The Influence of Audio Sensory Input on the Anxiety Level of the VR Users

Mahdiyeh sadat Moosavi
Arts et Metiers Institute of Technology
LISPEN, HESAM Université
F-71100 Chalon-Sur-Saone, France
mahdiyesadat.moosavi@ensam.eu

Frédéric MERIENNE
Arts et Metiers Institute of Technology
LISPEN, HESAM Université
F-71100 Chalon-Sur-Saone, France
frederic.merienne@ensam.eu

Suziah Sulaiman
Computer Information Sciences
Department, Universiti Teknologi
PETRONAS
Seri Iskandar, Perak, Malaysia
suziahs@utp.edu.my

Marcus Yung Chong
Electrical and Electronics Engineering
Department
Universiti Teknologi Petronas
Seri Iskandar, Perak, Malaysia
marcus_18002304@utp.edu.my

Christophe Guillet
Université de Bourgogne
LISPEN, UBFC
F-71100 Chalon-Sur-Saone
Christophe.Guillet@u-bourgogne.fr

Ainul Khairiyah Mohd Nazri
Electrical and Electronics Engineering
Department
Universiti Teknologi Petronas
Seri Iskandar, Perak, Malaysia
ainul_17008747@utp.edu.my

Dayang Rohaya Awang Rambli
Computer Information Sciences
Department, Universiti Teknologi
PETRONAS
Seri Iskandar, Perak, Malaysia
dayangrohaya.ar@utp.edu.my

***Abstract*— This study investigates subjects' behavior in visual audio conflict conditions in stressful and stressless environments. To increase subjects' stress and anxiety in subjects we placed subjects in a chaotic city and to decrease subjects' stress we benefit from a green virtual environment. On the other hand to study the effect of sound on physiological data we designed confilict condition in which in the City environment participants heared the sound of Forest and in the Forest nvironment they heared the sound of City. To evaluate our findings we benefit from the heart rate of the participants which is directly related to subjects' emotions and participants can not involve their opinion to change this feature autonomously.**

Keywords—Anxiety, Stress, visual-audio feedback, multisensory conflict, physiological data

I. INTRODUCTION

A. Applicable method for Treating Stress and Anxiety

Stress and anxiety have a tight connection and they have overlapping neural substrates of psychological states as well as intertwined behavioral and neural structures [1], [2], [3]. Dysfunctionality and hypersensitivities of the stress system are associated with anxiety disorders [3]. Due to the overlapping structure between stress and anxiety, placing in a stressful situation can intensify anxiety [4]. It has also been shown that anxiety and stress have an impact on individuals' physiological responses such as heart rate and skin conductance and blood pressure [5].

Virtual reality(VR) is one of the applicable and effective treatment approaches for conquering not only mental and cognitive disorders such as posttraumatic phobias [6], but also anxiety, and stress levels [5]. Furthermore, plenty of studies prove that virtual reality can be an effective medium. It can affect human emotion, which means that it can be designed differently to induce and elicit a certain type of emotion [7]; [8].

Different researches have shown that green environments like forests and nature alleviated tension and stress relief [9] [10]. And humans prefer a natural environment compared to an urban one. Naturalness, lightning, and weather type in a virtual environment have significant effects on users' emotional states. For instance, humans have a significant preference to stay in a light scene compared to a dark scene, as well as a sunny scene compared to an overcast scene [11]. [12] also showed that living in a city environment will negatively affect the brain responses to stress compared to towns and rural, and it makes people feel stressed easily [13] [14], experimented with a sunny forest as the calm environment and a roller coaster ride as the dynamic environment for a comparative study of mental states in 2D and 3D virtual environments using electroencephalogram (EEG). In addition, a scene with warm and bright light, peaceful audio, and natural environments are the most helpful in calming participants [15]

Providing an immersive and realistic environment helps to have more accurate sensory feedback. An immersive virtual environment (VE) conveys to participants that they are present there. Measuring the "sense of presence" of the VR participants is previously assessed by questionnaires. However most of the previous related works to the stress domain also accessed and obtained the result by purely using subjective questionnaires and surveys such as STAI survey, and sense of presence questionnaires, instead of using physiological data [8] [11]. Using subjective measurements is not an accurate and suitable method, because the subjective survey can be different and inconsistent for the participants. In addition, to fill the questionnaire participants should transform their body experience into a verbal judgment. Moreover, participants have to leave the environment and try to remember what they perceived during their presence in the VR environment.

On the other hand, the immersion degree can not only be differentiated by subjective feelings but can also be differentiated by physiological signals [16]. Participants' physiological characteristics (including skin conductress, heart rate, and blood pressure) state has a direct relation to

their degree of immersion in the VE. These measurements are independent of subjects' opinions and They cannot be controlled intentionally [5].

Previous research utilized both physiological [17]data and subjective questionnaires to access and gather the sense of the presence of the participants, however, it only gathered the skin electrodermal activity (EDA) for physiological data, and the experiment result shows poor relationships between EDA and self-reported presence from participants. (Meehan et al., 2002), showed that EDA might not be a good indicator of presence. Furthermore, a hypothesis from [18] explained that heart rate might be a better method to access the emotion of the participants. Hence, it is crucial to gather more physiological data such as heart rate, blood pressure, EDA, and so on, as well as subjective questionnaires and surveys to obtain the most accurate result.

B. Audiovisual Integration

Similar to the real world, in a virtual environment human beings perceive information through their sensory modalities such as vision, audio, and haptic. The notion of multisensory integration refers to the coherent and relevant interaction of the multi-sensory stimuli. Among all research on audio-visual integration, some studies indicated that audio-visual integration can lead to more intense emotional responses [19].

One of the most important questions in VR was always about, which sensory system is more dominant. To answer this question various studies focused on the impact of sound and audio on different fields of virtual reality such as mass and material perception, [7]or visual perception [20] or game experience [21]

It is a common belief that vision is the dominant sensory input system, however, in multisensory conflict conditions, it can be manipulated by other sensory systems such as audio [20] [22].

Our research is trying to study the behavior of the subject in a stressful and calm environment and analyze the sense of presence for the optimization of virtual immersion. Besides that this research wants to study the anxiety response of VR users to audio-visual stimuli and their reaction to visual-sound conflict conditions. To achieve these goals, our research focused more on physiological data such as heart rate

Thus, thanks to all the research available and after all the literature review to achieve our goal, we designed and built two emotional and affective immersive virtual environments, which are a chaotic City and a calm Forest. The city is disturbing and can induce the desired stressful and anxious emotion among participants; while the forest is calming and can induce the desired calmness in our participants

To induce a better sense of presence in the subjects in each environment, we tried to use two different human sensory inputs including visual, and audio elements.

To evaluate different subject behaviors, we benefit from not only subjects' physiological characteristics, but also the different questionnaires.

In this research we are trying to follow the following hypothesizes:

Study individual behaviors in a VR environment by Manipulating the sound of the environment and try to investigate subjects' reactions to the visual and audio stimuli.

II. METHODOLOGY

A. Subjects

15 healthy participants (6 Women and 9 men ranging in age from 19 to37 with an average of 26 ±5.44) participated in the experiment. All of them had normal vision and never presented with neurological or cognitive disorders. Subjects' heights ranged from 156 to 195 cm, and the mean height was 174.2±11.64 cm. Subjects from different backgrounds, either from inside or outside the university, agreed to participate voluntarily and without any compensation in the experiment.

B. Experimental Device and Environments

The VR environments (city and forest) were designed in Unity 3D software and it was carried out in a quiet room and participants were comfortably seated on a chair. We used the HTC Vive Pro headset, tracked by an HTC Vive camera. The headset's resolution was 1440*1600 pixels per eye with 110 degrees of field-of-view. The device was equipped with, a g-sensor, gyroscope, proximity, and eye comfort setting (IPD). Besides that, we utilized Empatica E4 bracelets for measuring physiological data. To induce or reduce subjects' anxiety we designed different environments such as cities and forests. The city environment is filled with stressful elements such as loud and chaotic city sound coming from cars, ambulance, rain, and also from the barking of a fierce dog. Moreover, the buildings are on fire, heavy rain, and a lot of running pedestrians makes the city looks very crowded imitating a very intense and chaotic city environment. The lighting of the city is set to be a moody and dim environment which let the participants experience nightlife in a chaotic city environment (see Fig.1).

On the other hand, the forest environment is designed with a calm forest element including the sound and the visuals of its surrounding. The environment is filled with the sound of the forest, lake, and also wind sound. Moreover, with the presence of the flying birds, participants will be able to hear the sound of flying birds around them in various directions. Besides, it has a hut with a lake view, another hut with a forest view, and also a furnished cabin where all of the elements are connected with a pathway. The lighting of the forest is designed to be a bright environment which lets participants experience a relaxing sunny day in the forest (see Fig. 2).

To make sure that each participant spends a fixed time in each environment and passes the same road and faces the same happening in the virtual scenes we developed an automatic navigation system using an animator system instead of using teleportation or other navigation systems. This system allows us to study the fixed effect of the environments on the subjects.

C. Experimental Protocol

Each subject was presented with a set of reference conditions followed by a sound conflict condition. Subjects spent 3.5 minutes per experimental phase. They rest about 2

Fig. 1. City environment: This environment is loaded with stressful elements like burning buildings, angry dogs, running people, dark smoky rainy weather, and traffic jams.

Fig. 2. Forest environment: In this environment green elements like trees and plants and other elements like water, mountains, etc, were used to induce calmness in the participants.

minutes before entering the other environment. Our experiment is divided into 2 different phases including reference conditions and conflict conditions.

Reference Condition

In this condition, participants did not perceive any conflict between the visual elements and audio In the city and forest scenes. When participants presented in this condition they perceived the visual elements like fire in the city, barking dogs, and running people with their relative sounds, in the same way, in the forest environment when they observed lake and birds they heard the relative sound of water and sound of the birds.

Conflict condition

To investigate subjects' physiological reactions to the audio and also to test the subject's sensitivity to the audio elements in the virtual environment, we designed audio conflict conditions where the sounds of the two environments were interchanged. In this condition, while subjects perceived the visual elements of the chaotic city environment they heard the audio of the forest scene and vice versa.

Pre-experiment condition

Before starting the experiment, we asked subjects to sit comfortably on a chair for 10 minutes in a calm environment. After that, we recorded the pre-experiment physiological data of the participants before entering the experiment. During this condition and recording the pre-experimental data we asked subjects to avoid performing sudden acts and movements.

III. RESULTS

In the following section, we will provide a detailed report of our findings by analyzing the heart rate (HR) data of the participants in different environments and conditions.

In an attempt to compare the different influences of environmental element effects (such as sound and visual) on the subject's HR, we plotted the HR mean of all subjects using a bar plot and also the individual's HR (dots) to monitor the subject's HR alternation due to different conditions.

Fig.3 shows the mean value of 12 subjects' heart rates that participated under different experimental conditions. In the reference condition, the mean value of HR in the city environment is (equal to 72.08) higher than the pre-experiment condition (equal to 69.7), while in the forest environment the mean value of HR of 12 participants decreased (equal to 68.6). On the other hand in the conflict condition, changing the audio of the experiment caused to change in the mean value of HR for 12 participants in the environments (see Fig.3 - Conflict). Compared to reference conditions, in the sound conflict, during the city environment, the subject's HR decreased (mean =70.57), while in the forest environment increased(71.53). However, the one-way repeated measure ANOVA test did not indicate any statistical differences between city and forest environments during different conditions (p-value > 0.05).

As Fig. 3 illustrates, some subjects' behaviors were modified during the conflict condition. Nonetheless, to monitor these changes in the individual's behavior we plotted the HR of each participant.

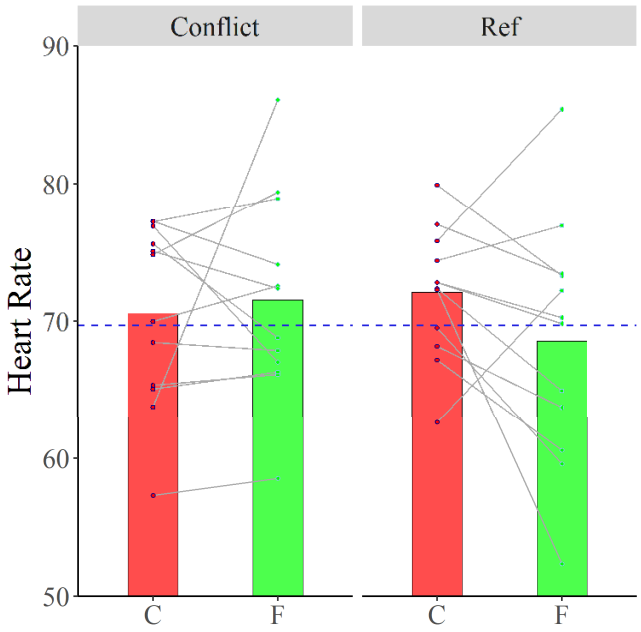

Fig. 3. Individual's and mean value of HR in the city (C) and forest (F) environment under experimental conditions. The gray line shows the trend of subjects' behavior in city and forest environments.

Fig. 4 shows typical HR trends produced by one random subject in different environments under experimental conditions. The highest trend for HR is related to the city environment in the Ref condition, while the lowest trend corresponds to the forest environment in the Ref condition. On the other hand, the effect of sound is significantly visible in the city environment during the conflict condition. However, the visual elements in the city environment caused to increase in the HR trend but the effect of sound is eye-catching in this condition compared to the city environment in the Ref condition. Similarly, in the forest environment under conflict condition sound effects leads to incrementing the HR trend.

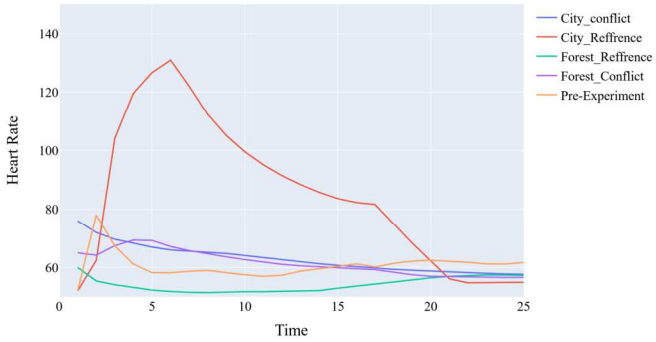

Fig. 4. The HR trend of one typical subject over time during different experimental conditions.

To investigate the influence of the sound on each subject's behavior Fig. 5 demonstrates an increasing or decreasing trend of HR means value during Ref or Conflict conditions. As Fig. 5 shows 58 % of the subjects in conflict conditions experienced higher HR values in the forest compared to the city, while in the Ref condition they experienced lower HR in the forest compared to the city. On the other hand, 48% of the subjects kept their trend and didn't react to the sound.

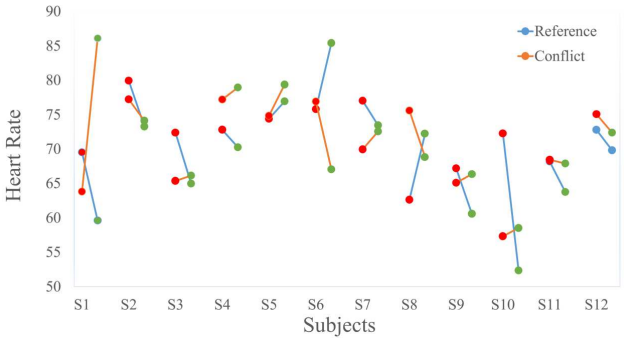

Fig. 5. Individual's HR behavior in the city (red dots) and forest (green dots). Blue lines indicate subjects' HR modification during reference condition and the red line represents HR transformation during sound conflict condition.

IV. Discussion

Due to the fact that virtual reality provides the opportunity to separate its users from the real world and transfer them to a virtual one, enhancing user experience in the virtual world has special importance. The main purpose of this research conducted to study subjects' behavior under multi-sensory conflict conditions to understand which sensory input is more dominant. Moreover, we were interested in pursuing the effect of environmental elements on subjects' stress and anxiety level.

There is a common belief that in a multi-sensory conflict condition visual sensory input is a dominant sensory system. However, some studies [22] [23]) showed that audio input can manipulate the visual sensory system. In this study, we compared the relative effect of sound on individuals' stress and anxiety level by means of the heart rate (HR) of the participants.

Previous studies showed that green environmental elements like the forest can reduce tension and anxiety [10], while [12] showed that the urban environment elevates human anxiety and stress levels. These findings are fundamental to our research. Therefore we tried to induce stress by means of urban environmental elements or decrease it by rural elements. To investigate the influence of the audio in the conflict condition we changed the audio of the city environment with the forest environment and vice versa.

To evaluate our findings we used HR as an objective measurement to quantify our results.

Although the result of the group analysis did not show statistical differences between different environments under both conditions, the trend of elevating HR in the city and demoting HR in the forest environment is attracting attention (see Fig. 3 – Ref). One of the reasons that we didn't observe significant differences between different environments could be due to the lack of participants.

However, the effect of the sound is significantly visible in subjects' HR. Fig.5 demonstrated that in conflict conditions subjects react differently to the sound modifications and 42% of them kept their previous behavior under the sound conflict condition while the rest of them were influenced by auditory manipulation. This finding shows that some subjects are more visual dependent in the multi-sensory conflict condition and mostly they rely on their visual input while some of them rely more on their auditory sensory system.

Our results are important findings and could be helpful for VR developers who want to increase the degree of immersion

in VR users. By profiling VR users into audio-dependent and visual-dependent groups, they can develop environments and focus on one sensory modality instead of two modalities and gain better results in elevating subjects' sense of presence.

V. Conclusion

There is a common belief that visual sensory input is dominant during a conflict condition. However, this study showed that during a conflict condition different subjects have a different reactions in the virtual environment. Some individuals are more sensitive to visual stimuli while others are more susceptible to audio. In better words, some people are more visual dependent while some others are more audio dependent. This finding is important for VR developers who are working on immersion topics. These results could help them to build and enhance VR environments according to the participants' sensory input dependency such that subjects feel a better sense of presence in the VR environments. On the other hand the manipulation visual input by audio is an intresing point which could be used designing illusory VR environment.

Acknowledgment

This work was supported by French government funding managed by the National Research Agency under the Investments for the Future program (PIA) grant ANR-21-ESRE-0030 (CONTINUUM).

References

[1] N. Daviu, T. Füzesi, D. G. Rosenegger, N. P. Rasiah, T.-L. Sterley, G. Peringod and J. S. Bains, "Paraventricular nucleus CRH neurons encode stress controllability and regulate defensive behavior selection," *Nature Neuroscience,* vol. 23, p. 398–410, 2020.

[2] G. P. Chrousos, "Stress and disorders of the stress system," *Nature reviews endocrinology,* vol. 5, p. 374–381, 2009.

[3] S. Seinfeld, I. Bergstrom, A. Pomes, J. Arroyo-Palacios, F. Vico, M. Slater and M. V. Sanchez-Vives, "Influence of music on anxiety induced by fear of heights in virtual reality," *Frontiers in psychology,* vol. 6, p. 1969, 2016.

[4] C. Grillon, R. Duncko, M. F. Covington, L. Kopperman and M. A. Kling, "Acute stress potentiates anxiety in humans," *Biological psychiatry,* vol. 62, p. 1183–1186, 2007.

[5] S. Min, J.-g. Moon, C.-H. Cho and G. J. Kim, "Effects of immersive virtual reality content type to mindfulness and physiological parameters," in *26th ACM Symposium on Virtual Reality Software and Technology*, 2020.

[6] J. L. Maples-Keller, C. Yasinski, N. Manjin and B. O. Rothbaum, "Virtual reality-enhanced extinction of phobias and post-traumatic stress," *Neurotherapeutics,* vol. 14, p. 554–563, 2017.

[7] R. W. Fleming, "Visual perception of materials and their properties," *Vision research,* vol. 94, p. 62–75, 2014.

[8] G. Riva, F. Mantovani, C. S. Capideville, A. Preziosa, F. Morganti, D. Villani, A. Gaggioli, C. Botella and M. Alcañiz, "Affective interactions using virtual reality: the link between presence and emotions," *Cyberpsychology & behavior,* vol. 10, p. 45–56, 2007.

[9] P. Grahn and U. A. Stigsdotter, "Landscape planning and stress," *Urban forestry & urban greening,* vol. 2, p. 1–18, 2003.

[10] M. Annerstedt, P. Jönsson, M. Wallergård, G. Johansson, B. Karlson, P. Grahn, Å. M. Hansen and P. Währborg, "Inducing physiological stress recovery with sounds of nature in a virtual reality forest—Results from a pilot study," *Physiology & behavior,* vol. 118, p. 240–250, 2013.

[11] F. Beute and Y. A. W. de Kort, "Let the sun shine! Measuring explicit and implicit preference for environments differing in naturalness, weather type and brightness," *Journal of Environmental Psychology,* vol. 36, p. 162–178, 2013.

[12] F. Lederbogen, P. Kirsch, L. Haddad, F. Streit, H. Tost, P. Schuch, S. Wüst, J. C. Pruessner, M. Rietschel, M. Deuschle and others, "City living and urban upbringing affect neural social stress processing in humans," *Nature,* vol. 474, p. 498–501, 2011.

[13] J. S. Yeung, Y. D. Wong and H. Xu, "Driver perspectives of open and tunnel expressways," *Journal of Environmental Psychology,* vol. 36, p. 248–256, 2013.

[14] P. Bilgin, K. Agres, N. Robinson, A. A. P. Wai and C. Guan, "A comparative study of mental states in 2D and 3D virtual environments using EEG," in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2019.

[15] J. Tarrant, J. Viczko and H. Cope, "Virtual reality for anxiety reduction demonstrated by quantitative EEG: a pilot study," *Frontiers in psychology,* vol. 9, p. 1280, 2018.

[16] C. E. Izard, "Basic emotions, relations among emotions, and emotion-cognition relations.," 1992.

[17] M. Meehan, B. Insko, M. Whitton and F. P. Brooks Jr, "Physiological measures of presence in stressful virtual environments," *Acm transactions on graphics (tog),* vol. 21, p. 645–652, 2002.

[18] W. A. IJsselsteijn, Presence in depth, Technische Universiteit Eindhoven Eindhoven, The Netherlands, 2004.

[19] F. Pan, L. Zhang, Y. Ou and X. Zhang, "The audio-visual integration effect on music emotion: Behavioral and physiological evidence," *PloS one,* vol. 14, p. e0217040, 2019.

[20] D. A. Bowman, E. T. Davis, L. F. Hodges and A. N. Badre, "Maintaining spatial orientation during travel in an immersive virtual environment," *Presence,* vol. 8, p. 618–631, 1999.

[21] K. Rogers, G. Ribeiro, R. R. Wehbe, M. Weber and L. E. Nacke, "Vanishing importance: studying immersive effects of game audio perception on player experiences in virtual reality," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018.

[22] L. Shams, Y. Kamitani and S. Shimojo, "What you see is what you hear," *Nature,* vol. 408, p. 788–788, 2000.

[23] H. McGurk and J. MacDonald, "Hearing lips and seeing voices," *Nature,* vol. 264, p. 746–748, 1976.

[24] A. Felnhofer, O. D. Kothgassner, M. Schmidt, A.-K. Heinzle, L. Beutl, H. Hlavacs and I. Kryspin-Exner, "Is virtual reality emotionally arousing? Investigating five emotion inducing virtual park scenarios," *International journal of human-computer studies,* vol. 82, p. 48–56, 2015.