 ISO/IEC JTC 1/SC 29/WG 2 N0038

**ISO/IEC JTC 1/SC 29/WG 2**

**MPEG Technical requirements   
Convenorship: SFS (Finland)**

**Document type:** Output Document

**Title:** Draft Submissions and Evaluation Procedures for the Haptics CfP – Phase 1

**Status:** Approved

**Date of document:** 2021-01-15

**Source:** ISO/IEC JTC 1/SC 29/WG 2

**Expected action:** none

**Action due date:** none

**No. of pages:** 33 (without cover page)

**Email of Convenor:** igor.curcio@nokia.com

**Committee URL:** <https://isotc.iso.org/livelink/livelink/open/jtc1sc29wg2>

**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 2**

**MPEG TECHNICAL REQUIREMENTS**

**ISO/IEC JTC 1/SC 29/WG 2 N0038**

**Online – January 2021**

|  |  |
| --- | --- |
| **Title:** | **DRAFT Submission and Evaluation Procedures for the Haptics CfP – Phase 1** |
| **Source:** | **WG 2 MPEG Technical Requirements** |
| **Editors:** | **Yeshwant Muthusamy, Chris Ullrich, Jamal Saboune (Immersion Corporation)**  **Philippe Guillotel, Fabien Danieau, Quentin Galvane (InterDigital Corporation)** |
| **Status:** | **Approved** |
| **Serial number** | **20122** |

1. **Table of Contents**

[1 Introduction 2](#_Toc61548898)

[2 Submission Procedures 3](#_Toc61548899)

[2.1 Access to Test Content 3](#_Toc61548900)

[2.2 Preparation of coded items 4](#_Toc61548901)

[2.3 Generation and Parsing of OHM Files 7](#_Toc61548906)

[3 Preparation of Test Material 8](#_Toc61548907)

[4 Objective Signal-based Tests 9](#_Toc61548908)

[4.1. PSNR 9](#_Toc61548913)

[4.2 Vibrotactile Haptics 9](#_Toc61548915)

[4.3 Kinesthetic Haptics 10](#_Toc61548916)

[5 Expressive Capabilities Tests 10](#_Toc61548917)

[6 Subjective Testing Procedures 13](#_Toc61548918)

[7 Evaluation Procedures 16](#_Toc61548919)

[7.1 Overview 16](#_Toc61548920)

[7.2 Statistical Analysis 16](#_Toc61548921)

[7.3 Figure of Merit 17](#_Toc61548922)

[7.3.1 Figure of Merit for the MUSHRA Tests 17](#_Toc61548923)

[7.3.2. Overall Figure of Merit 18](#_Toc61548935)

[7.4 Set 1 and Set 2 Performance and determination of RM0 19](#_Toc61548936)

[8 References 19](#_Toc61548937)

[ANNEX 1 – Test Protocol for Vibrotactile Haptics Tests (Tests 1.1 and 1.2) 20](#_Toc61548938)

[A1.1 Overview 20](#_Toc61548939)

[A1.2 System Calibration and Installation 20](#_Toc61548940)

[A1.3 User Study Protocol 20](#_Toc61548941)

[ANNEX 2 – Test Protocol for Kinesthetic Haptics Tests (Test 1.3) 23](#_Toc61548942)

[A2.1 Overview 23](#_Toc61548943)

[A2.2 Installation 23](#_Toc61548944)

[A2.2.1 3D System Touch device driver 23](#_Toc61548945)

[A2.2.2 Virtual Audio Cable 24](#_Toc61548946)

[A2.2.3 Kinesthetic driver 25](#_Toc61548947)

[A2.3 Usage 25](#_Toc61548948)

[A2.4 User Study Protocol 27](#_Toc61548949)

[ANNEX 3 – Test Subject Instructions 29](#_Toc61548950)

[A3.1 Overall instructions 29](#_Toc61548951)

[A3.1.1 User Interface 29](#_Toc61548952)

[A3.2 Training phase instructions 31](#_Toc61548953)

[A3.3 Test phase instructions 31](#_Toc61548954)

# Introduction

The Draft Call for Proposals on the Coded Representation of Haptics (Call) was issued at the 132nd MPEG meeting in October 2020. At the 133rd MPEG meeting a Revised Call was issued [1]. A Final Call is expected to be issued at MPEG 134 in April 2021 with the submissions evaluated at MPEG 136 in October 2021.

This document gives timeline and submission procedures for responding to the Call. Respondents are strongly urged to review this document so that their submissions are according to the guidelines and hence can be incorporated into the subjective testing process in a timely manner.

The following companies have registered the intent to participate in the Call. The rightmost columns indicate an intention to participate in Phase 1 or Phase 2 of the Call and to submit bitstreams and decoded signals for vibrotactile (V) and kinesthetic (K) tests.

**Table 1 - Registered Companies**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **ID** | **Company** | **Contact** | **Email** | **Phase** | | **Type** | |
| **1** | **2** | **V** | **K** |
| 1 |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |
|  |  |  |  | **Phase 1 Totals** |  |  |  |  |

# Submission Procedures

The following table gives the timeline for submissions to the Call. The two-phase submission of raw test scores gives the Test Administrator more time to perform data analysis.

**Table 2 – Timeline for Phase 1**

|  |  |
| --- | --- |
| **Meeting / Date** | **Action** |
| May 15, 2021 | Proponent Contact to receive Call website information from Test Administrator |
| July 5, 2021 | Proponent processed test items submitted to Call FTP site. |
| July 10, 2021 | Test items available to Test Labs on Call FTP site |
| July – September 2021 | Conduct evaluation tests |
| September 25, 2021 | Test Labs submit Test raw scores via email to Test Administrator |
| September 30, 2021 | Test Labs submit   * Contribution to 136th MPEG meeting describing their test setup. |
| 136th meeting, October 2021 | Proponent written documentation submitted as contribution to 136th MPEG meeting (see Call for Proposals for details on what documentation must be submitted and when it must be submitted). |
| 136th meeting, October 2021 | Selection of Reference Model 0 technology |
| 137th meeting, January 2022 | Proponent(s) must submit Reference Model 0 Working Draft text and Reference Software. |

**Test Administrator**

The Test Administrator for the Call shall be Dr. Yeshwant Muthusamy, whose contact information is:

Yeshwant Muthusamy

Immersion Corporation

Email: [ymuthusamy@immersion.com](mailto:ymuthusamy@immersion.com)

Phone: +1 469-583-2171

## Access to Test Content

Access to Call test items and necessary software requires proponent companies to execute agreements with the companies that provided the individual test items and software. On May 15, 2021, the proponent contact person shown in Table 1 will receive the license agreements by automated email from the Test Administrator. Upon electronically signing the agreements (the signed agreements will be automatically returned to the Test Administrator), the Administrator will send them an account activation link on ShareFile. Activating their account will give them access to the folders containing the test items and software. The proponents will have access to these folders for the entire duration of the Call.

The file format of the Original signals and file naming of the Original test items files shall be as in N40, “Encoder Input Format for MPEG Haptics” [2].

## Preparation of coded items

The following terms are defined to facilitate unambiguous specification of bitrates of proponent submitted compressed representations:

* **Nominal Bitrate** – the bitrate as specified in Table 3 and Table 4, e.g., 32 kb/s.
* **Actual Bitrate** – the bitrate for a specific compressed representation of an item submitted by a proponent, as specified in this section under heading “Bitrate of compressed representation.”

Proponents must participate in the tests for Set 1 and Set 2 Vibrotactile Haptics items, which are the Short and Long Effects, respectively, and (optionally) in the tests for Set 3 Kinesthetic Haptics items as well.

Proponents shall encode and decode each item separately, submitting a compressed representation and decoded waveforms for each item as specified in this section.

All proponent decoded WAV files shall be time-aligned to within +/- 1 sample relative to the original WAV files and shall be processed versions of the *complete* original WAV file (i.e., the processed version shall have no more or fewer samples at the beginning or end of the WAV file as compared to the original).

Proponent decoders shall not assume that object meta-data is known to the decoder. For example, if proponent decoders wish to decode and manipulate object signals based on object meta-data, then the proponent bitstreams must carry the object meta-data information, and this information shall be counted as part of the total compressed item bitrate.

Proponents shall submit the following for tests using Set 1 items:

**Table 3 – Submissions for Set 1 – Short Vibrotactile Effects**

|  |  |
| --- | --- |
| **Test** | **Submit** |
| Test 1.1 | Compressed representations for the target bitrates of  32 kb/s, 16 kb/s, 1 kb/s  Decoded WAV files and corresponding OHM files associated with each bitrate. |

Note: kb/s indicates units of 1000 bits per second.

Proponents shall submit the following for tests using Set 2 items:

**Table 4 – Submissions for Set 2 – Long Vibrotactile Effects**

|  |  |
| --- | --- |
| **Test** | **Submit** |
| Test 1.2 | Compressed representations for the target bitrates of  32 kb/s, 16 kb/s, 1 kb/s  Decoded WAV files and corresponding OHM files associated with each bitrate. |

Note: kb/s indicates units of 1000 bits per second.

**Table 5: Submissions for Set 3 – Kinesthetic Effects**

|  |  |
| --- | --- |
| **Test** | **Submit** |
| Test 1.3 | Compressed representations for the target bitrates of  32 kb/s, 16 kb/s, 1 kb/s  Decoded WAV files and corresponding OHM files associated with each bitrate. |

**Naming of compressed representation files for all tests**

Proponents shall submit compressed representation for all bitrates in Tests 1.1 and 1.2, and Test 1.3 if they are participating in the optional Set 3 Tests.

Proponents shall submit using the following naming procedure.

Naming of compressed representation:

<test>\_<id>\_<haptics\_type>\_<bitrate>\_<item\_name>.<ext>

Where

<test> shall be one of:

Test1-1

Test1-2

Test1-3

<id> is the proponent ID field from Table 1 - Registered Companies

<haptics\_type> is vib (vibrotactile) or kin (kinesthetic)

<bitrate> is in kb/s and constrained to be one of 32, 16, 1.

<item\_name> is as shown in the column with heading “Name” in Tables 3 through 7 of Annex 5 in the Call.

<.ext> can be one of .ahap, .ivs, or .wav. Alternatively, <.ext> could indicate a proprietary file format.

Examples are shown here:

Test1\_1\_<id>\_vib\_16\_RedBullFly.<ext>

Test1-3\_<id>\_kin\_1\_BikeRiding.<ext>

**Bitrate of compressed representation**

For Test 1.1 through Test 1.3, for each test item and each target bitrate, proponents shall submit as part of their written documentation a measure of actual bitrate as:

* Average bitrate per coded item, calculated as (total\_bits/item\_duration) in kb/s.

Proponents shall clearly explain how to calculate the coded bitrate from the submitted compressed representation. If such bitrate calculation is beyond simple computations based on file size, proponent shall provide a utility as both source-code and executable that can be used in the bitrate computation.

**Naming of compressed representation zip archives**

The compressed representations for all items in Test 1.1 through Test 1.3 shall be put in a single zip archive named as:

<test>\_<id>\_<haptics\_type>\_<bitrate>.zip

where

<test>, <id>, <haptics\_type>, and <bitrate> are as in the naming of the compressed representation files.

Examples are shown here:

Test1-1\_<id>\_vib\_1.zip

Test1-3\_<id>\_kin\_16.zip

Zip archives should unpack into a folder named as

<test>\_<id>\_<haptics\_type>\_<bitrate>

**Latency issues**

All compressed representations and associated encoder/decoder technology shall comply with the Latency restrictions specified in Annex 1 of the Call. Note that using a per-file-based compression technology such as “zip” as part of a proprietary compressed representation would obviously not satisfy the Latency requirement. Additionally, two-pass coding would not satisfy the Latency requirement.

**Naming of decoded WAV files for Test 1.1 through Test 1.3**

Proponents shall submit decoded waveforms (as .WAV files) suitable for feeding to the actuator. The decoded .WAV files shall contain content sampled at 8 kHz and 16 bits.

Proponents shall submit using the following naming procedure.

Naming of decoded WAV files:

<test>\_<id>\_<haptics\_type>\_<bitrate>\_<item\_name>\_<padding>.wav

<test> shall be one of:

Test1-1

Test1-2

Test1-3

<id> shall be the company ID listed in Table 1 - Registered Companies, shown above.

<haptics\_type> shall be vib (for vibrotactile) or kin (for kinesthetic)

<bitrate> shall be nominal bitrate in kb/s, i.e., 32, 16, or 1.

<item\_name> is as shown in the column with heading “Name” in Tables 3 through 6 of Annex 5 in the Call.

<padding> is either pad or nopad, depending on whether there are 1000 ms of leading and trailing silences or not, respectively.

Examples are shown here:

Test1-1\_<id>\_vib\_32\_Inflate\_nopad.wav

Test1-2\_<id>\_vib\_1\_LexusF\_pad.wav

Test1-3\_<id>\_kin\_16\_BigBuckBunny\_pad.wav

**Submission of zip archives**

Each registered proponent shall receive a website address and a unique username/password valid at that website address for use in submission, which will log into the root of the proponent upload directory. All zip archives shall be put at the root of the proponent’s directory.



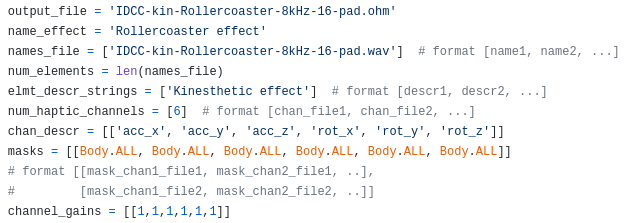
## Generation and Parsing of OHM Files

The OHM files (described in [2]) can be generated and parsed using the following software package.

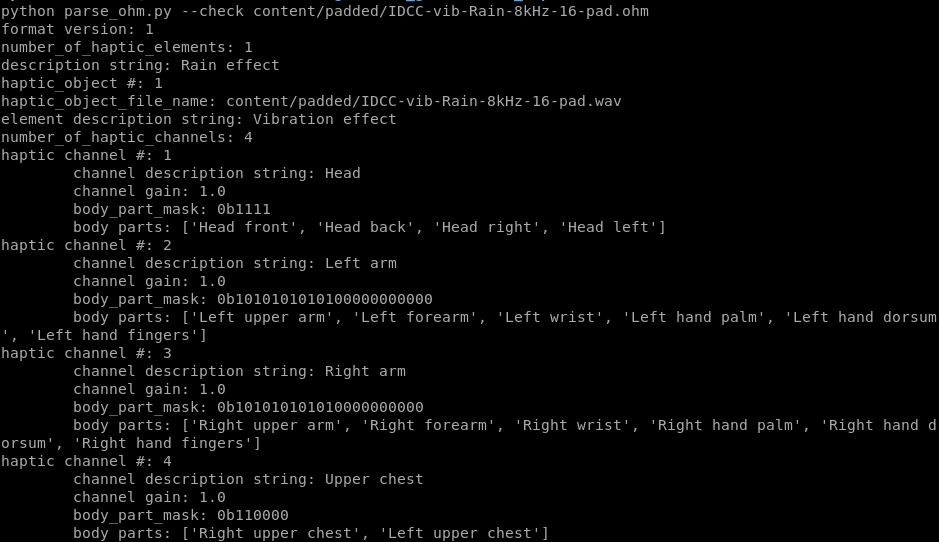
The code is located on a git repo: <https://github.com/InterDigitalInc/MPEG_Haptics>

It contains two python files: generate\_ohm.py and parse\_ohm.py. It requires python 3 (tested with python 3.8.5). parse\_ohm.py requires the optional libraries soundfile and matplotlib in order to display the signals.

Simply run ‘python generate\_ohm.py’ to generate an ohm file. Parameter should be modified within the python file in the main function. Several examples are provided such as:



To parse an ohm file run ‘python parse.ohm <file\_name>’. <file\_name> being the path to the file. For instance:

Adding the option –check will display a figure with the signal, such as that shown in Figure 3 of [2].

# Preparation of Test Material

All tests will incorporate an open reference and a hidden reference. There will be no low-pass or mid-band anchors used due to the low frequencies of the haptic signals themselves.

The Test Administrator will rename or “blind” all test material so that the proponent ID is replaced by the string “sysN,” where N is an integer. Hidden reference will be similarly blinded.

**Test Material Format**

On the date that test items are available to Test Labs, as indicated in Table 2 – Timeline for Phase 1, the Test Administrator will send the Test Labs the test content license agreements by automated email. Upon electronically signing the agreements (the signed agreements will be automatically returned to the Test Administrator), the Administrator will send them an account activation link on ShareFile. Activating their account will give them access to the folders containing the test items.

The test material for each test will not be placed in zip archives, but instead will be available uncompressed.

The TestLabs folder will be organized with the following directory structure:

* Originals/Vibrotactile\_Short\_Effects/<contentprovidername>-vib-<contentname>-<samplingrate>-<bitspersample>-pad.wav
* Originals/Vibrotactile\_Long\_Effects/<contentprovidername>-vib-<contentname>-<samplingrate>-<bitspersample>-pad.wav
* Originals/Kinesthetic\_Effects/<contentprovidername>-kin-<contentname>-<samplingrate>-<bitspersample>-pad.wav
* <test>/Vibrotactile\_Short\_Effects/<test>\_<sys>\_<bitrate>\_<item\_name>\_pad.wav
* <test>/Vibrotactile\_Long\_Effects/<test>\_<sys>\_<bitrate>\_<item\_name>\_pad.wav
* <test>/Kinesthetic\_Effects/<test>\_<sys>\_<bitrate>\_<item\_name>\_pad.wav

# Objective Signal-based Tests



## PSNR

The Peak Signal-to-Noise Ratio (PSNR) is commonly used to measure the quality of reconstruction of lossy compression codecs (e.g., for audio or image compression). It is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation.

PSNR is used to evaluate both vibrotactile and kinesthetic signals and is part of the objective signal-based tests described in Section 4.6.1 of the Call [1]. The signal is the original data, and the noise is the error introduced by the compression (or any other processing).

A python code to compute the PSNR is provided. It requires a Python 3.6 environment with the numpy and soundfile packages installed. These packages can be installed with the following commands:

> conda install numpy

> conda install soundfile

The PSNR can then be computed with the following command:

> python psnr.py ORIGINAL\_FILE COMPRESSED\_FILE

where ORIGINAL\_FILE is the input file provided with the test platform, and COMPRESSED\_FILE is the output file provided by the proponents.

The files should have the same number of samples. An error message is returned if this is not the case.

The output is the corresponding PSNR value in dB.

The PNSR python code may be downloaded from this site: <https://github.com/InterDigitalInc/MPEG_Haptics>



## Vibrotactile Haptics

The STSIM algorithm [4], part of the objective signal-based tests described in Section 4.6.1 of the Call [1], is provided as source/object code for MATLAB. Based on pre-testing and PCM coding considerations, the algorithm constants have been set as follows:

Fs = 8000Hz

η = 0.6

frameLen = 200

The reference signal should be loaded into MATLAB as x. The test signal should be loaded into MATLAB as variable y. Once this is done, the STSIM algorithm can be executed simply by running the main.m code. The objective similarity measure is stored in the output variable stsim.

The STSIM MATLAB code may be downloaded by registered proponents from the ShareFile site maintained by the Test Administrator, as described in Section 2.1.

## Kinesthetic Haptics

The HSSIM algorithm [5], part of the objective signal-based tests described in Section 4.6.1 of the Call [1], is provided as source/object code for MATLAB. The algorithm parameters have been set based on prior testing by the IEEE 1918.1 standards group as follows:

c = 1.4

α = 1.0

β = 1.0

γ = 0.5

The reference signal should be loaded into MATLAB as x. The test signal should be loaded into MATLAB as variable y. Once this is done, the HSSIM algorithm can be executed simply by running the main.m code. The objective similarity measure is stored in the output variable hssim.

The HSSIM MATLAB code may be downloaded by registered proponents from the ShareFile site maintained by the Test Administrator, as described in Section 2.1.

# Expressive Capabilities Tests

The goal of the expressive capabilities tests is to ensure that any proposal meets the basic requirements of the Call. This is needed because some aspects of the system are not measurable using purely signal based or MUSHRA evaluation within a reasonable time.

The expressive capabilities tests will consist of a series of questions that are answered yes/no independently by 3 expert reviewers drawn from the administrator and other non-proponent experts in haptics. *Only the proponent’s submission will be utilized in answering these questions and no other information is permissible*. A score of 1 point will be awarded for each ‘yes’ response. The average of the three reviews will be used to assign an Expressive Capabilities Test score to each submission.

Proponents must provide a document that explains how their submission satisfies each question as part of their submission. The document should ideally clarify which components of the submission are addressed by the coded representation itself and which are addressed by their API.

The questions are directly derived from the requirements found in section A1.1 of the Call.

|  |  |  |  |
| --- | --- | --- | --- |
| **Phase 1 Requirement** | **Capability Test** | **Coding** | **System Level** |
| **Coding Leve**l | |  |  |
| High Quality | Does the coding permit wide-band signals (independent frequency and amplitude)? |  |  |
| Update Rate | Does the coding support presentation sample rates of up to 1kHz? |  |  |
| Playback Transcoding | Does the coded representation enable transcoding? |  |  |
| Target hardware | Does the coded representation support signaling of intended rendering capabilities? |  |  |
| Physical placement | Can the coded representation support mapping of signals for a specific user-relative actuator placement? |  |  |
| Track counts | Does the coded representation support multiple simultaneous tracks? |  |  |
| Tracks | Does the coded representation support multiple versions of the same haptic effects for different endpoints? |  |  |
| Compression | Does the coded representation support lossy or lossless compression? |  |  |
| Multi-actuator systems | Does the coded representation support multi-actuator systems? |  |  |
| Mixing | Can the coded representation be mixed with other coded representations during presentation? |  |  |
| Haptic modality descriptor | Does the coding have modality descriptors that distinguish at least between ‘vibrotactile’ and ‘kinesthetic’? |  |  |
| **System Level** | |  |  |
| Modulation | Does the system support user modulation of overall sensation? |  |  |
| Arbitrary temporally driven hardware | Does the coding system support synthesis for general haptic hardware? |  |  |
| Synchronization | Does the system enable perceptual synchronization? |  |  |

A sample submission is shown below with the completed version of these questions for the Apple Core Haptics system.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Phase 1 Requirement** | **Capability Test** | **Coding** | **System Level** | **Justification /Evidence** |
| **Coding Level** | |  |  |  |
| High Quality | Does the coding permit wide-band signals (independent frequency and amplitude)? | Yes | Yes | Core Haptics can synthesize and render wideband signals from descriptive effects described using AHAP. |
| Update Rate | Does the coding support presentation sample rates of up to 1kHz? | Yes | Yes | Event timing is specified using 32-bit float values in AHAP and CoreHaptics can synthesize/render at 1kHz. |
| Playback Transcoding | Does the coded representation enable transcoding? | Yes | Yes | AHAP is a descriptive format and does not provide device-specific coding, it must always be synthesized and transcoded when rendered.  Distinct iPhone and external game controllers are supported, given identical AHAP inputs. |
| Target hardware | Does the coded representation support signaling of intended rendering capabilities? | No | Yes | API calls allow routing of an AHAP file to a specific device type, but this is not present in the coded representation. |
| Physical placement | Can the coded representation support mapping of signals for a specific user-relative actuator placement? | No | Yes | API calls allow for left/right motor selection on game controllers. (createEngineWithLocality), but this is not present in the coded representation. |
| Track counts | Does the coded representation support multiple simultaneous tracks? | Yes | Yes | Multiple simultaneous effects may be specified in AHAP directly using the structured JSON format. |
| Tracks | Does the coded representation support multiple versions of the same haptic effects for different endpoints? | No | No | Core Haptics can be used to manage multiple AHAP files depending on the target device, but the coded representation does not reflect this. Developers must explicitly manage the different devices. |
| Compression | Does the coded representation support lossy or lossless compression? | No | No | Compression is not contemplated in this encoding. |
| Multi-actuator systems | Does the coded representation support multi-actuator systems? | No | Yes | Locality related functionality in the API allows programmatic specification of target actuators but not in the coded representation. |
| Haptic modality descriptor | Does the coding have modality descriptors that distinguish at least between ‘vibrotactile’ and ‘kinesthetic’? | No | No | No modality descriptors are present. |
|  |  |  |  |  |
| **System Level (CoreHaptics)** | |  |  |  |
| Modulation | Does the system support user modulation of overall sensation? | No | Yes | Global modulation is part of the Core Haptics API. |
| Arbitrary temporally driven hardware | Does the coding system support synthesis for general haptic hardware? | No | Yes | Since the coding is purely descriptive, it can be interpreted and synthesized for any target hardware. |
| Synchronization | Does the system enable perceptual synchronization? | No | Yes | API calls explicitly enable synchronized audio/haptic presentation. |
| Mixing | Can the coded representation be mixed with other coded representations during presentation? | No | Yes | Multiple simultaneous effects are supported. Mixing is not specified in the coding standard. |
|  |  |  |  |  |
| Total |  | 4 | 11 |  |

# Subjective Testing Procedures

A number of Test Labs have been identified that have the requisite test setups. Those labs are indicated in Table 6, below.

**Table 6 – Labs Participating in MUSHRA Tests**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **TL\_ID** | **Company** | **Contact Name** | **Email** |
| 1 | TL\_1 | Immersion Corporation | Jamal Saboune | jsaboune@immersion.com |
| 2 | TL\_2 | TU Delft (tentative) |  |  |
| 3 | TL\_3 | McGill University (tentative) |  |  |

**Allocation of Tests to Test Labs**

For each of Set 1 and Set 2 subjective tests, Test Labs have committed to participating in the tests.

The Test Administrator will create FTP site passwords for each test (i.e., 9 in total). The Test Administrator will give the passwords to each Test Lab for the tests in which they participate. In addition, registered proponents that submit technology will receive passwords for the content set (i.e., Set 1 and/or Set 2) for which they submit. Finally, Test Labs that have submitted all test data for which they have committed (as shown in the table) may volunteer to run additional tests in which case the Test Administrator will give the passwords to that Test Lab. It is understood that the minimum number of test subjects for each test is 8.

**Table 7 – Test Labs Participating in Set 1 or Set 2 Tests. Number under each Test Lab ID is the minimum number of subjects that will participate in the test.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Test** | **Immersion** | **<Test Lab 2>** | **<Test Lab 3>** | **….** |
| Test1-1 | 8 |  |  |  |
| Test1-2 | 8 |  |  |  |
| Test1-3 |  |  |  |  |

**MUSHRA Testing Programs**

The Test Labs can use any presentation system that complies with the MUSHRA methodology, but it is highly recommended that the ARL STEP (<https://www.audioresearchlabs.com/arl-step/>) be used, to ensure compatibility in test input and output formats across Test Labs. Test Labs using the ARL STEP system are urged to read the “STEP\_ReadMe.doc” file in the zip archive of this document for an overview of how to use STEP Version 1.0 in the Haptics tests. An ARL STEP Session file (\*.asi) will be provided for each test.

**Test Protocol for Vibrotactile Haptics MUSHRA Tests**

The test protocol for the Vibrotactile Haptics (Tests 1.1 and 1.2) are described in Annex 1.

**Test Protocol for Kinesthetic Haptics MUSHRA Tests**

The test protocol for the Kinesthetic Haptics (Test 1.3) is described in Annex 2.

**Test Subject Training**

The test subject instructions for training and testing are shown in Annex 3.

**Score Submission Format**

Test Labs will submit the listening test raw scores to the Test Administrator on the date indicated in Table 2 – Timeline for Phase 1. Raw scores shall be submitted via the Excel spreadsheet found in the zip archive of this document. Results for each test shall be submitted as a separate spreadsheet file using the Test\_ID\_score\_template.xlsx file. See the example spreadsheet for score submission, “ExampleScoreSubmission.xlsx”

**Excel Spreadsheet**

Spreadsheet files shall be named as:

<test>\_<tl\_id>\_scores.xls

Where

<test> shall be one of:

|  |
| --- |
| Test1-1 |
| Test1-2 |
| Test1-3 |

<tl\_id> is the Test Lab ID shown in Table 6 – Test Labs Participating in MUSHRA Tests.

**ARL STEP score files**

Test Labs using the ARL STEP testing software can *additionally* submit a zip archive of the raw subject scores. These should be named as:

<subject\_id> txt

They should be put in a zip archive named as

<test>\_<tl\_id>.zip

which unpacks to the directory

<test>\_<tl\_id>

**Description of test setup**

Test Labs shall describe their test setup in a contribution to 136th MPEG meeting

# Evaluation Procedures

## Overview

When subjective testing is complete, there will be raw scores for each proponent participating in a given Set for the following tests:

The tests for Set 1 and Set 2 items are:

**Table 8 – Vibrotactile Haptics Tests**

|  |  |  |
| --- | --- | --- |
| **<test>** | **Test** | **No. of Items** |
| Test1-1 | Short Effects (Set 1) | 9 |
| Test1-2 | Long Effects (Set 2) | 24 |

The tests for Set 3 items are:

**Table 9 – Kinesthetic Haptics Tests**

|  |  |  |
| --- | --- | --- |
| **<test>** | **Test** | **No. of Items** |
| Test1-3 | Kinesthetic Effects | 10 |

## Statistical Analysis

**Post-Screening**

In order to enhance the significance of the subjective test results, post-screening of the test subject scores is done as follows: all data associated with subjects whose score for the hidden reference stimuli was below 90 are removed from the corresponding test.

**Data Analysis**

Data analysis will be done using two methods:

* Excel Pivot table, computing mean and 95% confidence interval (CI) for each system under test.
* ANOVA using a linear model of factors and interactions. This will produce the same mean scores as with the Excel Pivot table and a, typically smaller, common 95% CI for the systems under test.

The ANOVA will be used to check assumptions for pooling:

* Normal error distribution with approximately equal variance across factor levels (e.g., across Test Labs) can be pooled.

It is expected that the ANOVA will use the following linear models, although these may change to exclude factors and interactions that are not significant, and to incorporate factors and interactions that are found to be significant:

Test 1.1 Score = TL + TL\*Sys + Sig + Sys + Sig\*Sys + Error

Test 1.2 Score = TL + TL\*Sys + Sig + Sys + Sig\*Sys + Error

Test 1.3 Score = TL + TL\*Sys + Sig + Sys + Sig\*Sys + Error

Factors are:

TL Test Lab

Sig Test items

Sys System under test

.

A follow-on to the ANOVA will report the result of the Tukey Honestly Significant Difference (HSD) test for the systems under test. (which is a function in the “R” statistical analysis package [3]).

## Figure of Merit

### Figure of Merit for the MUSHRA Tests

Tests 1.1 through 1.3 assess proponent haptic technology at very different operating and test modes. Hence, it does not make sense to pool all subjective scores and calculate a grand mean and associated 95% confidence interval as a single figure of merit.

Instead, the relative merits of each test are reflected in a set of weights used to combine the grand mean for a proponent system in each test or test operating mode (e.g., Test1-1, 16 kb/s) as an overall score for a proponent system.

Table 10 and Table 11, below, gives the weights used for combining subjective scores into a single Figure of Merit, for the Vibrotactile (Set 1 and Set 2) and Kinesthetic Haptics (Set 3) tests, respectively.

A Figure of Merit () is computed for each system under test for Test Components 1 through 6. It is computed as follows:

* The mean score (M) for each system under test is computed for each Test Component, where the Test Components are enumerated in column 1 of Table 10, below.
* For each Test Component, the highest mean score is identified (Mmax) and the value Mnorm = 100 – Mmax is computed.
* For each Test Component, the mean scores are scaled as Mscaled = M + Mnorm. In this way, the highest mean score is translated to 100.
* For each Test Component, the lower (L) and upper (U) value of the confidence interval is scaled using the offset Mnorm as Lscaled = L + Mnorm and Uscaled = U + Mnorm.
* The FoM for a system under test is the sum of the weighted Mscaled.
* In addition, for a system under test, the sum of the weighted Lscaled and Uscaled are computed.

In summary, the FoM is computed as:

where *i* is the index over Test Component and *Wi* are the weights shown in Table 10.

Similarly, the Figure of Merit for the optional Kinesthetic Haptics (Set 3) tests would be as follows:

where *i* is the index over Test Components *Wi* are the weights shown in Table 11.

**Table 10 – Figure of Merit weights for the Vibrotactile Haptics Tests**

|  |  |  |  |
| --- | --- | --- | --- |
| **Test**  **Component** | **Test** | **Bitrate and Configuration** | **Weight** |
| 1 | Test 1.1 | 32 kb/s | 10% |
| 2 | Test 1.1 | 16 kb/s | 15% |
| 3 | Test 1.1 | 1 kb/s | 25% |
| 4 | Test 1.2 | 32 kb/s | 10% |
| 5 | Test 1.2 | 16 kb/s | 15% |
| 6 | Test 1.2 | 1 kb/s | 25% |
|  | Total |  | 100% |

**Table 11: Figure of Merit weights for the Kinesthetic Haptic Tests**

|  |  |  |  |
| --- | --- | --- | --- |
| **Test**  **Component** | **Test** | **Bitrate and Configuration** | **Weight** |
| 1 | Test 1.3 | 32 kb/s | 20% |
| 2 | Test 1.3 | 16 kb/s | 30% |
| 3 | Test 1.3 | 1 kb/s | 50% |
|  | Total |  | 100% |

For each test or test operating mode (i.e., Test Component), all Test Lab scores are pooled to compute the grand mean and 95% confidence interval on the mean for each system under test. If feasible, an ANOVA will be used to check that the variance across the factor “Test Lab” is relatively constant such that pooling across the Test Labs is statistically appropriate.

MPEG WG 2 experts shall evaluate Proponent submissions, taking into account all submitted information and also the outcome from the subjective tests. Subjective test information to consider include at least the calculated Figure of Merit, its associated weighted confidence interval, and whether a system under test has consistent performance across tests and test items and whether a system under test has a confidence interval that overlaps with other systems under test for the individual tests (indicating similar performance in a statistical sense).



### Overall Figure of Merit

Given the three sets of disparate evaluation tests on a proponent’s submission: Objective Signal-based Tests, Expressive Capabilities Tests, and the Subjective MUSHRA Tests, it is not feasible to arrive at a single overall figure of merit that would make technical sense. For this reason, the results of these three sets of tests for all proponent submission will be provided to the WG in tabular form. The WG will take into consideration all aspects of the submissions and tests results to determine, preferably by consensus, or a vote (if necessary), on the submission that is deemed to be the RM0. In any case, the Core Experiments phase of the CfP will provide ample opportunity for the WG to collaboratively augment or fine-tune the RM0.

## Set 1 and Set 2 Performance and determination of RM0

Quoting from Section 4.9 of the Call:

Submissions shall be evaluated, taking into account all submitted information including subjective test results. Based on this information, a single submission that is best for both Test Set 1 and Test Set 2 will be selected as the RM0 technology.

Subsequent Core Experiments will be used to arrive at an integrated architecture, based on consensus of MPEG WG2.

# References

1. N37, Revised Draft Call for Proposals on the Coded Representation of Haptics – Phase 1
2. N40, Revised Draft Encoder Input Format for MPEG Haptics
3. The R Project for Statistical Computing (<http://www.r-project.org>)
4. R. Hassen and E. Steinbach, "Subjective Evaluation of the Spectral Temporal SIMilarity (ST-SIM) Measure for Vibrotactile Quality Assessment," in *IEEE Transactions on Haptics*, vol. 13, no. 1, pp. 25-31, 1 Jan.-March 2020, doi: 10.1109/TOH.2019.2962446.
5. R. Hassen and E. Steinbach, “HSSIM: An Objective Haptic Quality Assessment Measure for Force-Feedback Signals” - 2018 Tenth International Conference on Quality of Multimedia Experience

# ANNEX 1 – Test Protocol for Vibrotactile Haptics Tests (Tests 1.1 and 1.2)

## A1.1 Overview

The rendering pipeline from the MUSHRA STEP software to the device mockup follows up the following scheme that has been tested with Windows 10 Enterprise:



## A1.2 System Calibration and Installation

In order to guarantee acceptable driving voltage levels, the system should be calibrated through the output volume level of the testing station. To do so, the calibration .wav file (‘setup\_calibration.wav’) should be played back on the testing station using any suitable media player (e.g., Audacity), while making sure the USB audio interface is the output audio device. Using a scope, the test administrators need to make sure the peak-to-peak voltage measured at the input of the mockup device (i.e., output of the audio amplifier) is 4V and adjust the testing station’s volume output accordingly. Please note that this step needs to be performed when installing a new testing station and/or when the testing station’s volume level is modified.

When starting the STEP software, the test administrators need to make sure the USB audio interface is chosen as the audio device (Figure 1).

Graphical user interface, application

Description automatically generated

## A1.3 User Study Protocol

1. The participant sits comfortably on a chair and faces the test computer.
2. The Mockup device is to be placed on the table in front of the user, at a reasonable distance from them, and on the dominant arm side as to avoid muscle fatigue (Figure 2).A picture containing text, indoor, computer, electronics

   Description automatically generated
3. The user should hold the phone with both hands in a landscape mode (Figure 3), while making sure the device rests comfortably in the palms of both hands.

A picture containing person, hand

Description automatically generated

1. The user should not exercise excessive pressure on the mockup device as to not influence their perception of the tested stimuli, but at the same time should hold the device with enough pressure to keep it from falling.
2. The user might rest their arms on their lap, on the chair arms, or on the table to avoid excessive strain on their muscles during the test.
3. After experiencing each variant of the test signal the user can hold the device in their non-dominant hand resting on their lap, on the chair arms or on the table if needed (Figure 4) and reach out with their dominant hand to provide the score through the STEP interface.

A picture containing text, person

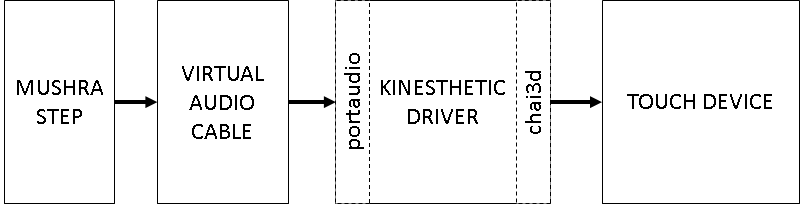
Description automatically generated

1. After experiencing the different test signals for a given trial, the user can rest the device on the table.

# ANNEX 2 – Test Protocol for Kinesthetic Haptics Tests (Test 1.3)

## A2.1 Overview

The platform is based on a 3D System Touch. It is a 3DoF force-feedback device able to apply 3.3N along each axis. It relies on a USB 2.0 connection. The rendering pipeline from the MUSHRA STEP software to the Touch device is described in this Annex. It has been tested with Windows 10 Enterprise 1909.

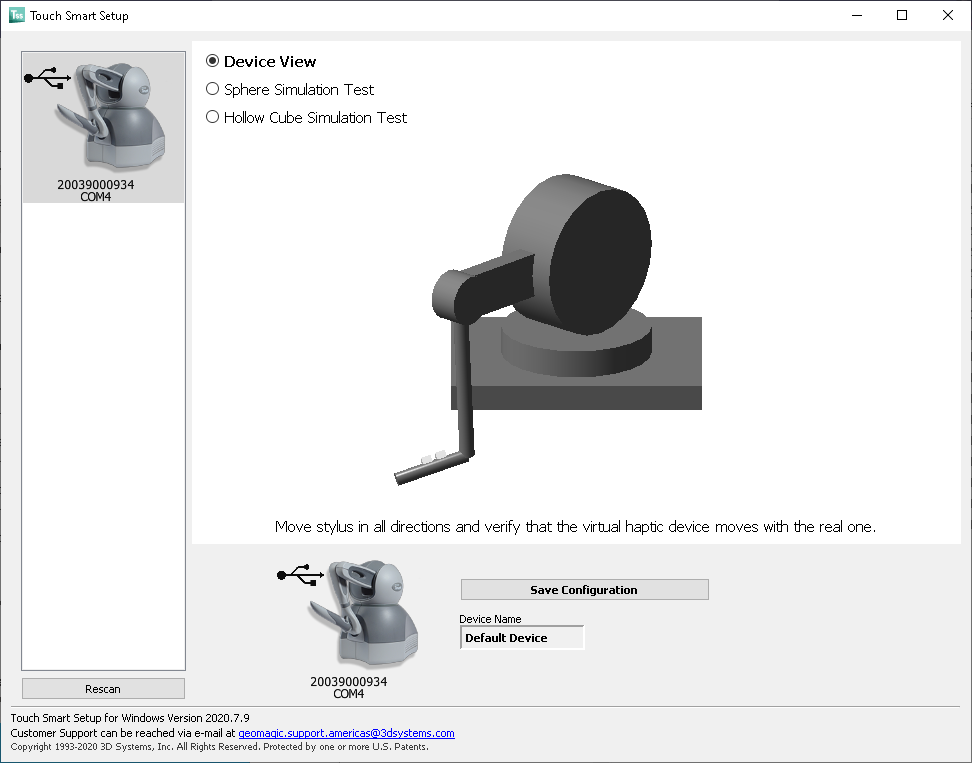


## A2.2 Installation

### A2.2.1 3D System Touch device driver

Download and install **Touch Device Driver v2020.7.9** and **OpenHaptics v3.5** <https://support.3dsystems.com/s/article/Haptic-Device-Drivers>

If successfully installed the *Touch Smart Setup* should be available on the desktop. The detection of the device can then be tested (see screenshot below).

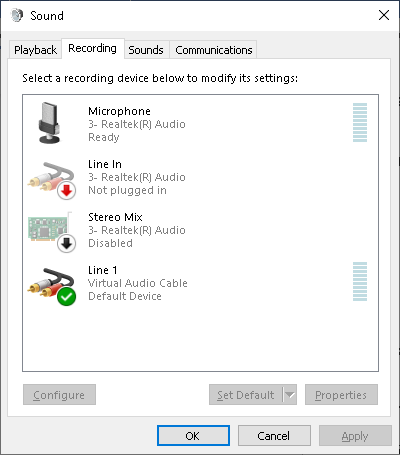
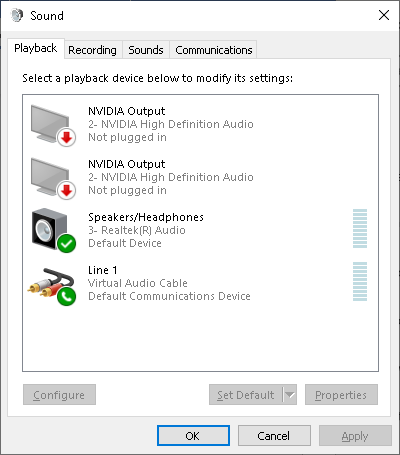


Examples located in C:\OpenHaptics\Developer\3.5.0\examples\bin\x64 should also work.

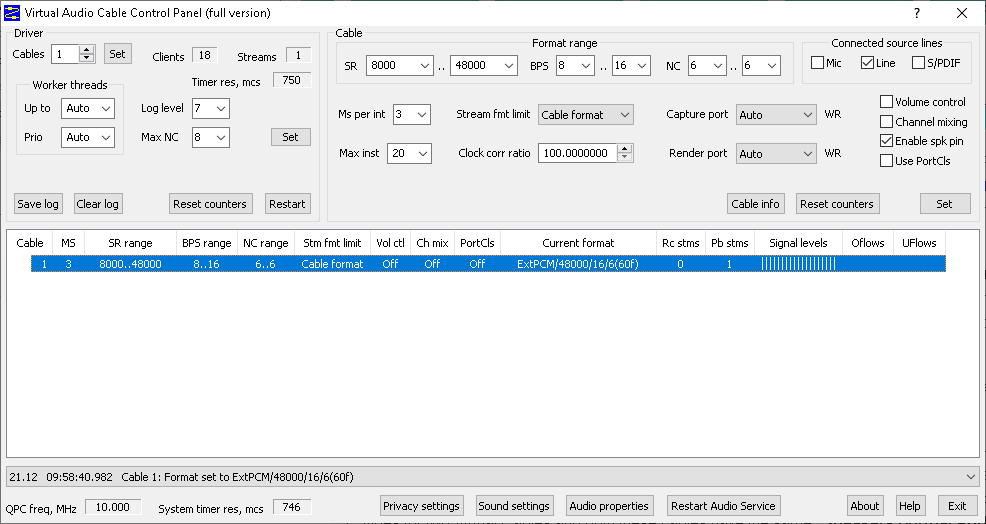
### A2.2.2 Virtual Audio Cable

Download and install the *Virtual Audio Control Cable* <https://vac.muzychenko.net/en/download.htm>.

This ‘virtual cable’ connects STEP output to the Force feedback driver input. If successfully installed, ‘Line 1’ should be visible in the audio devices panel.



The *Virtual Audio Cable Control Panel* should also be available. When the output sound is sent to the virtual cable, the signals levels are activated.



The Number of Channels (NC) should be specified at 6 (input and output). ‘Channel mixing’ should be deactivated and ‘Enable spk pin’ activated.

### A2.2.3 Kinesthetic driver

This piece of software requires Windows 10, Visual Studio 2019, cmake, and git. It requires two libraries, *portaudio* for capturing audio data and *chai3d* for rendering forces on the Touch device, both of which will be automatically compiled.

Git clone <https://github.com/InterDigitalInc/MPEG_Haptics>

Step by step instructions to compile the code are listed in the file README.md, in the *kinesthetic\_RefSetup* folder.

Basically, use cmake to build the visual project file.

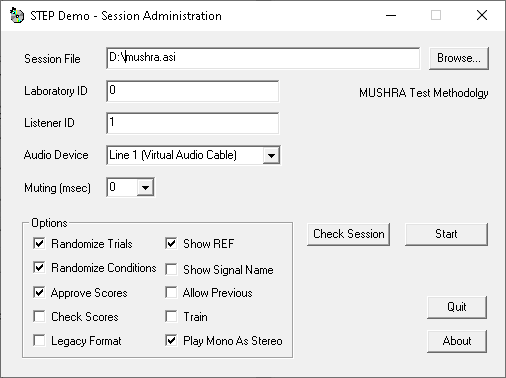
cmake .. -DCMAKE\_TOOLCHAIN\_FILE=D:\vcpkg\scripts\buildsystems\vcpkg.cmake

Then open the visual studio project KINE\_STEP\_TEST.sln and build all (Release 64bits). It should build **kine\_step.exe**

## A2.3 Usage

1. Run MUSHRA STEP and select the provided \*.asi file.

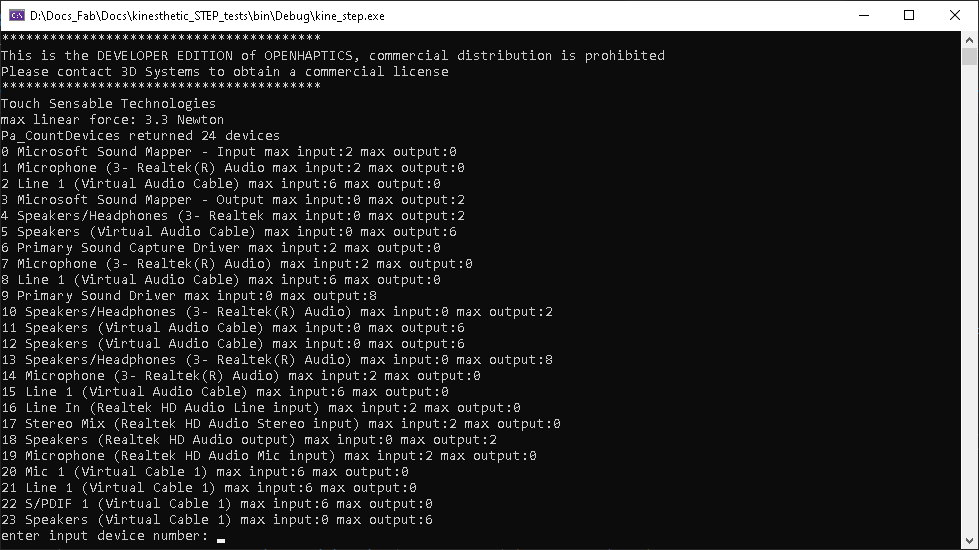
**Select Virtual audio cable as Audio Device**.



1. Plug the Touch device

run *kine\_step.exe*

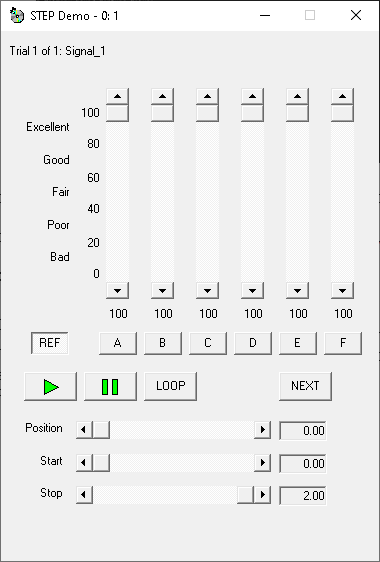
select interface **Line 1 (Virtual Audio Cable) max input:6 max output:0** by entering its id number and press enter (number 2 in this example).



*kine\_step.exe* is now running and listening to audio input from STEP. Stimulus are rendered on the Touch device.

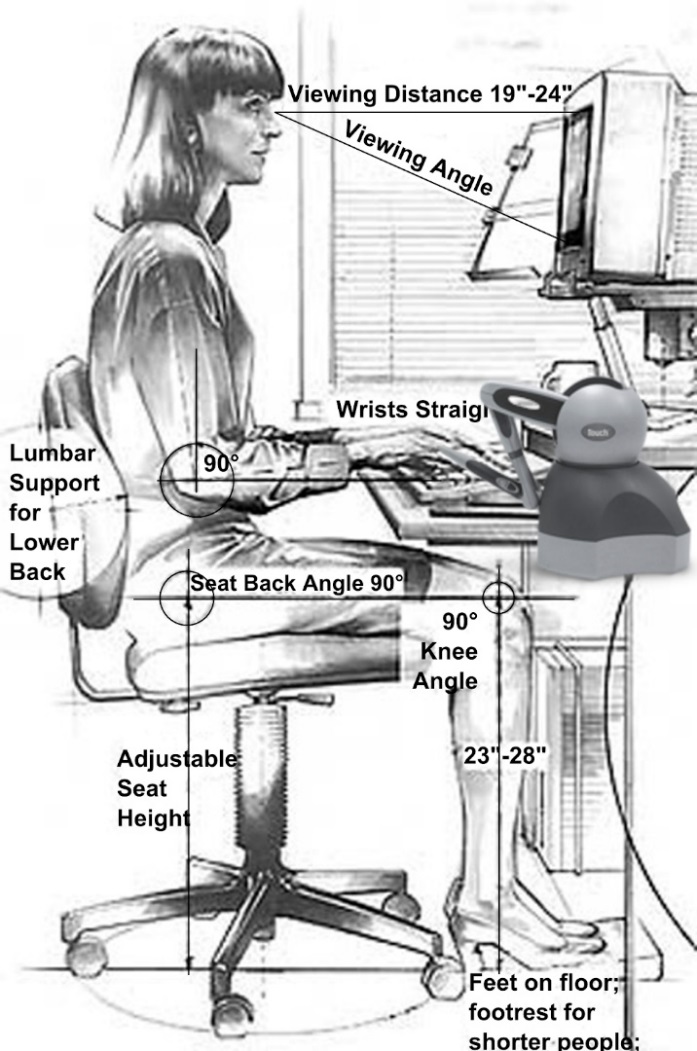
1. The test is ready to be performed. Start the session from the MUSHRA step interface. Pressing Play button would activate the Touch device.

Then follow instructions in Annex 3 to perform the full tests.



## A2.4 User Study Protocol

1. Participant sits comfortably on a chair and faces the test computer as in figure below



1. The Touch device is on the left or right of the screen depending on the participant’s dominant hand, at a reasonable distance to prevent muscle fatigue of the arm (see figure below).
2. The participant holds the stylus of the device in the dominant hand. Arm is NOT resting on the table during playback of a stimulus, and the participant should NOT apply any force. This way the arm is free to move while the signal is played.
3. Before starting a stimulus, the participant’s hand must be at the middle of the device workspace. A guide must be displayed (see picture, red dot is the center of the workspace).
4. When the signal is played, the participant may rest his/her arm between stimuli as long as needed. Answers are entered in STEP at that moment.
5. The Participant goes through the entire session as usual with the STEP software. Make sure that item 4. is respected before playing a stimulus.



Recommended handling of the Geomagic Touch stylus

# ANNEX 3 – Test Subject Instructions

## A3.1 Overall instructions

The MPEG WG 2 wishes to assess the subjective quality of various technologies that are able to compress and render haptics. You will feel the haptics effects in your hand when you hold the handheld mockup device (for vibrotactile haptics) or when you hold the stylus in your hand (for kinesthetic haptics).

This test will use a Multi-Stimulus test methodology, which has the advantage of displaying all stimuli (as processed by all systems under test) for a given test item. Hence, you are able to directly compare the stimuli in the course of giving a grade to each.

The vibrotactile haptics tests (Tests 1.1 and 1.2) will use a custom-built hardware platform with a mobile handheld device form factor (called “the mockup”). The protocol for handling the mockup is described in Annex 1 above. The kinesthetic haptics test (Test 1.3) will use the GeoMagic Touch, the protocol for handling it is described in Annex 2 above.

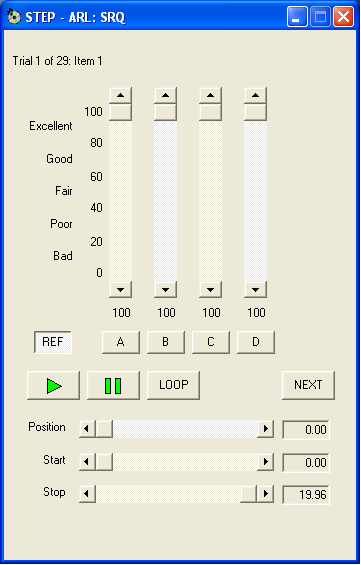
### A3.1.1 User Interface

The Figure below shows an example user-interface presenting one test item as processed by four systems under test. The buttons represent the open reference (REF), which is always displayed at the bottom left, and all the systems under test, including the three codecs under test and the hidden reference, which are displayed as letter buttons. Above each button, with the exception of the button for the open reference, a slider permits you to grade the quality of the test item on a continuous quality scale. The descriptors associated with the scale are:

|  |  |
| --- | --- |
| **Descriptor** | **Grade Range** |
| “excellent” | 80-100 |
| “good” | 60-80 |
| “fair” | 40-60 |
| “poor” | 20-40 |
| “bad” | 0-20 |

For each of the test items, the systems under test are randomly assigned to the buttons. In addition, the order of presenting the test items is randomized for each listener.

Note that the haptic signals (for each system and the open reference) are time aligned such that you are able to switch instantly and seamlessly from one haptic signal to another. Clicking on the *Loop* button plays the signal continuously. The horizontal *Position* slider indicates the instantaneous position in the signal waveform. Grabbing and moving the *Start* slider alters the start point for waveform looping, and similarly moving the *Stop* slider alters the end point, thus permitting a “loop and zoom” function that is particularly powerful for subjective evaluation. You are asked to rate the processed signals (as compared to the open reference) by grabbing and moving the vertical sliders above their corresponding letter buttons. **Any perceived difference from the open reference is required to be rated down**. Please note that one of the test systems is always the hidden reference and should be rated 100. When you are satisfied with the ratings, click on the *Next* button to go on to the next trial. When the last trial is scored, the MUSHRA window is replaced by the *Administrator* window. Please notify the test administrator that you have completed the test session.



## A3.2 Training phase instructions

The training phase will be conducted individually, that is without an instructor. You will be participating in a simulated test whose graphical interface is exactly like the actual test. The training will be shorter, and the scores will not be used in the final tally. However, the training items are selected to represent the range of impairments that you will encounter in the actual test. Please feel the effects of the training signals to understand how processed signals feel relative to the reference.

You will have up to one hour of training before undertaking the first test.

In the course of experiencing these signals, please:

* Feel a test item several times.
* Rate overall haptic quality, not merely the quality of a portion of a test item (for example, only the portion in which the effect is the most prominent).
* Note that if two systems have a different perceived quality with respect to the Reference, they should receive a different score.

## A3.3 Test phase instructions

The test phase will be carried out individually in test sessions each lasting about 30 minutes. In each trial, you will experience several versions of the test items, each processed by a different system under test.

Please remember the following:

You are asked to judge the “Basic Haptic Quality” of the versions of the test item in each trial as compared to the open reference. Any perceived difference from the open reference is required to be rated down. Please note that one of the test systems is always the hidden reference. You are required to identify this hidden reference and give it a rating of 100. Note that it is perfectly acceptable to give more than one system a score of 100.

Components of “Basic Haptic Quality” are related to any and all differences between the reference and the coded/decoded test item, including.

* strength of the effect,
* sharpness of the transients,
* roughness/smoothness of sensation,
* timing of overall effect,
* timing of effect elements,
* ringing/buzziness of the effect,
* distinctness of effect elements

Please note that attributes mentioned above are not necessarily applicable to every haptic item presented to you. Therefore, the relation between Basic Haptic Quality and attributes regarding the grading scale depend entirely on you. Also, your score should reflect your overall personal opinion.

The assessment is to be done on a scale from 0 to 100, as shown here:

