Virtual reality applications are on the rise and touching numerous domains, including healthcare, training, and gaming, to name a few. Existing experiences, however, are not fully immersive, as only two senses (audio-visual) are stimulated. To overcome this limitation, olfactory and haptic devices are emerging, thus making multisensory immersive experiences a reality. To date, however, little is known about the impact that each stimulated sense has on the overall experience, as well as on the user’s sense of e.g., realism, immersion, and engagement. In this pilot, we aim to answer this question. Using a multisensory pod, sixteen participants were immersed in a 2.5-minute virtual world, where smells, vibroacoustic, and somatosensory stimuli (i.e., wind and heat) were presented, in addition to 360-degree video and surround sound. Using two wearable devices, we kept track of the user’s heart rate, breathing rate, skin temperature, blood volume pulse, and electrodermal activity while they were immersed. In this paper, we report the impact that stimulating different senses had on the users’ overall experience, sense of presence, immersion, realism, flow, cybersickness, and emotional states, both subjectively, as well as objectively using features extracted from the wearable devices. Additional Key Words and Phrases: Virtual reality, quality of experience, multisensory, olfaction, haptics.

1 INTRODUCTION

Technological advances in computer graphics hardware, communications (e.g., 5G wireless networks), and immersive media software have caused a re-emergence of extended reality (xR) applications. Users can now experience virtual reality (VR) fully untethered over wireless networks, thus increasing their sense of presence. In fact, recent reports suggest that the xR field was hardly affected by the COVID-19 pandemic and approximately 30% compounded growth rates are projected for the next few years [7]. Existing applications, however, still rely on stimulating just two senses: vision and audition. As such, we have yet to provide users with fully immersive experiences, akin to those promised by the so-called “metaverse.”

To overcome this limitation, several innovations are emerging in scent diffusion and in haptic devices, many of which can be fully integrated within the VR headsets (e.g., the OVR Technology’s ION2 scent device) or can be easily paired with xR devices (e.g., the TESLASUIT haptic suit). While the ultimate goal is to increase the user’s experience, limited knowledge exists on the

Both authors contributed equally to this research.

Published in accordance with the terms of the Creative Commons Attribution 4.0 International Public License (CC BY 4.0). Permission to reproduce or distribute this work, in part or in whole, verbatim, adapted, or remixed, is granted without fee, provided that the appropriate credits are given to the original work, not implying any endorsement by the authors or by SBC.

ACM IMX’22, June 22–24, 2022, Aveiro, Portugal
© 2022 Brazilian Computing Society.
impact of stimulating different senses, both qualitatively and quantitatively, not only in terms of overall experience, but also in terms of e.g., sense of presence, immersion, and user engagement. As highlighted by Perkis et al. (2020), these human influential factors are extremely crucial to assure acceptable user experience levels [13]. Ultimately, it is the experience that immersive technologies provide to the users that will dictate the success or failure of such technology [1]. As such, quantifying the impact of multisensory experiences is of crucial importance.

Vast literature exists on so-called multi-sensorial media (or mulsemedia), where olfactory and somatosensory stimuli have been added to on-screen media and their impact on overall quality of experience (QoE) quantified, both subjectively (e.g., [5]) and objectively via eye tracking and heart rate measurements (e.g., [8]). Recently, experiments have shown that when multiple senses are stimulated, one may reduce the quality of the VR video without serious detrimental effects to overall QoE [4]. Moreover, adding smells to VR content showed to significantly improve sense of presence, while addition of wind showed no significant improvements [12]. To the best of the authors’ knowledge, however, no study exists that quantifies the impact of stimulating multiple senses (in VR) on different human influential factors (HIFs), which have been shown to be crucial for immersive media experiences [13]. These HIFs include, but are not limited to: sense of realism, presence, immersion, flow, engagement, and emotion.

In this paper, we aim to fill this gap. We describe a pilot experiment in which participants were exposed to immersive media content within a multisensory pod, where smells, vibroacoustic stimuli, wind, and heat were presented in synchrony with 360-degree VR video and sound. Participants wore wearable devices to monitor several biometric signals and reported their HIF ratings across four conditions: video only, audio-visual, audio-visual-smells, and audio-visual-smells-haptic. We describe the differences observed in the subjective ratings, as well as in the physiological measures and provide insights on which modalities contribute the most towards overall QoE.

2 MATERIALS AND METHODS

Sixteen (6 females) healthy participants (27 ± 7.46 years) consented to participate in this pilot study that had Ethical approval from the authors’ institution. A SENSIKS multisensory pod (SENSIKS, Netherlands) was used (see Fig. 1a). The pod provides surround sound, low-frequency acoustic vibrations via a subwoofer placed under the seat, four fans to simulate wind, several heating elements surrounding the sides, back, and front of the user to simulate heat, scent diffusion where up to six different aromas can be diffused, and 360° video was presented to the participants via an Oculus Quest VR headset. Time synchronization of all sensory elements was performed via software. The 2.5-minute virtual environment was comprised of two scenarios. The first was based on a fictional outer space experience (Fig. 1b), followed by a forest environment (Fig. 1c), both with a background female voice promoting a reflective existential message.

The experiment was conducted under four conditions/sessions, namely: video only (VO), audio-visual (AV), audio-visual-smells (AVS), and audio-visual-smells-haptics (AVSH). As the condition names suggest, each condition had varying number and types of sensory stimuli. In the AVSH condition, the vibroacoustic, heater, and wind simulations were enabled. The ordering of the conditions was counterbalanced across participants to avoid any ordering biases. Participants also wore a BioHarness3 chest-strap (Zephyr, USA), which collected electrocardiogram at a 250 Hz sample rate and breathing curves at 18 Hz sample rate. Moreover, an Empatica E4 wristband (Empatica, USA) was used to measure skin temperature, electrodermal activity, and blood volume pulse signals. From these signals, several conventional time domain, frequency domain, and statistical descriptors features were extracted from the five signal modalities, such as heart rate, breathing rate, electrodermal reactions, to name a few. Table 1 shows a summary of all the extracted features from each signal modality. At the end of each session, participants rated their perceived levels of
Fig. 1. (a) Multisensory pod showing A-fans, B-heating elements, and C-scent diffuser, (b) visual of the outerspace and (c) forest experiences.

Fig. 2. Questions and scale used for the subjective rating survey.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – Did the experience seem realistic to you?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>2 – Did you have the sense of “being there” in the virtual environment?</td>
<td></td>
</tr>
<tr>
<td>3 – Did you feel immersed in the virtual environment?</td>
<td></td>
</tr>
<tr>
<td>4 – Were you involved with the virtual experience?</td>
<td></td>
</tr>
<tr>
<td>5 – Did you lose the sense of time while experiencing the environment?</td>
<td></td>
</tr>
<tr>
<td>6 – Did you enjoy being in the virtual environment?</td>
<td></td>
</tr>
<tr>
<td>7 - Did you feel disoriented and/or nauseous during the interaction with the virtual environment?</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2. Questions and scale used for the subjective rating survey.

realism, presence, immersion, engagement, flow, emotion, and cybersickness using 5-point scales in a customized questionnaire similar to the one displayed on Figure 2.
Table 1. Extracted features by each signal modality.

<table>
<thead>
<tr>
<th>Device</th>
<th>Modality</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bioharness 3</td>
<td>Breathing curve</td>
<td>mean, standard deviation, range, skewness, kurtosis, mean of 1st diff, spectral energy from 0-0.2 Hz, 0.2-0.4 Hz, 0.4-0.6 Hz, 0.6-0.8 Hz, and 0.8-1.0 Hz, spectral energy ratio between 0.05 - 0.25 Hz and 0.25-0.5 Hz, the frequency with maximum spectral energy, centroids.</td>
</tr>
<tr>
<td></td>
<td>ECG</td>
<td>mean RR, standard deviation RR, coefficient of variation, RMSDD, pNN50, mean of 1st diff, mean of absolute 1st diff normalized, standard deviation of absolute 1st diff, low-frequency power (0.04-0.15 Hz), high-frequency power (0.15-0.4 Hz), very-low-frequency power (0-0.04 Hz), normalized low-frequency power, normalized high-frequency power, the ratio of low-frequency power and high-frequency power, total power.</td>
</tr>
<tr>
<td>Empatica E4</td>
<td>BVP</td>
<td>maximum, minimum, spectral energy from 0-0.5 Hz, 0.5-1.0 Hz, 1.0-1.5 Hz, 1.5-2.0 Hz, and 2.0-2.5 Hz, spectral energy ratio between 0.04-0.15 Hz and 0.15-0.5 Hz.</td>
</tr>
<tr>
<td></td>
<td>EDA</td>
<td>mean, standard deviation, mean of 1st difference, mean of negative 1st difference, spectral energy from 0-0.2 Hz, 0.2-0.4 Hz, 0.4-0.6 Hz, 0.6-0.8 Hz, and 0.8-1.0 Hz from the phasic component of EDA signal.</td>
</tr>
<tr>
<td></td>
<td>Skin temperature</td>
<td>mean, standard deviation, range, mean of 1° difference, minimum, maximum, skewness, kurtosis, delta temperature, spectral energy from 0-0.01 Hz, and 0.1-0.2 Hz.</td>
</tr>
</tbody>
</table>

3 RESULTS AND DISCUSSION

Figure 3 depicts the average ratings obtained across all participants for each of the seven influential factors. One-way repeated-measures ANOVA tests were performed to compare the significant differences obtained for each HIF across the four test conditions. With the exception of cybersickness, stimulating additional senses (conditions AVS and AVSH) showed significant improvements across the six remaining HIFs, relative to the VO and/or AV conditions, namely: realism $F(3, 45) = 7.67, p < 0.001$; presence $F(3, 45) = 12.59, p < 0.001$; immersion $F(3, 45) = 12.13, p < 0.001$; engagement $F(3, 45) = 8.07, p < 0.001$; flow $F(3, 45) = 6.6, p < 0.001$; emotion $F(3, 45) = 9.45, p < 0.001$; cybersickness $F(3, 45) = 0.96, p = 0.42$. Next, post-hoc multiple comparisons between groups tests were performed for each HIF and three different significance levels were explored: 95%, 99%, and 99.9%; these are represented by one, two, or three asterisks, respectively, in Figure 3.
As can be seen, multisensory experiences showed significant improvements across all HIFs, except cybersickness. Introducing smells showed the greatest improvements in realism, presence, immersion and emotion. Introducing haptic feedback, in addition to smells, brought further increases (though not significantly) to sense of realism, presence, and flow, but showed minimal impact on immersion, engagement, and emotion. Aside from cybersickness, stimulating all four senses showed limited benefits in engagement and flow ratings. This could be explained by the experience itself, as it was passive, thus users had little engagement with the environment, thus not affecting flow. Overall, stimulating four senses seems to be crucial to improve the sense of realism, presence and immersion, three of the most important factors driving immersive media QoE [13]. Lastly, while the environment itself was not prone to induce cybersickness, the negligible effects seen by introducing smells was promising and corroborated those reported by Narciso et al. [12].

Finally, we explored the differences seen in the biometric features per condition and across conditions. The chest-strap data from three participants had to be excluded from this analysis due to poor quality recordings. Moreover, the wristband data from one participant had to excluded as issues with the device battery resulted in an incomplete session recording. With these data removed from the analysis, the following features showed significant ($p < 0.05$) Pearson correlations with realism (low frequency component of BVP signal, $r=-0.33$; mean RR interval, $r=-0.42$), presence (mean RR interval, $r=-0.67$ for condition C and -0.78 for condition D), immersion (high frequency power of EDA signal, $r=-0.31$), engagement (low frequency to high frequency ratio measure of HRV, $r=0.37$), flow (low frequency power of breathing curve, $r=-0.48$) and emotion (low frequency to high frequency ratio measure of HRV, $r=0.33$). The scatter plots in Figure 4 exhibit these behaviours. These findings suggest that wearables can potentially be used to provide quantitative insights on perceived QoE. In the near future, they could be used to adjust multisensory environments in real-time to maximize QoE per user [10].
Future Research Directions

As for future research directions, work will concentrate on exploring additional biological signal modalities to improve and/or optimize the determination of the HIFs described herein. For example, via an instrumented VR headset, as the one described by Cassani et al. (2020) [2] and depicted in Figure 5, additional modalities may become available, including electroencephalography (EEG),
Fig. 5. Instrumented VR headset with (a) 16 dry ExG electrodes and (b) a wireless, portable bioamplifier.

electro-oculography (EOG) and facial electromyography (EMG). Previous research has shown that these modalities can be useful to monitor, for example, saccadic eye movements from EOG [9] and their relationship to sense of presence in virtual reality [10], as well as indices related to mental workload, enjoyment, engagement and cybersickness from EEG (e.g., [6, 11, 14]) and detect facial gestures from the EMG signals (e.g., [3]).

5 CONCLUSION

In this paper, we have explored the effects of stimulating different senses in a multisensory immersive experience. Significant improvements in perception of realism, sense presence and immersion, engagement, flow, and emotion were seen once smells and haptic feedback was introduced, with no negative effects in cybersickness. Several biometric features also showed differences across conditions, thus suggesting that, in the future, objective measures of multisensory immersive media QoE could be achieved in real-time with the help of wearables. As instrumented VR headsets become available, other modalities, such as electroencephalography and eye gaze, may provide additional insights into the overall user experience.

ACKNOWLEDGMENTS

The authors wish to acknowledge Dr. Raymundo Cassani and Dr. Abhishek Tiwari for sharing codes to extract the biometric features, as well as the Natural Sciences and Engineering Research Council of Canada for funding this research.

REFERENCES


