

<p>ITU – Telecommunications Standardization Sector STUDY GROUP 16 Question 6 Visual, audio and signal coding ("VCEG")</p> <hr/> <p>73rd Meeting: 17-26 January 2024, Teleconference</p>	<p>Document: VCEG-BU04-v1</p>
--	-------------------------------

Question: Q6/SG16 (VCEG)

Editors: Jonathan Pfaff (Fraunhofer HHI) Email: jonathan.pfaff@hhi.fraunhofer.de
Gary Sullivan (Dolby Laboratories) gary.sullivan@dolby.com
United States & Rapporteur halford@musc.edu
Study Group 16 Question 6)

Jonathan Halford (Medical
University of South Carolina &
Co-chair DICOM WG32)

Title: Call for Proposals on the coding of biomedical waveform data

Purpose: Call for Proposals

Status: Approved 25 January 2024 by Q6/16

1 Introduction

This document is a Call for Proposals (CfP) on the coding of biomedical waveform data. A need for the coding of such time-based neurophysiology signal data has been reported to Q6/16 in the liaison statement [SG16-TD103/Gen](#) from DICOM WG32. It is noted that there is no well-accepted compressed coding format for biomedical waveform data such as electrocardiography (ECG), electroencephalography (EEG), and electromyography (EMG) signals. It has since been demonstrated in the response [VCEG-BU03](#) to the Call for Evidence (CfE) [VCEG-BT07](#) on the coding of biomedical waveform data that compression technology exists with significantly higher compression performance than the identified benchmark set. The submission and evaluation criteria have followed the requirements laid out in the draft A.1 justification produced at the 10–21 July 2023 meeting of ITU-T SG16, as attached in [Annex A](#) of this document. Responses to this Call are therefore requested for consideration at the Q6/16 meeting of 12-19 July 2024 in Sapporo, Japan. Moreover, it is also possible for ITU-T members to submit initial responses for consideration at the SG16 meeting of 15–26 April 2024 in Rennes, France. Additional information submitted beyond July 2024 may also be considered.

2 Purpose and procedure

The purpose of this CfP is to collect and evaluate coding technology for biomedical waveform data. Companies and organizations that have identified or developed such technology are invited to submit a response to this Call.

To evaluate a proposed compression technology, bit rates will be traded off against distortion measures as specified in section 4. Moreover, in a later stage, DICOM experts are expected to evaluate whether data compressed by a proposed technology in the coding conditions specified in section 5 could have clinically-relevant differences or result in the same medical diagnoses

(performed either by human experts or by machines) as the original data. These evaluations are expected to follow the protocol which was developed by the DICOM WG32 group for an assessment of the impact of artificially added signal noise on medical diagnoses. Based on the results of the tests and based on technical aspects (like, e.g., computational complexity, memory requirements, minimum structural delay) the course of action regarding the proposed technologies will be decided.

Descriptions of responses shall be registered as input contributions to the Q6/16 meeting planned in Sapporo, Japan, 12–19 July 2024. Moreover, member proponents willing to provide preliminary versions of their responses to be discussed already at the SG16 meeting of 15–26 April 2024 in Rennes, France, are requested to register the description of their initial responses as input contributions to the SG16 meeting of 15–26 April 2024 in Rennes, France. Proponents also need to attend the meetings to which they submit their responses in order to present them. Further information about the SG16 meeting in Rennes can be found at <https://itu.int/md/T22-SG16-COL-0007>.

It is the intent of Q6/16 to ensure that any relevant experts who wish to respond to this Call will be able to do so and to contribute to the evaluation of responses to the maximum extent feasible within the policies and operational rules of ITU-T. Participants who want to submit a response to this call and are not currently ITU-T members should contact the Q6/16 Rapporteur Gary J. Sullivan (see Contacts in section 7 of this document) to request an invitation to participate in the July 2024 meeting. The Rapporteur plans to be supportive of such requests in order to facilitate the submission and discussion of responses by all interested parties who may choose to join in the work. Remote participation access to the July 2024 meeting is also expected to be provided (on a best-effort basis).

Q6/16 is expected to propose the creation of a work item in the SG16 work programme at the April meeting if sufficient support by members and/or other standardization organizations is expressed.

3 Timeline

The timeline for the Call for Proposals for proponents who want their initial responses to be discussed at the April 2024 (Rennes) meeting is as follows:

2024-04-02	Upload of document describing the submitted technology via member contribution to the SG16 document system (see the above-referenced Collective Letter for further information about this meeting)
2024-04-15/26	Consideration of responses at the meeting of SG16 in April 2024 (Rennes)

The timeline for the Call for Proposals for proponents who submit their response to the July 2024 (Sapporo) meeting is as follows:

2024-07-02	Upload of bitstreams and decoder software
2024-07-05	Upload of document describing the submitted technology
2024-07-12/19	Consideration of responses at the meeting of Q6/16 in July 2024 (Sapporo)

4 Error measures

Two error measures are employed to objectively evaluate the compressed representations of the test data. The input sequences are specified in section 5. Let N be the number of channels and let M be the number of samples per channel of an input sequence. Furthermore, let $a_{i,j}$ be the j -th sample (with $0 \leq j < M$) of channel i (with $0 \leq i < N$) and let $\tilde{a}_{i,j}$ be the corresponding reconstructed sample after decoding a bitstream. The maximum absolute error (MAE) is then defined as:

$$MAE = \max\{|a_{i,j} - \tilde{a}_{i,j}| \mid 0 \leq i < N, 0 \leq j < M\}.$$

Moreover, for m_i defined as the mean value of the input data for the i -th channel, i.e.

$$m_i = \frac{1}{M} \cdot \sum_{j=0}^{M-1} a_{i,j},$$

the percentage root mean square distortion (PRD) is defined as:

$$PRD = 100\% \cdot \sqrt{\frac{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (a_{i,j} - \tilde{a}_{i,j})^2}{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (a_{i,j} - m_i)^2}}.$$

Please note that, in contrast to some definitions found in the literature, this definition of the PRD includes a mean-removal in order to be invariant towards constant signal-shifts. In order to take different variance-ranges in different channels into account, the channel-normalized percentage root mean square distortion (CPRD) shall be defined as:

$$CPRD = \frac{100\%}{N} \cdot \sum_{i=0}^{N-1} \sqrt{\frac{\sum_{j=0}^{M-1} (a_{i,j} - \tilde{a}_{i,j})^2}{\sum_{j=0}^{M-1} (a_{i,j} - m_i)^2}}.$$

In the latter sum, all indices i for which:

$$\sum_{j=0}^{M-1} (a_{i,j} - m_i)^2 = 0$$

are to be excluded.

5 Test data and coding conditions

The input sequences to be tested are specified in three categories as specified in subsections 5.2 to 5.4. Nine working points (WP0 to WP8) are defined. Here, the first working point (WP0) defines a lossless compression while the last eight working points (WP1 to WP8) are defined in terms of restrictions on the bitstream size. The latter is measured by the number of bits per sample (BPS), defined as:

$$BPS = \frac{\text{\#number of bits in the bitstream}}{N \cdot M},$$

where N is the number of channels and M is the number of samples per channel of a given input sequence. The last eight working points target compression technologies which, for a given maximal BPS, minimize the PRD.

The working points are defined as follows:

- WP0: $MAE = 0$ (lossless)
- WP1: $BPS \leq 4.0$
- WP2: $BPS \leq 3.5$
- WP3: $BPS \leq 3.0$
- WP4: $BPS \leq 2.5$
- WP5: $BPS \leq 2.0$
- WP6: $BPS \leq 1.5$
- WP7: $BPS \leq 1.0$
- WP8: $BPS \leq 0.75$

One bitstream shall be produced for each input sequence and working point so that the conditions specified for the working point are fulfilled. Submission of bitstreams for the lossless working point WPO is not mandatory but is highly encouraged.

5.1 Availability and format of the test data

All test data used for the CfP can be downloaded from the following location:

Server: <ftp.hhi.fraunhofer.de>

Login: dicom

Password: yX5GUw.Zn

The files are provided in the European Data Format (EDF). See Bob Kemp, Alpo Värri, Agostinho C. Rosa, Kim D. Nielsen and John Gade. "[A simple format for exchange of digitized polygraphic recordings](#)". *Electroencephalography and Clinical Neurophysiology*, 82 (1992): 391-393. See also <https://www.edfplus.info/>.

5.2 Electroencephalography (EEG) signals

Name of dataset: EEG dataset containing interictal epileptiform discharges and seizures

FTP-file: MUSC_Dataset_E.zip

Number of input sequences: 41

5.3 Electrocardiography (ECG) signals

Name of dataset: MIT-BIH Arrhythmia Database

FTP-file: MIT_ECG_Dataset.zip

Number of input sequences: 48

5.4 Electromyography (EMG) signals

Name of dataset: Dataset for multi-channel surface electromyography (sEMG) signals of hand gestures

FTP-file: MENDELEY_Dataset.zip

Number of input sequences: 40

6 Requested content of submissions

Proponents are requested to submit a technical description of the proposed technology sufficient for full conceptual understanding and generation of equivalent performance results by experts to the meeting where the evaluation is performed.

Proponents should implement their proposed technology in software and include information about the used programming language in their proposal document.

Proponents should also upload bitstreams for all input sequences and working points of at least one of the three categories specified in sections 5.2 to 5.4 to the ftp server specified in section 5.1 by the date specified in the timeline of section 3. The access data for uploading the bitstreams can be obtained from the CfP coordinators of section 7 upon request.

Proponents shall report *PRD* and *CPRD* values for each bitstream along with the number of bits per sample (*BPS*).

Proponents are encouraged to provide a draft versions of possible specification texts of their submission which includes at least a bitstream syntax and semantics.

For each category and each of the eight working points WP-1 to WP-8, the average of the *BPS*, *PRD* and *CPRD* values of all associated bitstreams shall be reported.

For each input sequence, a graph shall be provided that shows the *PRD* values over the *BPS* values for the eight working points WP-1 to WP-8 and connects the working points by linear or some other interpolation method. Moreover, for each input sequence, a second graph shall be provided that shows the *CPRD* values over the *BPS* values for the eight working points WP-1 to WP-8 and connects the working points by linear or some other interpolation method.

Proponents are requested to provide binary executable decoders to decode the submitted bitstreams and a description of the necessary runtime environment. Provision of source code for encoding and decoding is encouraged but not required. The provided decoder software must be capable of decoding the bitstreams and storing the decoded data in the same format as the test sequence or in a similar uncompressed raw format. When a different uncompressed raw format is used, proponents must provide software or at least a clear description to show how to readily extract the channel data from the decoded files. Submissions of software shall include permission to enable the software to be copied and used at least for purposes of evaluation of the submissions. In the event that ITU-T chooses to develop a standard on this subject at a later stage, it is likely that source code would be expected to be provided for encoding, decoding, and experimentation, and potentially to be used as the basis of implementations of the standard, and such software would need to conform to the ITU-T Software Copyright Guidelines. As with prior projects of Q6/16, the development of software source code to be published and available under a "permissive software licence" is likely to be undertaken. The development of a standard would also be subject to the Common Patent Policy for ITU-T/ITU-R/ISO/IEC. Further information on IPR management in ITU-T is found at <https://itu.int/en/ITU-T/ipr/>.

7 Contacts

Jonathan Pfaff (CfP coordinator and chair of Q6/16 AHG on coding of biomedical waveform data), Gary J. Sullivan (Rapporteur Q6/16), Jonathan Halford (DICOM WG 32 co-chair).

Coordinator Email: jonathan.pfaff@hhi.fraunhofer.de; halfordj@muscc.edu; gary.sullivan@dolby.com

Annex A:

Draft A.1 justification for potential new Recommendation on biomedical and waveform signal coding

Question:	Q6/16	Proposed new ITU-T Recommendation	Online or Venue, day-day month year
Reference and title:	Biomedical and waveform signal coding		
Base text:	N/A	Timing:	YYYY-MM
Editor(s):	Name, membership, e-mail address	Approval process:	AAP
Scope:			
<p>A compressed coding format for medical waveform data (e.g. neurophysiology, electrocardiography, and so on) will be defined. This compressed format is targeted towards medical applications in DICOM and other organizations</p>			
Summary (provides a brief overview of the purpose and contents of the Recommendation, thus permitting readers to judge its usefulness for their work):			
<p>Currently, there lacks a codec with sufficient compression capability for neurophysiology and electrocardiography data and other similar types of biomedical waveform signals. Current audio codecs have significant limitations when applied to biomedical waveform data due to the application of psychoacoustic masking and limitations on channel number and block size. As such, a new standard that can support efficient lossless and near-lossless compression and transmission specifically for biomedical waveform data is almost certainly needed.</p> <p>Neurophysiology data consists of 16-24 bit data with between 21 and several hundred channels sampled at between 256 and 40K Hz. Clinical scalp EEG is typically 21 channels sampled at between 256 and 1000 Hz. Clinical intracranial EEG consists of between 64 and several hundred channels sampled with a bit depth of 16 or 24 bits at around 2000 Hz. Research on intracranial EEG (human and animal neurophysiology) involves up to several hundred channels sampled at a bit depth of 16 or 24 bits and a sampling rate up to approximately 40K Hz.</p> <p>The codec will need to meet the following minimum requirements:</p> <ol style="list-style-type: none"> 1. Be able to support 16 to 24 bits (possibly higher) per sample. 2. Be able to support a large range of sampling rates (500–2000 sample/s for clinical use, up to 40k samples/s for research). 3. Be able to support a large number of channels (up to hundreds). 4. Be able to support channels with the same or different sampling rates. 5. Be able to support lossy, near-lossless and lossless compression. A lossless-only approach will not produce better than a 2-3X compression ratio, which is not sufficient. 6. Be able to support blocking and indexing for rapid access within large datasets. Block size should be optimized, and an index to blocks can be stored outside of encapsulated bitstream (such as in fragments/frames encoding in DICOM). 7. Have a mode where independent decoding of channels is supported. 8. Associated metadata should be supported (possibly stored outside of the bitstream). 			
Relations to ITU-T Recommendations or to other standards (approved or under development):			
TBD			
Liaisons with other study groups or with other standards bodies:			
DICOM WG32			
Supporting members that are committing to contributing actively to the work item:			
Fraunhofer HHI, others TBD			