Emotional Beasts: Visually Expressing Emotions through Avatars in VR

Guillermo Bernal
MIT Media Lab
75 Ahmerst Ave.
Cambridge, MA 02143, USA
gbernal@media.mit.edu

Pattie Maes
MIT Media Lab
75 Ahmerst Ave.
Cambridge, MA 02143, USA
pattie@media.mit.edu

Abstract
With the advances in Virtual Reality (VR) and physiological sensing technology, even more immersive computer-mediated communication through life-like characteristics is now possible. In response to the current lack of culture, expression and emotions in VR avatars, we propose a two-fold solution. First, integration of bio-signal sensors into the HMD and techniques to detect aspects of the emotional state of the user. Second, the use of this data to generate expressive avatars which we refer to as Emotional Beasts. The creation of Emotional Beasts will allow us to experiment with the manipulation of a user's self-expression in VR space and as well as the perception of others in it, providing some valuable tools to evoke a desired emotional reaction. As this medium moves forward, this and other tools are what will help the field of virtual reality expand from a medium of surface-level experience to one of deep, emotionally compelling human-to-human connection.

CHI'17 Extended Abstracts, May 06-11, 2017, Denver, CO, USA
ACM 978-1-4503-4656-6/17/05.
http://dx.doi.org/10.1145/3027063.3053207

Author Keywords
Self-expression; social virtual reality; avatars; bio-signals; affective computing.

ACM Classification Keywords
J.5 [Arts and humanities]: Performing arts;

Introduction
In Western culture, self-expression is valued as a powerful sign of individual freedom, involving the projection of one’s own thoughts and ideas into the world [21]. With the rise of Social Virtual Reality, new possibilities of expression and communication are emerging; people can now download applications like Altspace VR [2] to socialize or keep in touch with friends and family in a digital lobby through their Virtual Reality devices, thus creating new social practices. It is important in such environments to still

Figure 1: Nica Cave, Soundsuit 2012

Figure 2: Kohei Nawa. Trans 2012

support aspects of traditional, social face-to-face interaction.

Current virtual reality expeiences lack the ability for a user to accurately convey the subtle clues that make us human. Research aimed at enhancing the ability for people to connect through computer mediated communication needs to be grounded in psychology, neuroscience and the arts. Psychological studies and visual analysis of body movements show that human movement and expression differs from other forms of movement and expression it is the only visual stimulus we have experience in both perceiving and producing [17]. Charles Darwin devoted Chapter 13 of his 1872 *The Expression of the Emotions in Man and Animals* to complex emotional states including self-attention, shame, shyness, modesty and blushing. He described blushing as "... the most peculiar and most human of all expressions." [3]

Non-verbal communication signals are primitive signals communicated between people through unconscious behaviors that contain information about our intentions, goals, and values [1]. For example, several different psychological and psycho-physiological mechanisms for blushing have been hypothesized by Crozier: *"An explanation that emphasizes the blush's visibility proposes that when we feel shame we communicate our emotion to others and in doing so we send an important signal to them. It tells them something about us. It shows that we are ashamed or embarrassed, that we recognize that something is out of place…. The blush makes a particularly effective signal because it is involuntary and uncontrollable*." [4] By observing these unconscious behaviors, we can gain insight into circumstances without the use of verbal-communication, such as how well a meeting is going [1] or how successfully a deal will be made [17]. Furthermore, theories linking the response to motion visual stimuli to mirror motion-visual neurons allow a deeper understanding of how we comprehend others' emotions, expressions and empathy [6].

Analyzing and decoding human emotion is a non trivial problem, as we are exploring multi-modal means of emotional communication. Current physiological signal driven systems are able to automatically detect and recognize emotions in 81% of the cases using just four physiological signals(EMG, GSR, BVP,RR) [1] [32] [33], which is at a level comparable to machine recognition of facial and vocal expressions. It is important to keep in mind that the malleability of virtual self can become a powerful tool through which one can insert, augment, diminish, twist, or invert how someone feels.

In this paper, we present Emotional Beasts, a concept for affective expression in social VR. Through the use of VR headsets that have been altered to accommodate physiological sensors, we were able to implement a working prototype. We discuss the collection and integration of physiological data to perceive human effect. The motivation of this work is to investigate the development of new platforms for social interaction in VR, both in collocated and distant environments. We propose a DABA (dynamic adaptation-based avatar) as a platform in which to provide the user the experience

[1] Electromyogram (EMG), Blood volume pressure (BVP), Skin Conductivity (GSR) and Respiration Waves (RR)

Creative Arts Expression: Refers to participation in a range of activities that allow for creative and imaginative expression, such as music, art, creative movement, and drama

Figure 3: Virgen de la candelaria festival, Peru 2014

Figure 4: 32 Neuron setup. Image courtesy of Noitom Inc.

of sharing and exchanging emotional expressions with other users. Our methods pull the avatar design away from the uncanny valley and into the 'expressive' domain.

Related Work

We look at *Creative Arts Expression* as a motivation for VR experiences, and our work builds on the efforts of many others from the field of Human Computer Interaction (HCI). Researchers in HCI have explored novel methods to recover the face of the user in real-time [29] as well as to detect facial expressions on the visible part of the face and use it to change the upper face model of an avatar accordingly. This method requires the user to have a webcam in front of them to learn the user appearance offline and build a 3D textured model of his/her head from a series of pictures. Another relevant project is PhysioVR [12] an open-source software tool developed to facilitate the integration of physiological signals measured from consumer wearable devices into VR applications. Social virtual environments have been investigated through the formation of avatar-based systems [5,13] and using these we are able to examine social channels such as gaze behavior [15] and their impact on communication quality [16].

In the area of dance and choreography, artists have been using computer technology in creation and performance since the 1960s. Merce Cunningham is widely regarded as the first to use computer software to choreograph. In collaboration with Simon Fraser University, the LifeForms software was developed and used to create *Trackers* (1992), the first of many of his works to be choreographed using computer software. The software enabled Cunningham to clearly conceptualize movement sequences in digital time and space, as a kind of "visual idea generator" [20].

The avatars developed for this exploration (Figure 6,8) draw inspiration from artists like Nick Cave and Kohei Nawa. Artist Nick Cave is well known for his Soundsuits [21]: full-body sized sculptures, often worn as costumes and performed in. Cave's Soundsuits (Figure 1) are made of everything from collected and repurposed buttons, to wooden sticks, beaded baskets, doilies and sequined fabric. Soundsuits conceal the wearer's identity leaving the viewer with no indication of race, gender, or age. His suits are often described as a joyful secondary skin and offer a liberating experience that frees the wearer from inhibitions. Cave's work is often associated with 'folkloric' costumes seen in Africa and Latin America. These 'folkloric' costumes are big part of cultural expression (Figure 3) using synthetic or hand-made materials in their costumes, wearing masks and mustaches or with a hooked nose and no facial hair, stooping low to the ground and stomping hard while representing a harvest scene, or keeping an upright posture and waving a handkerchief while dancing to a coastal tune, performers give form to these 'ethnic/racial' identities.

Artist Kohei Nawa's work [22] presents a series of sculptures that explore the transformation of the human form through digital sculpting techniques (Figure 2), manipulation, and mesh smoothing. His workflow is based on texture mapping, whereby the collected data is manipulated to generate fluid 3D surfaces. The work evokes a sense that these forms are not part of this world, but live parallel to it.

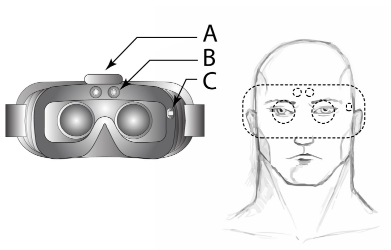

Figure 5: The HMD is fitted with electrodes on the plate mask for physiological data collection. A) Front end analog circuit board with bluetooth connection. B) PPG sensor for Heart Rate data collection. C)Electrode for GSR data collection.

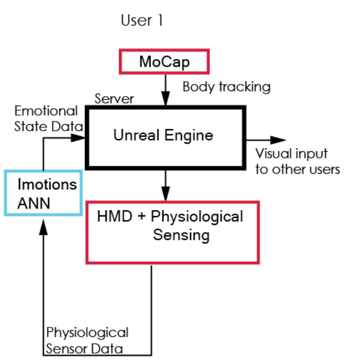

Figure 6: Flow of information in system modules

Proof-of-Concept System

Based on the inspiration and analysis of these previous works, we created a multimodal VR system that provides the user with second skins, with the aim to provide anonymity and free the user from any inhibition. Our system (Figure 6) is composed of three sub-modules: i) a user-sensing module which collects data from the user (e.g. head position, body position, hand position, GSR, PPG) and displays the virtual environment to the user, ii) an affective module, a sotware layer used for emotion classification and iii) a physics-based graphics engine module which is responsible for generating the responsive second skin and consolidates all the data, allowing the user to interact with the virtual environment.

Bio-sensing and HMD integration

To create a seamless experience, we have integrated several bio-signal sensors into the face plate of an HTC VIVE VR headset and utilized the Shimmer3 sensor for emotion-sensing [27]. For the collection of Galvanic Skin Response, dry electrodes were positioned on the forehead area due to the fact that it is one of the areas most dense with sweat glands. GSR data reflects emotional arousal, but in order to identify positive or negative valence, it becomes necessary to complement GSR with other biosensors. For the heart rate, a PPG (photoplethysmogram) sensor, which senses the rate of blood flow by utilizing light to monitor the heart's pumping action, was placed in the temple region of the user (Figure 5). This is done to get insights into the respondents' physical state, anxiety and stress levels (arousal), and used to determine how changes in their physiological state relate to their actions and decisions.

Based on the work from Anil, et.al [31], we have begun to explore the use of a neural network to estimate four kinds of emotions; angry, stressed, neutral and happy, these are based on the Lang model [35]: two dimensions of valence and arousal (Figure 11). While much of the detailed evaluation of the recorded data continues, the early results suggest that this setup yields accuracy and synchronized data correlation.

Motion to Emotion

Our system uses the Perception Neuron suits [24] to track the movements of the users and control their rigged avatars. Perception Neuron is an inertial sensor based motion-capture system; that means that each sensor, called a Neuron, measures its own orientation and acceleration using a gyroscope, a magnetometer and an accelerometer (Figure 4). On top of this, each Neuron also does its own drift correction. This creates a skeleton data that is then broadcasted over a local network. After this point is where the Unreal Engine [26] integration comes into play, as it will connect to the broadcast and turn the binary values into usable numbers for the game engine. The plugin provides a function to import motion capture data issued by the Perception Neuron to Unreal Engine to drive the movement of the appropriate avatar.

Avatar Aesthetics

It was important to be able to feed the information from the physiological reading into a system that would map those readings into visual space.

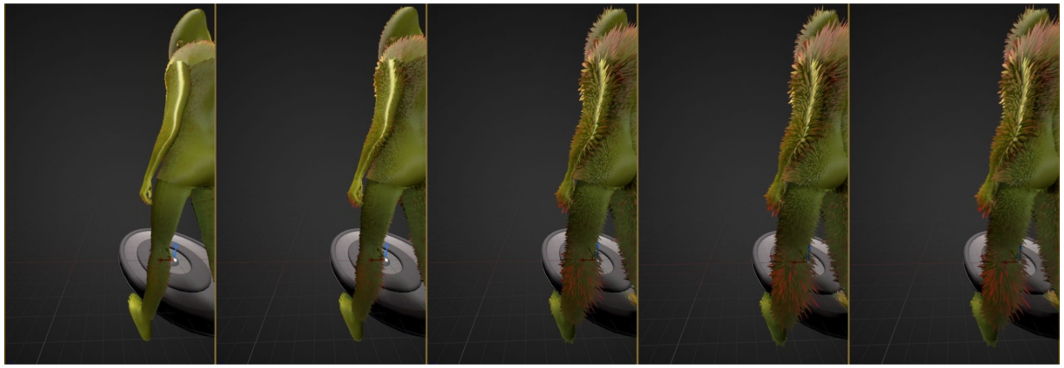

Figure 7: Depiction of the range of possible transformations and qualities for fur forms of expression. The material sample shows the a form of expression and variations transformations in color and growth.

Figure 8: sequence of frames of two fur based avatars interacting with each other. Note the strands growth as the users emotional state changes

Fur

One of the methods in which an avatar can visually express emotion to other users in the VR environment is through a fur system made of individual splines that grow between two boundaries, referred to as shells. The inner shell is representative of the body surface of the avatar, while the outer shell is the furthest extent to which the splines of the fur can grow. We shaped the outer shell to achieve a desired spline length and then applied the blendshape function to the splines to create a 3-dimensional affect and thus creating a fur-like surface on the avatar. (Figure 8)

The splines have the ability to grow or contract between the extents of the shells to visually express the emotions of the avatar; emotions such as joy and sadness can be expressed through the growth and the contractions of the fur, respectively. Preferences can be adjusted by the user to achieve whichever form of visual expression they prefer, as the system is not limited to strictly fur growth and contraction. For the purpose of this experiment, we evaluated the four emotional states show in Lang's model (Figure 11) by assigning emotional values between 0 and 1. When comparing neutral and happy, neutral was assigned a value of 0 and happy was assigned a value of 1. When the avatar interpreted an emotional value closer to 0, the splines remained close to the inner shell and therefore resembled a colored mesh. Alternatively, when the emotional value increased and read closer to 1, the splines would grow to longer lengths, depicting a joyous expression (Figure 7).

Particle System

Another strategy we have experimented with is a particle system-based avatar through the use of PopcornFX [25], a 3D real time FX plugin for Games & interactive applications. The primary job of the particle system itself is to control the behavior of particles, while the specific look and feel of the particle system as a whole is often controlled by way of materials.

The material of the particle can depict various ranges of aesthetics, ranging from dim to bright, and from blue to red. Working with the emotional scale from 0 to 1, arousal would be depicted through the particle system by converting the particles from dim to bright; as the user experiences a higher degree of arousal, the particles increase in brightness until they reach the

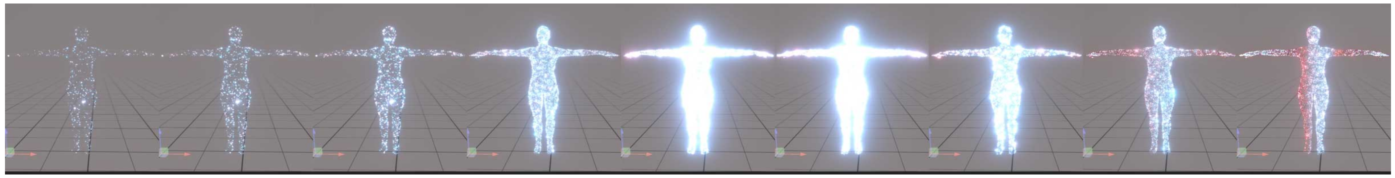

Figure 9: Depiction of the range of possible transformations and qualities for a particle system form of expression. The figure shows the particle system's variations in particle size, density, brightness, and color.

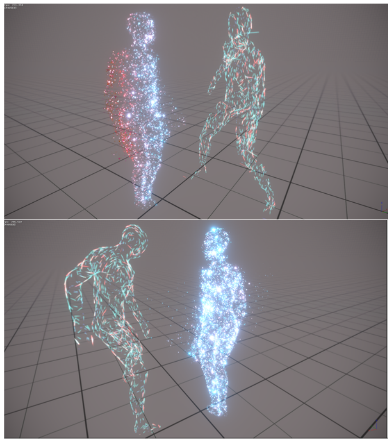

Figure 10: sequence of frames of two particles based avatars interacting with each other. Note the color and brightness increases as the users emotional state changes

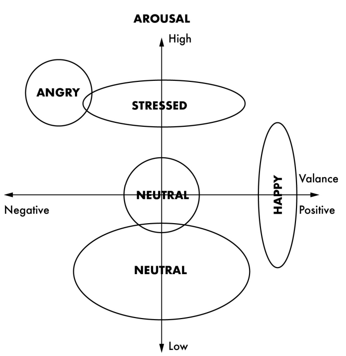

Figure 11: Discretization of arousal-valence based on Lang's model.

maximum at a value of 1 (Figure 9). In a similar fashion, if the user experiences a high degree of anger, the particles convert from a blue shade to that of red, making it clear to other avatars within the VR environment that this particular user is experiencing great frustration. As such, we have created colorful and vibrant illustration sequences; the user is able to watch Van Gogh-like strokes swirl and coil into seemingly living, breathing compositions, expressing vivid motifs soured from the sky, stars, universe and beyond (Figure 10). These effects help immerse viewers in virtual environments by adding detail to them, as well as by making them more attractive.

Discussion

We have described a method for generating expressive and previously unseen VR avatars that project non-verbal cues from the wearer's emotional state. Early results of these explorations have also been presented. An evaluation indicates that the avatars exhibit novel, surprising and inspiring expressive qualities and as initially suspected, the engagement and reaction from those that have tried the setup are described as joyful; a very liberating experience that makes the user want to dance. We argue that the use of collaborative virtual environments which incorporate emotionally expressive avatars has the potential to engender expression of emotions amongst the users of such environments. Emotional responses are unique and the field is constantly growing; currently there is no clear solution as to how to properly detect emotion. We would like to use state of the art machine learning to further explore emotional response, however it is quite intensive and challenging. Nonetheless, it is an interesting area of research that we would like to persue.

Conclusion

The creation of Emotional Beasts allowed us to experiment with the manipulation of a user's self-expression in VR space and as well as the perception of others in it, providing some valuable tools to evoke a desired emotional reaction. As this medium moves forward, this and other tools are what will help the field of virtual reality expand from a medium of surface-level experience to one of deep, emotionally compelling human-to-human connection.

Acknowledgements

This work was partially funded by Samsung. The views and conclusions contained herein are those of the authors and should not be interpreted as representing the official policies or endorsements, either expressed or implied, of Samsung.

References

1. Pentland, A. S. 2008. Honest Signals. MIT Press.
2. AltspaceVR 2015. Social software for virtual reality environments. Launched its initial product in May 2015.
3. Darwin, Charles, Paul Ekman, and Phillip Prodger. *The expression of the emotions in man and animals*. Oxford University Press, USA, 1998.
4. Crozier, R. 2010. "The Puzzle of Blushing", *The Psychologist*, Vol 23. No 5, May 2010, pp.390-393.
5. Bente, Gary, et al. "Avatar-Mediated Networking: Increasing Social Presence and Interpersonal Trust in Net-Based Collaborations." Human communication research 34.2 (2008): 287-318.
6. Rizzolatti, G., Fogassi, L., Gallese, V.: Mirrors in the mind. Scientific American 295, 54–61 (2006)
7. Boone, R.T., Cunningham, J.G.: Children's decoding of emotion in expressive body movement: the development of cue attainment. Developmental psychology 34, 1007–1016 (1998)
8. De Meijer, M.: The contribution of general features of body movement to the attribution of emotions. Journal of Nonverbal Behavior 13, 247–268 (1989)
9. Wallbott, H.G.: Bodily expression of emotion. (1998)
10. Schiphorst, T. A Case Study of Merce Cunningham's Use of the Lifeforms Computer Choreographic System in the Making of Trackers . M.A. Thesis, Simon Fraser
11. S. Rank and C. Lu, "PhysSigTK: Enabling engagement experiments with physiological signals for game design," in Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on, 2015, pp. 968–969.
12. Muñoz, John Edison, et al. "PhysioVR: A novel mobile virtual reality framework for physiological computing." e-Health Networking, Applications and Services (Healthcom), 2016 IEEE 18th International Conference on. IEEE, 2016.
13. Steed, Anthony, and Ralph Schroeder. "Collaboration in Immersive and Non-immersive Virtual Environments." Immersed in Media. Springer International Publishing, 2015. 263-282.
14. Bailenson, Jeremy N., et al. "The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction." Presence: Teleoperators and Virtual Environments 15.4 (2006): 359-372.
15. Roberts, David, et al. "Communicating eye-gaze across a distance: Comparing an eye-gaze enabled immersive collaborative virtual environment, aligned video conferencing, and being together." 2009 IEEE Virtual Reality Conference. IEEE, 2009.
16. Garau, Maia, et al. "The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment." Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 2003.
17. Beale, Christian Peter Russell. "Affect and Emotion in Human-Computer Interaction."
18. Maddux, W.; Mullen, E.; and Galinsky, A. 2007. Chameleons bake bigger pies and take bigger pieces: Strategic behavioral mimicry facilitates negotiation outcomes. Journal of Experimental Social Psychology 44(2):461–468.
19. Aron-Rubin, J. (2005). Child art therapy. Hoboken, NJ: Wiley and Sons, Inc.
20. Schiphorst, T. A Case Study of Merce Cunningham's Use of the Lifeforms Computer Choreographic System in the Making of Trackers . M.A. Thesis, Simon Fraser University, 1993.

21. Soundsuit. 2012. Nick Cave. website. (15 July 2013.). Retrieved January 06, 2017 from https://www.google.com/culturalinstitute/beta/u/0/asset/soundsuit/HwGmQj45rDcBHQ
22. TRANS. 2012. Kohei Nawa. website. (15 July 2012.). Retrieved January 06, 2017 from https://http://kohei-nawa.net/works/trans
23. Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. Journal of Neuroscience Methods, 190, 80–91.
24. Perception Neuron. Perception Neuron Axis Forum . https://neuronmocap.com. (2016)
25. PopcornFx. http://www.popcornfx.com. (2016)
26. Games, Epic. "Unreal engine 3." URL: http://www. unrealtechnology. com/html/technology/ue30. shtml (2006).
27. Shimmer. http://www.shimmersensing.com. (2016)
28. iMotions. https://imotions.com. (2016)
29. Burgos-Artizzu, Xavier P., et al. "Real-time expression-sensitive HMD face reconstruction." *SIGGRAPH Asia 2015 Technical Briefs*. ACM, 2015.
30. R.W. Picard, Affective Computing, MIT Press, London, England, 1997
31. Anil K. Jain, Jianchang Mao, K.M. Mohiuddin, "Artificial Neural Networks: A tutorial" IEEE Computer Society 1996, March 1996 Vol. 29, No. 3
32. Vyzas, E., and Picard, R. W. (1999).Online and Offline Recognition of Emotion Expression from Physiological Data, submitted to Workshop on Emotion-Based Agent Architectures, Int. Conf. on Autonomous Agents, Seattle, WA.
33. Picard, Rosalind W. "Affective Computing for HCI." HCI (1). 1999.
34. Lee, ChungK, et al. "Using neural network to recognize human emotions from heart rate variability and skin resistance." 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference. IEEE, 2006.
35. Lang, Peter J., Michael Davis, and Arne Öhman. "Fear and anxiety: animal models and human cognitive psychophysiology." Journal of affective disorders 61.3 (2000): 137-159.