Respire: Virtual Reality Art with Musical Agent Guided by Respiratory Interaction

Kivan Çtatar, Mirjana Prpa, Philippe Pasquier

Leonardo Music Journal, Volume 29, 2019, pp. 19-24 (Article)

Published by The MIT Press

Respire

Virtual Reality Art with Musical Agent Guided by Respiratory Interaction

KIVANÇ TATAR, MIRJANA PRPA AND **PHILIPPE PASQUIER**

ABSTRACT

Respire is an immersive art piece that brings together three components: an immersive virtual reality (VR) environment, embodied interaction (via a breathing sensor) and a musical agent system to generate unique experiences of augmented breathing. The breathing sensor controls the user's vertical elevation of the point of view under and over the virtual ocean. The frequency and patterns of breathing data guide the arousal of the musical agent, and the waviness of a virtual ocean in the environment. *Respire* proposes an intimate exploration of breathing through an intelligent mapping of breathing data to the parameters of visual and sonic environments.

THE IMMERSIVE ENVIRONMENT OF *RESPIRE*

Respire (2018) is built on our previous work *Pulse.Breath.Water* (2016), which immerses the user in a virtual environment depicting an ocean, and the user traverses the environment using their breathing (Fig. 1). The environment, made with Unity and presented using the HTC Vive headset, evokes a dark, gloomy atmosphere with elements like fog and waves that wrap around the user [1]. Ambiguous visuals and a lack of focal objects stimulate the user to engage in the process of making sense of the scene. Ambiguity in the design of the interactive artifacts engages users to project their own values and experiences in the process of making meaning of the visual stimulus [2]. *Respire* also exercises a minimalist color scheme to elicit the beholder's share effect. Kandel [3] proposes *beholder's share* as a process of the user's projection of previous experiences in sense-making of an ambiguous scene. By immersing the user in an ambiguous environment, *Respire* aims to give the user a canvas to "paint" their own experiences. The immersion avoids imposing an explicit narrative and the constraints of a story. Hence the environment empowers the user to curate their experiences.

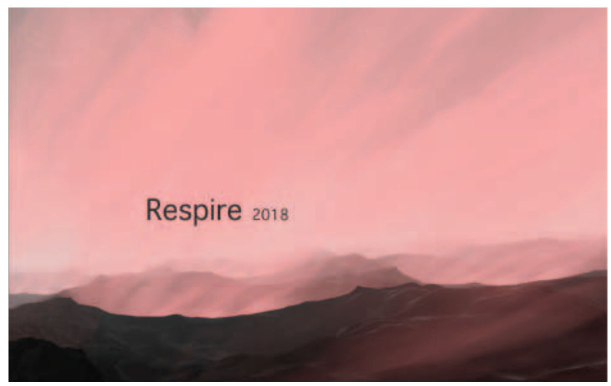

Fig. 1. The virtual environment visuals in *Respire*. (© Mirjana Prpa)

BODY AND BREATH IN IMMERSIVE VIRTUAL ENVIRONMENTS

Cartesian dualism approaches a human being as a "thinking thing" that is divorced from bodily experience. Questioning Cartesian dualism sparks new discourses on what it means "to be" in the environment. The examples of embodied interaction challenge the Cartesian separation between subject and object, and this separation translates into artistic practices as a separation between the artwork and the audience. Embodied interaction [4] emphasizes the value of engaging our bodies in interaction and transcends Cartesian dualism by focusing on the body along with the mind as a united medium for experience of the environment.

What could that link between this united medium and the environment be? Artists explored breathing as a connector between the body and virtual environments generated using computational means. By positioning the body in the center of the artwork, and employing breath in an embodied interaction paradigm, artists succeeded in creating that tangible

Kıvanç Tatar (post-doctoral researcher), School of Interactive Arts and Technology, Simon Fraser University, 250-13450 102nd Avenue, Surrey, British Columbia, V3T 0A3, Canada. Email: ktatar@sfu.ca. ORCID: 0000-0003-4133-8641.

Mirjana Prpa (Ph.D. candidate), School of Interactive Arts and Technology, Simon Fraser University, 250-13450 102nd Avenue, Surrey, British Columbia, V3T 0A3, Canada. Email: mprpa@sfu.ca. ORCID: 0000-0003-0368-6581

Philippe Pasquier (artist, educator, researcher), School of Interactive Arts and Technology, Simon Fraser University, 250-13450 102nd Avenue, Surrey, British Columbia, V3T 0A3, Canada. Email: pasquier@sfu.ca. ORCID: 0000-0001-8675-3561.

See www.mitpressjournals.org/toc/lmj/29 for supplemental files associated with this issue.

 doi:10.1162/LMJ_a_01057

yet invisible link. For instance, Sonia Cillari's *As an artist, I need to rest* (2009) explores how a body can be a source of artificial life through breathing [5]. Cillari employs breathing data to generate a virtual environment of feathers and maps the breathing patterns to the movement of the feathers in the virtual environment.

Likewise, Char Davies's pioneering piece *Osmose* (1995) is an immersive virtual environment presented on a head-mounted display [6]. The user navigates movement in the virtual environment with breathing and body balance. The breathing controls the elevation whereas the body balance changes the horizontal 2D direction. This mapping resembles the experience of diving, and, likewise, we are inspired by diving phenomena in the breathing interaction of *Respire*. Davies juxtaposes two ideas: the immateriality of computer-generated worlds and the body-felt phenomena elicited by those environments. The sense of virtual presence afforded by VR opened a dynamic space for artistic explorations, as demonstrated in Davies's piece [7]. *Respire* maintains that space for artistic exploration by introducing agent architecture that reacts to breathing.

ARTIFICIAL INTELLIGENCE, MULTI-AGENT SYSTEMS AND MUSICAL AGENTS

All humans breathe, consciously or unconsciously. A critical element of being alive, breathing continues even when we do not attend to the act of breathing. Our control of breathing shifts depending on our attention. Inspired by this mechanism, *Respire* is intended to use advanced computational tools for an intelligent mapping from breathing to movement, sound and visuals so as to elicit attention and mindfulness. AI and Multi-Agent Systems (MAS) provide such tools for computational creative applications [8].

The agent paradigm appears in many disciplines, including social sciences, philosophy, cognitive science and computer sciences. In computer science, an agent is an autonomous system that initiates actions to respond to its environment in timely fashion [9]. MAS studies the agent architectures for computational applications. Musical agents are artificial agents that automatize musical creative tasks [10]. *Respire*'s architecture implements this intelligent mapping using a musical agent system.

AFFECT RECOGNITION

Affect recognition focuses on designing computational models that can estimate the affective state of a piece of content; the content can be, for example, an image, a video, a human body posture, a sound or a music piece. Two main types of affective models appear in affective computing: discrete and continuous [11]. *Respire* implements a two-dimensional continuous affect model previously presented by Tatar and Pasquier [12]. Dimensional affect estimation models generate a bounded, continuous output to which we apply signal processing, mapping and generative algorithms.

Respire computationally generates the visual and sonic environments using two separate frameworks: a VR system and a musical agent generating the sonic environment. These frameworks utilize affective dimensions in the system architectures. The generative content and reactive behaviors of these systems use affective dimensions as a high-level cross-medium paradigm for intelligent mapping. This enables a human-readable parametrization of two systems generating visuals and audio separately.

SYSTEM DETAILS

Respire aims to bridge the virtual presence and innate experiences of the user. The artwork builds upon breath-based embodied interaction and utilizes a breath controller (Thought Technology ProComp2 with respiration harness) for the user's vertical position in the scene, simultaneously allowing for exploration of the virtual environment and breathing patterns (Fig. 2). This mapping of breathing to vertical elevation in a resemblance to diving guides the interaction such that the audience does not need to learn anything new to participate in the artwork. We have compared this mapping with other options and found the current mapping to be intuitive from the user's perspective [13]. Also, rapid breathing creates a more eventful sonic environment and is reflected in and ocean surface filled with waves. Less eventful breathing (slow-paced breathing) calms the ocean surface and generates a calmer sonic environment. The pleasantness dimension calculated by the musical agent controls the color of the sky in the environment (see Fig. 2), changing the sky in a range between black (low pleasantness values) to white (high pleasantness values).

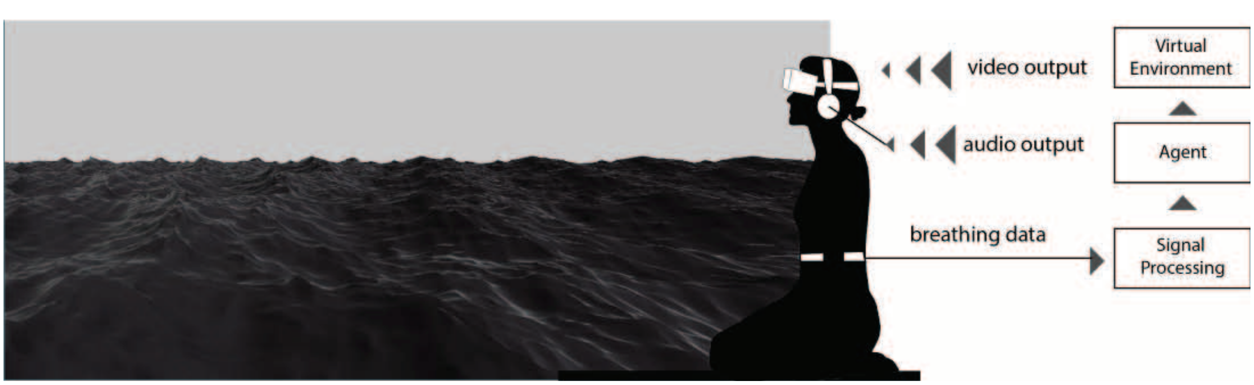

Fig. 2. The system architecture of *Respire.* (© Mirjana Prpa)

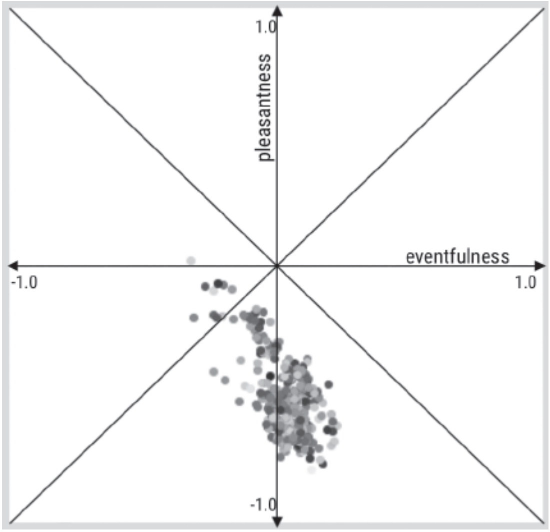

Fig. 3. The affective labels of audio samples in the musical agent's *memory*. Each dot represents one audio sample. (© Kıvanç Tatar)

The musical agent is developed in Cycling74's Max, and the communication between Max and Unity utilizes UDP-based OSC. The breathing sensor data is passed to Max using M+M middleware [14]. Below, we delve into the agent architecture, which consists of five main modules: *memory*, *perception*, *goal*, *action* and *postprocessing*.

Sound Memory

The agent's *memory* consists of audio samples and symbolic data on pleasantness and eventfulness of the samples (Fig. 3). This type of corpus in a musical agent is known as a hybrid corpus [15]. In line with *Respire*'s aesthetic choice of ambiguity, we aimed for a curation of abstract and ambient sounds for the musical agent's *memory*. We focused on quartal and quintal harmonies in piano recordings. The quartal and quintal harmony theory uses the musical interval of fourths and fifths, thus generating ambient sounds that avoid tensions and resolutions of the tonal harmony. In line with the visual aesthetics, this harmonic choice avoids imposing an explicit narrative on the user. To obscure the recordings' source material, we suppressed the initial attacks of sounds using long fade-in durations. Then we applied time stretch and pitch shifting to increase the number of samples in the sound memory. For the pitch-shifting, we applied intervals of fourths and fifths to stay within the quartal harmony. Then we automatically labeled each sound using an affect estimation algorithm for sound. The labels are vectors with two dimensions: average pleasantness and average eventfulness of an audio sample. The details of this affect estimation algorithm are previously published, where the multivariate linear regression model is trained on ground-truth data [16]. The negative correlation between the eventfulness and pleasantness in Fig. 3 has been observed in our previous studies. In sound studies, the affective dimensions valence and arousal are exchanged with eventfulness and pleasantness, because a sound doesn't feel an emotion; it stimulates one. For example, there is no concept of a happy sound (excluding anthropomorphism), but some sounds trigger positive emotions in humans.

Signal Processing of Breathing Data

The *perception* module recognizes the frequency of user breathing by the wavelet transform of the breathing amplitude stream (Fig. 4a). The wavelet transform outputs the spectrum of a signal. The breathing frequencies can go as low

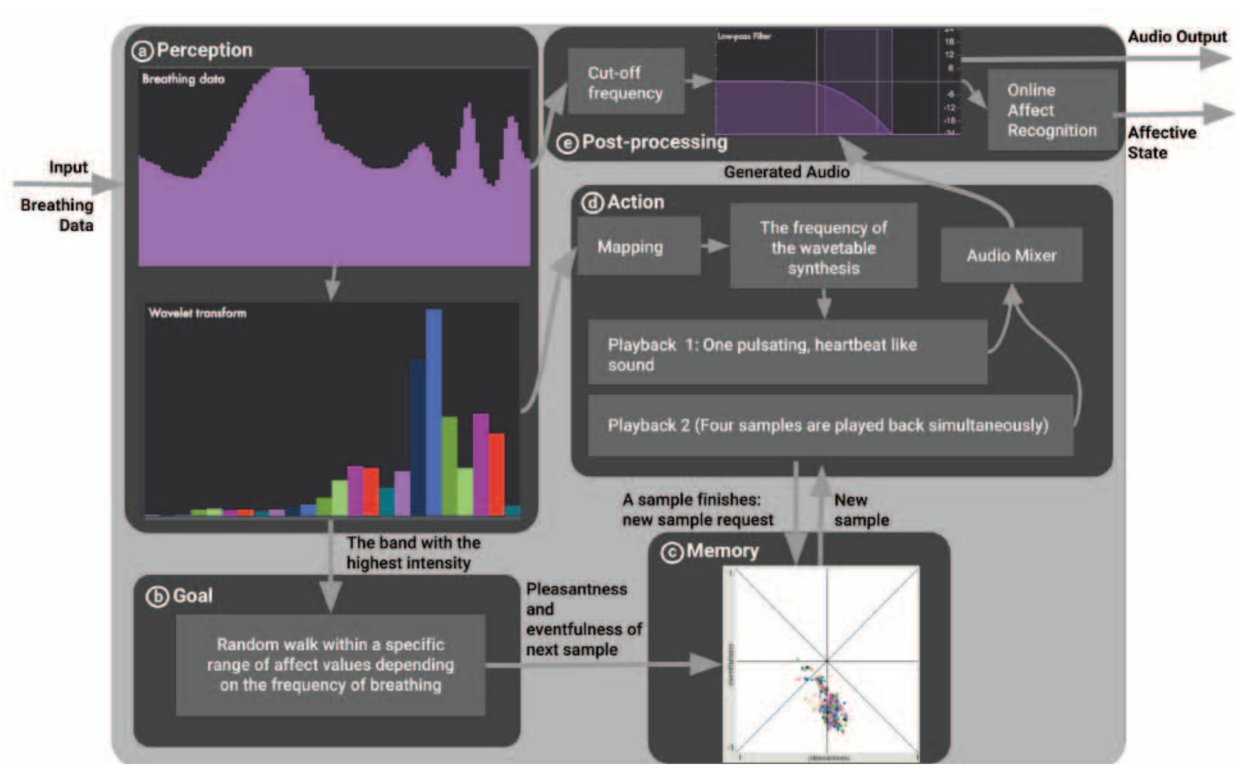

Fig. 4. The musical agent in *Respire*. (© Kıvanç Tatar)

as 0.03 Hz. As the window size increases, we introduce longer delays to the system. The wavelet transform addresses this by using different window sizes to calculate each band, which provides for the detection of sudden changes in the signal while calculating low frequencies. In our implementation, the wavelet minimum frequency is 0.03 Hz, the maximum frequency is 2 Hz, the carrier frequency is 0.06 Hz and there are 4 bands per octave. Hence, the output of the wavelet transform is the power of 24 bands. Lastly, the *perception* module outputs the band with the highest power.

Generative Algorithm of Musical Agent

The *goal* module (Fig. 4b) generates a vector with two dimensions, eventfulness and pleasantness, to select a sound from the *memory* (Fig. 4c) to be played by the *action* module (Fig. 4d). The *memory* module chooses the audio sample with the closest Euclidean distance to the 2D vector (pleasantness and eventfulness) generated by the *goal* module. The *goal* module generates the eventfulness values by using a mapping between the wavelet frequency with the highest power and the eventfulness of audio samples. The lowest and highest wavelet frequency bands are mapped to the lowest and highest eventfulness values of the agent's *memory*, respectively. Using these maximum values (Fig. 3), the frequency bands of breathing are mapped to 24 eventfulness values with equal distance. The agent applies a 2D random walk around one of these 24 eventfulness values so that the 24 wavelet bands correspond to 24 different areas in 2D affective space in Fig. 3.

The *action* module incorporates two playback engines (Fig. 4d). The first engine applies wavetable synthesis to generate a heartbeat-like sound. The prominent frequency of breathing data is mapped to the looping frequency of the wavetable synthesis. Hence, the pulsation sound in the sonic environment slows down as the user breathes more slowly. Similarly, the pulsation speeds up following acceleration in breathing. The lower-frequency bands slow down the pulsation sound to a point that the pulsation morphs into an ambient, pad-like sound. The second playback engine in the *action* module includes four voices. Each voice is active at all times. When the *goal* module chooses an audio sample, this playback engine plays the given sample. When a sample finishes, the action module requests a new sample from the *goal* and *memory* module.

We developed the *postprocessing* module to further enhance the interaction between the user and the audio environment by introducing a low-pass filter (Fig. 4e). The amplitude of breathing controls the cutoff frequency of this filter. The cutoff frequency increases as the user breathes in and vice versa. This mapping resembles the relation between musical pitch and movement [17]. Hence, as the user breathes in and out, the timbre of the sonic environment oscillates between a muddy, low-frequency-prominent audio environment and a full-spectrum audio environment. This mapping is meant to enhance the user's feeling of submersion.

Lastly, the agent applies an affect estimation algorithm to estimate the eventfulness and pleasantness of the generated audio environment (Fig. 4e). The estimation algorithm is the online version of the estimation algorithm that is used to label the audio in the agent's *memory*. The output of the online affect estimation is a vector with two dimensions: eventfulness and pleasantness. These values are further used to control the parameters of the virtual reality environment (see the System Details section and Fig. 2).

EXHIBITIONS

Respire is the continuation of our previous artwork *Pulse.Breath.Water*. Both artworks share the same system design, and we improved the visual environment in *Respire* by exploring different color schemes and virtual lights and adding fog to improve the ambiguous aesthetics. We presented *Pulse.Breath.Water* in three collective exhibitions. The first exhibition, *Scores + Traces: exposing the body through computation*, took place at the One Art Space gallery in Manhattan, NY [18], in March 2016. The theme of the exhibition was movement and computation, and the exhibition brought a new perspective on how to incorporate movement theories in computational arts.

After the *Scores + Traces* exhibition, we were invited by the Regina Miranda & Actors/Dancers Company for a collaboration to create and produce a piece titled *P.O.E.M.A* (Percursos Organizados Entre Movimentos Aleatórios; in English: Organized Paths among Aleatory Movements) for the cultural program at the 2016 Rio Olympics, which was shown for five weeks in summer 2016.

P.O.E.M.A is a choreographic installation that incorporates contemporary dance to *Pulse.Breath.Water* (see Figs 5–7). The piece was exhibited in a 10-×-10m room, and the

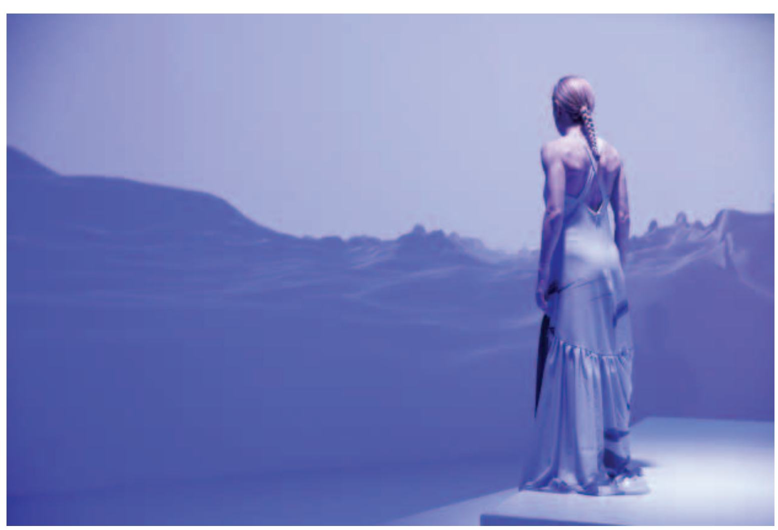

Fig. 5. A dancer in still position in *P.O.E.M.A*, 2016. (© Kıvanç Tatar. Photo © Adriano Fagundes.)

Fig. 6. A user with VR headset, a dancer and spectators in the exhibition space of *P.O.E.M.A*, 2016. (© Kıvanç Tatar. Photo © Adriano Fagundes.)

main challenge was creating a space that would bring the virtual environment of *Pulse.Breath.Water* and the dancers together. There would be several spectators in the space while one audience member was in the virtual environment (Fig. 6). The view of the user in the virtual environment was projected on one wall. To expand the 2D projection of the virtual environment to the 3D space of the dancers, we utilized a white colored space with a light design inspired by the work of James Turrell (Fig. 7). This light design blended the 2D projection of the virtual environment with the 3D space of dancers. The audio followed this approach of blending, and we expanded the stereo output of the sonic environment to quadrophonics by using spectral spatialization.

Three dancers joined the project and began performing one at a time. The dancers were changed daily to allow them rest after several hours of performance. The dancers interacted with the virtual environment using a set of 200 choreographic cells, snippets of movements. These choreographic cells were the dancers' vocabulary for reaction to the emerging behaviors initiated by the interaction between the user and the virtual environment. During changes of the user of the virtual environment, the visuals went to black, the lights were dimmed, the dancers remained in a static posture and there were two video loops of three dancers on the side walls.

The third exhibition of *Pulse.Breath.Water*, at the MUTEK Mixed Realities VR Exhibition (November 2016), included 40 virtual reality artworks covering interactive immersive environments, narrations within immersive environments and 360 videos [19]. We heard that *Pulse.Breath.Water* brought a new perspective to the exhibition by provoking the idea of exploring a virtual presence using breathing as a way of interacting with the immersive environment. Later, we revisited the visuals and created *Respire*. We exhibited this new version at the *CHI 2018* VR exhibition in Montreal in April 2018 [20] and *Digital Carnival 2018* in Vancouver in August 2018 [21].

Fig. 7. A dancer in movement in *P.O.E.M.A*, 2016. (© Kıvanç Tatar. Photo © Adriano Fagundes.)

FUTURE STEPS

The overall interaction design and the system architecture of *Respire* stand closer to reactivity in comparison to interactivity. Tatar and Pasquier [22] clarify that interactivity of agents involves proactivity, that is, planning future actions and interacting with other human, biological or artificial agents. A further step could be to research how the introduction of proactive behaviors would affect the ambiguity and meaning-making of *Respire*'s user experience. For example, in an interactive scenario, periods of prolonged rapid breathing could lead to a calming of the water surface and less musical agent activity, which in turn could lead to a change in the user's breathing. However, our previous exhibition experiences showed that the audience interacted with *Respire* in two main phases: a playful phase where the effect of the breathing and the sensor was explored, followed by a calmer phase where the user started moving less than in the initial phase. Although these findings are speculative, they initiate a direction for further research on reactive and interactive scenarios.

Acknowledgments

This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Social Sciences and Humanities Research Council of Canada (SSHRC).

References and Notes

1 K. Tatar, M. Prpa and P. Pasquier, *Respire*: www.kivanctatar.com/respire (accessed 26 September 2018).

2 W.W. Gaver, J. Beaver and S. Benford, "Ambiguity as a Resource for Design," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)* (New York: ACM, 2003) pp. 233–240.

3 E. Kandel, *Reductionism in Art and Brain Science: Bridging the Two Cultures* (New York: Columbia Univ. Press, 2016).

4 P. Dourish, *Where the Action Is* (MIT Press, 2004).

5 NIMkartchannel, *Sonia Cillari: As an Artist, I Need to Rest (2009)*: www.youtube.com/watch?v=GcA6Yw2QER0.

6 C. Davies, "OSMOSE: Notes on Being in Immersive Virtual Space," *Digital Creativity* **9**, No. 2, 65–74 (January 1998).

7 M. Slater, "Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments," *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **364**, 1535, 3549–3557 (December 2009).

8 P. Pasquier et al., "An Introduction to Musical Metacreation," *Computers in Entertainment* **14** (2017) pp. 1–14: www.doi.org/10.1145/2930672.

9 M. Wooldridge, *An Introduction to MultiAgent Systems* (Chichester, U.K.: Wiley, 2009).

10 K. Tatar and P. Pasquier, "Musical Agents: A Typology and State of the Art towards Musical Metacreation," *Journal of New Music Research* **48**, No. 1, 1–50 (2018).

11 J. Panksepp, *Affective Neuroscience: The Foundations of Human and Animal Emotions* (Oxford: Oxford Univ. Press, 1998).

12 K. Tatar and P. Pasquier, "MASOM: A Musical Agent Architecture Based on Self Organizing Maps, Affective Computing, and Variable Markov Models," in *Proceedings of the 5th International Workshop on Musical Metacreation (MUME 2017)* (Atlanta, GA, U.S.A.: MUME, 2017).

13 M. Prpa et al., "The Pulse Breath Water System: Exploring Breathing as an Embodied Interaction for Enhancing the Affective Potential of Virtual Reality," in *HCI International 2017, Proceedings* (Vancouver: Springer, 2017).

14 "M+M Architecture": *H+ Technology* www.hplustech.com/blogs/news/m-m-middleware (accessed 31 December 2018).

15 Tatar and Pasquier [10].

16 J. Fan, M. Thorogood and P. Pasquier, "Automatic Soundscape Affect Recognition Using A Dimensional Approach," *Journal of the Audio Engineering Society* **64**, No. 9, 646–653 (2016): www.doi.org/10.17743/jaes.2016.0044.

17 B. Sievers et al., "Music and Movement Share a Dynamic Structure That Supports Universal Expressions of Emotion," *Proceedings of the National Academy of Sciences* **110**, No. 1, 70–75 (2013).

18 "Scores+Traces: Exposing the body through computation," *One Art Space* (2016, March 10): www.oneartspace.com/2016/03/10/scorestraces-exposing-the-body-through-computation (accessed 16 July 2019).

19 "MUTEK_IMG Mixed Realities," MUTEK: www.mutek.org/img/2016/artworks (accessed 4 January 2018).

20 M. Prpa et al., "Respire: A Breath Away from the Experience in Virtual Environment," *CHI EA '18 Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal: ACM, 2018) pp. 1–6.

21 "Digital Carnival 2018," *Cinevolution Media*: www.cinevolutionmedia.com/digital-carnival-twines-ancestral-wisdom-and-digital-innovation-aug-31-sep-1-2018 (accessed 26 September 2018).

22 Tatar and Pasquier [10] p. 42.

Manuscript received 18 July 2018.

KIVANÇ TATAR *is "a worker in rhythms, frequencies and intensities," playing trumpet and electronics, composing experimental music, performing audiovisuals and researching creative artificial intelligence for music and interactive media. His career aims to integrate science, technology, engineering, interactive arts, contemporary arts and design to research interdisciplinary topics to create transdisciplinary knowledge.*

MIRJANA PRPA *researches immersive virtual environments that support mindfulness, embodied interaction and regulation of breathing and affective states.*

PHILIPPE PASQUIER *works on creative AI and generative systems. He is a scientist specialized in artificial intelligence, a multidisciplinary artist, an educator and a community leader.*