application combining multiple real-time video sources

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 XX.4-1). 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.

XX.5 USE CASE 5: 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.

XX.6 EXAMPLE OF DICOM REAL-TIME VIDEO IMPLEMENTATION

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

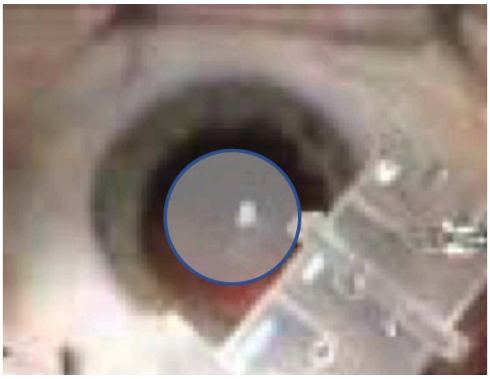


Figure XX.6-1: Example of implementation for Augmented reality based on optical image

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 XX.6-1.

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 XX.6-2, 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.

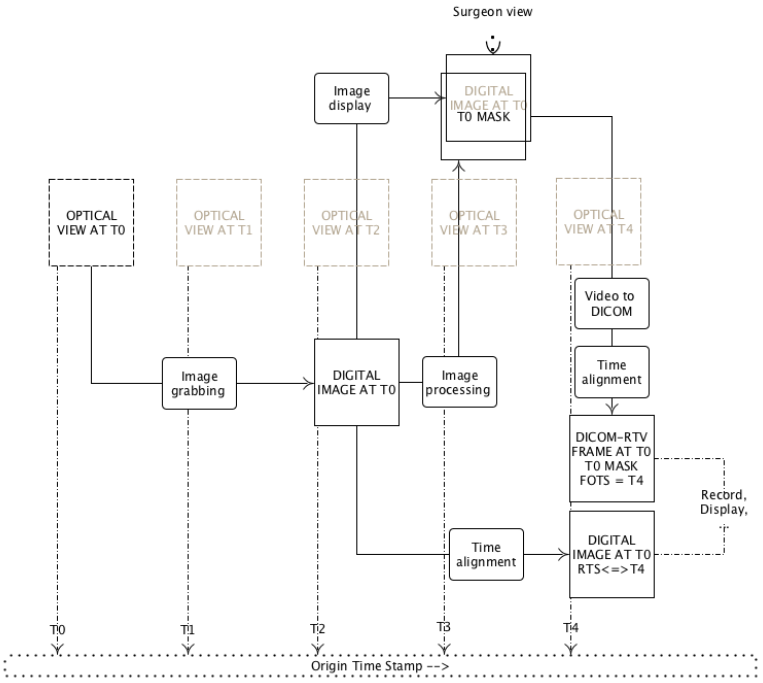


Figure XX.6-2: Example of implementation for Augmented reality based on optical image

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 XX.6-3.

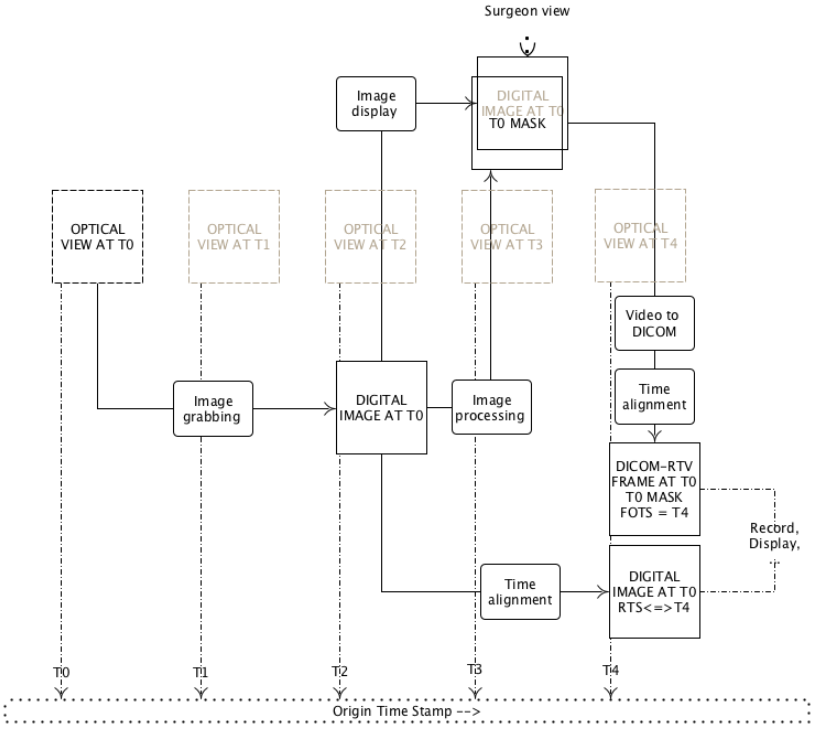


Figure XX.6-3: Example of implementation for Augmented reality based on digital image

XX.7 STORAGE CONSIDERATION

XX.7.1 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).

XX.7.2 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 XX.7.1.

- 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.

XX.8 EXAMPLE OF ENGINEERING IMPLEMENTATION

An example of implementation of the Video-to-DICOM converter presented in the use cases XX.1 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 (DICOM-RTV Meta Information) as well as dynamic payload (DICOM Current Frame Functional Groups Module) for every frame, and including additionally the static payload (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 (DICOM-RTV Meta Information) as well as dynamic payload (DICOM Current Frame Functional Groups Module) for every sample, and including additionally the static payload (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 (DICOM Rendition 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 XX.1-1.

XX.9 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.X.X Stereoscopic 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.

YY 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 YY-1 below:

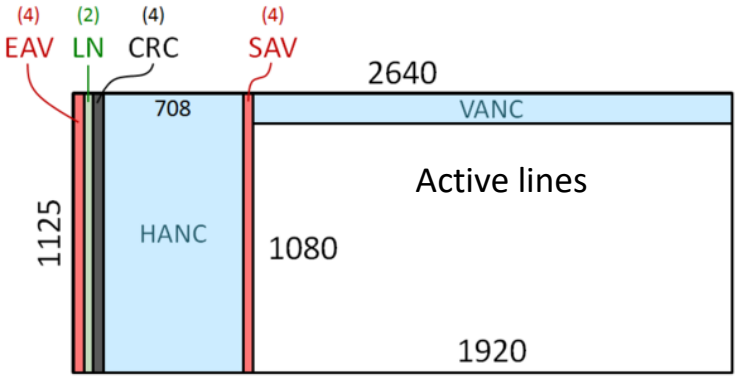


Figure YY-1 Structure of a High Definition SDI signal

Note

Acronyms on the Figure YY-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.

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:

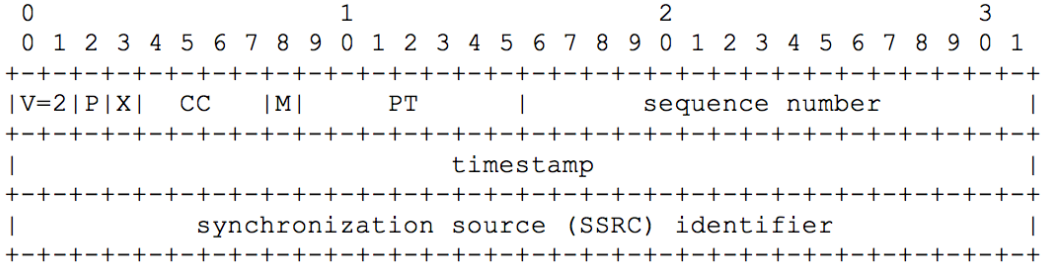


Figure YY-2 RTP Header

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:

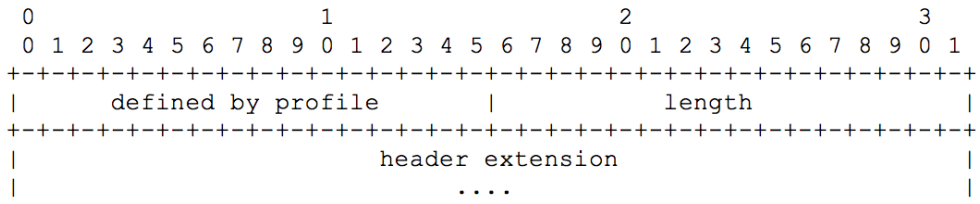


Figure YY-3 RTP Header Extension

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:

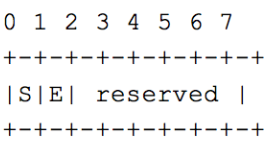


Figure YY-4 RTP Grain Flags

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.

509
510
511
512
513

514 **Add a new NEMA Standards Publication PS 3.X-20xx**

515 **Digital Imaging and Communications in Medicine (DICOM)**

516 **Part X: 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 Conformance

An implementation claiming conformance to PS3.X 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.X. The structure of Conformance Statements is specified in PS3.2.

3 Normative References

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

[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*

4 Terms and 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.

DICOM-RTV Service Element

Essence	Video, audio or data, typically intended for consumption, 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].

5 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

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).

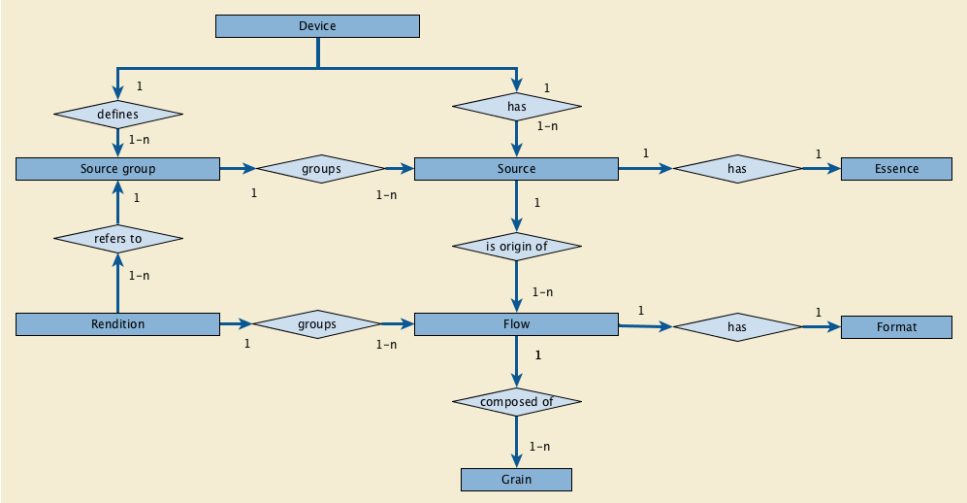


Figure 6-1. Real World diagram of DICOM-RTV

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. Such 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.

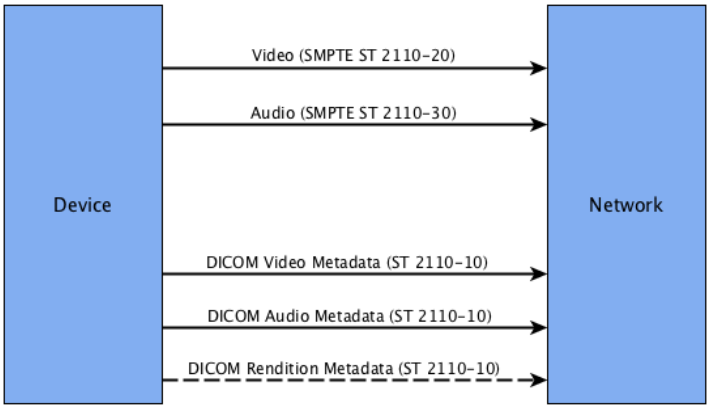


Figure 6-2. Interaction Diagram

SMPTE ST 2110-10 provides end-to-end network transport functions for applications transmitting real-time data. End systems generate the content in RTP packets and/or consume the content of received RTP

packets respecting SMPTE ST 2110-10. An end-system can act as one or more Sources in a particular RTP session.

A device can provide and/or consume content. A device that provides content has one or more Sources that can be of various Essences (e.g., Video and Audio). A Source is the origin of one or more Flows. Flows coming from the same Source are the representation of the same content in different resolutions and coding. A device provides content whether the presence or not of a device consuming the Flow(s) related to the Source(s) of this sending device. A device that consumes content can subscribe/unsubscribe to available Flows without the need of any session.

A DICOM Metadata flow, describing the context and content of the Flow, shall be associated to each of these Flows. However the same 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, referencing all Flows provided by one device may also be transported.

6.2 TRANSPORT

6.2.1 RTP Header

All Essences shall be transported with RTP respecting SMPTE ST 2110-10 which requires that each Flow is described by a SDP object which specifies its content as well as connection details enabling the receiver to start 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 only way for associating multiple Essences belonging to a same stream. The presence of such information in the SDP implies that it is contained in the RTP Extended Header present in the first packet of a grain (video frame, audio sample, metadata set…). It makes it possible to automatically associate and temporarily synchronize two Flows from their content.

By definition, all the Flows respecting SMPTE ST 2110-10 are synchronized and associated together very precisely by means of a common reference to the Universal Time, using PTP mechanism, with a nanoseconds precision.

The RTP header, for the video and the audio data Essences, shall respect the SMPTE ST 2110-20 and ST 2110-30, respectively.

The RTP header, for the DICOM Metadata Essence, shall respect the SMPTE ST 2110-10. As an application of the principles defined in SMPTE ST 2110-10, it shall contain the information above. The clock rate shall be identical to the one defined for referenced Essence (e.g., the clock rate of the DICOM Video Metadata shall be identical to the one of the video Essence it refers). The following additional constraints apply:

extension (X): 1 bit	Shall be set to 1.
marker (M): 1 bit	Is set to 1 when the RTP packet is carrying the last Essence of the dataset. It shall be set to 0 for all other packets.
payload type (PT)	The payload type belongs to range 96-127. The payload type number cannot conflict with any other currently active Essences. Since SMPTE does not allow static number to be reserved, applications must be prepared to dynamically allocate this number. It is recommended to avoid numbers frequently used for audio (97) and video (96), and, for example, the value 104 may be used for DICOM Metadata Essence. The value shall be associated to the media type “application” and the subtype “dicom” in the SDP.

Note

An example of SDP is below (bold is mandatory):

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

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

- PTP Sync Timestamp
- PTP Origin Timestamp
- Source Identifier
- Flow Identifier
- Optionally, Grain Duration
- Grain Flags

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

6.2.2 Payload

The payload consists of two parts, a payload header and a DICOM dataset compliant with real-time communication.

The DICOM dataset is made of three parts, starting with the Meta Information, followed by the dynamic part containing information that varies over time (e.g., Origin Timestamp of the frame, Position of a probe, circle defining the eye), and the static part containing information that doesn't vary over time (e.g. Patient Name, Modality, …).

The transmission rate of the dynamic part shall be identical to the rate of the associated Flow (e.g., one dataset per frame). The transmission rate of the static part shall be at least 1Hz.

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 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 a RTP session the Data Set representing a SOP Instance related to a DICOM IOD.

Figure 7-1 illustrates the encapsulation of a DICOM dataset in the RTP packet. 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.

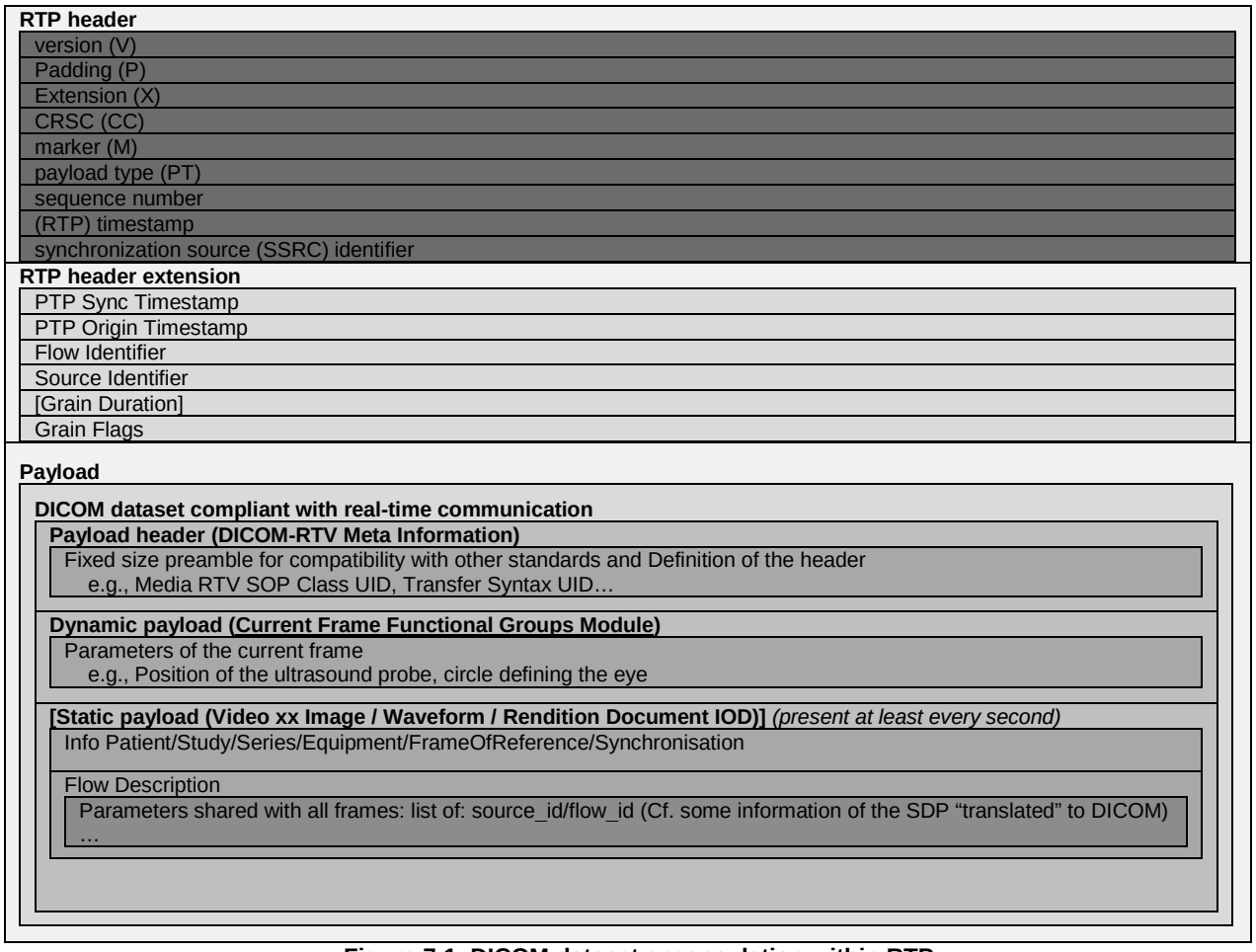


Figure 7-1. DICOM dataset encapsulation within RTP

7.1 DICOM-RTV META INFORMATION

The RTV Meta Information includes identifying information on the encapsulated Data Set. This header shall be present in every DICOM metadata RTP packet.

In case the DICOM dataset (Payload header, Dynamic payload and Static payload) is too big for one single IP packet, the Payload header has to be repeated in every packet and the Dynamic and/or Static payload(s) ha/s(ve) to be cut at a Data Element boundary (one single Data Element cannot be split on two packets).

Note

682 The group number of the DICOM-RTV Meta Information attributes is lower than the one of other
683 attributes in order to place the DICOM-RTV Meta Information at the beginning of the payload, like
684 it is done in DICOM PS 3.10.

685 **Table 7.1-1. DICOM-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.
RTV Meta Information Group Length	(kkkk,ee01)	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
RTV Meta Information Version	(kkkk,ee02)	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.
Media Storage SOP Class UID	(0002,0002)	1	Uniquely identifies the SOP Class associated with the Data Set. SOP Class UIDs allowed for media storage are specified in section 7.2 STANDARD SOP CLASSES.
Media Storage SOP Instance UID	(0002,0003)	1	Uniquely identifies the SOP Instance associated with the Data Set placed in the RTP payload and following the 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 Source Identifier	(kkkk,ee03)	3	The UUID of the RTP source that sends the DICOM Real-Time Video metadata Flow.
RTV Flow Identifier	(kkkk,ee04)	3	The UUID of the DICOM Real-Time Video metadata Flow.
RTV Flow RTP Sampling Rate	(kkkk,ee05)	3	The rate of the dynamic part of the DICOM Real-Time Video metadata Flow, the same as the bulk-data Flow rate.
RTV Flow Actual Frame Duration	(kkkk,ee06)	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

Attribute Name	Tag	Type	Attribute Description
			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	xxxxxxx1	Real-Time Video Endoscopic Image IOD
Video Microscopic Image Real-Time Communication	xxxxxxx2	Real-Time Video Microscopic Image IOD
Video Photographic Image Real-Time Communication	xxxxxxx3	Real-Time Video Photographic Image IOD
Audio Waveform Real-Time Communication	xxxxxxx4	Real-Time Audio Waveform IOD
Rendition Document Real-Time Communication	xxxxxxx5	Rendition Document IOD

8 SECURITY CONSIDERATIONS

It is expected that the OR environment is secure such that no additional security mechanism is required. Security aspects are out of scope of the present document.

It is expected that the OR environment provides the protection needed to ensure the essential performance and integrity.

There are common security solutions (e.g., IPSEC, VLAN mechanisms, TLS) that might work with the proposed specification, but they may be unable to satisfy the performance requirements.

701
702
703
704
705
706

Changes to NEMA Standards Publication PS 3.3-20xx

Digital Imaging and Communications in Medicine (DICOM)

Part 3: Information Object Definitions

TO BE DONE AFTER APPROVAL OF THE SUPPLEMENT: APPEND THE NEW IOD(s) IN THE TABLE(s) OF THE SECTION A.1.2

Add a new section A.32.x Real-Time Video Endoscopic Image IOD

A.32.x Real-Time Video Endoscopic Image IOD

A.32.x.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, thanks to the Real-Time Video service.

A.32.x.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.x-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.X1	M
	Acquisition Context	C.7.6.14	M

IE	Module	Reference	Usage
	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
	SOP Common	C.12.1	M
	Common Instance Reference	C.12.2	M
	Stereoscopic Acquisition	C.X.X	C - Required if this flow contains a stereoscopic pair
	Current Frame Functional Groups	C.7.6.X2	M

A.32.x.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.x.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.x.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.x.3.3 Anatomic Region Sequence

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

A.32.x.3.4 Current Frame Functional Groups Module

The Current Frame Functional Groups Module shall include the Real-Time Video Endoscopic Functional Group Macro (C.7.6.X3). It shall be placed in the dynamic payload.

Table A.32.x-2 Functional Groups Macros

Functional Group Macro	Section	Usage
Real-Time Video Endoscopic Image	C.7.6.X3	M
Frame Content	C.7.6.16.2.2	M
Interlaced Video	C.7.6.X9	C – Required if the referenced video is interlaced (vs. progressive)
Frame Relevance	C.7.6.X6	U
Camera Position	C.7.6.X7	U

A.32.x.3.5 Stereoscopic Acquisition Module

The Stereoscopic Acquisition Module is defined in Table C.X.X.

A.32.x.3.6 Time Distribution Protocol

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

Add a new section A.32.y Real-Time Video Microscopic Image IOD

A.32.y Real-Time Video Microscopic Image IOD

A.32.y.1 Real-Time Video Microscopic Image IOD Description

The Real-Time Video Microscopic Image IOD specifies the Attributes of Real-Time Video Microscopic Images, transmitted in real-time, thanks to the Real-Time Video service. It includes both imaging of specimens and direct microscopic imaging of the patient (e.g., perioperative microscopy). Microscopic Images with Slide Coordinates shall not be encoded with this IOD.

A.32.y.2 Real-Time Video Microscopic 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.

Note

The video shall not include audio channel. In case such channel is present for acquiring patient voice or physiological sounds, healthcare professionals comment, or environment sounds, it shall be transported using another Real-time IOD (e.g., Real-Time Audio Waveform IOD).

Table A.32.y-1. Real-Time Video Microscopic 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

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	Frame of Reference	C.7.4.1	C – Required if multiple instances are spatially related or if the Camera Position Functional Group Macro or the Capsulorhexis 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.X1	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
	Frame Extraction	C.12.3	C - Required if the SOP Instance was created in response to a Frame-Level retrieve request
	Stereoscopic Acquisition	C.X.X	C - Required if this flow contains a stereoscopic pair
	Current Frame Functional Groups	C.7.6.X2	M

A.32.y.3 Real-Time Video Microscopic Image IOD Content Constraints

A.32.y.3.1 Modality

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

A.32.y.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.y.3.3 Current Frame Functional Groups Module

The Current Frame Functional Groups Module shall include the Real-Time Video Microscopic Functional Group Macro (C.7.6.X4). It shall be placed in the dynamic payload.

Table A.32.y-2 Functional Groups Macros

Functional Group Macro	Section	Usage
Real-Time Video Microscopic Image	C.7.6.X4	M
Frame Content	C.7.6.16.2.2	M
Interlaced Video	C.7.6.X9	C – Required if the referenced video is interlaced (vs. progressive)
Frame Relevance	C.7.6.X6	U
Camera Position	C.7.6.X7	U
Capsulorhexis	C.7.6.X8	U

A.32.y.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.32.z Real-Time Video Photographic Image IOD

A.32.z Real-Time Video Photographic Image IOD

A.32.z.1 Real-Time Video Photographic Image IOD Description

The Real-Time Video Photographic Image IOD specifies the attributes of VL Multi-frame photographic Images.

A.32.z.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. The Frame of Reference IE is not a component of this IOD.

Note

1. The video shall not include audio channel(s) for acquiring patient voice or physiological sounds, healthcare professionals' commentary, or environmental sounds, which has(ve) to be transported in a separate IOD.

2. 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.

812
813

Table A.32.z-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
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.X1	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
	Stereoscopic Acquisition	C.X.X	C - Required if this flow contains a stereoscopic pair
	Current Frame Functional Groups	C.7.6.X2	M

814

A.32.z.3 Real-Time Video Photographic Image IOD Content Constraints

A.32.z.3.1 Modality

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

A.32.z.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.z.3.3 Current Frame Functional Groups Module

The Current Frame Functional Groups Module shall include the Real-Time Video Photographic Image Functional Group Macro (C.7.6.X5). It shall be placed in the dynamic payload.

Table A.32.z-2 Functional Groups Macros

Functional Group Macro	Section	Usage
Real-Time Video Photographic Image	C.7.6.X5	M
Frame Content	C.7.6.16.2.2	M
Interlaced Video	C.7.6.X9	C – Required if the referenced video is interlaced (vs. progressive)
Frame Relevance	C.7.6.X6	U
Camera Position	C.7.6.X7	U

A.32.z.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.x Real-Time Audio Waveform IOD

A.34.x Real-Time Audio Waveform IOD

A.34.x.1 Real-Time Audio Waveform IOD Description

The Real-Time Audio Waveform IOD is the specification of one or two channel digitized audio signals, transmitted in real-time, thanks to the Real-Time Video service.

A.34.x.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 below the Series IE.

A.34.x.3 Real-Time Audio Waveform IOD Module Table

Table A.34.x-1 specifies the Modules of the Real-Time Audio Waveform IOD.

Table A.34.x-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
	Clinical Trial Study	C.7.2.3	U
Series	General Series	C.7.3.1	M
	Clinical Trial Series	C.7.3.2	U
Frame of Reference	Synchronization	C.7.4.2	M
Equipment	General Equipment	C.7.5.1	M
	Enhanced General Equipment	C.7.5.2	M
Waveform	Waveform Identification	C.10.8	M
	Real-Time Bulk Data Flow	C.7.6.X1	M
	Acquisition Context	C.7.6.14	M
	SOP Common	C.12.1	M
	Current Frame Functional Groups	C.7.6.X2	M

A.34.x.4 Real-Time Audio Waveform IOD Content Constraints

A.34.x.4.1 Modality

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

A.34.x.4.2 Waveform Sequence

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

A.34.x.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.x.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.x.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”.

865 **A.34.x.4.6 Waveform Sample Interpretation**

866 The value of Waveform Sample Interpretation (5400,1006) in each Waveform Sequence Item shall be SB or SS.
867
868 **A.34.x.4.7 Current Frame Functional Groups Module**

869 The Current Frame Functional Groups Module shall be placed in the dynamic payload.

870

871 **A.34.x.4.8 Time Distribution Protocol**

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

873

874 **Add an new section A.35.x Rendition Document IOD**

875 **A.35.X Rendition Document IOD**

876 **A.35.X.1 Rendition Document IOD Description**

877 The Rendition Document IOD associates a group of time-synchronized Flows produced for a simultaneous
878 presentation, transported using DICOM-RTV.
879
880 **A.35.X.2 Rendition Document IOD Entity-Relationship Model**

881 This IOD uses the E-R Model in Section A.1.2, with only the SR Document IE below the Series IE.
882
883 **A.35.X.3 Rendition Document IOD Module Table**

884 Table A.35.X-1 specifies the Modules of the Rendition Document IOD.

885 **Table A.35.X-1. Rendition Document IOD Modules**
886

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
Frame Of Reference	Synchronization	C.7.4.2	M

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

A.35.X.3.1 Rendition Document IOD Content Constraints

A.35.X.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.X.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.X-2 specifies the relationship constraints of this IOD.

Table A.35.X-2. Relationship Content Constraints for Rendition 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.X.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 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.

The details on the way for encoding Pixel Data in the rest of the section only applies if Pixel Data is present (7FE0,0010). In case the Photometric Interpretation relates to the referred video flow (Real-Time Video), the way of encoding pixels shall respect the underlying SMPTE standards.

Defined Terms:

...

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

C.7.6.X1 REAL-TIME BULK DATA FLOW MODULE

The module describes the reference to the pixels/waveforms that are not contained within the DICOM dataset but conveyed in the associated streams, as described in PS 3.X in section 6.1. 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 multi-media 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.

Table C.7.6.X1-1 specifies the Attributes for the Real-Time Bulk Data Flow Module.

Table C.7.6.X1-1 Real-Time Bulk Data Flow Module

Attribute Name	Tag	Type	Attribute Description
Real-Time Flow Sequence	(gggg,ee07)	1	The source and the Flows providing Bulk Data of this instance
>Source Identifier	(gggg,ee08)	1	UUID of the referenced Source. See 10.xx.1.1.
>Flow Identifier Sequence	(gggg,ee09)	1	Sequence identifying the Flows provided by the Source
>>Flow Identifier	(gggg,ee10)	1	UUID of the referenced Flow. See 10.xx.1.2.
>>Flow Transfer Syntax UID	(gggg,ee11)	1	UID of the encoding method of the referenced Flow See 10.xx.1.3.
>>Flow RTP Sampling Rate	(gggg,ee12)	1	Sampling rate in Hertz used by RTP for generating

			timestamp See 10.xx.1.4.

C.7.6.X1.1 Real-Time Bulk Data Flow Module Attributes

C.7.6.X1.1.1 Source Identifier

The Source Identifier 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.X1.1.2 Flow Identifier

The Flow Identifier is a Universally Unique Identifier (UUID). The value is 16 bytes 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.X1.1.3 Flow Transfer Syntax UID

The Flow Transfer Syntax UID 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 (kkkk,ee05), respectively, of the DICOM-RTV Meta Information Header.

C.7.6.X1.1.4 Flow RTP Sampling Rate

The Flow RTP Sampling Rate shall be the one defined in the SDP of the corresponding Flow.

Add New Module: Current Frame Functional Groups Module

C.7.6.X2 Current Frame Functional Groups Module

Table C.7.6.X2-1. defines the Attributes related to the current frame when the IOD is transported using Real-Time Communication.

Note

The group number of the attributes of the Current Frame Functional Groups Module, relative to the “dynamic” payload, is lower than the one of other attributes in order to be placed before the “static” payload, but higher than the attributes of the DICOM-RTV Metadata Information, in order to be placed after the “header” payload, like it is done in DICOM PS 3.10.

Table C.7.6.X2-1 Current Frame Functional Groups Module Attributes

Attribute Name	Tag	Type	Attribute Description
Current Frame Functional Groups Sequence	(hhhh,ee14)	1	Sequence that contains the Functional Groups Sequence Attributes corresponding to the current frame or audio sample. Only one Item shall be included in this Sequence.
>Frame Origin Timestamp	(gggg,ee15)	1	This timestamp contains the capture time of the payload content for this frame or audio sample. It is relative to Time Distribution Standard (gggg,ee13).
>Include one or more Functional Group Macros.			**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.X2.1 Current Frame Functional Groups Module Attributes

C.7.6.X2.1.1 Frame Origin Timestamp

This field contains 10 bytes conforming with IEEE 1588:2008 (PTPv2) standard representing seconds and nanoseconds since Epoch, defined as 1 of January, 1970, at 00:00:00 TAI (International Atomic Time) and UTC (Universal Coordinated Time), which were the same. The 6 first bytes contain the number of seconds, and the 4 last bytes, the number of nanoseconds. It shall comply with Time Source, Time Distribution Protocol and Time Distribution Standard values defined in C.7-7. Synchronization Module Attributes.

It must match with the Origin Timestamp for this frame, contained within the RTP extended header of the payload content of this frame, in the case this one is present. It shall be used for post-synchronizing different content payloads (e.g., video and corresponding audio) after they have been recorded. If not present, the RTP Timestamp, part of regular RTP header, is derived from Frame Origin Timestamp and Flow RTP sampling rate, and is used to pair content from different Flows (payload flow with metadata flow), through a time alignment mechanism.

C.7.6.X2.1.2 Functional Group Macros

Example of one Functional group macro that could be included in all frames of the video IOD transported using Real-Time Communication:

Table C.7.6.X2-2 Functional Groups Macros

Functional Group Macro	Section	Usage
Real-Time Video Endoscopic Image	C.7.6.X3	U
Frame Relevance	C.7.6.X6	U

Add New Macro: Real-Time Video Endoscopic Image Macro

C.7.6.X3 Real-Time Video Endoscopic Image Macro

Table C.7.6.X3-1 specifies the attributes of the Real-Time Video Endoscopic Image Functional Group Macro.

Table C.7.6.X3-1 Real-Time Video Endoscopic Image Functional Group Macro Attributes

Attribute Name	Tag	Type	Attribute Description
Light Brightness Ratio	(gggg,ee19)	3	The light brightness ratio, expressed in percentage. See Section C.7.6.X3.1.1 for further explanation.

C.7.6.X3.1 Real-Time Video Endoscopic Image Macro Attributes

C.7.6.X3.1.1 Light Brightness Ratio

Brightness (0018,1182) of the light illuminating the scene, expressed as a ratio between 0 and 100, 100 meaning the light is at its maximum value.

Add New Macro: Real-Time Video Microscopic Image Macro

C.7.6.X4 Real-Time Video Microscopic Image Macro

Table C.7.6.X4-1 specifies the attributes of the Real-Time Video Microscopic Image Functional Group Macro.

Table C.7.6.X4-1 Real-Time Video Microscopic Image Functional Group Macro Attributes

Attribute Name	Tag	Type	Attribute Description
Light Brightness Ratio	(gggg,ee19)	3	The light brightness ratio, expressed in percentage. See Section C.7.6.X4.1.1 for further explanation.
Focal Distance	(0018,1182)	3	Focal distance of the lens, in mm. See Section C.7.6.X4.1.2 for further specialization.
Zoom Factor	(0028,0031)	3	The amount of magnification applied to each pixel in the image, specified by a numeric pair: row value (delimiter) column value. See Section C.7.6.X4.1.3 for further explanation.

C.7.6.X4.1 Real-Time Video Microscopic Image Macro Attributes

C.7.6.X4.1.1 Light Brightness Ratio

Brightness (0018,1182) of the light illuminating the scene, expressed as a ratio between 0 and 100, 100 meaning the light is at its maximum value.

C.7.6.X4.1.2 Focal Distance

Focal Distance (0018,1182) for Image data is the focal distance, in mm, measured from the front face of the sensor to the focus.

C.7.6.X4.1.3 Zoom Factor

Zoom Factor (0028,0031) is the magnification factor that was used during the acquisition. When this attribute is not given, it is assumed to be 1.0\1.0.

Add New Macro: Real-Time Video Photographic Image Macro

C.7.6.X5 Real-Time Video Photographic Image Macro

Table C.7.6.X5-1 specifies the attributes of the Real-Time Video Photographic Image Functional Group Macro.

Table C.7.6.X5-1 Real-Time Video Photographic Image Functional Group Macro Attributes

Attribute Name	Tag	Type	Attribute Description
Focal Distance	(0018,1182)	3	Focal distance of the lens, in mm. See Section C.7.6.X5.1.1 for further specialization.
Zoom Factor	(0028,0031)	3	The amount of magnification applied to each pixel in the image, specified by a numeric pair: row value (delimiter) column value. See Section C.7.6.X5.1.2 for further explanation.

C.7.6.X5.1 Real-Time Video Photographic Image Macro Attributes

C.7.6.X5.1.2 Focal Distance

Focal Distance (0018,1182) for Image data is the focal distance, in mm, measured from the front face of the sensor to the focus.

C.7.6.X5.1.3 Zoom Factor

Zoom Factor (0028,0031) is the magnification factor that was used during the acquisition. When this attribute is not given, it is assumed to be 1.0\1.0.

Add New Macro: Frame Relevance Macro

C.7.6.X6 Frame Relevance Macro

Table C.7.6.X6-1 specifies the attributes of the Frame Relevance Functional Group Macro, related to the relevance of current frame in regards to the clinical use of information.

Table C.7.6.X6-1 Frame Relevance Functional Group Macro Attributes

Attribute Name	Tag	Type	Attribute Description
Frame Relevance Group Sequence	(hhhh,ee17)	1	Sequence that contains the Functional Groups Sequence Attributes corresponding to the current frame or audio sample. Only one Item shall be included in this Sequence.
>Image Relevance	(gggg,ee16)	3	Specify if the current frame is relevant for clinical use. Enumerated value: **OFF** video channel if not containing any relevant pixel **OUT** the image is captured outside the patient **IN** the image is captured inside the patient

C.7.6.X6.1 Frame Relevance Macro Attributes

C.7.6.X6.1.1 Image Relevance

The attribute specifies if the current image capture is active and if its content has been acquired from inside or outside patient.

Add New Macro: Camera Position Macro

C.7.6.X7 Camera Position Macro

Table C.7.6.X7-1 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.X7-1 Camera Position Functional Group Macro Attributes

Attribute Name	Tag	Type	Attribute Description
Render Projection	(0070,1602)	1	Projection style. Enumerated Values: **ORTHOGRAPHIC** **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.
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.

Attribute Name	Tag	Type	Attribute Description
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.
Pixel Spacing	(0028,0030)	3	Physical distance in the imaging target (patient, specimen, or phantom) between the centers of each pixel, specified by a numeric pair - adjacent row spacing (delimiter) adjacent column spacing in mm. See Section 10.7.1.3 for further explanation of the value order.

Add New Macro: Capsulorhexis Macro

C.7.6.X8 Capsulorhexis Macro

Table C.7.6.X8-1 specifies the attributes of the Capsulorhexis Functional Group Macro related to the position of eye as detected in the current frame.

Table C.7.6.X8-1 Capsulorhexis Functional Group Macro Attributes

Attribute Name	Tag	Type	Attribute Description
Source Frame Origin Timestamp	(gggg,ee20)	1	This timestamp contains the capture time of the payload content for the frame used as the source for calculation of the eye parameters. It is relative to Time Distribution Standard (gggg,ee13).
Pupil Size	(0046,0044)	3	The horizontal diameter measurement of the pupil, in mm.
Corneal Size	(0046,0046)	3	The horizontal diameter measurement of the cornea, in mm.
Corneal Vertex Location	(0046,0202)	3	Location of the corneal vertex. Given as column\row. Column is the horizontal offset and row is the vertical offset. Image relative position specified with sub-pixel resolution such that the origin at the Top Left Hand Corner (TLHC) of the TLHC pixel is 0.0\0.0, the Bottom Right Hand Corner (BRHC) of the TLHC pixel is 1.0\1.0, and the BRHC of the BRHC pixel is Columns\Rows (see Figure C.10.5-1). The values must be within the range 0\0 to Columns\Rows. This location shall anchor the corneal vertex at the x, y and z coordinates of 0.0, 0.0, 0.0, in mm. See Section C.8.30.3.1.4 for further explanation.
Circle Size	(gggg,ee21)	1	The horizontal diameter measurement of the circle annotation. Given as column. Image relative dimension specified with sub-pixel resolution. The

Attribute Name	Tag	Type	Attribute Description
			values must be within the range 0 to Columns.
Circle Vertex Location	(gggg,ee22)	1	Location of the circle annotation vertex. Given as column\row. Column is the horizontal offset and row is the vertical offset. Image relative position specified with sub-pixel resolution such that the origin at the Top Left Hand Corner (TLHC) of the TLHC pixel is 0.0\0.0, the Bottom Right Hand Corner (BRHC) of the TLHC pixel is 1.0\1.0, and the BRHC of the BRHC pixel is Columns\Rows (see Figure C.10.5-1). The values must be within the range 0\0 to Columns\Rows.
Pixel Spacing	(0028,0030)	3	Physical distance in the imaging target (patient, specimen, or phantom) between the centers of each pixel, specified by a numeric pair - adjacent row spacing (delimiter) adjacent column spacing in mm. See Section 10.7.1.3 for further explanation of the value order.

Add New Macro: Interlaced Video Macro

C.7.6.X9 Interlaced Video Macro

Table C.7.6.X9-1 specifies the attributes of the Interlaced Video Functional Group Macro related to the parity of the current frame. This macro is mandatory when the referenced video is interlaced (the Flow Transfer Syntax UID (gggg,ee11) has a value which specifies the video is interlaced (e.g., SMPTE ST 2110-20 Uncompressed Interlaced Active Video)).

Table C.7.6.X9-1 Interlaced Video Functional Group Macro Attributes

Attribute Name	Tag	Type	Attribute Description
Frame Lines Parity	(gggg,ee23)	1	Parity of the current frame Enumerated Values: **ODD** this frame contains only odd lines (e.g., lines 1, 3, …) **EVEN** this frame contains only even lines (e.g., lines 2, 4, …)

Amend Table C.7-7. Synchronization Module Attributes

Table C.7-7. Synchronization Module Attributes

Attribute Name	Tag	Type	Attribute Description
Synchronization Frame of Reference UID	(0020,0200)	1	UID of common synchronization environment. See Section C.7.4.2.1.1.
Synchronization Trigger	(0018,106A)	1	Data acquisition synchronization with external equipment

Attribute Name	Tag	Type	Attribute Description
			Enumerated Values: **SOURCE** this equipment provides synchronization channel or trigger to other equipment **EXTERNAL** this equipment receives synchronization channel or trigger from other equipment **PASSTHRU** this equipment receives synchronization channel or trigger and forwards it **NO TRIGGER** data acquisition not synchronized by common channel or trigger
Trigger Source or Type	(0018,1061)	3	Specifies equipment ID of trigger source and/or type of trigger
Synchronization Channel	(0018,106C)	1C	Identifier of waveform channel that records the synchronization channel or trigger, see Section C.7.4.2.1.3. Required if synchronization channel or trigger is encoded in a waveform in this SOP Instance
Acquisition Time Synchronized	(0018,1800)	1	Acquisition DateTime (0008,002A) synchronized with external time reference. Enumerated Values: **Y** **N** See Section C.7.4.2.1.4
Time Source	(0018,1801)	3	ID of equipment or system providing time reference
Time Distribution Protocol	(0018,1802)	3	Method of time distribution used to synchronize this equipment. Enumerated Values: **NTP** Network Time Protocol **IRIG** Inter Range Instrumentation Group **GPS** Global Positioning System **SNTP** Simple Network Time Protocol **PTP** IEEE 1588 Precision Time Protocol
Time Distribution Standard	**(gggg,ee13)**	**3**	**Standard used for the time delivered by the Time Source (0018,1801).** **Enumerated Value** **UTC: all timestamp such as FrameOriginTimeSource are expressed in UTC** **TAI: all timestamp such as FrameOriginTimeSource are expressed in TAI**
NTP Source Address	(0018,1803)	3	IP Address of NTP, SNTP, or PTP time source. IPv4 addresses shall be in dotted decimal (e.g., 192.168.1.1). The IPv6 addresses shall be in colon separated hexadecimal (e.g., 12:34:56:78:9a:bc:de:f0). Note Identity of this value in two instances acquired

Attribute Name	Tag	Type	Attribute Description
			contemporaneously implies a common time base. The NTP Source Address might not persist over time.

C.7.4.2.1.2 Time Source ~~and~~, Time Distribution Protocol and Time Distribution Standard

Time may originate with a primary source (e.g., a national standards bureau) and be distributed through a chain of secondary distribution systems until reaching the imaging equipment. Time Distribution Protocol (0018,1802) specifies the immediate (last link) method used by the equipment to receive time from the immediately prior Time Source (0018,1801). It does not specify the ultimate time reference from which the Time Source may derive its synchronization.

Note

~~The time value distributed through the specified Time Distribution Protocol may need to be corrected to align with UTC. For example, GPS does not compensate for leap seconds.~~

If Time Distribution Standard (gggg,ee13) is present, whether or not a correction for leap seconds has been applied, is explicitly defined. Otherwise, the time value may need to be corrected to align with whatever standard is being used (e.g., compensate for leap seconds).

Add New Module: Stereoscopic Acquisition Module

C.X.X Stereoscopic Acquisition Module

Table C.X-X. defines the Attributes related to the current frame when the IOD is transported using Real-Time Communication.

Table C.X-X Stereoscopic Acquisition Module Attributes

Attribute Name	Tag	Type	Attribute Description
Stereo Pairs Present	(0022,0028)	1	The multi-frame pixel data consists of left and right stereoscopic pairs. See Section C.X.X.1.1 for further explanation. Enumerated Values: YES NO

C.X.X.1 Stereoscopic Acquisition Module Attributes

C.X.X.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 PS 3.5-20xx

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

This section describes the constraints applying to pixel data carried into DICOM-RTV Flow (separated from metadata Flow) and fully described in SMPTE ST2110-20.

The following table describes the valid values for attributes

- Samples per pixel (0028,0002)
- Bits Allocated (0028,0100)
- Bits Stored (0028,0101)
- High Bit (0028,0102)

Table X.1: constraints applicable to attributes describing pixel data

samples per pixel	Bits Allocated	Bits Stored	High bit
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 X.2: List of supported SMPTE ST 2110-20 {sampling system, colorimetry}

SMPTE ST 2110-20		DICOM Photometric Interpretation
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

The following table lists the unsupported combination:

Table X.3: List of supported SMPTE ST 2110-20 {sampling system, colorimetry}

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.x Transfer Syntax for SMPTE ST 2110-20 Uncompressed Progressive Active Video

This Transfer Syntax is used to specify the use of Uncompressed Video pixels carried in a separate Flow as described by 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.X.

10.t Transfer Syntax for SMPTE ST 2110-20 Uncompressed Interlaced Active Video

This Transfer Syntax is used to specify the use of Uncompressed Video pixels carried in a separate Flow as described by 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.Y.

10.t.1 Interlaced vs. Progressive video (Informative)

Interlaced video enables to transmit video with a smaller bandwidth. A 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.y Transfer Syntax for SMPTE ST 2110-30 PCM Digital Audio

This Transfer Syntax is used to express the constraints applying to audio channel data carried in a separate Flow as described by SMPTE ST 2110-30 standard. The main parameters of the transfer syntax are described in the Annex A.Z.

10.z Transfer Syntax for SMPTE ST 2110-30 Digital Waveform

This Transfer Syntax is used to express the constraints applying to waveform channel data carried into DICOM-RTV Flow (separated from metadata Flow) and fully described in SMPTE ST2110-30. The main parameters of the transfer syntax are described in the Annex A.Z.

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

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

This Transfer Syntax is used in DICOM-RTV Metadata Flow in order to clearly describe the accompanying Video Flow compatible with SMPTE ST2110-20.

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 pixel formats:

- SMPTE ST 2110-20 YCbCr-4:4:4 sampling system
  Photometric Interpretation (0028,0004) shall be YBR_FULL (see Table X.1)
- SMPTE ST 2110-20 RGB sampling system
  Photometric Interpretation (0028,0004) shall be RGB (see Table X.2)

- SMPTE ST 2110-20 YCbCr-4:2:2 sampling system
  Photometric Interpretation (0028,0004) shall be YBR_FULL_422 (see Table X.3)
- SMPTE ST 2110-20 YCbCr-4:2:0 sampling system
  Photometric Interpretation (0028,0004) shall be YBR_PARTIAL_420 (see Table X.4)

Table X.1: DICOM attributes for different color resolution in YCbCr-4:4:4 sampling system

SMPTE ST 2110-20 YCbCr-4:4:4 BT601	DICOM Attributes (Photometric Interpretation YBR_FULL)			
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

Table X.2: DICOM attributes for different color resolution in RGB sampling system

SMPTE ST 2110-20 RGB BT601	DICOM Attributes (Photometric Interpretation RGB)			
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

Table X.3: DICOM attributes for different color resolution in YCbCr-4:2:2 sampling system

SMPTE ST 2110-20 YCbCr-4:2:2 BT601	DICOM Attributes (Photometric Interpretation YBR_FULL_422)			
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

Table X.4: DICOM attributes for different color resolution in YCbCr-4:2:0 sampling system

SMPTE ST 2110-20 YCbCr-4:2:2 BT601	DICOM Attributes (Photometric Interpretation YBR_PARTIAL_420)

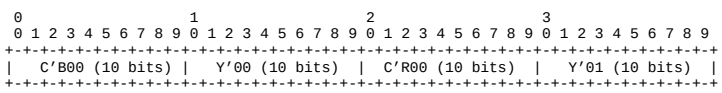
Bit Depth	Samples per Pixel (0028,0002)	Bits Allocated (0028,0100)	Bits Stored (0028,0101)	High Bit (0028,0102)
8	3	8	8	7
10	3	16	10	9
12	3	16	12	11
16	3	16	16	15

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

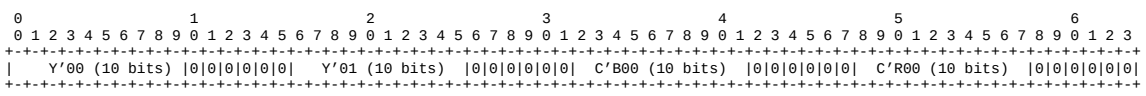
Note

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



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

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

This Transfer Syntax is used in DICOM-RTV Metadata Flow in order to clearly describe the accompanying Video Flow compatible with SMPTE ST2110-20.
The parameters are similar to the ones described in the SMPTE ST 2110-20 Uncompressed Progressive Active Video (Annex A.X), but the frames are interlaced, the first frame containing only odd lines and the next frame containing only even lines.

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

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

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

DICOM attributes

- Number of Waveform Channels (003A,0005) is limited to 15
- Number of Waveform Samples (003A,0010) is restricted
- 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 Z.1: AES67 and DICOM sampling frequency

AES67 Sampling Frequency	Sampling frequency (0003,001A)
44.1 kHz	44100

48 kHz*	48000
96 kHz	96000

* 48 kHz should be preferred

Table Z.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 Z.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
16	16	1,2	48	96,192
24	24	1,2	48	144,288

SMPTE ST 2110-30 is based on AES67, and restricts the audio Flow:

- Sampling frequency is either 44.1 kHz, 48 kHz or 96 kHz, 48 kHz being the preferred 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

Add New Section to Annex A : SMPTE ST 2110-30 Digital Waveform

A.T : SMPTE ST 2110-30 DIGITAL WAVEFORM Transfer Syntax

This Transfer Syntax is used in DICOM-RTV Metadata Flow in order to describe the accompanying Flow compliant with SMPTE ST 21110-30 carrying Digital Signal (such as Heart Rate or Breath count).

The limitation applicable is just that every frame has a limit due to UDP transport.

DICOM attributes

- Numbers of Waveform Channels (003A,0005)
- Number of Waveform Samples (003A,0010)
- Sampling frequency (003A,001A)
- Waveform Bits Stored (003A,021A)
- Waveform Bits Allocated (5400,1004)
- Waveform Sample Interpretation (5400,1006)

are still applicable with some accommodations specified below:

- Numbers of Waveform Channels * Number of Waveform Samples * Waveform Bits Allocated shall be < available payload in the IP packet.

- Available payload in the IP packet shall be 1,388.

Changes to NEMA Standards Publication PS 3.6-20xx

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	
…					
(gggg,ee07)	Real-Time Flow Sequence	RealTimeFlowSequence	SQ	1	
(gggg,ee08)	Source Identifier	SourceIdentifier	OB	1	
(gggg,ee09)	Flow Identifier Sequence	FlowIdentifierSequence	SQ	1	
(gggg,ee10)	Flow Identifier	FlowIdentifier	OB	1	
(gggg,ee11)	Flow Transfer Syntax UID	FlowTransferSyntaxUID	UI	1	
(hhhh,ee14)	Current Frame Functional Groups Sequence	CurrentFrameFunctionalGroupsSequence	SQ	1	
(gggg,ee15)	Frame Origin Timestamp	FrameOriginTimestamp	OB	1	
(gggg,ee13)	Time Distribution Standard	TimeDistributionStandard	CS	1	
(gggg,ee12)	Flow RTP Sampling Rate	FlowRTPSamplingRate	UL	1	
(kkkk,ee01)	RTV Meta Information Group Length	RTVMetaInformationGroupLength	UL	1	
(kkkk,ee02)	RTV Meta Information Version	RTVMetaInformationVersion	OB	1	
(kkkk,ee03)	RTV Source Identifier	RTVSourceIdentifier	OB	1	
(kkkk,ee04)	RTV Flow Identifier	RTVFlowIdentifier	OB	1	
(kkkk,ee05)	RTV Flow RTP Sampling Rate	RTVFlowRTPSamplingRate	UL	1	
(kkkk,ee06)	RTV Flow Actual Frame Duration	RTVFlowActualFrameDuration	IS	1	

Tag	Name	Keyword	VR	VM	
(gggg,ee16)	Image Relevance	ImageRelevance	CS	1	
(hhhh,ee17)	Frame Relevance Group Sequence	FrameRelevanceGroupSequence	SQ	1	
(gggg,ee19)	Light Brightness Ratio	LightBrightnessRatio	UL	1	
(gggg,ee20)	Source Frame Origin Timestamp	SourceFrameOriginTimestamp	OB	1	
(gggg,ee21)	Circle Size	CircleSize	UL	1	
(gggg,ee22)	Circle Vertex Location	CircleVertexLocation	UL	1	
(gggg,ee23)	Frame Lines Parity	FrameLinesParity	CS	1	

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

Add new UIDs to Annex A.

UID Value	UID Name	UID Type	Part
xxxxxx6	SMPTE ST 2110-20 Uncompressed Progressive Active Video	Transfer Syntax	PS3.5
xxxxxx9	SMPTE ST 2110-20 Uncompressed Interlaced Active Video	Transfer Syntax	PS3.5
xxxxxx7	SMPTE ST 2110-30 PCM Digital Audio	Transfer Syntax	PS3.5
xxxxxx8	SMPTE ST 2110-30 Digital Waveform	Transfer Syntax	PS3.5
xxxxxx1	Video Endoscopic Image Real-Time Communication	SOP Class	PS3.X
xxxxxx2	Video Microscopic Image Real-Time Communication	SOP Class	PS3.X
xxxxxx3	Video Photographic Image Real-Time Communication	SOP Class	PS3.X
xxxxxx4	Audio Waveform Real-Time Communication	SOP Class	PS3.X
xxxxxx5	Rendition Document Real-Time Communication	SOP Class	PS3.X

Changes to NEMA Standards Publication PS 3.16-20xx

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**
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 XXX "Real-Time Video Rendition Titles"		

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

CID XXX Real Time Video Rendition Titles

Type: **Extensible**
Version: **2017mmdd**

Table CID XXX. Real Time Video Rendition Titles

Coding Scheme Designator	Code Value	Code Meaning
DCM	Sup202_aa01	RTV Rendition
DCM	Sup202_aa02	RTV Audio and Video Rendition
DCM	Sup202_aa03	RTV Stereo Video Rendition

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

Code Value	Code Meaning	Definition	Notes
Sup202_aa01	RTV Rendition	DICOM objects communicated in time-synchronized flows using DICOM-RTV, intended for a simultaneous presentation	
Sup202_aa02	RTV Audio and Video Rendition	Audio and Video DICOM objects communicated in time-synchronized flows using DICOM-RTV, intended for a simultaneous presentation	
Sup202_aa03	RTV Stereo Video Rendition	Two Video DICOM objects communicated in time-synchronized flows using DICOM-RTV, intended for a stereo video	

Changes to NEMA Standards Publication PS 3.2-20xx

Digital Imaging and Communications in Medicine (DICOM)

Part 2: Conformance

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

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

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.

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.

X.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

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

Table X.1-1. Network Services

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

X.2 Table of Contents

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

X.3 Introduction

X.3.1 Revision History

Table X.3.1-1. Revision History

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

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

See example text in Section A.3.

X.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.

X.4 Networking

X.4.1 Implementation Model

X.4.1.1 Application Data Flow

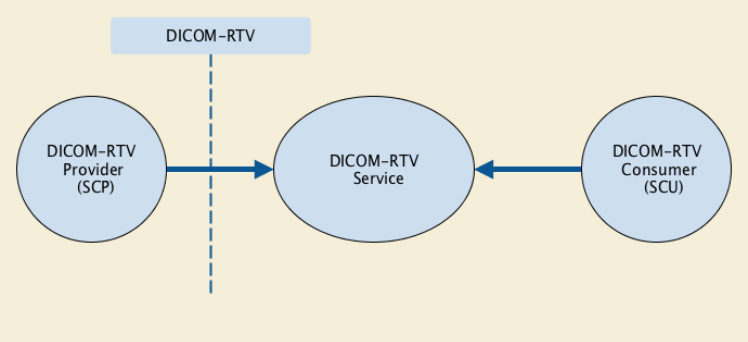


Figure X.4.1-1. Application Data Flow Diagram

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).

X.4.1.2 Functional Definition of AEs

X.4.1.2.1 Functional Definition of RTV Service Application

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

X.4.2 AE Specifications

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

X.4.2.1 DICOM-RTV Application Entity Specifications

X.4.2.1.1 SOP Classes

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

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

SOP Class Name	SOP Class UID	SCU	SCP
Video Photographic Image Real-Time Communication	xxxxxx3	No	Yes

Some restrictions applies on the Real-Time Communications:

Table X.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 X.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.

X.4.2.2.4 Connection Policies

X.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.

X.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.

X.4.3 Network Interfaces

X.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.

X.4.3.2 Additional Protocols

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

X.4.3.3 IPv4 and IPv6 Support

This product supports both IPv4 and IPv6 connections.

X.4.4 Configuration

X.4.4.1 DICOM-RTV Interface

The EXAMPLE-RTV-SERVICE is configured to define some parameters expressed in the SDP object. By default, the port number used for the video is 96 and the port used for DICOM-RTV metadata is 104.

X.5 Media Interchange

Not applicable.

X.6 Support of Character Sets

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

X.7 Security

Not Applicable.

X.8 Annexes

X.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.

X.8.2 Data Dictionary of Private Attributes

No private attributes is provided.

X.8.3 Coded Terminology and Templates

See conformance claim for EXAMPLE-INTEGRATED-MODALITY.

X.8.4 Standard Extended / Specialized / Private SOP Classes

Not Applicable.

X.8.5 Private Transfer Syntaxes

Private transfer syntaxes are not supported.

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

Y 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.

Y.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

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

Table Y.1-1. Network Services

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

Y.2 Table of Contents

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

Y.3 Introduction

Y.3.1 Revision History

Table Y.3.1-1. Revision History

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

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

See example text in Section A.3.

Y.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.

Y.4 Networking

Y.4.1 Implementation Model

Y.4.1.1 Application Data Flow

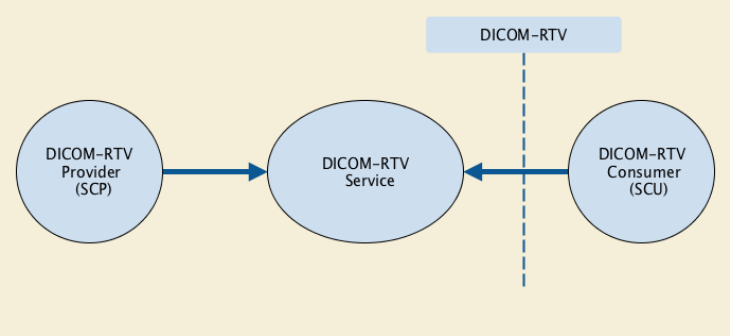


Figure Y.4.1-1. Application Data Flow Diagram

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.

Y.4.1.2 Functional Definition of AEs

Y.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.

Y.4.2 AE Specifications

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

Y.4.2.1 DICOM-RTV Application Entity Specifications

Y.4.2.1.1 SOP Classes

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

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

SOP Class Name	SOP Class UID	SCU	SCP
Video Photographic Image Real-Time Communication	xxxxxx3	Yes	No
Audio Waveform Real-Time Communication	xxxxxx4	Yes	No
Rendition Document Real-Time Communication	xxxxxx5	Yes	No

Some restrictions applies on the Real-Time Communications:

Table Y.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	48000

Table Y.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 set depending on the one of the sent video.

Y.4.2.2.4 Connection Policies

Y.4.2.2.4.1 General

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

Y.4.2.2.4.2 Number of Connections

EXAMPLE-RTV-DISPLAY is consuming multicast communication.

Y.4.3 Network Interfaces

Y.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.

Y.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.

Y.4.3.3 IPv4 and IPv6 Support

This product supports both IPv4 and IPv6 connections.

Y.4.4 Configuration

Y.4.4.1 DICOM-RTV Interface

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

Y.5 Media Interchange

Not applicable.

Y.6 Support of Character Sets

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

Y.7 Security

Not Applicable.

Y.8 Annexes

Y.8.1 IOD Contents

Not Applicable.

Y.8.2 Data Dictionary of Private Attributes

No private attributes is provided.

Y.8.3 Coded Terminology and Templates

Not Applicable.

Y.8.4 Standard Extended / Specialized / Private SOP Classes

Not Applicable.

Y.8.5 Private Transfer Syntaxes

Private transfer syntaxes are not supported.