 ISO/IEC JTC 1/SC 29/WG 2 N0037

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

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

**Document type:** Output Document

**Title:** Revised Draft Call for Proposals on the Coded Representation of Haptics – 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:** 32 (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 N0037**

**Online - January 2021**

|  |  |
| --- | --- |
| **Title:** | **Revised DRAFT Call for Proposals on the Coded Representation of Haptics – 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)**  **Thomas Begeot (Actronika)**  **Camille Moussette (Apple)** |
| **Status**  **Serial number** | **Approved**  **20121** |

Table of Contents

[1 Introduction 2](#_Toc61547562)

[2 Definitions 3](#_Toc61547563)

[3 Haptics CfP Workflow 3](#_Toc61547564)

[4 Timetable and Procedures 4](#_Toc61547565)

[4.1 Overview 4](#_Toc61547566)

[4.2 Envisioned Standard 6](#_Toc61547567)

[Phase 1 6](#_Toc61547568)

[4.3 Register 6](#_Toc61547569)

[4.4 Get test items 7](#_Toc61547570)

[4.5 Submit Coded Materials 7](#_Toc61547571)

[4.6 Objective Tests 8](#_Toc61547572)

[4.6.1 Signal-Based Tests 8](#_Toc61547573)

[4.6.2 Expressive Capabilities Tests 10](#_Toc61547574)

[4.7 Subjective Tests 10](#_Toc61547575)

[4.8 Submit Documentation 13](#_Toc61547576)

[4.9 Participate in Evaluation and Selection 14](#_Toc61547577)

[Core Experiments 14](#_Toc61547578)

[4.10 Core Experiments 14](#_Toc61547579)

[4.11 Verification Tests 14](#_Toc61547580)

[4.12 Further Information 14](#_Toc61547581)

[5 References 15](#_Toc61547582)

[ANNEX 1 - Requirements 16](#_Toc61547583)

[A1.1 Phase 1 (Basic Haptics) 16](#_Toc61547584)

[A1.2 Phase 2 (Advanced Haptics) 17](#_Toc61547585)

[ANNEX 2 - Application Scenarios 18](#_Toc61547586)

[A2.1 Phase 1 (Basic Haptics) 18](#_Toc61547587)

[A2.2 Phase 2 (Advanced Haptics) 18](#_Toc61547588)

[ANNEX 3 - Hardware Configurations 20](#_Toc61547589)

[A3.1 Vibrotactile Test Platform 20](#_Toc61547590)

[Introduction 20](#_Toc61547591)

[Block diagram 20](#_Toc61547592)

[Haptic actuator 20](#_Toc61547593)

[Vibrotactile Test Platform Specification 21](#_Toc61547594)

[Module mockup 21](#_Toc61547595)

[Pricing 23](#_Toc61547596)

[A3.2 Kinesthetic Test Platform 24](#_Toc61547597)

[Introduction 24](#_Toc61547598)

[Device Specifications 24](#_Toc61547599)

[Block Diagram 25](#_Toc61547600)

[API 25](#_Toc61547601)

[Pricing 25](#_Toc61547602)

[ANNEX 4 – Haptic Effects and Test Signal Categorization 26](#_Toc61547603)

[A4.1 Introduction 26](#_Toc61547604)

[A4.2 Haptic Design Taxonomy from Immersion 26](#_Toc61547605)

[A4.3 Test Signal Categorization for MUSHRA Tests 27](#_Toc61547606)

[ANNEX 5 - Phase 1 Test Material 28](#_Toc61547607)

# Introduction

Haptics provide an additional layer of entertainment and sensory immersion to the user. Therefore, the user experience and enjoyment of media content, be it in ISOBMFF files or streams such as ATSC 3.0 broadcasts, streaming games, and mobile advertisements can be significantly enhanced by the judicious addition of haptics to the audio/video content. To that end, haptics has been proposed as a first-order media type, akin to audio and video, in ISOBMFF. Further, haptics has also been proposed as an addition to the MPEG-DASH standard to signal the presence of haptics in the MP4 segments to the DASH streaming clients. Lastly, the MPEG-I Phase 2 use cases have been augmented with haptics [11] resulting in a set of haptic-specific requirements for MPEG-I Phase 2 [10]. All these proposals are in various stages of the MPEG standardization process.

These ongoing haptics standardization efforts highlight the need for standardizing a coded representation of haptics. A standard haptics coding format (and associated standardized decoder) will facilitate the incorporation of haptics into the ISOBMFF, MPEG-DASH, and MPEG-I standards, making it easier for content creators as well as media/streaming content providers to incorporate haptics and improve the overall user experience.

This document is the Call for Proposals (CfP) on a Coded Representation of Haptics – Phase 1. It addresses the Phase 1 (Basic Haptics) requirements and applications scenarios, listed in Annexes A1.1 and A2.1, respectively. The Phase 2 (Advanced Haptics) requirements and application scenarios, listed in Annexes A1.2 and A2.2, respectively, will be the subject of a separate CfP, to be issued at a later date, and will build on technology selected from this Call. Phase 2 requirements and application scenarios are listed in this document to provide prospective proponents a roadmap of both phases and to enable them to plan their responses appropriately.

Section 2 defines terms; Section 3 describes the workflow of the CfP; Section 4 presents the timeline for and procedures used in the submission to the Call; Annex 1 presents the requirements that will be satisfied by the proposals; Annex 2 presents the application scenarios that should be addressed by the proposals; Annex 3 defines the hardware configurations that will be used for evaluating the candidates for standardization; Annex 4 describes the haptic effects of interest and the categorization of test signals, and Annex 5 describes the test material to be used.

# Definitions

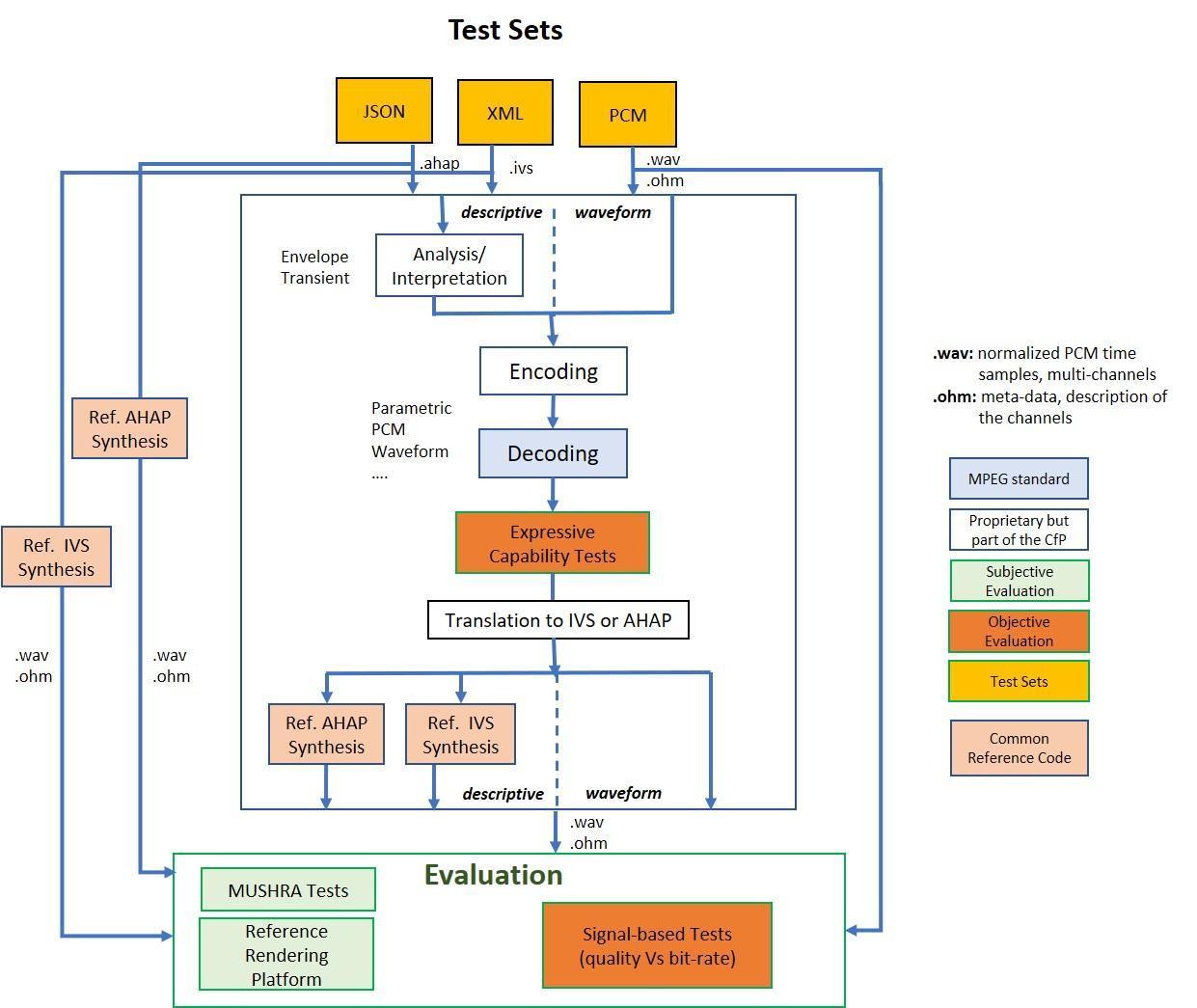
1. **Channel based input**, where each channel is intended to drive a distinct haptic actuator location or type in the target device. Channels can be any mixture of spatial (left/right, etc.) or actuator type (amplitude only, wideband, etc.).
2. **PCM**, where the signal is coded using a pulse coded modulation with a specific (fixed) bit rate per sample (typically 1 or 8 bits/sample).
3. **JSON/descriptive**, where the haptic signal is coded using a descriptive, structured representation such as JavaScript Object Notation or other similar XML or structured coding of a signal.
4. **Transcoding**, where the decoded actuator drive signal is transformed to provide a more suitable actuator drive signal for a specific playback system.

# Haptics CfP Workflow

Figure 1 below shows the haptics standardization workflow in this CfP from the input test signals to the evaluation. As is the norm with all MPEG standardization, the coded representation and decoder will be standardized. The encoding and rendering implementations are outside the scope of MPEG, best left to vendor differentiation. That said, we will need a reference rendering platform to test and evaluate the candidate proposals. Annex 3 contains descriptions of the reference rendering platforms used for vibrotactile and kinesthetic haptics.

Further, the test set (see Annexes 4 and 5) will contain a combination of JSON, XML, and PCM + OHM haptic files. Respondents to the CfP are free to use any encoding mechanism of their choice but must be able to ingest test data in all these three formats. As shown in Figure 1, respondents proposing a non-PCM coded representation must employ a synthesis means to synthesize PCM data from the decoded representation.

Finally, PCM equivalents of all the test signals will be used for the objective tests and as references for the MUSHRA tests. To ensure a level-playing field for all proponents, common reference synthesis will be made available to all proponents, as shown in Figure 1. All decoded signals will be evaluated using the expressive capability tests to ensure that translation does not mask the expressive capabilities of the proponent’s coding.



**Figure 1: Haptics CfP Workflow**

# Timetable and Procedures

## Overview

This Call for Proposals shall be conducted in two phases: Phase 1 and Phase 2. Phase 1, the subject of this document, is a first call for technology that satisfies the Basic Haptics Requirements (see Annex A1.1). Technology selected at the end of Phase 1 is designated Reference Model 0 (RM0). Phase 2 is a second submission of technology that should extend RM0 technology in an integrated manner, addressing the Advanced Haptics Requirements (see Annex A1.2). Technology selected at the end of Phase 2 is designated RM1. Because it builds on RM0, Phase 2 must have a submission deadline after the Phase 1 RM0 technology has been selected and RM0 documentation (as text and source code) is available to MPEG. Phase 2 technologies will be assessed for the 3D object interaction and VR video calling scenarios. Phase 2 timelines and procedures will be described in greater detail in a separate CfP, to be issued at a later date.

A timetable for Phase 1 of the Call for Proposals relative to specific MPEG meetings is given in the following table.

|  |  |
| --- | --- |
| **Meeting / Date** | **Action** |
| 132nd meeting, October 2020 | Issue Draft Call for Proposals on the Coded Representation of Haptics – Phase 1 |
| 133rd meeting, January 2021 | Issue Revised Draft Call for Proposals on the Coded Representation of Haptics – Phase 1  Issue Draft Submission and Evaluation Procedures for Haptics  Issue Draft MPEG Haptics Core Experiment Methodology |
| 134th meeting, April 2021 | Issue Call for Proposals on the Coded Representation of Haptics – Phase 1  Issue Submission and Evaluation Procedures for Haptics  Issue MPEG Haptics Core Experiment Methodology  Test platform available to proponents |
| Prior to May 15th, 2021 | Proponents must register intention to participate in Call |
| July 5, 2021 | Proponent processed test items due |
| July – September 2021 | Conduct subjective evaluation tests (extended time for testing allocated due to COVID considerations) |
| Contributions to 136th meeting | Proponent written documentation due |
| 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, subject to possible technology merge procedure. |

The following steps are envisioned for the standardization of the new technology:

* All proposals shall be prepared in accordance with the requirements set forth in this document and in the Encoder Input Format document [3].
* All proposals shall be evaluated using the procedure described in [4]. An important component of the evaluation process will be a subjective test to assess the quality of items coded by the proposed technology.
* It is anticipated that at the 136th MPEG meeting, the Phase 1 submitted technologies will be evaluated according to [4]. All submitted information will be considered and selection of a Reference Model 0 technology will be by the consensus of MPEG WG 2.
* It is expected that at the 137th MPEG meeting, the proponent(s) of the technology designated as Reference Model 0 will submit a detailed technical description, bitstream syntax and decoding semantics and reference source code for the encoding and decoding process. Source code shall be in C++ or ANSI-C with C++ compatible header files. A decoder compiled from the source code shall decode the proponent submitted bitstreams and produce the associated proponent submitted waveforms.
* After RM0 technology technical description and source code is available, a collaborative phase will start that aims to improve upon RM0 technology using the MPEG Haptics Core Experiment Methodology [5]. The CE phase can span one or more meeting cycles, depending on the number and complexity of the CEs to be conducted.
* Prior to the conclusion of the standardization process, MPEG WG 2 will conduct a formal verification test and generate a report that characterizes the performance of the technology.

## Envisioned Standard

It is envisioned that the final Haptic coding standard:

1. Shall support PCM-based waveform inputs.
2. Shall support text-based descriptive inputs (e.g., JSON or XML).
3. Should have a unified architecture to the greatest extent possible.
4. Should re-use existing MPEG technology wherever possible. Alternate technology can be used if it provides performance substantially better than MPEG technology or provides functionality not possible with MPEG technology or with a simple extension to MPEG technology.

# Phase 1

## Register

Register by May15th, 2021 an intention to participate in the Call. Registering intent is not binding and registered parties are not required to submit proposals. However, parties that do not register will not be able to submit proposals. Register by sending an email to [the](about:blank) contacts listed in Section 4.12. Email should indicate:

* Company name
* Contact name and contact email address
* Whether the intent is to submit technology that will process just Sets 1 and 2 (both vibrotactile haptics) or Set 3 (kinesthetic haptics) as well.

After registration, the proponent will receive a “ProponentID” for use in submission of coded materials. Also, processing Sets 1 and 2 (both vibrotactile haptics) in Phase 1 is mandatory, but processing Set 3 (kinesthetic haptics) is optional.

## Get test items

The test items used in the evaluation process shall be made available in a format as defined in [4]. Test items are divided into three sets: Set 1, Set 2, and Set 3. The items are described in the following table and are listed in Annex 5.

|  |  |
| --- | --- |
| Test Set | Description |
| Set 1 | A representative set of Short Effect signals. A Short Effect is an effect with a simple envelope (e.g., ADSR), where the envelope is the salient characteristic of the effect |
| Set 2 | A representative set of Long Effect signals. A Long Effect is an effect where the envelopes and placement of its subcomponents are the salient characteristics of the effect |
| Set 3 (Optional) | A representative set of kinesthetic signals, including force signals, acceleration, or movement. |

## Submit Coded Materials

Submit by July 5th, 2021 the following: the compressed representation and corresponding decoded haptics files (\*.wav) for the signal files associated with both of Set 1 and Set 2 test items (as well as Set 3, if processing kinesthetic haptics) for each test in the Phase 1 subjective evaluation. The compressed representation must be a unique file format. Total bitrate of the compressed representation shall be calculated as (total information bits for test item)/(test item duration). The decoded haptics files shall conform to the format defined in [4] and be time-aligned within 1 sample relative to the reference or reference rendering. The specifics of naming the submitted files are also defined in [4].

## Objective Tests

### Signal-Based Tests

To compare the performance of the proposed technology, objective tests will be conducted based on three different metrics: Peak Signal to Noise Ratio (PSNR), Spectral Temporal SIMilarity (ST-SIM) and HapticS SIMilarity Index Measure (HSSIM).

The three metrics are used to compare the input signal with the decoded signal and their formulations are detailed below.

1. **PSNR**

PSNR is used to evaluate both vibrotactile and kinesthetic signals and has the following formulation:

***with:***

***where*** *v* is the modality (e.g., position, force, velocity, orientation, …)

1. **ST-SIM - For Vibrotactile Signals**

The Spectral Temporal SIMilarity (ST-SIM) [9] measure is a method for measuring the similarity between two vibrotactile signals. The ST-SIM measure can be viewed as a quality assessment measure of one of the signals being compared, provided the other signal is regarded as of “perfect” quality (i.e., considered the reference). ST-SIM is formulated as follows:

where is a weighting factor that emphasizes the importance of the two similarity components. S-SIM and T-SIM are respectively a spectral measure based on the human detection probability function and a temporal measure based on Mean Subtracted Contrast Normalized (MSCN) coefficients. The two input signals (original and distorted ) are divided intoframes and locally evaluated using these two similarity measures.

The spectral similarity of the k-th local frame is computed as

where and are the detection probability densities of the k-th local frame of the frequency spectrum for input signals and is the number of frequency samples.

The probability density functions and are calculated with a cumulative normal distribution function of the input signals as:

The temporal measure T-SIM, is calculated as:

where

,

whereis a small constant that prevents instabilities when the denominator value is close to zero, , , , and are the local means and standard deviations of a given frame, and is the number of samples in a frame.

1. **HSSIM - For Kinesthetic Signals**

This metric proposed in [6] is used to assess the quality of force feedback signals and is computed as follows:

where *N* is the number of samples in the perceptual quality map and the Minkowski power. is computed locally with a sliding window over the signal with the following formula. The size of the window is 4% of the signal length.

α>0, β>0, and γ>0 are parameters used to adjust the relative importance of the three comparison components (typically α = β = 1 and γ = 0.5). L, C and S are defined as :

, ,

Where , , , and are local means and standard deviations of the two perceptually transformed signals (original signal) and (compressed signal). , , and are the local standard deviations and correlation coefficient between original signals and . , , and are positive stabilizing constants with , , , , and the dynamic range of the original force signal (0N to 5N).

Finally, and are the perceived force signals after applying Stevens power law:

and

where is a scaling constant and is the power value.

The details of these test metrics (software, inputs, outputs, etc.) are provided in the Submission and Evaluation Procedures for Haptics document [4].

### Expressive Capabilities Tests

In addition to signal-based objective evaluation, proposed coded representations will be evaluated to ensure that they are able to satisfy the Phase 1 requirements through coded signaling in addition to PCM signaling. In particular, coded representations must:

1. Enable device independent representations in support of playback transcoding
2. Enable multiple simultaneous tracks that can be mixed and modulated by the presentation engine
3. Enable haptic modality descriptors, at least able to distinguish between vibrotactile and kinesthetic modalities with support for additional future modalities.

The details of these test criteria will be provided in [4].

## Subjective Tests

**Overview**

The timeline and procedures for conducting the subjective tests that compare the performance of proposed technology is given in [4].

The performance of submissions to Phase 1 of this Call will be evaluated using two subjective tests for vibrotactile signals and one subjective test for kinesthetic signals.

**Test 1.1**

This subjective test is meant to assess the performance of submitted technology for **Short Effects** when used in the mobile handset application scenario and in which inertial actuators are used to present the haptic program.

|  |  |
| --- | --- |
| Test Methodology | MUSHRA |
| Presentation | Handheld mobile device form factor |
| Reference | The PCM original item is the rendered reference |
| Actuator Mounting | Handheld (custom-built hardware; see Annex A3.1) |
| Test Items | The 9 items of Set 1 |
| Bit Rates | The bit rates per channel shall be the following:  32 kb/s  16 kb/s  1 kb/s |
| Restrictions | None |
| Requirements addressed | High quality  Update rate  Synchronization  Playback transcoding  Compression  Actuators  Modulation |

**Test 1.2**

This subjective test is meant to assess the performance of submitted technology for **Long Effects** when used in the mobile handset application scenario and in which inertial actuators are used to present the haptic program.

|  |  |
| --- | --- |
| Test Methodology | MUSHRA |
| Presentation | Handheld mobile device form factor |
| Reference | The PCM original item is the rendered reference |
| Actuator Mounting | Handheld (custom-built hardware; see Annex A3.1) |
| Test Items | The 24 items of Set 2 |
| Bit Rates | The bit rates shall be the following:  32 kb/s  16 kb/s  1 kb/s |
| Restrictions | None |
| Requirements addressed | High quality  Update rate  Synchronization  Playback transcoding  Compression  Actuators  Modulation |

**Test 1.3 (Optional)**

This subjective test is meant to assess the performance of submitted technology for kinesthetic effects when applied on the user’s hand.

|  |  |
| --- | --- |
| Test Methodology | MUSHRA |
| Presentation | Force-feedback device held in one hand |
| Reference | The PCM original item is the rendered reference |
| Actuator Mounting | Handheld (Geomagic Touch; See Annex A3.2) |
| Test Items | The 10 items of Set 3 |
| Bit Rates | The bit rates per channel shall be the following:  32 kb/s  16 kb/s  1 kb/s |
| Restrictions | None |
| Requirements addressed | High quality  Update rate  Synchronization  Playback transcoding  Compression  Actuators  Modulation |

## Submit Documentation

Submit as contributions to the 136th MPEG meeting:

* A written description of the technology having sufficient detail to permit technical discussions.
* Evidence of the performance of the technology, as outlined in [4].

All proponents shall submit a written description. Proponents that are MPEG members shall register these documents as contributions to the 136th MPEG meeting and send title and author information to the contacts listed in Section 4.12 prior to the close of contribution upload. Proponents that are not MPEG members shall email the documents to the contacts in Section 4.12 prior to Sept 30th, 2021, so that the documents can be uploaded and registered as contributions. The documents should be written in Microsoft Word.

The results of the subjective Tests 1.1, 1.2, and 1.3 will be available as a contribution to the 136th MPEG meeting.

## Participate in Evaluation and Selection

Attend the 136th MPEG meeting (details on meeting location and date will be communicated via email to parties that are not MPEG members). It is strongly urged that experts familiar with the proposed technology attend in order to allow discussions on details of the proposals.

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. Details of the evaluation of the submissions and selection of the RM0 technology can be found in [4].

As contributions to the 137th MPEG meeting, proponents of selected technology (RM0) submit:

* Full source code for conformant encoder and normative decoder, according to [5].
* Written description as bitstream syntax, decoding semantics and decoding description.

# Core Experiments

## Core Experiments

The RM0 technology selected shall be the basis for subsequent core experiments. Core Experiments (CE) will be conducted according to [5]. It is envisioned that the CE process will use two sets of test items (one each for vibrotactile and kinesthetic haptics) and this will be determined prior to the start of the CE phase of standardization.

**Prior to the CD stage of the standard**

## Verification Tests

The performance of the new technology shall be measured via a formal subjective test, to be carried out prior to the Committee Draft stage of the standardization process. An acceptable level of performance, as judged by the consensus of MPEG WG 2, must be achieved in order for the technology to progress in the standardization process.

## Further Information

For any questions related to this Call for Proposals or associated evaluation procedures please contact:

Dr. Yeshwant Muthusamy

Senior Director, Standards

Immersion Corporation

Phone: +1 469-583-2171

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

Dr. Igor Curcio

Convenor of ISO/IEC JTC1/SC29/WG2 on MPEG Technical Requirements

Bell Labs Distinguished Member of Technical Staff, CTO

Nokia Technologies, Tampere, Finland

email: [igor.curcio@nokia.com](mailto:igor.curcio@nokia.com)

# References

1. ITU-R Recommendation BS.1534-3, “Method for the Subjective Assessment of Intermediate Sound Quality (MUSHRA)”, International Telecommunications Union, Geneva, Switzerland, 2015
2. ITU-R Recommendation BS.1116-1, “METHODS FOR THE SUBJECTIVE ASSESSMENT OF SMALL IMPAIRMENTS IN AUDIO SYSTEMS INCLUDING MULTICHANNEL SOUND SYSTEMS”
3. MPEG WG2 N40, Revised Draft Encoder Input Format for MPEG Haptics
4. MPEG WG2 N38, Draft Submission and Evaluation Procedures for Haptics
5. MPEG WG2 N39, Draft MPEG Haptics Core Experiment Methodology
6. 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
7. Seifi, H., Zhang, K., and MacLean, K.E., “VibViz: Organizing, visualizing and navigating vibration libraries,” in *Proc. of IEEE WorldHaptics (WHC'15)*, 2015.
8. VibViz dataset: <https://www.cs.ubc.ca/~seifi/VibViz/main.html>
9. 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.
10. N19511, MPEG-I Phase 2 Requirements
11. N19513, MPEG-I Phase 2 Haptics Use Cases

# ANNEX 1 - Requirements

The detailed requirements identified in N19511 [10] were simplified to the following set, in order to simplify the process of evaluating submissions and to explicitly extract near-term standards ready requirements (Phase 1) and those that represented advanced functionality (Phase 2). Phase 1 requirements focus on coding of time-dependent haptic signals and are suitable for coding of timed-haptic experiences that may be synchronized with audio and/or video media. Phase 2 requirements focus on interactive haptic signals that are useful for coding fully immersive, interactive XR experiences. In general, Phase 2 requirements are intended to support user-object interactions common in such experiences. Phase 1 requirements are suitable for near-term usage and are intended to enable efficient coding of haptics.

## A1.1 Phase 1 (Basic Haptics)

Requirements related to basic haptics content.

1. **High quality**: The coded representation shall support coding of haptic signals that preserve frequency and amplitude independently.
2. **Update rate:** Coded representation shall provide decoded data to a presentation engine sufficient to be able to generate a drive signal between 5 and 1000Hz.
3. **Synchronization**: coding shall enable perceptual synchronization with audio and video.
4. **Playback transcoding**: Shall enable re-rendering or synthesis of coded signals for distinct target playback hardware.
   1. Includes target hardware capabilities (e.g., frequency response)
   2. Includes physical placement and arrangement relative to user
   3. Includes differing channel counts between encoded and rendered experience
5. **Tracks**: Shall support multiple simultaneous tracks and multiple encoded versions of a single track.
6. **Compression**: Shall support lossy or lossless compression.
7. **Actuators**:
   1. Shall support playback on arbitrary temporally driven haptic hardware.
   2. Shall support playback on multi-actuator systems.
   3. Shall support localization of rendering devices relative to a user.
8. **Mixing**: Shall support mixing of multiple simultaneous tracks on a single output device. Smooth transitions are required.
9. **Modulation**: Shall support user or application-controlled modulation of the overall haptic sensation.
10. **Haptic modality descriptor per track:** Shall support modality descriptors such as ‘vibrotactile’ or ‘kinesthetic’ with support for at least ‘vibrotactile’ in Phase 1.

## A1.2 Phase 2 (Advanced Haptics)

Requirements related to advanced haptic experiences.

1. **Surface Properties**: Shall support 2D (or higher) spatial haptic effect encoding.
2. **Material properties**: Shall support association of spatio-temporal haptic effects with 3D objects in a scene graph. In particular, surface properties such as viscoelastic surface compliance and thermal flux can be encoded.
3. **Latency**: Shall have sufficiently low latency to ensure real-time and interactive user experience. Typically, in the order of 1 ms.
4. **Avatar representation**: Shall support representation of the haptic interactive elements of the user’s virtual representation.
5. **Interaction model**: Shall support at least one interaction model such as:
   1. Touch-to-feel: avatar representation must intersect with objects to generate haptic feedback.
   2. Look-to-feel: Avatar representation must look at specific objects to generate haptic feedback.
   3. Ambient sensation: Global haptic sensations for a set of avatars in some volume.

# ANNEX 2 - Application Scenarios

The following the most important MPEG-I haptics application scenarios [11]:

## A2.1 Phase 1 (Basic Haptics)

**Haptics with audio/video content -** A haptic track designed to be played in sync with audio/video media. This is basic haptic functionality that is subsumed by other scenarios but is worth calling out separately.

**Haptics associated with point cloud media** - A haptic track may be associated with a set of point cloud media data. Subsets of the point cloud may be indexed against specific objects/participants. This haptic track can be activated, deactivated, or modulated based on the user’s viewpoint.

**Haptics associated with a capture device** – A haptic track may be associated with a capture device such as a sensor or camera that captures Tactile Essence (SMPTE st21001-2017). This haptic track can be activated, deactivated, or modulated based on the user’s viewpoint.

**Smooth haptic transitions** – When transitioning between haptics associated with point cloud media, video media, and capture devices, the haptic tracks may be mixed to minimize tactile artifacts and ensure a smooth transition.

**User configurable haptics** – In addition to a default haptic schema, the user may choose among alternative schemas. For example, the user may only activate haptics during a first-person perspective, enable only haptics associated with point clouds, etc.

## A2.2 Phase 2 (Advanced Haptics)

**Haptics associated with user-selected 3D object** – The player selected for 2D overlay activates the haptic track associated with that player, regardless of that player’s position relative to the viewer’s perspective.

**Haptic profiles are associated with objects**

1. The haptic profile may be based on the 3D object’s geometry. For example, when part of a user’s avatar interacts with the object by colliding with it, a haptic effect may be displayed that:

* Signifies the collision
* Prevents the user’s interaction gesture from crossing the boundary of the 3D object

1. The haptic profile may be based on the 3D object’s surface features.
   * These surface features may be explicitly defined at design time such as in the case with a haptic texture being associated with a 3D object
   * Alternatively, the surface features may be derived from other object attributes such as its surface geometry, applied textures/shaders, or virtual material, visual appearance, context, past interactions, etc.
2. A virtual thermal profile may be created based on the virtual material. For example, the virtual heat flux (sensed by the human body through thermoreception) of a virtual wood material, presented through the user through thermal feedback, can be distinct from the virtual heat flux associated with a virtual metal material.

**Smooth haptic transitions** – When transitioning between interacting with one part of a room and another part, or one object in the room and another object, haptics associated with each will be mixed to minimize tactile artifacts and ensure a smooth transition.

**VR video calling -** When interacting via a VR video call, users may send haptic effects to each other in the following ways:

1. Through gesture, for example:
   1. By touching the user’s avatar image (representing the VR user) or the region of the video that includes the image of the person (representing the video user).
   2. By touching objects in the remote user’s environment
2. By attaching and sending an external media element with an associated haptic track or haptic effect such as a haptic sticker, GIF, animation, video clip, or virtual object.

# ANNEX 3 - Hardware Configurations

## A3.1 Vibrotactile Test Platform

## Introduction

Actronika is providing a haptic actuator that can be driven with class-D amplifiers or equivalents.

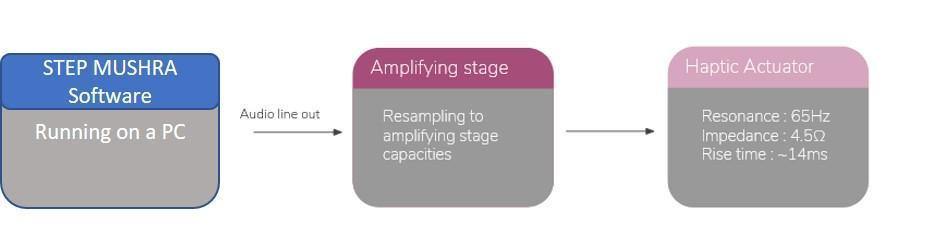
For the purpose of the CfP, Actronika proposes the following amplifying stage : MAX9744

20W Stereo amplifier - to be powered with 9V.

Reference Adafruit 1752, Mouser 485-1752.

**Important Note:** The details of this hardware test platform, including pricing, are being provided in the CfP merely for informational purposes. Proponents are **NOT** required to purchase this platform as a prerequisite for responding to the CfP. That said, proponents wishing to purchase a test platform for their own internal testing, prior to submission of their technology, are free to purchase either this platform or any other platform with comparable specifications (as described in Table 1 below).

## Block diagram

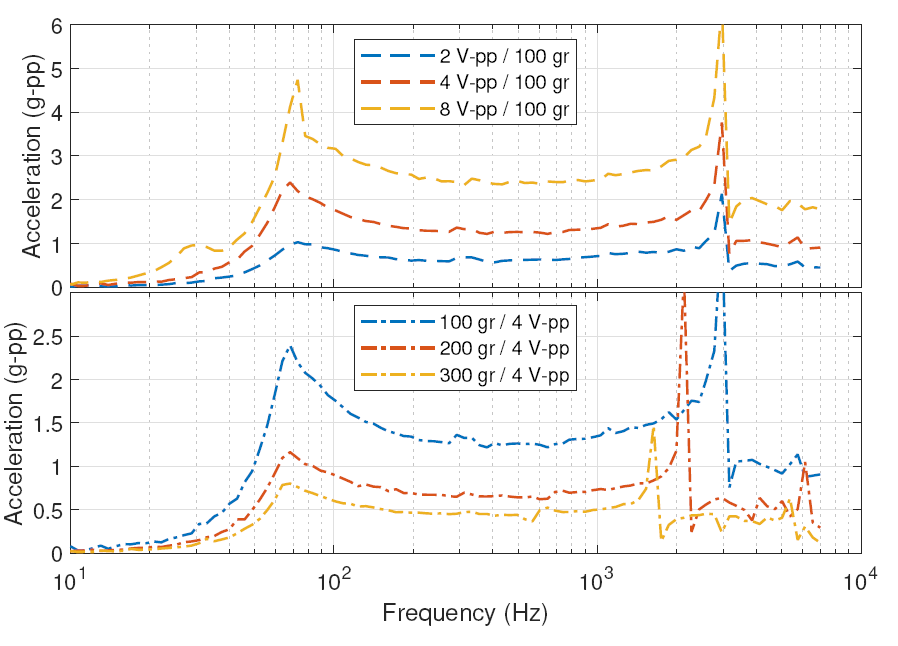


## Haptic actuator

* Standard actuator - HFBA121238

|  |  |
| --- | --- |
| Technology | Large bandwidth voice coil |
| Dimensions | 11.5x12x37.7 mm3 |
| Resonance frequency | 65 Hz |
| Weight | 8.7 grams |

* Bandwidth and testing



The usable range of this actuator is at least 65-300 Hz.

Testing tolerances:

* 60-200Hz - +/- 10%
* 200-800 Hz - +/-7%

## Vibrotactile Test Platform Specification

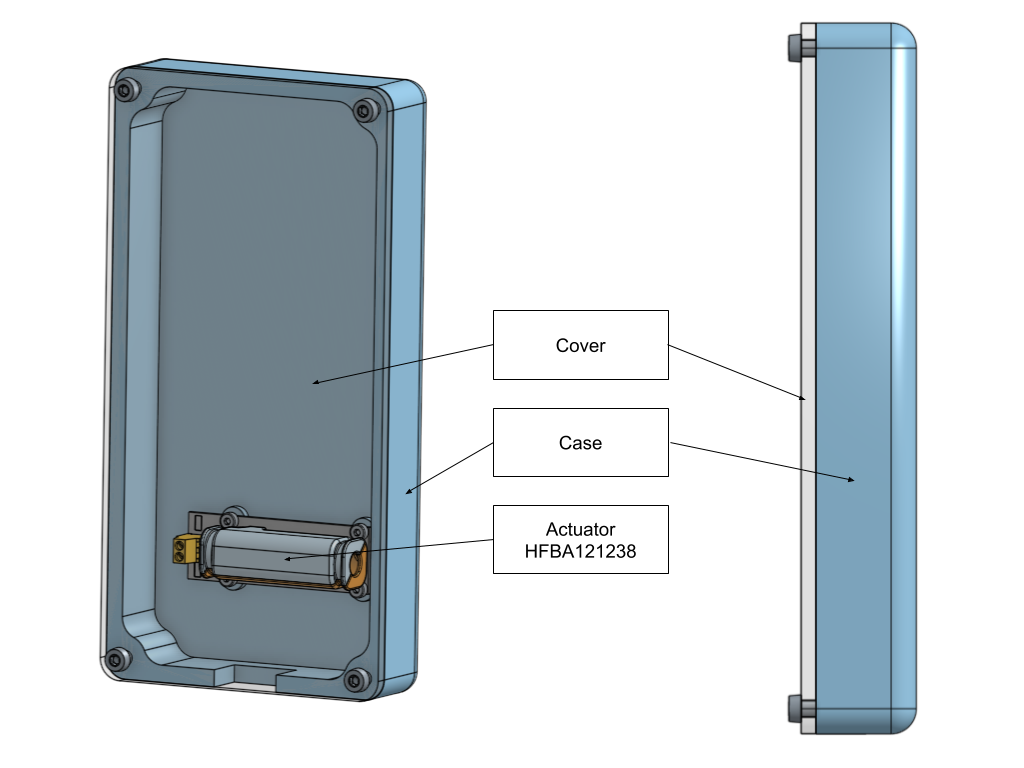
**Table 1: Vibrotactile Test Platform Specifications**

|  |  |
| --- | --- |
| Technology | Large bandwidth voice coil |
| Dimensions | 11.5x12x37.7 mm3 |
| Resonance frequency | 65 Hz |
| Mass | 8.7 grams |
| Manufacturing tolerance | * 60-200Hz - +/- 10% * 200-800 Hz - +/-7% |
| Bandwidth (100g test mass) | At least 1G-pp (peak-peak) of uniaxial acceleration from 65-300Hz. |

## Module mockup

The module mockup is custom made for the purpose of that proposal. It allows fixing the actuator within a casing of a generic shape. It consists of a main case and a transparent cover fixed with screws. Orientation marking will be done so the object is tested in the same orientation.

Overview



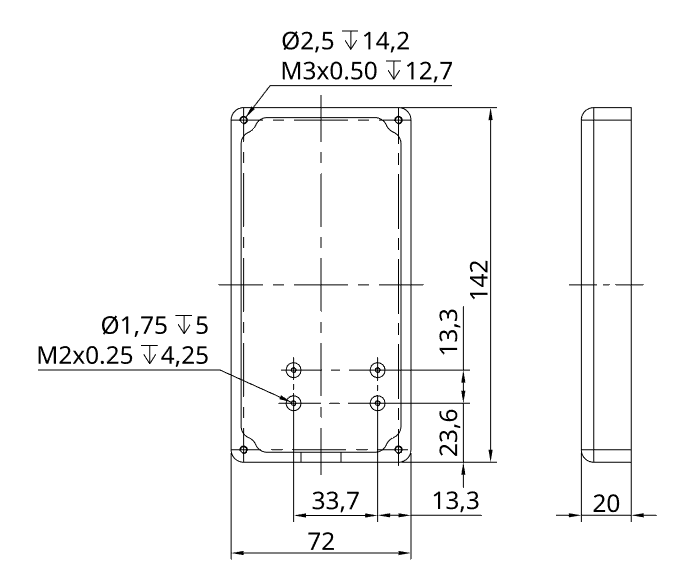
Technical specifications

|  |  |
| --- | --- |
| Dimensions (case + cover) | 142x72x22 |
| Case (material) | ABS |
| Case (mass) | 67 grams\* |
| Cover (material) | Plexiglass |
| Cover (mass) | 23 grams\* |
| Actuator assembly (mass) | 11 grams |
| Total mass | ~100 grams |

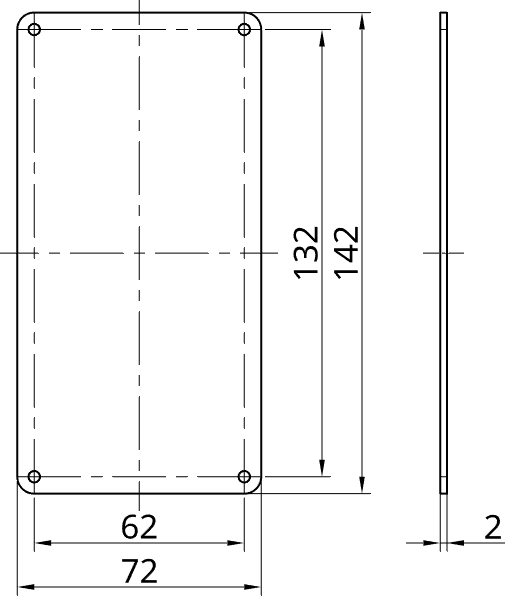
\*Mass estimations from the CAD software. The final design might vary from the presented above.

Technical drawing

* Casing



* Cover



## Pricing

|  |  |  |
| --- | --- | --- |
| **Component** | **Qty** | **Unit price** |
| Actuators | 2 | 10 EUR\* |
| Amplifying stage | 1 | 16,90 EUR |
| Module mockup | 1 | 200 EUR\* |
| **TOTAL** |  | **236,90 EUR**\* |

\*Pricing without shipping and taxes

All above components, beside module mockup will be available before January 2021. For more information about acquiring this platform, please contact Thomas Begeot at [thomas.begeot@actronika.com](mailto:thomas.begeot@actronika.com).

## A3.2 Kinesthetic Test Platform

## Introduction

The 3DSystem Geomagic Touch provides 3DoF force-feedback and inputs 6DoF position through a handheld stylus. For the purpose of the CfP, it is proposed for evaluating kinesthetic content.

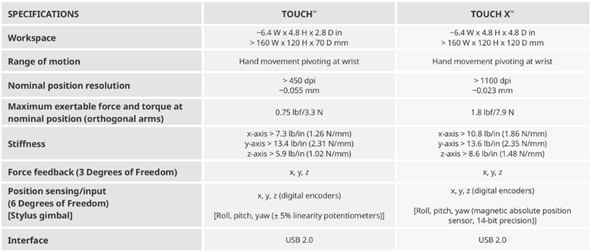
**Important Note:** The details of this hardware test platform, including pricing, are being provided in the CfP merely for informational purposes. Proponents are NOT required to purchase this platform as a prerequisite for responding to the CfP. That said, proponents wishing to purchase a test platform for their own internal testing, prior to submission of their technology, are free to purchase either this platform or any other platform with comparable specifications (as described in Table 2 below).



## Device Specifications

The device is connected to a computer with an USB 2.0 interface. Table 2 below presents its specifications. More information can be found on the constructor website (<https://www.3dsystems.com/haptics-devices/touch/specifications>)

**Table 2: Kinesthetic Test Platform Specifications**



## Block Diagram

The content of a wav file will be playback by the MUSHRA test platform that outputs the signal to an ASIO audio driver. This signal will be captured by a provided piece of software and will be rendered on the Touch device.



**MUSHRA Platform**

## API

The open-source API CHAI3D C++ will be used to set the force to be rendered by the device. The documentation can be found here: https://www.chai3d.org/documentation/getting-started.

## Pricing

The cost of the device is about $2,300.

# ANNEX 4 – Haptic Effects and Test Signal Categorization

## A4.1 Introduction

The goal of this Annex is to define content materials that will cover all the sensations provided by vibrotactile feedback in the application scenarios of interest in this CfP. Defining a comprehensive taxonomy of haptic effects is a difficult challenge that has only been partially answered in the literature [7,8]: vibrotactile content may be classified according to physical properties, sensory properties (i.e., roughness), emotional perspective (pleasantness, arousal, etc.), usage examples (alarm, message, ..) or metaphoric representation (heartbeat, explosion, etc.). This categorization could potentially help select a variety of representative signals. However, the taxonomy described in [7,8] has the following caveat: a large part of the taxonomy is not applicable to evaluating coded representations of haptics (the focus of this CfP). More to the point, only a small subset of them is representative of the kinds of haptic effects likely to be encountered by users in MPEG-I application scenarios described in Annex 2. However, we are unable to include any content from the VibViz database [8] in our suite of test signals due to licensing issues. As such, we will be using test signals extracted from other content that is more representative of the application scenarios of interest.

In A4.2, we describe the taxonomy of haptic effects used by Immersion Corporation that has some overlap with that described in [7,8] and provide illustrative examples. In A4.3, we describe the categorization of test signals used for the MUSHRA tests.

Annex 5 has the full listing of test content from Immersion, InterDigital, Apple, and Actronika from which test signals for the MUSHRA tests will be extracted, using the categorization in A4.3.

## A4.2 Haptic Design Taxonomy from Immersion

The Immersion test content (A5.1) is built using a specific set of well-defined effect types, which are described in this section and are generally useful for vibrotactile haptic experiences. The test content provided by Immersion was built using these haptic design concepts and implemented using IVS or PCM effect encodings using coding primitives (described in Section 5 of [3]) that may not map directly to the conceptual design element taxonomy. An overview of the design taxonomy is provided here to ensure that final test data provides sufficient coverage of commonly used effect types.

|  |  |  |  |
| --- | --- | --- | --- |
| **Design Element** | **Description** | **Example** | **Test Content Exemplar** |
| Pulse | Short, strong, and crisp, often used for button clicks or abrupt collisions. | Pops and clicks on virtual buttons and animations. | IMMR-vib-pulse01-8Khz-16-nopad |
| Tick | Very short and crisp, nearing the edge of perceptual threshold. | Tick of a number wheel picker. | IMMR-vib-tick02-nopad |
| Fill | Tightly clustered group of variegated pulses and ticks. | Explosion | IMMR-vib-fill03-8Khz-16-nopad |
| Sweep | A single tactile event longer than a pulse but not sustained, usually with a magnitude that either fades in, fades out, or both. | Car racing past fixed observer | IMMR-vib-sweep02-8Khz-16-nopad |
| Grain | A very short duration, very low magnitude effect that only takes on meaning when played as part of a sequence of grains. | UI: Drawing on virtual paper surface | IMMR-vib-grain02-8Khz-16-nopad |
| Sustained (random) | Vibration texture with components that have random or pseudo- random magnitudes and timing | Road texture | MMR-vib-sust01-441Khz-16-nopad |
| Sustained (periodic) | Vibration texture that loops consistent magnitudes and timing | Rollercoaster track as rollercoaster ascends the ramp | MMR-vib-sustper01-8Khz-16-nopad |

## A4.3 Test Signal Categorization for MUSHRA Tests

Based on the discussion of the haptic effects above, the following categorization will be used to divide the test signals (from all test content) for the MUSHRA tests:

* **Short**: A short effect with a simple envelope (e.g., ADSR), where the envelope is the salient characteristic of the effect. Could range anywhere from 20 ms to 1000 ms.
* **Long**: A long effect where the placement and envelopes of its subcomponents are the salient characteristics of the effect. Could range anywhere from 1000 ms to over 5000 ms.

There will be one set of MUSHRA tests for each category of effects. Each MUSHRA test session will be limited to 30 minutes, to prevent test subject fatigue.

Annex 5 has the listing of test content from which test signals for the two MUSHRA tests will be selected.

# ANNEX 5 - Phase 1 Test Material

Please note that the generation and preparation of the actual test signals to be used in the MUSHRA tests is ongoing and is expected to be completed on April 15, 2021. Test content will be made available at no charge to registered proponents upon signing of appropriate license agreements with the content owners. If there is any issue with licensing content for the purposes of this CfP evaluation, that content will be excluded from the evaluation.

The tables below show the source content to be used in Tests 1.1, 1.2, and 1.3.

Each test signal will have 1000 ms of preceding silence and 1000 ms of trailing silence so that the effect is properly isolated and is perceptually clear.

All the test signals will have an associated .ohm metadata file, as described in [3].

**Table 3: (Vibrotactile) Short Test Content Using Immersion Taxonomy**

**(usage license required for all test content)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ID** | **Content Type** | **File Name** | **Duration (seconds)** | **File Format** | **Originator** |
| IMMS1 | UI | IMMR-vib-tick02-nopad | 0.02 | .ivs | Immersion Corporation |
| IMMS2 | UI | IMMR-vib-tick03-nopad | 0.02 | .ivs | Immersion Corporation |
| IMMS3 | UI | IMMR-vib-tick04-nopad | 0.02 | .ivs | Immersion Corporation |
| IMMS4 | Media | IMMR-vib-fill01-441Khz-16-nopad | 0.98 | .wav | Immersion Corporation |
| IMMS5 | Media | IMMR-vib-grain02-8Khz-16-nopad | 0.02 | .wav | Immersion Corporation |
| IMMS6 | Media | IMMR-vib-sweep03-8Khz-16-nopad | 0.94 | .wav | Immersion Corporation |
| IMMS7 | Media | IMMR-vib-sweep04-8Khz-16-nopad | 0.96 | .wav | Immersion Corporation |
| IMMS8 | UI | IMMR-vib-tick01-8Khz-16-nopad | 0.05 | .wav | Immersion Corporation |

**Table 4: (Vibrotactile) Long Test Content Using Immersion Taxonomy**

**(usage license required for all test content)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ID** | **Content Type** | **File Name** | **Duration (seconds)** | **File Format** | **Originator** |
| IMML1 | Alerts | IMMR-vib-ringtone-nopad | 2.4 | .ivs | Immersion Corporation |
| IMML2 | Media | IMMR-vib-transformers-nopad | 254 | .ivs | Immersion Corporation |
| IMML3 | Media | IMMR-vib-fill02-8Khz-16-nopad | 4.14 | .wav | Immersion Corporation |
| IMML4 | Media | IMMR-vib-fill03-8Khz-16-nopad | 2.11 | .wav | Immersion Corporation |
| IMML5 | Media | IMMR-vib-grain01-441Khz-16-nopad | 1.66 | .wav | Immersion Corporation |
| IMML6 | UI | IMMR-vib-pulse01-8Khz-16-nopad | 1.3 | .wav | Immersion Corporation |
| IMML7 | Media | IMMR-vib-sust01-441Khz-16-nopad | 4.38 | .wav | Immersion Corporation |
| IMML8 | Media | IMMR-vib-sustper01-8Khz-16-nopad | 2.53 | .wav | Immersion Corporation |
| IMML9 | Media | IMMR-vib-sweep01-441Khz-16-nopad | 6.65 | .wav | Immersion Corporation |
| IMML10 | Media | IMMR-vib-sweep02-8Khz-16-nopad | 3.39 | .wav | Immersion Corporation |
| IMML11 | Media | IMMR-vib-sweep05-8Khz-16-nopad | 2.05 | .wav | Immersion Corporation |
| IMML12 | Media | IMMR-vib-sweep06-8Khz-16-nopad | 3.15 | .wav | Immersion Corporation |

**Table 5: (Vibrotactile) Long Test Content from InterDigital/Actronika**

**(usage license required for all test content)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ID** | **Name** | **Description** | **Duration (seconds)** | **File Format** | **Originator** |
| IDC1 | IDCC-vib-HeartBeat-8Khz-16-nopad | Heartbeat pattern. 2 beats per sec. | 3" | .wav | InterDigital |
| IDC2 | IDCC-vib-Rain-8Khz-16-nopad | Four channels\*, rain like effect. | 4" | .wav | InterDigital |
| IDC3 | IDCC-vib-Carpet-8Khz-16-nopad | One channel, finger rubbing carpet surface. | 4" | .wav | InterDigital |
| IDC4 | IDCC-vib-Paper-8Khz-16-nopad | One channel, finger rubbing the surface of a paper. | 4" | .wav | InterDigital |
| IDC5 | IDCC-vib-Towel-8Khz-16-nopad | One channel, finger rubbing the surface of a towel. | 4" | .wav | InterDigital |
| ACT1 | ACTK-vib-pantheon-grandstarfall-8Khz-16-nopad | Energy channeling followed by a charge to its target. Made for a custom demonstration of League of Legend. | 7" | .wav | Actronika |

\* This content includes 4 channels, channel 3 (right arm) must be selected for the tests.

**Table 6: Kinesthetic Test Content from InterDigital**

**(usage license required for all test content)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ID** | **Name** | **Description** | **Duration (seconds)** | **File Format** | **Originator** |
| IDC6 | IDCC-kin-ForceYFast-8kHz-16-nopad | Quick up down force (force along y axis) | 2” | .wav | InterDigital |
| IDC7 | IDCC-kin-ForceYSlow-8kHz-16-nopad | Slow up down force (force along y axis) | 4” | .wav | InterDigital |
| IDC8 | IDCC-kin-ForceXFast-8kHz-16-nopad | Quick left right force (force along x axis) | 2” | .wav | InterDigital |
| IDC9 | IDCC-kin-ForceXSlow-8kHz-16-nopad | Slow left right force (force along x axis) | 4” | .wav | InterDigital |
| IDC10 | IDCC-kin-ForceZFast-8kHz-16-nopad | Quick front back force (force along z axis) | 2” | .wav | InterDigital |
| IDC11 | IDCC-kin-ForceZSlow-8kHz-16-nopad | Slow front force (force along z axis) | 4” | .wav | InterDigital |
| IDC12 | IDCC-kin-HorseRiding-8kHz-16-nopad | Acceleration signal (x,y,z) | 5” | .wav | InterDigital |
| IDC13 | IDCC-kin-BikeRiding-8kHz-16-nopad | Acceleration signal (x,y,z) | 5” | .wav | InterDigital |
| IDC14 | IDCC-kin-Rollercoaster-8kHz-16-nopad | Position signal (x,y,z) | 5” | .wav | InterDigital |
| IDC15 | IDCC-kin-BigBuckBunny-8kHz-16-nopad | Position signal (x,y,z) | 5” | .wav | InterDigital |

**Table 7: (Vibrotactile) Test Content from Apple, Inc.**

**(usage license required for all test content)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ID** | **Name** | **Description** | **Duration (seconds)** | **File Format** | **Originator** |
| APPL1 | APPL-vib-boing-nopad.ahap | A continuous event tweaked by sloped parameter curves, emphasized by a precisely placed transient event, creating the feeling of a spring or rubber band. | 0.265 | .ahap | Apple Inc |
| APPL2 | APPL-vib-drums-nopad.ahap | A sequence of haptic events paired with a custom audio file 'Drums.wav' | 2.6 | .ahap | Apple Inc |
| APPL3 | APPL-vib-gravel-nopad.ahap | A sequence of tightly spaced transient events and parameter variations to create a gravel-like texture. | 3.0 | .ahap | Apple Inc |
| APPL4 | APPL-vib-heartbeats-nopad.ahap | Three organic heartbeats over three seconds, made using precisely spaced transient events at varying parameters. | 2.255 | .ahap | Apple Inc |
| APPL5 | APPL-vib-inflate-nopad.ahap | An effect that builds in sharpness and intensity. | 1.7 | .ahap | Apple Inc |
| APPL6 | APPL-vib-oscillate-nopad.ahap | A mixture of two continuous events, shifting the sharpness of one to create a smooth oscillating feel. | 3.0 | .ahap | Apple Inc |
| APPL7 | APPL-vib-rumble-nopad.ahap | A sequence of transient events with decreasing sharpness, to create a precise rumble, reminiscent of driving over uneven ground. | 1.52 | .ahap | Apple Inc |
| APPL8 | APPL-vib-sparkle-nopad.ahap | A combination of transient and continuous events to create a pop with trailing sparkles. | 1.02 | .ahap | Apple Inc |