Increased Affect-Arousal in VR can be Detected from Faster Body Motion with Increased Heart Rate

Patrice Robitaille
École de technologie supérieure
Montreal, Canada
patrice.robitaille.1@ens.etsmtl.ca

Michael J. McGuffin
École de technologie supérieure
Montreal, Canada
michael.mcguffin@etsmtl.ca

ABSTRACT

We instrumented an immersive VR platform with physiological (heart rate and electrodermal activity) sensors to investigate the use of movement data and physiological data to automatically detect changes in affect (emotional state). 12 users were asked to complete four blocks of tasks requiring them to hit moving targets while standing and moving about. One of the four blocks (in counterbalanced order) was designed to be stressful (S), while the other blocks were designed to be calm (C). The motions required of the users were the same in both conditions; only the visual and audio feedback were different across the S and C conditions. Users' self-scored arousal in the S condition was significantly higher. We analyzed the recorded motions by segmenting out 2747 "fast motions", i.e., intervals of time where the sum of the speed of the hands was above a threshold. A simple machine learning algorithm (a decision tree) could learn to classify these fast motions as either calm or stressed, with ≈80% accuracy, using only two features: the maximum speed achieved during the motion, and the heart rate at the moment of maximum speed, where both features were normalized. If only the maximum speed feature is used (i.e., with no physiological data), ≈70% accuracy is achieved.

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**; Empirical studies in HCI; • **Computing methodologies** → *Motion capture*;

KEYWORDS

Virtual reality, affective computing, body movement analysis, motion capture, mocap, physiological response

ACM Reference Format:
Patrice Robitaille and Michael J. McGuffin. 2019. Increased Affect-Arousal in VR can be Detected from Faster Body Motion with Increased Heart Rate. In *Proceedings of ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games (I3D 2019).* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Virtual reality (VR) applications offer a wide variety of experiences, eliciting a range of emotional responses. One of the main uses of VR is the simulation of scenarios that are sometimes stressful, such as military battles, emergency response training, and games. When a user is under stress, they are less able to tell the system explicitly what to do through traditional interaction techniques.

Automatic detection of the user's affect (such as stress) could be useful for (1) triggering automatic offers of help, (2) throttling task difficulty in games or simulations to avoid user frustration, and (3) measuring user response, as a complement to other metrics (like task time, errors, subjective ratings) to evaluate VR scenarios.

Certain changes in the user's emotions might be detectable by merely analyzing the movements of head and hands, using information already available from standard VR hardware. However, there is little previous work on affect detection in immersive VR (i.e., using a head-mounted display), and it is unclear how much information can be gleaned from the motion of head and hands.

We had users undergo two experimental conditions in VR, one of which was designed to induce stress. Self-assessment questionnaires found significantly different levels of arousal between the two conditions. Arm motions were often faster in the "stressed" condition, and heart rate was often higher. A machine learning algorithm could correctly classify arm motions into one of the two conditions with 70% accuracy when only using information about arm speed. When also using heart rate information, accuracy was 80%. We report additional features computed from motion that were not useful for classification, to inform future designers.

2 BACKGROUND

2.1 Presence

VR experiences can be characterized by the degree of *presence* felt by the user [39], which can influence emotional state. A virtual environment that produces a heightened feeling of presence should be able to elicit emotional responses and behavioral changes [35]. Presence can be measured with subjective, behavioral and physiological methods [14, 38, 42]. Our experiment collected all three kinds of information.

Subjective measures rely on post-experiment questionnaires, such as the widely used Witmer and Singer questionnaire [41, 42], which we also used in our study. Kober and Neuper [22] review additional questionnaires.

Previous studies have found a strong relationship between user behavior and perceived level of presence [15, 27]. The more a participant feels present, the more their behavioral response to virtual stimuli resembles behavior in a real environment. For example, the more present a user feels, the more likely they are to duck to avoid virtual objects hitting them.

I3D 2019, May 2019, Montreal, Canada

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Physiological responses can also reveal emotional states [24, 28]. For example, heart rate and EDA (electrodermal activity) increase when a person is under stress. Wiederhold et al. [40] found a correlation between physiological measures (heart rate and EDA) and presence as measured with questionnaires. Meehan et al. [28] used a virtual elevated environment to induce vertigo and stress in participants while measuring heart rate, EDA, and skin temperature, and demonstrated that heart rate serves as a measure of presence.

Unlike previous work, our study found an association between affect-arousal, heart rate, and motion when using a VR HMD (head-mounted display) while standing.

2.2 Affective Computing

Affect has been estimated using keyboard and mouse event data [23], a modified keyboard and mouse that detect pressure [13], and various physiological signals (EEG, heart rate, EDA, eye gaze) [7, 9, 25, 30–32]. Several surveys exist [4, 10, 12, 18, 19, 43].

Our work is more closely related to the detection of affect from body postures and movements. Some previous work has involved asking actors (such as dancers) to enact different poses and movements that exemplify various emotional states [1, 5, 17, 21, 33, 34]. Once such postures or movements are captured, algorithms can be trained to recognize them. Our work instead used real users performing spontaneous movements.

Other previous work has leveraged information from head movements [6, 11, 29] or the sitting posture of a user's back [36].

The most relevant previous works we are aware of are [20, 37], which are also the only works we know that detect affect from full body posture or motion, without actors. Kleinsmith et al. [20] classified affective states from static postures, while Savva et al. [37] classified affective states from movements over time. In both works, data was captured from users standing and playing video games with a Nintendo Wii, and the input features to the learning algorithm were joint angles (and quantities related to these, such as angular speed) that had been captured with a full-body mocap (motion capture) system. Users experienced different affective states throughout the gameplay in an uncontrolled manner, and these were extracted and labelled by human judges. Kleinsmith et al. trained a Multilayer Perceptron (MLP) to classify the affective states, achieving an overall accuracy of ≈60%, while Savva et al. trained a Recurrent Neural Net (RNN), achieving an overall accuracy of ≈61%. Both works made clever use of human judges to establish a benchmark accuracy rate, showing that the accuracy of the machine algorithms was comparable to human judges. In our work, by contrast, our data is labelled by the different gameplay conditions under which the data were collected, without human judgment, establishing a ground truth.

We are aware of no previous work that attempts to detect affect from body motion while the user is wearing an HMD. Compared to previous work, our present study is the only one that makes use of a VR system with HMD and hand-held controllers, obviating the need for full-body mocap. In addition, ours is the only study that captured body motion along with physiological data (heart rate and EDA).

Other related work includes FACETEQ [26], a facial expression and emotion recognition device made for VR headsets.

3 EXPERIMENT

3.1 Apparatus

We used an HTC Vive virtual reality system, which includes a head-mounted display (HMD) and two hand-held controllers, along with three extra Vive Trackers. The Vive Trackers were attached to a belt around the user's pelvis and to the shoe laces of the user's shoes. We wanted to capture an approximation of full body motion, because we were unsure which motion data would be useful. The 3D position, orientation, linear and angular velocity of the HMD, of each hand and foot, and of the pelvis were captured at 100 Hz. The Vive was connected to a desktop PC with a 3.4 GHz i7-2600K CPU, and Nvidia GTX 970 graphics card, running Microsoft Windows 10 (64 bits). Output was rendered at ≈60 fps, as estimated by Unity 5.5[1]. Users stood and moved within a 2.5×2.5 meter area. The Vive provides a "chaperone" feature that fades in virtual walls (in the form of vertical blue grids) when the user comes too close to the boundary of the square area.

Heart rate was measured with a Polar H10 belt[2] at ≈60 Hz. Electrodermal Activity (EDA), also called Galvanic Skin Response (GSR) or skin conductance, was captured with a Shimmer Consensys GSR[3] connected to the index and middle fingers of the user's non-dominant hand. EDA values were measured at ≈50 Hz.

3.2 Participants

12 users (6 women, 6 men), aged 20-30 years (average 25.2, standard deviation (s.d.) 3.1) were recruited. All were right-handed and had previous experience with VR. None reported any vision, muscular or cardiovascular disorders.

Before donning the HMD or other equipment, each user was given a Titmus stereo acuity test to evaluate their depth perception. The test involves looking at 13 sets of shapes and judging which shape is at a different depth from the other shapes. The difficulty of the test increases which each set of shapes, with disparity varying from 400 arc seconds to 20 arc seconds. On average, users completed 11.75 (out of 13) sets successfully (s.d. 2.26).

3.3 Experimental Design

After the stereo acuity test, users put on the HMD and other equipment and underwent a calibration procedure. Each user then underwent several blocks of tasks, involving fast, swatting motions through the air to hit moving targets with virtual tennis rackets.

The experiment involved two conditions (Figure 1), a "calm" condition C, and a "stressed" condition S designed to induce stress and/or fear. Each user completed four blocks of tasks, where three of the blocks were calm, and one was stressed. The ordering of conditions was counter-balanced, with one quarter of users undergoing the ordering (C,C,C,S), another quarter undergoing (C,C,S,C), another (C,S,C,C), and another (S,C,C,C). Users were not told about the stressed condition beforehand, so it occurred as a surprise.

Blocks were separated by breaks of several minutes during which the user would sit down to answer a questionnaire about the block they had just experienced, and allowing the user's heart rate to return to normal. Each user's heart rate was monitored during each

[1]https://unity3d.com/unity
[2]https://www.polar.com/ca-en/products/accessories/h10_heart_rate_sensor
[3]http://www.shimmersensing.com/products/gsr-optical-pulse-development-kit

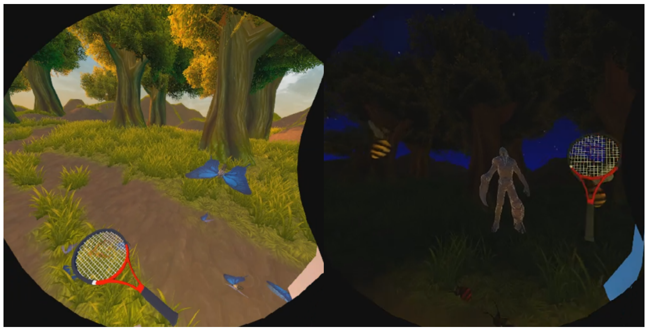

Figure 1: Left: in condition C (calm), users were asked to hit blue butterflies in a forest in daylight. Right: In condition S (stressed), users were asked to hit bees in a forest at night surrounded by monsters.

Figure 2: The butterfly in the calm condition (a), the bee in the stressed condition (b), and the bee in its angry state (c).

break to ensure that it had returned to a baseline level before the user was allowed to proceed with the next block.

3.4 Condition C: calm

In condition C, users were placed in a virtual forest in daylight with wind and bird sounds. Users were asked to hit 50 moving targets (flying in place with some random deviation) represented by blue butterflies (Figure 2, a) making a smashing sound when hit. Each target appeared in a predefined place around the user. At first, targets appeared one-by-one, each one appearing after the previous target had been hit. After hitting 25 targets, subsequent targets appeared in pairs simultaneously. After 40 targets were hit, targets appeared in triplets simultaneously until all 50 targets are spawned.

3.5 Condition S: stressed

In condition S, users were placed in the same virtual forest but at night. An audio track played strange noises, howling wolves, and people screaming in the distance. Also, monsters appeared in the distance looking directly at the user and then disappeared without making noise when the user looked away. As the condition progressed, a monster appeared closer and closer to the user, closing in. The user did not need to interact with the monsters; they were only present to induce fear.

Again, the user had to swat at 50 moving targets, but now they were represented as virtual bees making buzzing noises, with two possible states: when a bee was yellow (Figure 2, b), she was harmless, but after 3 seconds, the bee would switch to an angry state

Condition	Valence	Arousal	Dominance
Calm	7.16 (1.46)	3.86 (2.37)	5.66 (2.16)
Stressed	6.66 (2.18)	6.25 (2.67)	5.16 (2.41)

Table 1: Means (standard deviations) rating of the affective dimensions of valence, arousal and dominance on a 9-point Likert scale.

Realism	5.63 (0.42)
Ability to act	5.83 (0.58)
Interface quality	5.97 (0.90)
Ability to examine	5.58 (0.71)
Self-evaluation of performance	6.33 (0.55)
Audio fidelity	5.72 (0.95)
Haptic fidelity	3.46 (1.96)
Overall presence	5.59 (0.41)

Table 2: Means (standard deviations) for the 7 categories of questions used to measure presence on a 7-point Likert scale, across both conditions

(Figure 2, c) with a red color, shooting stingers toward the user. If the user was hit by a stinger, the user's entire field-of-view would flash with a transparent red color that dissolved away over 0.5 seconds. See the companion video for a demonstration of the stimuli.

The number and distribution of insects, and the movements required of the user to swat and eliminate the insects, was the same in both conditions, with the same algorithm used to position insects. Only the audio and visual feedback was different between the two conditions, with the intention of making condition S more stressful.

4 RESULTS

4.1 Subjective Impressions

During the pauses between blocks, users answered a Self-Assessment Manikin 9 scales test [3] used to rate the affective dimensions of valence, arousal and dominance (Table 1). (To illustrate these dimensions with some examples, pleasure is associated with positive valence, boredom with low arousal, fear with negative valence and low dominance, anger with negative valence and high dominance.) Likert scores were analyzed using t-tests. Arousal varied significantly ($t(11) = 2.93$, $p = 0.00521$) between the two conditions, but the other dimensions did not ($p > 0.05$).

After completing all blocks, users completed the standard Witmer and Singer presence questionnaire [41, 42], translated to French[4]. Overall, the level of presence (Table 2) was greater than the middle of the range, suggesting that our experiment should be able to elicit emotional and behavioral changes [15, 27].

4.2 Movement and Physiological Data

We sought a way to automatically detect changes in affect from only the movement data and/or physiological data. We conjectured that, under condition S, users might exhibit moments of backing up or other hesitant motions, frantic head rotation, trembling, or

[4]French version: http://w3.uqo.ca/cyberpsy/docs/qaires/pres/QEP_vf.pdf

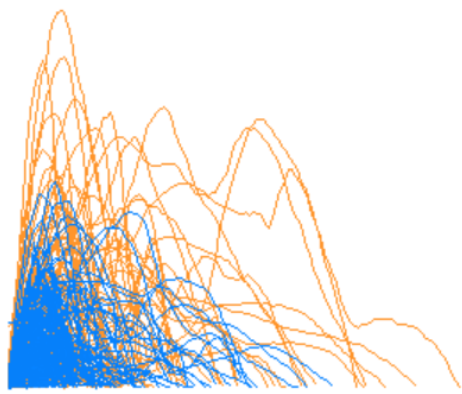

Figure 3: The *fast motions* for one user. The vertical axis is the sum of the speeds of the hands (prior to normalizing), and the horizontal axis is the time since the start of each *fast motion*. Blue is calm, orange is stressed.

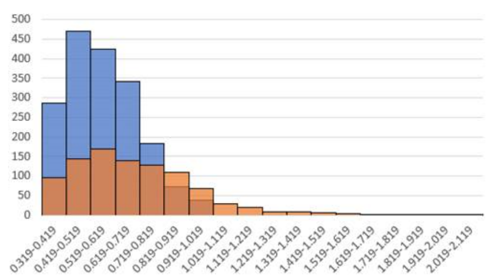

Figure 4: Superimposed histograms of max speed of the hands. The vertical axis is the number of *fast motions* extracted, and the horizontal axis is the normalized max speed of each *fast motion*. Blue is calm, orange is stressed.

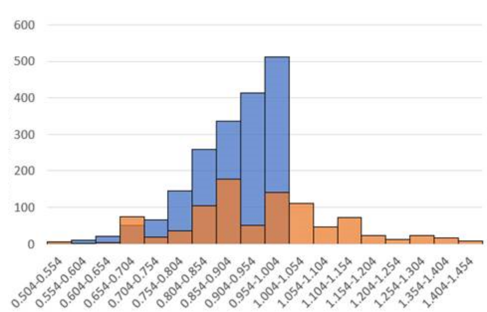

Figure 5: Superimposed histograms of heart rates. The vertical axis is the number of *fast motions* extracted, and the horizontal axis is the normalized heart rate at the max speed of each *fast motion*. Blue is calm, orange stressed.

inefficient limb motions. We watched replays of the captured movements animating a stick figure avatar for clues of differences between behavior in the C and S conditions. One of us watched these animations while trying to guess if they had been recorded in the

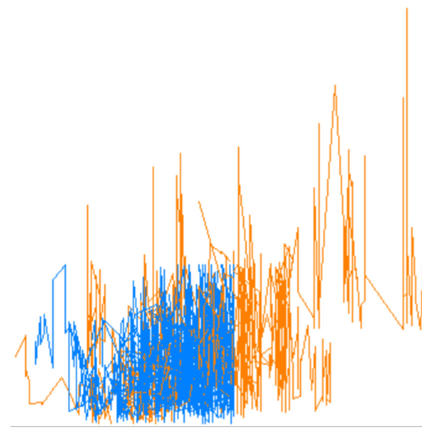

Figure 6: Superimposed parametric plots of the evolution of *fast motions* over time. Each block of each user is plotted as a polygonal curve, where each *fast motion* is a single point. The vertical axis is normalized max speed of the *fast motion*, and the horizontal axis is the normalized heart rate at max speed. Blue is calm, orange stressed.

		Predicted	
		Calm	Stressed
Actual	Calm	64.4%	1.6%
	Stressed	17.3%	16.8%

		Predicted	
		Calm	Stressed
Actual	Calm	62.9%	2.9%
	Stressed	25.7%	8.5%

Figure 7: Confusion matrices. Left: when using the two features normalized max speed and normalized heart rate at max speed, the accuracy is 81.1%. Right: when only using the first feature (i.e., no physiological data), accuracy is 71.4%.

C or the S conditions. The only difference we noticed was possibly faster hand motion in the S condition.

To investigate further, we segmented out intervals of time where the sum of the speed of the hands was above a threshold of 4 meters/second, and called these intervals *fast motions*. 2747 such motions were extracted over all users. The motions of one user are shown in Figure 3. We used the sum of the speed of the hands so that either user's hand could trigger the start of a *fast motion*. We computed several features of each *fast motion*: the maximum speed achieved during the motion, the maximum acceleration achieved, the heart rate at the moment of maximum speed, the duration of the motion, and the *curviness* of the motion which we defined as the ratio of the total distance traveled during the motion divided by the straight-line distance from the starting point to the ending point of the motion. (Our intuition was that curviness would serve as a measure of inefficiency, and that users might have less efficient movements when stressed.)

EDA values were found to be useless for distinguishing conditions. With three of the users, EDA was usually higher in the stressed condition; with four of the users, our instrument had failed to capture a useful EDA signal; and with the five other users, the EDA values were neither consistently higher nor consistently lower in the stressed condition.

We visualized the features of the *fast motions* in a variety of ways and, using the Accord.NET 3.8[5] C# library, we trained decision trees on various subsets of features. We found that the most promising features for classifying *fast motions* were max speed and heart rate. We then normalized these by dividing by the maximum values for each user across that user's three calm blocks. (Note that, by definition, the normalized max speed and normalized heart rate cannot exceed 1.0 in the calm conditions; they can only exceed 1.0 in the stressed conditions. In a practical application, such a normalization that is tailored to each user could be done if a sample is first collected of the user's "typical" calm motions. If it is not feasible to collect a sample of the particular user, then the normalization could be based on a threshold averaged over several users who were tested prior to deploying the software.)

We found that training a decision tree with just these two normalized features resulted in a tree containing only 2 thresholds (one for each feature) and yielded encouraging accuracy. The results in Figure 7 were obtained by training a decision tree on a random 50% of the *fast motions* and testing on the remaining 50%, repeating this 10 times and averaging. We note that, even without physiological data, using only the motion of the hands (which is already available on any VR platform with controllers) enables a recognition accuracy above 70%. We also note that Figure 7 shows the rates of false positives (calm motions misclassified as stressed) to be below 3%, meaning that the classification is conservative: responding to a user under stress would only be triggered if the algorithm is confident that the user is stressed. In a real-world implementation, our use of decision trees could be replaced with easy-to-implement thresholding heuristics.

5 CONCLUSION AND FUTURE DIRECTIONS

We have demonstrated that the maximum speed achieved by arm movements in VR can provide a reliable signal that a user is in a state of increased arousal, which could be useful in future affective computing scenarios. If available, heart rate data can also be used to increase accuracy.

Many possibilities remain for future work. As is visible in Figure 3, many of the *fast motions* contain multiple peaks and could be further segmented into finer grain motions. Literature on analysis of motion capture data [2, 8] could provide ideas for better ways to segment and classify the motion data into meaningful intervals. Tasks that require more complicated sequences of motions, such as the game Beat Saber where users must hit multiple blocks in mid-air using different slicing directions, might elicit signatures of stress beyond those found in our study. These signatures could include multiple closely-spaced movements of short duration (possibly indicating jerking or rapidly changing directions), or longer intervals of low speed separated by bursts of quick movements (indicating hesitation prior to executing motions, whereas a more relaxed user could be expected to more gradually prepare each successive stroke). More advanced ways to visualize the motion data [16] could also lead to new insights on how to analyze it. A larger motion dataset might allow machine learning algorithms to be tuned to recognize hesitant movements (backing up), head shaking, trembling, or other signatures of stress. Finally, future VR headsets may be equipped with the ability to measure pupil size; pupillary analysis could therefore be used to increase the accuracy of automatic affect detection.

[5]http://accord-framework.net/

ACKNOWLEDGMENTS

Many thanks to Luc Tremblay, Asma Ghandeharioun, Gierad Laput, Sheldon Andrews, David Labbé for advice, and NSERC for funding.

REFERENCES

[1] Paolo Alborno, Stefano Piana, Maurizio Mancini, Radoslaw Niewiadomski, Gualtiero Volpe, and Antonio Camurri. 2016. Analysis of intrapersonal synchronization in full-body movements displaying different expressive qualities. In *Proc. International Working Conference on Advanced Visual Interfaces (AVI)*. 136–143.

[2] Jernej Barbič, Alla Safonova, Jia-Yu Pan, Christos Faloutsos, Jessica K Hodgins, and Nancy S Pollard. 2004. Segmenting motion capture data into distinct behaviors. In *Proc. Graphics Interface (GI)*. 185–194.

[3] Margaret M Bradley and Peter J Lang. 1994. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry* 25, 1 (1994), 49–59.

[4] Rafael A Calvo and Sidney D'Mello. 2010. Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing* 1, 1 (2010), 18–37.

[5] Ginevra Castellano, Santiago D Villalba, and Antonio Camurri. 2007. Recognising human emotions from body movement and gesture dynamics. In *International Conference on Affective Computing and Intelligent Interaction*. Springer, 71–82.

[6] Sidney D'Mello, Rick Dale, and Art Graesser. 2012. Disequilibrium in the mind, disharmony in the body. *Cognition & Emotion* 26, 2 (2012), 362–374.

[7] Anders Drachen, Lennart E Nacke, Georgios Yannakakis, and Anja Lee Pedersen. 2010. Correlation between heart rate, electrodermal activity and player experience in first-person shooter games. In *Proc. ACM SIGGRAPH Symposium on Video Games*. 49–54.

[8] Ajo Fod, Maja J Matarić, and Odest Chadwicke Jenkins. 2002. Automated derivation of primitives for movement classification. *Autonomous Robots* 12, 1 (2002), 39–54.

[9] Kiel Gilleade, Alan Dix, and Jen Allanson. 2005. Affective videogames and modes of affective gaming: assist me, challenge me, emote me. In *Proc. DiGRA (Digital Games Research Association) Conference: Changing Views–Worlds in Play*.

[10] Hatice Gunes and Maja Pantic. 2010. Automatic, dimensional and continuous emotion recognition. *International Journal of Synthetic Emotions* 1, 1 (2010).

[11] Hatice Gunes and Maja Pantic. 2010. Dimensional emotion prediction from spontaneous head gestures for interaction with sensitive artificial listeners. In *Intelligent Virtual Agents*. Springer, 371–377.

[12] Hatice Gunes, Massimo Piccardi, and Maja Pantic. 2008. From the lab to the real world: Affect recognition using multiple cues and modalities. In *Affective Computing*. InTech.

[13] Javier Hernandez, Pablo Paredes, Asta Roseway, and Mary Czerwinski. 2014. Under pressure: sensing stress of computer users. In *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*. 51–60.

[14] Brent E Insko. 2003. Measuring presence: Subjective, behavioral and physiological methods. (2003).

[15] Brent Edward Insko, M Meehan, M Whitton, and F Brooks. 2001. *Passive haptics significantly enhances virtual environments*. Ph.D. Dissertation. University of North Carolina at Chapel Hill.

[16] Sujin Jang, Niklas Elmqvist, and Karthik Ramani. 2014. GestureAnalyzer: visual analytics for pattern analysis of mid-air hand gestures. In *Proc. ACM Symposium on Spatial User Interaction (SUI)*. 30–39.

[17] Asha Kapur, Ajay Kapur, Naznin Virji-Babul, George Tzanetakis, and Peter Driessen. 2005. Gesture-based affective computing on motion capture data. *Affective Computing and Intelligent Interaction* (2005), 1–7.

[18] Michelle Karg, Ali-Akbar Samadani, Rob Gorbet, Kolja Kühnlenz, Jesse Hoey, and Dana Kulić. 2013. Body movements for affective expression: A survey of automatic recognition and generation. *IEEE Transactions on Affective Computing* 4, 4 (2013), 341–359.

[19] Andrea Kleinsmith and Nadia Bianchi-Berthouze. 2013. Affective body expression perception and recognition: A survey. *IEEE Transactions on Affective Computing* 4, 1 (2013), 15–33.

[20] Andrea Kleinsmith, Nadia Bianchi-Berthouze, and Anthony Steed. 2011. Automatic recognition of non-acted affective postures. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 41, 4 (2011), 1027–1038.

[21] Andrea Kleinsmith, Tsuyoshi Fushimi, and Nadia Bianchi-Berthouze. 2005. An incremental and interactive affective posture recognition system. In *International Workshop on Adapting the Interaction Style to Affective Factors*. 378–387.

[22] Silvia Erika Kober and Christa Neuper. 2013. Personality and presence in virtual reality: does their relationship depend on the used presence measure? *International Journal of Human-Computer Interaction* 29, 1 (2013), 13–25.
[23] Agata Kołakowska. 2013. A review of emotion recognition methods based on keystroke dynamics and mouse movements. In *International Conference on Human System Interaction (HSI)*. IEEE, 548–555.
[24] Peter J Lang, Mark K Greenwald, Margaret M Bradley, and Alfons O Hamm. 1993. Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology* 30, 3 (1993), 261–273.
[25] Changchun Liu, Pramila Agrawal, Nilanjan Sarkar, and Shuo Chen. 2009. Dynamic difficulty adjustment in computer games through real-time anxiety-based affective feedback. *International Journal of Human-Computer Interaction* 25, 6 (2009), 506–529.
[26] Ifigeneia Mavridou, James T McGhee, Mahyar Hamedi, Mohsen Fatoorechi, Andrew Cleal, Emili Ballaguer-Balester, Ellen Seiss, Graeme Cox, and Charles Nduka. 2017. FACETEQ interface demo for emotion expression in VR. In *IEEE Virtual Reality (VR)*. 441–442.
[27] Michael Meehan. 2001. *Physiological reaction as an objective measure of presence in virtual environments*. Ph.D. Dissertation. University of North Carolina at Chapel Hill.
[28] Michael Meehan, Brent Insko, Mary Whitton, and Frederick P Brooks Jr. 2002. Physiological measures of presence in stressful virtual environments. *ACM Transactions on Graphics (TOG)* 21, 3 (2002), 645–652.
[29] Hamed Monkaresi, Rafael A Calvo, and M Sazzad Hussain. 2012. Automatic natural expression recognition using head movement and skin color features. In *Proc. International Working Conference on Advanced Visual Interfaces (AVI)*. ACM, 657–660.
[30] Lennart Nacke. 2009. *Affective ludology: Scientific measurement of user experience in interactive entertainment*. Ph.D. Dissertation. Blekinge Institute of Technology.
[31] Lennart Nacke and Craig A Lindley. 2008. Flow and immersion in first-person shooters: measuring the player's gameplay experience. In *Proc. Conference on Future Play: Research, Play, Share*. ACM, 81–88.
[32] Lennart Erik Nacke, Michael Kalyn, Calvin Lough, and Regan Lee Mandryk. 2011. Biofeedback game design: using direct and indirect physiological control to enhance game interaction. In *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*. 103–112.
[33] Stefano Piana, Alessandra Stagliano, Antonio Camurri, and Francesca Odone. 2013. A set of full-body movement features for emotion recognition to help children affected by autism spectrum condition. In *IDGEI International Workshop*.
[34] Stefano Piana, Alessandra Staglianò, Francesca Odone, and Antonio Camurri. 2016. Adaptive body gesture representation for automatic emotion recognition. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 6, 1 (2016), 6.
[35] Giuseppe Riva, Fabrizia Mantovani, Claret Samantha Capideville, Alessandra Preziosa, Francesca Morganti, Daniela Villani, Andrea Gaggioli, Cristina Botella, and Mariano Alcañiz. 2007. Affective interactions using virtual reality: the link between presence and emotions. *CyberPsychology & Behavior* 10, 1 (2007), 45–56.
[36] Jyotirmay Sanghvi, Ginevra Castellano, Iolanda Leite, André Pereira, Peter W McOwan, and Ana Paiva. 2011. Automatic analysis of affective postures and body motion to detect engagement with a game companion. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 305–311.
[37] Nikolaos Savva, Alfonsina Scarinzi, and Nadia Bianchi-Berthouze. 2012. Continuous recognition of player's affective body expression as dynamic quality of aesthetic experience. *IEEE Transactions on Computational Intelligence and AI in Games* 4, 3 (2012), 199–212.
[38] Mel Slater, Beau Lotto, Maria Marta Arnold, and Maria V Sanchez-Vives. 2009. How we experience immersive virtual environments: the concept of presence and its measurement. *Anuario de psicología* 40, 2 (2009).
[39] Mel Slater, Martin Usoh, and Anthony Steed. 1994. Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 3, 2 (1994), 130–144.
[40] Brenda K Wiederhold, DP Jang, Mayumi Kaneda, Irene Cabral, Yair Lurie, T May, IY Kim, Mark D Wiederhold, and SI Kim. 2001. An investigation into physiological responses in virtual environments: an objective measurement of presence. *Towards Cyberpsychology: Mind, Cognitions and Society in the Internet Age* (2001).
[41] Bob G Witmer, Christian J Jerome, and Michael J Singer. 2005. The factor structure of the presence questionnaire. *Presence: Teleoperators and Virtual Environments* 14, 3 (2005), 298–312.
[42] Bob G Witmer and Michael J Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments* 7, 3 (1998), 225–240.
[43] Zhihong Zeng, Maja Pantic, Glenn I Roisman, and Thomas S Huang. 2009. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31, 1 (2009), 39–58.