

# Quality of Experience in UHD-1 Phase 2 television: the contribution of UHD+HFR technology

Vedad Hulusic\*, Giuseppe Valenzise<sup>†</sup>, Jean-Charles Gicquel<sup>‡</sup>, Jérôme Fournier<sup>‡</sup> and Frédéric Dufaux<sup>†</sup>

\*LTCI, Télécom ParisTech, Université Paris-Saclay, 75013 Paris, France

<sup>†</sup>L2S, UMR 8506, CNRS - CentraleSupélec - Université Paris-Sud, 91192 Gif-sur-Yvette, France

<sup>‡</sup>Orange Labs, 35512 Cesson-Sévigné Cedex, France

**Abstract**—A key factor to determine the quality of experience (QoE) of a video is its capability to convey the large spectrum of perceptual phenomena that our eyes can sense in real life. In order to meet this demand, the recent DVB UHD-1 Phase 2 specification employs new video features, such as higher spatial resolutions (4K/8K) and High Frame Rate (HFR). The first enables larger field of view and level of details, while the second offers sharper images of moving objects going well beyond the current frame rates. While the contribution of each of these technologies to QoE has been investigated individually, in this paper we are interested to study their interaction, and in quantifying the benefits to users from their combination. To this end, we conduct a subjective test on compressed UHD+HFR content on a recent display capable of reproducing 100 pictures per second at 2160p resolution, with the goal to assess the increase in QoE of UHD and HFR with respect to conventional video, both individually and in combination. The results indicate that for content with fast motion, at higher bitrates the combination of UHD and HFR significantly improves the QoE compared to that obtained when these features are used individually.

## I. INTRODUCTION

In order to further increase the quality of experience (QoE), emerging video formats aim at offering immersive video features: on one hand, by enabling higher interactivity, such as in free-viewpoint and 360-degrees video [1], [2]; and on the other hand, by providing a better reproduction fidelity, e.g., by increasing the video resolution and field of view (FOV), or the range of reproducible luminance and dynamic range [3].

Ultra high-definition TV (UHDTV) systems augment realism by adding to existing high-definition video (HDTV) new features such as higher spatial resolution (and thus a larger FOV), high dynamic range (HDR), wide color gamut and high frame rate (HFR). The coding and transmission requirements of each feature, as well as their *individual* contribution to improve QoE, have been widely studied in the past few years [4], [5], [6]. However, the *combined* effect of two or more of them has not been investigated so far, mainly due to the lack of display devices able to jointly reproduce two or more of these functionalities. While the initial phase of UHDTV was limited to increasing spatial resolution to 4K ( $3840 \times 2160$ ), phase 2 of UHDTV deployment aims at adding the other features on top of it. In this paper, we study and quantify the QoE improvement obtained by having two UHDTV features – HFR (at 100 Hz) and 4K resolution – compared to the HDTV baseline, UHDTV phase 1 (4K only), and HDTV with high frame rate.

The effect of frame rate on QoE and its relationship with large fields of view have been studied both from a psycho-perceptual [7], [8], [9] and, more recently, from a QoE perspective [10], [11], [12], [4]. From the psycho-perceptual point of view, a number of studies have attempted to measure and model the spatio-temporal contrast sensitivity function, which describes the human sensitivity to contrast at various spatial and temporal frequencies [8]. The portion of FOV occupied by the display plays a role in the perception of contrast, e.g., humans are more sensitive to flicker in extra-foveal vision. An insufficient frame rate may result in temporal aliasing and strobing artifacts, leading to juddery motion or multiple imaging. On the other hand, acquisition devices have finite integration times, which produces motion blur when an object moves faster than the shutter speed. Similarly, when an object moves on the display, it stays still for a short amount of time, while the retina tracks objects continuously, resulting in the perception of smeared edges – a phenomenon known as “retinal slip” [13], [11]. Recently, Noland has applied classical sampling theory to find the minimum frame rate, for different viewing distances (and thus FOVs), required to remove any temporal aliasing from video display [13]. She shows that the required frame rate increases linearly with spatial dimension, and that a frame rate of about 140 Hz would be sufficient for UHD-1 4K video viewed at 1.5 screen heights. However, she also reports that if eye tracking of objects is taken into account, the minimum frame rate becomes sensibly higher.

The results of a number of subjective studies investigating the individual contribution to QoE of UHD 4K spatial resolution and HD HFR are reported in [4]. At a viewing distance of 1.5 screen heights (i.e., the recommended viewing distance for 4K resolution), the video quality of 4K images is half a grade higher on a five-point quality scale compared to HD images. As for HFR, the improvement in scores between 60 and 120 Hz, and 120 and 240 Hz are 0.46 and 0.23, respectively. These results lead to the conclusion that a frame frequency of at least 120 Hz is desirable for UHDTV to improve its moving picture quality. Kuroki et al. [11] found that a frame rate of 250 Hz would guarantee the absence of blurring and retinal slip on most video content. Internal (unpublished) tests conducted jointly by OrangeLabs, EBU and BBC show that, for HD content, 100 Hz video results in QoE improvements of 10 to 30 points in a 100-point quality scale. Even when video is compressed at low bitrate (3 Mbps), HD 100 Hz content

achieved good quality scores with a 15 points difference on average over 50 Hz video. It is important to mention that all the subjective studies on HFR/UHD QoE report that quality improvements are strongly content dependent [10], [12].

In this paper we investigate the joint effect of high spatial and temporal resolution on visual quality, which is one key component of QoE, by conducting a subjective test using compressed UHD+HFR video sequences. To our knowledge, this is the first time that the joint contribution to QoE of these two technologies is evaluated on complex stimuli. This is made possible by the recent introduction on the market of a display (described in Sec. II-C) capable of reproducing 4K video at 100 Hz. Our contribution is twofold: i) we measure the gain in quality obtained by UHD+HFR video, compared to standard HD (50 Hz), HD at 100 Hz, and UHD-1 Phase 1 (4K resolution at 50 Hz); ii) by employing sequences compressed with HEVC [14] at different bitrates, we estimate which are the minimal bitrates such that UHD+HFR provides a significant gain in QoE compared to the existing solutions.

The rest of the paper is organized as follows: in Section II the test methodology is introduced and the details discussed; in Section III the results of the subjective study are presented, followed with the short discussion on the findings and conclusions in Section IV.

## II. SUBJECTIVE EVALUATION

### A. Design

In the study, the participants' perceptual responses were evaluated in a pseudo randomized design. The independent variables were image content, spatio-temporal resolution (*Res\_freq*) and bitrate. The *Res\_freq* condition consisted of video resolution (number of pixels per frame) and frame rate (number of frames per second). The dependent variable was the rating of the perceived video quality. The adjusted double-stimulus continuous quality-scale method (DSCQS) was used [15]. The participants were asked to evaluate the overall quality of each presented sequence on continuous vertical scales. The scales were divided into five equal lengths, corresponding to the standard ITU five-point quality scale. The associated terms *Excellent*, *Good*, *Fair*, *Poor* and *Bad* were included for general guidance.

Four scenes, with significant differences in image statistics, all lasting for 10 seconds, were evaluated in the study. All of them were displayed using four spatio-temporal configurations: 1080p-50Hz, 1080p-100Hz, 2160p-50Hz and 2160p-100Hz. The sequences were encoded using 5 bitrates, which correspond to realistic bandwidth allocation for HD and UHD material, see Table I. 100Mbps was the highest bitrate that the hardware used in the study could support, and it was used as the reference case for each tested condition.

For each participant a playlist has been generated with a random sequence order, not allowing for a consecutive display of the same sequence and the test/reference order was changed randomly. Each playlist contained three dummy trials at the beginning of the first session in order to stabilize the participant's opinion and the hidden reference for each content

TABLE I: Video conditions with their corresponding bitrates. The abbreviated names as used later in the paper.

| <i>Res_freq</i><br>condition | Abbreviated<br>name | Bitrates (Mbps) |   |   |    |    |
|------------------------------|---------------------|-----------------|---|---|----|----|
|                              |                     | 1.5             | 3 | 6 | 12 | 24 |
| 1080p-50Hz                   | RF1                 | 1.5             | 3 | 6 | 12 | 24 |
| 1080p-100Hz                  | RF2                 | 1.5             | 3 | 6 | 12 | 24 |
| 2160p-50Hz                   | RF3                 |                 | 3 | 6 | 12 | 24 |
| 2160p-100Hz                  | RF4                 |                 | 3 | 6 | 12 | 24 |

(reference sequence compared to itself). The scores issued from the dummy sequences were not used for data analysis. Each video in the sequence pair was displayed twice, and the voting was allowed after seeing both videos once. Each trial was preceded by the sequence number, while each video was preceded by letters 'A' or 'B' indicating first and the second test video, each time displayed on a black background. Gray background was substituted by the black to avoid some flicker caused by technical limitations of the display.

### B. Participants

27 participants (16M/11F), aged between 16 and 45, with an average age of 25.7, were recruited for the experiment. All of them were tested with the Essilor Ergovision apparatus for the visual acuity (near and far vision: at least 10/10), hyperopia, astigmatism and phoria, and all showed at least the visual acuity of 10/10 and no other vision disorders. All participants were external to the affiliated institutions and were naïve to the purpose of the experiment.

### C. Apparatus

The experiments were conducted in a test room with the middle-gray walls and D65 illuminant. The background luminance was set to  $28 \text{ cd/m}^2$ , that is around 8.2% of the display peak luminance. Viewing distance was fixed to 1.6 times the height of the screen (128 cm). The test was performed on a 65 inches, LG OLED B6 display with HFR capabilities (prototype provided by LG for the project's studies on HFR) and peak luminance of  $340 \text{ cd/m}^2$ . During the whole duration of the test, *Standard display* mode was activated and *TruMotion* mode disabled to avoid some visual artifacts on moving objects at 50 Hz.

Time measurements and calibration for the display were carried out using the ELDIM equipment (OPTIScope-SA and associated measurement software). Electro-Optical Transfer Function (EOTF) was in accordance with ITU-R BT.1886 recommendation for HD and UHD resolutions as well as 50 Hz and 100 Hz frame rates.

### D. Stimuli

Four test scenes *Bicycle*, *Gasquet-Adv*, *Rugby-Mele* and *Euro2016-2* were selected from a pool of 9 proprietary videos, based on spatial and temporal perceptual information metrics (see Fig. 1) and content semantics. All video sequences were captured at 100fps and exported with the following video format (2160p-100Hz):  $3840 \times 2160$  image resolution;

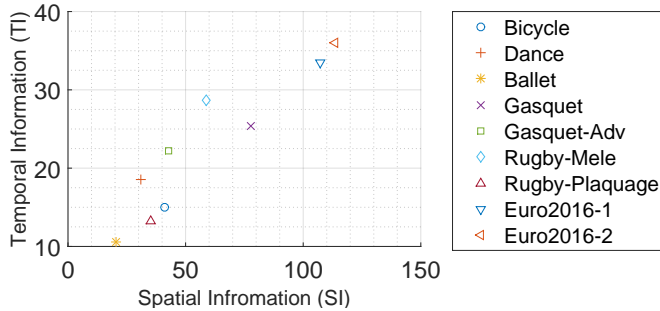


Fig. 1: Spatial perceptual information (SI) and temporal perceptual information (TI) computed according to the R-REC-BT.500-13 recommendation document.

YUV 4:2:2 10-bits components; Gamma 2.4; and Rec.2020 color gamut. Subsequently, other three conditions: 1080p-50Hz, 1080p-100Hz and 2160p-50Hz were generated. The frame rate reduction was performed using pair-based frame averaging. The videos were downsized from 2160p to 1080p using the Lanczos3 kernel. They were then coded using the ATEME Titan video encoder (version 3.8) with the following parameters: HEVC video encoder (Main-10 profile); constant video bitrate; Rec.2020 color gamut. The Intra period was fixed to 1 second, corresponding to realistic zapping duration in a broadcasting scenario. Due to the interface bandwidth, the reference sequences were presented using video bistreams compressed at 100 Mbps.

#### E. Procedure

All videos were evaluated by each participant. Due to the complete duration of around 70 minutes, the evaluation was performed in three sessions. Prior to the experiment, the participants were explained the concepts of spatial resolution, temporal frequency, compression and the experimental procedure. Viewers were asked to assess the overall video quality using the video quality scale. They were told that the rendering of textures, details as well as motion could help them to adjust their opinion about the overall video quality. They were then showed four example trials, using two sequences not used in the study, at different *Bitrate* and *Res\_freq* conditions. These were used in order to demonstrate the magnitude of the effect of different independent variables, and to familiarize themselves with the experimental procedure. The user evaluation was performed on the printed copies of the quality-rating forms.

### III. RESULTS

In order to analyze the collected data, firstly the difference scores were computed by subtracting the rated quality assigned to a test video from the quality assigned by the same subject to the corresponding reference sequence. With the computed difference scores the descriptive analysis was performed, computing the mean values with corresponding confidence intervals, see Fig. 2. The data was analyzed per

scene, condition and bitrate for better understanding of the outcomes. Screening of the observers, as proposed in the R-REC-BT.500-13 recommendation document [15], resulted with one outlier, that was excluded from the further analysis.

To check for the significance between the DMOS values between four different scenes, four *Res\_freq* conditions, and four common bitrates, the repeated-measures factorial ANOVA test was conducted. In order to test for the sphericity, that is to test the hypothesis that the variances of the differences between conditions are equal, the Mauchly's test is performed. In case of violation of this assumption, and since even very small departures from sphericity produce large biases in the F-test, the degrees of freedom are corrected using Greenhouse-Geisser estimates of sphericity for assessing the observed F-ratio. Furthermore, to control for the Type I error rate in pairwise comparisons, Bonferroni correction is utilized. For the rest of the paper, the *Res\_freq* conditions will be referred to using the abbreviated names, see Table I.

#### A. Cumulative results

The results of the Mauchly's test revealed that it was violated for the main effect of *Scene* and *Res\_freq*,  $\chi^2(5) = 28.13, p < .001$ ; main effect of *Bitrate*,  $\chi^2(5) = 29.22, p < .001$ , and for the interaction effect between the *Scene* and *Res\_freq*,  $\chi^2(44) = 72.18, p = .006$ . Therefore, the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. All effects are reported as significant at  $p < .05$ . There were significant main effects of *Scene*,  $F(1.88, 47.01) = 68.76, p < .001$ ; *Res\_freq*,  $F(1.67, 41.79) = 39.3, p < .001$ ; and *bitrate*,  $F(1.65, 41.32) = 279.83, p < .001$ . A significant interaction effect between all the variables was also reported. The multiple comparison with the Bonferroni correction indicates that the significant main effect reflects a significant difference between RF1 and RF2 ( $p < .001$ ), RF1 and RF4 ( $p < .001$ ), RF2 and RF3 ( $p < .001$ ), and between RF3 and RF4 ( $p < .001$ ) conditions. However, to better understand these effects, the data has to be further analyzed per scene and bitrate.

#### B. Per scene results

Due to the significant effect of the scene, and in accordance with the R-REC-BT.500-13 recommendation document [15], the results were further analyzed per scene. The same statistical test was performed on four *Res\_freq* conditions and four bitrates. The summary of the t-test performed on each pair of *Res\_freq* conditions (pairwise comparison) is displayed in Table II.

1) *Bicycle scene*: Mauchly's test revealed that the sphericity was violated for the main effect of *Bitrate*,  $\chi^2(5) = 20.99, p = .001$  and interaction between the *Bitrate* and *Res\_freq*,  $\chi^2(44) = 82.64, p < .001$ . The test of within-subject effects shows that the effect of *Bitrate*  $F(2.06, 51.46) = 77.48, p < .001$ , and the interaction effect between the two factors  $F(5.29, 132.26) = 2.72, p = .02$  were found as significant. The pairwise comparisons for the main effects of *Res\_freq* did not show any significant difference between

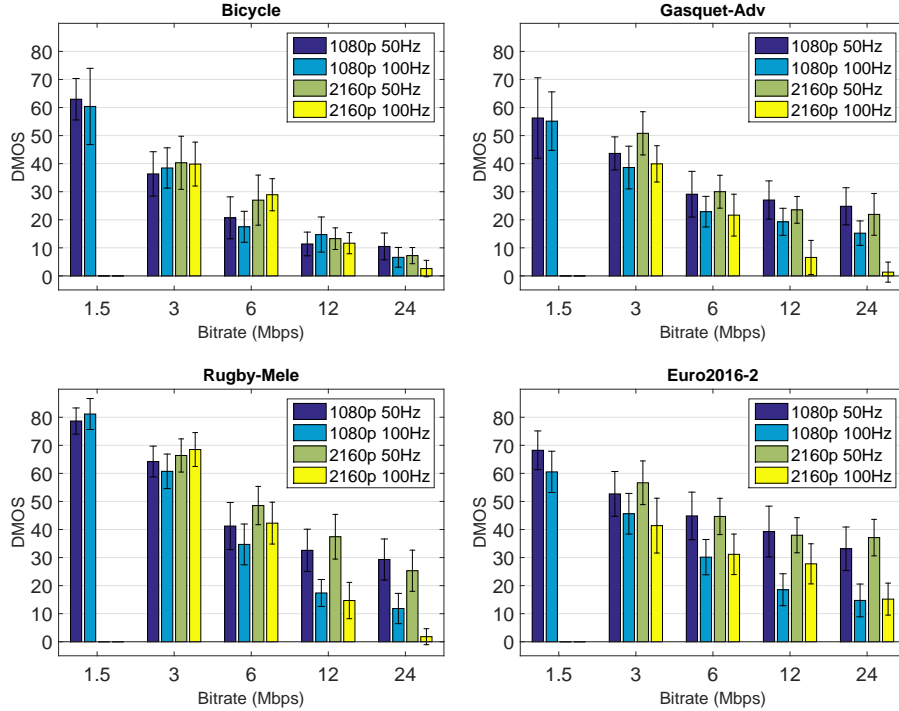


Fig. 2: Comparison of difference mean opinion score (DMOS) across different *scenes*, *bitrates* and *Res\_freq* conditions, with the 95% confidence intervals.

the conditions. The same test found a significant difference between each level of bitrate at  $p < .001$ .

2) *Gasquet-Adv scene*: Mauchly's test for the Gasquet-Adv scene data revealed that the sphericity was violated for the main effect of *Bitrate*,  $\chi^2(5) = 15.62, p = .008$  and interaction between the two factors,  $\chi^2(44) = 79.96, p = .001$ . The main effects of *Res\_freq*,  $F(3, 75) = 25.25, p < .001$ , *bitrate*  $F(2.27, 56.67) = 78.52, p < .001$ , and the interaction effect between the two factors  $F(5.22, 130.54) = 3.53, p = .004$  were significant. The pairwise comparisons for the main effects of *Res\_freq* corrected using a Bonferroni adjustment showed that there is a significant difference ( $p < .05$ ) between all *Res\_freq* condition pairs, except between the RF1 and RF3 conditions. The same test has not found a significant difference between conditions at 12 and 24 Mbps. For all other pairs, a significant difference between bitrates was found at  $p < .05$ .

3) *Rugby-Mele*: For the Rugby-Mele scene, the sphericity was not violated for any effect. The main effects of *Res\_freq*  $F(3, 75) = 23.93, p < .001$ , *Bitrate*  $F(3, 75) = 212.41, p < .001$  and the interaction effect between the two factors,  $F(9, 225) = 8.2, p < .001$  were significant. The pairwise comparisons for the main effects of *Res\_freq* showed that there is a significant difference ( $p < .05$ ) between RF1 and RF2, RF1 and RF4, RF2 and RF3, and between RF3 and RF4 conditions. A significant difference between each level of bitrate was found at  $p < .001$ .

4) *Euro2016-2 scene*: Mauchly's test revealed that the sphericity was violated for the main effects of *Res\_freq*,

TABLE II: A summary of the pairwise comparison per scene. The fields are marked if the row condition is significantly better than the column condition. Different markers and colors are used for each scene: Gasquet-Adv ( $\bullet$ ), Rugby ( $\triangle$ ) and Euro2016-2 ( $\star$ ). There were no significant differences between the conditions for the Bicycle scene.

| <i>Res_freq</i> | 1080p<br>50Hz             | 1080p<br>100Hz | 2160p<br>50Hz             | 2160p<br>100Hz |
|-----------------|---------------------------|----------------|---------------------------|----------------|
| 1080p-50Hz      | n/a                       |                |                           |                |
| 1080p-100Hz     | $\bullet \star \triangle$ | n/a            | $\bullet \star \triangle$ |                |
| 2160p-50Hz      |                           |                | n/a                       |                |
| 2160p-100Hz     | $\bullet \triangle \star$ | $\bullet$      | $\bullet \triangle \star$ | n/a            |

$\chi^2(5) = 24.17, p < .001$  and *Bitrate*,  $\chi^2(5) = 18.31, p = .003$ , and for the interaction between the two factors,  $\chi^2(44) = 63.41, p = .033$ . The test of within-subject effects show that both the effects of *Res\_freq*,  $F(1.89, 47.25) = 19.15, p < .001$  and *Bitrate*  $F(1.98, 49.53) = 50.04, p < .001$  were significant. The interaction effect between the two factors was not found as significant. The pairwise comparisons for the main effects of *Res\_freq* corrected using a Bonferroni adjustment showed that there is a significant difference ( $p < .05$ ) between RF1 and RF2, RF1 and RF4, RF2 and RF3, and between RF3 and RF4 conditions. The same test found a significant difference between each level of bitrate at  $p < .01$ .

TABLE III: A summary of the pairwise comparison per scene and per bitrates. The fields are marked if the row condition is significantly better than the column condition. Different markers and colors are used for each scene: Bicycle ( $\diamond$ ), Gasquet-Adv ( $\bullet$ ), Rugby ( $\triangle$ ) and Euro2016-2 ( $\star$ ). ‘n/a’ indicates that a pair of conditions was not compared.

|        |             | 1080p<br>50Hz                      | 1080p<br>100Hz                     | 2160p<br>50Hz                      | 2160p<br>100Hz                     |
|--------|-------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|
| 3Mbps  | 1080p-50Hz  | n/a                                |                                    |                                    |                                    |
|        | 1080p-100Hz |                                    | n/a                                |                                    |                                    |
|        | 2160p-50Hz  |                                    |                                    | n/a                                |                                    |
|        | 2160p-100Hz |                                    |                                    | $\star$                            | n/a                                |
|        | REF         | $\diamond \bullet \triangle \star$ | $\diamond \bullet \triangle \star$ | $\diamond \bullet \triangle \star$ | $\diamond \bullet \triangle \star$ |
| 6Mbps  | 1080p-50Hz  | n/a                                |                                    |                                    | $\diamond$                         |
|        | 1080p-100Hz | $\star$                            | n/a                                | $\bullet \triangle \star$          | $\diamond$                         |
|        | 2160p-50Hz  |                                    |                                    | n/a                                |                                    |
|        | 2160p-100Hz |                                    |                                    |                                    | n/a                                |
|        | REF         | $\diamond \bullet \triangle \star$ | $\diamond \bullet \triangle \star$ | $\diamond \bullet \triangle \star$ | $\diamond \bullet \triangle \star$ |
| 12Mbps | 1080p-50Hz  | n/a                                |                                    |                                    |                                    |
|        | 1080p-100Hz | $\triangle \star$                  | n/a                                | $\triangle \star$                  | $\star$                            |
|        | 2160p-50Hz  |                                    |                                    | n/a                                |                                    |
|        | 2160p-100Hz | $\bullet \triangle$                | $\bullet$                          | $\bullet \triangle$                | n/a                                |
|        | REF         | $\diamond \bullet \triangle \star$ | $\diamond \bullet \triangle \star$ | $\diamond \bullet \triangle \star$ | $\diamond \triangle \star$         |
| 24Mbps | 1080p-50Hz  | n/a                                |                                    |                                    |                                    |
|        | 1080p-100Hz | $\bullet \triangle \star$          | n/a                                | $\triangle \star$                  |                                    |
|        | 2160p-50Hz  |                                    |                                    | n/a                                |                                    |
|        | 2160p-100Hz | $\diamond \bullet \triangle \star$ | $\bullet \triangle$                | $\bullet \triangle \star$          | n/a                                |
|        | REF         | $\diamond \bullet \triangle \star$ | $\bullet \triangle \star$          | $\diamond \bullet \triangle \star$ | $\star$                            |

### C. Per scene and per bitrate results

As expected, the *bitrate* effect was found as significant for all but one tested conditions. Therefore, the results per scene and per bitrate are further discussed, in order to see if there is a particular pattern, from which we could draw some further conclusions. The summary of the t-test performed on each pair of *Res\_freq* conditions (pairwise comparison) is displayed in Table III.

1) *3 Mbps*: Mauchly’s test revealed that the sphericity was not violated for the main effect of *Res\_freq* at 3 Mbps for any scene. The test of within-subject effects showed that the main effect of *Res\_freq* is significant for *Gasquet-Adv*,  $F(3, 75) = 4.35$ ,  $p = .007$ , and *Euro2016-2*,  $F(3, 75) = 4.28$ ,  $p = .008$  scenes. The pairwise comparisons for the main effects of *Res\_freq* at 3 Mbps showed that there is a significant difference ( $p < .05$ ) only between RF3 and RF4 conditions for *Euro2016-2* scene.

2) *6 Mbps*: At 6 Mbps, the sphericity for the main effect of *Res\_freq* was not violated for any scene. The test of within-subject effects showed that the main effect of *Res\_freq* was found as significant for all scenes at  $p < .05$ . The pairwise comparisons for the main effects of *Res\_freq* showed that there is a significant difference ( $p < .05$ ) between RF1 and RF4 and RF2 and RF4 pairs for the *Bicycle* scene, between RF2 and RF3 conditions for the *Gasquet-Adv* and *Rugby-Mele* scenes, and between RF1 and RF2, RF2 and RF3, RF3 and RF4 pairs for the *Euro2016-2* scene.

3) *12 Mbps*: Testing for the sphericity at 12 Mbps, Mauchly’s test revealed the violation for the *Bicycle*,  $\chi^2(5) = 13.32$ ,  $p = .021$ . The test of within-subject effects showed a significant effect of *Res\_freq* for *Gasquet-Adv*,  $F(3, 75) = 14.26$ ,  $p < .001$ , *Rugby*,  $F(3, 75) = 16.52$ ,  $p < .001$ , and *Euro2016-2*,  $F(3, 75) = 11.55$ ,  $p < .001$  scenes. The pairwise comparisons for the main effects of *Res\_freq* showed that there are significant differences ( $p < .05$ ) between RF4 and all other cases for the *Gasquet-Adv*, between RF1 and RF2, RF1 and RF4, RF2 and RF3, and between RF3 and RF4 for the *Rugby* scene. For the *Euro2016-2* scene, RF2 condition was found as significantly better than any other condition.

4) *24 Mbps*: At 24 Mbps, the sphericity for the main effect of *Res\_freq* was not violated for any scene. The test of within-subject effects showed that the main effect of *Res\_freq* is significant for all the scenes at  $p < .05$ . The pairwise comparison showed that there is a significant difference between the RF1 and RF4 conditions for *Bicycle*. At the same time, for the other three scenes, a significant difference was found for most of the conditions, see Table III. Furthermore, the same test showed that there is only a significant difference ( $p = .001$ ) between the DMOS of the RF4 and the hidden reference for the *Euro2016-2* scene, which has the highest spatial and temporal complexity (see Fig. 1). For the other three contents the QoE of UHD+HFR video at 24 Mbps is at the same level as the reference at 100 Mbps.

## IV. DISCUSSION AND CONCLUSIONS

While the current display technology is not capable of simultaneously reproducing all the features that will compose UHD-1 Phase 2, in this work we have evaluated the joint effect on QoE of high frame rate and 4K spatial resolution. Our initial results show that there is a significant effect of the content, confirming previous findings [4] on HFR and UHD video evaluated individually. As expected, there was a significant preference, for some contents, of RF4 over all other, and RF2 over RF1 conditions. Also, this preference depends on the bitrate at which the sequences are coded. Not surprisingly, only for sufficiently high bitrate (higher than 12 Mbps) the UHD and HFR functionalities together provide significantly better visual quality than the other modalities, at least for two contents (see Table III). The results of the comparison between the RF3 and RF4 conditions showed that for three fast-paced scenes there is a significant gain by adding HFR, and that for most of the content the quality delivered at 24 Mbps is indistinguishable from the reference at 100 Mbps. This was not the case only for the *Euro2016-2* scene that had the highest SI and TI values, and so the highest coding complexity (see Fig. 1). On the other hand, we observed that for 3Mbps there is no significant gain by adding UHD and HFR, except for one case with *Euro2016-2* scene. Conversely, at 6Mbps, for the *Bicycle* scene, that was the slowest-paced content, both the HD (RF1 and RF2) conditions were found as higher quality compared to the UHD conditions. This is not surprising, since low bitrates are not sufficient for UHD-HFR video and compression artifacts become predominant.



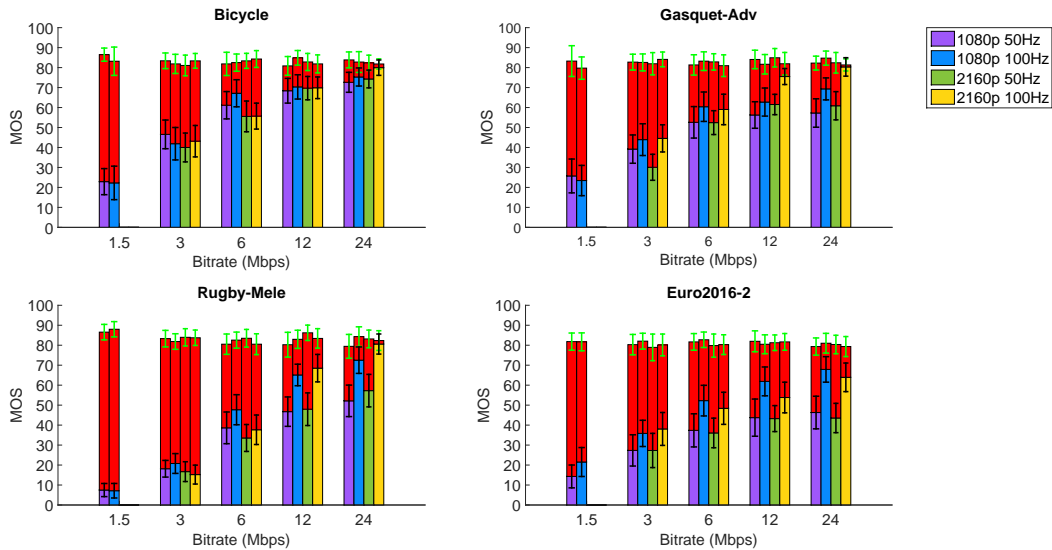


Fig. 3: MOS values for the test sequences per condition and for the reference sequences (red) with the 95% confidence intervals.

Finally, for fast-paced content at 6Mbps, frame rate is found as more important than resolution. However, we cannot claim, in general, that frame rate matters more than spatial resolution, on the basis of the results of this study. In fact, our HD stimuli were obtained by spatial decimation of UHD captured videos, and not by original acquisition at HD resolution. Ideally, future extensions of this work should rather employ stimuli directly acquired at different temporal/spatial resolutions.

To provide further analysis, we report in Fig. 3 MOS values on a 100-point scale, for both the reference and the tested conditions, as obtained through the DSCQS evaluation. From MOS values, we can observe that, indeed, at 3 Mbps all the tested modalities provide poor quality. Interestingly, for all the test sequences the average score for the reference oscillates between 80 and 90. This is partially due to the testing methodology, where the reference was hidden, and to the use of compressed reference, although at a very high bitrate (100 Mbps). However, it might also imply that 100 Hz on a 4K display viewed at 1.5 screen heights is not sufficient, confirming results of previous studies that suggest frame rates of 150 Hz or higher for these viewing conditions [11], [13].

Displays capable of reproducing UHD, HFR and HDR video are expected to appear on the market in the next few months. The next step of this research will be to evaluate the overall gain in QoE of UHD-1 Phase 2 at its full potential.

#### ACKNOWLEDGMENT

The work presented in this document was supported by Région Île de France, in the framework of the FUI 4EVER2 project.

#### REFERENCES

[1] M. Yu, H. Lakshman, and B. Girod, "Content adaptive representations of omnidirectional videos for cinematic virtual reality," in *Proceedings*

of the 3rd International Workshop on Immersive Media Experiences. ACM, 2015, pp. 1–6.

[2] T. Ebrahimi, S. Foessel, F. Pereira, and P. Schelkens, "JPEG pleno: Toward an efficient representation of visual reality," *IEEE MultiMedia*, vol. 23, no. 4, pp. 14–20, 2016.

[3] F. Dufaux, P. Le Callet, R. Mantiuk, and M. Mrak, *High Dynamic Range Video: From Acquisition, to Display and Applications*. Academic Press, 2016.

[4] ITU-R, "The present state of ultra-high definition television," ITU-R Recommendation BT.2246-6, March 2017.

[5] A. Gabriellini, "HFR and video compression," *EBU Tech-I Magazine*, vol. 19, no. 47, March 2014.

[6] M. Rerabek, P. Hanhart, P. Korshunov, and T. Ebrahimi, "Subjective and objective evaluation of hdr video compression," in *9th International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM)*, 2015.

[7] J. Laird, M. Rosen, J. Pelz, E. Montag, and S. Daly, "Spatio-velocity CSF as a function of retinal velocity using unstabilized stimuli," in *SPIE Electronic Imaging 2006*, 2006, pp. 605 705–605 705.

[8] P. G. Barten, *Contrast sensitivity of the human eye and its effects on image quality*. SPIE press, 1999, vol. 72.

[9] A. B. Watson and A. J. Ahumada, "The pyramid of visibility," *Electronic Imaging*, vol. 2016, no. 16, pp. 1–6, 2016.

[10] M. Emoto, Y. Kusakabe, and M. Sugawara, "High-frame-rate motion picture quality and its independence of viewing distance," *Journal of Display Technology*, vol. 10, no. 8, pp. 635–641, 2014.

[11] Y. Kuroki, T. Nishi, S. Kobayashi, H. Oyaizu, and S. Yoshimura, "A psychophysical study of improvements in motion-image quality by using high frame rates," *Journal of the Society for Information Display*, vol. 15, no. 1, pp. 61–68, 2007. [Online]. Available: <http://dx.doi.org/10.1889/1.2451560>

[12] R. Salmon, T. Borer, M. Pindoria, M. Price, and A. Sheikh, "Higher frame rates for television," *SMPTE Motion Imaging Journal*, vol. 123, no. 4, pp. 50–55, May 2014.

[13] K. Noland, "The application of sampling theory to television frame rate requirements," *BBC Research & Development White Paper*, vol. 282, 2014.

[14] G. J. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Transactions on circuits and systems for video technology*, vol. 22, no. 12, pp. 1649–1668, 2012.

[15] ITU-R, "Methodology for the subjective assessment of the quality of television pictures," ITU-R Recommendation BT.500, January 2012.