

# Real and virtual body percussionists interaction

Elisabetta Bevacqua  
ENIB, Lab-STICC  
25, rue Claude Chappe  
29200, Brest, France  
bevacqua@enib.fr

Gireg Desmeulles  
ENIB, Lab-STICC  
25, rue Claude Chappe  
29200, Brest, France  
desmeulles@enib.fr

## ABSTRACT

A first approach to an autonomous virtual agent able to play body percussion with real body percussionists is presented. The agent is autonomous in the sense that it can recognize the artists' *calls* and react to them by playing back a prerecorded sequence. The agent architecture is described focusing mainly on the artists' *calls* recognition module. This work, still under construction, produced two artistic performances which were presented in front of an audience.

## CCS CONCEPTS

•**Human-centered computing** → **Virtual reality**; *Sound-based input / output*; •**Applied computing** → *Performing arts*;

## KEYWORDS

Human-virtual agent interaction, body percussion, virtual reality, arts and sciences

### ACM Reference format:

Elisabetta Bevacqua and Gireg Desmeulles. 2017. Real and virtual body percussionists interaction. In *Proceedings of MOCO '17, London, United Kingdom, June 28-30, 2017*, 4 pages.  
DOI: <http://dx.doi.org/10.1145/3077981.3078048>

## 1 INTRODUCTION

Arts and sciences (A&S) projects face similar epistemological approaches, stimulate creativity and offer an original form of interdisciplinarity. When being combined with Virtual Reality (VR), artistic requirements and emphasis on sensitivity and emotions re-evaluate concepts of realism, credibility, immersion, co-presence... and ask for new technical requirements. The A&S project we present here mixes body percussion and Virtual Reality.

Body percussion is a musical genre which consists in playing rhythmic patterns generated by hitting, clapping and rubbing the human body. Various sounds with different intensity and tone compose infinite poly-rhythms combinations. Since the body is considered as an instrument, the body percussion allows to focus on embodied music. Artists we worked with use oral transmission to share and compose the musics/dances and most of the rhythms

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
*MOCO '17, London, United Kingdom*

© 2017 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
978-1-4503-5209-3/17/06...\$15.00  
DOI: <http://dx.doi.org/10.1145/3077981.3078048>

come from oral tradition. A sound is not only a note written on a music sheet or a key pressed on a keyboard, but it is considered as a gesture ([4]). In collaboration with the theatrical group Teki-toi<sup>1</sup>, who practices body percussion, we aim at implementing an autonomous virtual agent system able to play together with the artists. The final goal consists in providing the body percussionists a tool that they can use to produce artistic performances, involving virtual and real interacting dancers, in front of an audience. This ambitious goal requires to address different technical and scientific problems, such as rhythmic patterns recognition, virtual agent animation and the respect of the musical ear sensibility which constrains the whole system to a temporal precision which is shorter than 5 milliseconds. Furthermore, within this project we want also to explore a ternary interaction between artists, audience and virtual human. Such an interaction, mediated by vision and hearing, must emerge naturally during performance which can take place outside the research laboratory.

In the next section we provide some examples of scientific researches which find in arts a fertile ground for testing and application. Then, a description of our system is presented. Finally, two examples of applications which produced artistic performances are reported.

## 2 HUMAN-MACHINE INTERACTION IN ARTS & SCIENCES

Several research works which link A&S have been conducted in the last years, particularly since the appearance of more and more affordable sensor devices and software able to track and analyze human movements and gestures. Moreover, artistic domain provides a wide and heterogeneous set of possible applications in which complex subjects, such as human-machine real-time interaction, synchronization, multimodal signals generation and recognition, etc., can be developed and tested.

For example, the Gesture Follower [3], a real-time gesture recognition system, was used in an interactive music performance, the "augmented string quartet". In this performance the players' bowing gestures were recognized and synchronized with digital sounds during a concert [2].

Nijholt et al. [6] implemented an anticipation model allowing a virtual agent to generate its behavior according to the expected human behavior. The proposed architecture has been used in artistic contexts: for example, (i) a virtual dancer able to dance with a human (who moves on a dance pad) and to adapt to the rhythm of the music and to the user's movements; (ii) a virtual orchestra conductor able to conduct a group of musicians and to perceive their tempo in order to make them play faster or slower when they

<sup>1</sup>Stéphane Paugam, Florent Clause <http://www.vivrelemonde.fr>

do not respect the partition tempo; (iii) a virtual fitness trainer who can respect the music tempo and adapt its behavioral expressivity to that showed by the human.

Similarly, within the INGREDIBLE project<sup>2</sup>, Bevacqua et al. [1] studied bodily interaction between a virtual agent and a real user in an artistic context. They proposed a whole agent architecture from user's gestures capture, analysis and recognition to agent behavior decision and synthesis. This architecture was applied in two distinct contexts: (i) a theatrical exercise in which two players imitated each other's upper-body movements but introduced subtle changes by proposing, from time to time, new movements; (ii) a fitness session, which consisted of a virtual agent able to perform fitness movements with a human. The two interactants could vary their speed, for example, by accelerating to motivate each other, or they could adapt to the speed of the other, for instance, by slowing down if the partner could not keep the pace.

All the previous cited work show how linking A&S can be profitable to both domains by providing a wide set of possible applications. In the present work, collaborating with body percussionists provides us with an interesting application field where we can study user-agent interaction which must respect strong constraints, such as perfect synchronization and believable autonomous agent behavior. Moreover, the resulting system must be reliable enough to allow the artists to produce a performance that can be shown in front of an audience.

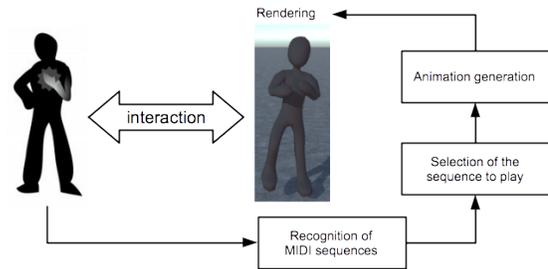
## 2.1 Dwende

The title of the artistic performance is "Dwende" relative to the Hispanic word "duende" which is used in flamenco dance to describe a particular disposition close to trance. When an artist, musician, dancer or singer, reaches this disposition, their performance becomes easy and fluid and does not need any effort. The notion of duende can be seen as related to the better known notion of *flow* [5]. This concept describes a mental state of intense concentration in which a person can be when they are fully engaged in an activity. Such a state, when attained, generates a deep satisfaction. The concept of flow, which is linked to motivation, is studied both in the artistic domain and in computer ergonomics. Particularly, in video games domain, it is considered as an important element of an engaging interaction between the human and the virtual environment. Artists, such as body percussionists or dancers, try to attain the flow in collective practice in order to create and transmit emotions. Within the artistic project presented in this paper, we look for the flow in order to evaluate the possibility of an accurate coupling and synchronization between human and virtual environment. Figure 1 shows the VR system architecture realized in this work.

## 3 MODEL DESCRIPTION

We define the behavior of virtual humans by copying the explicit rules applied during the collective practice of body percussion.

The musical/dance pieces are composed of several patterns which are a superposition of complementary rhythms, called a *poly-rhythm*. A body percussionist can, at any time, propose to change pattern by emitting a "call", that is a specific sequence of sounds which is known by all dancers. The tempo of the call sets the tempo



**Figure 1: System architecture. Human (left) interacts with virtual agent (right).**

of the following pattern. Thus, the artist proposing the call can modify the collective tempo by playing it slower or faster. In our system we want the agent to be able to recognize and respond correctly to the calls by playing the desired pattern. However, as a first approach and as required by the body percussionists, we want that just the artists can emit calls. So, during the performance very few improvisations happen, and they are all decided by the body percussionists. That means that:

- the agent does not improvise patterns;
- patterns are played if a call occurs or if it is planned;
- a pattern can be played over and over until a call occurs or it can be played for a predefined number of time.

## 3.1 States machine and fast prototyping

We can see each human (real or virtual) behave like a state machine sensible to occurrence of musical calls. For such a reason, we design the virtual human behavior through a finite state machine. Each state loops a rhythmic pattern (that is, a musical gesture) and a transition occurs if:

- the pattern has been played as many time as expected,
- a musical call is heard.

The states machine enables the interaction between real and virtual body percussionists during the performance. Based on the emitted calls, the agent has to synchronize its own machine with the others, tuning its tempo. When a call is recognized at the instant  $t_r$ , all players sensible to this call have to react at the instant  $t_s$ . The delay  $t_r - t_s$  has to be long enough to handle the transition (see figure 2). For example, playing at 120bpm, this delay usually last 0.5s (minus recognition time which is negligible in comparison with). According to the duration of the delay, different animations of the same pattern can be selected to allow a smoother transition between two consecutive patterns (see section 3.3 for more details).

According to our A&S purpose, an interaction has to be considered not only during the final performance but during all the previous creation steps. The artists compose music and choreography in action using a trial and error method. However, when such a composition must be done with a virtual agent, that is a computer application, each trial implies behavior coding time that ruins composition dynamics. To avoid such an interdisciplinary bottle neck which causes a loss of time, a fast prototyping tool is needed to enable interactive creation. That is a tool which allows

<sup>2</sup><http://www.ingredibile.fr>

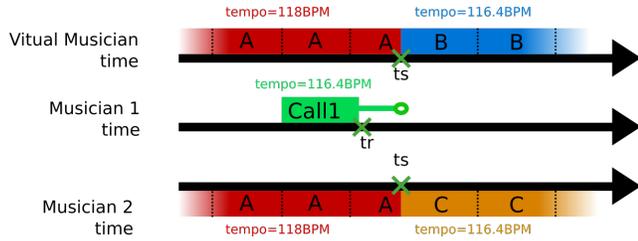


Figure 2: A, B and C represent patterns played by musicians. When a call is played, the system recognizes it at the instant  $t_r$ . All players sensible to this call have to react at the instant  $t_s$ , tuning their tempo.

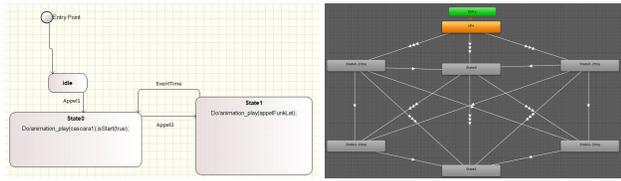


Figure 3: Behavior state machine (left) and Animator (right)

us to modify easily the state machine in the code. Through an UML modeling application such as Modelio<sup>3</sup>, we can change, remove or add new states and the tool we implemented translates the UML state machine both in the state machine used in our virtual agent application and in an *Animator Controller* (see figure 3). The latter is a state machine, provided by the game development platform Unity3D<sup>4</sup>, which determines and merges the animations that the agent must play at a given moment. We use Unity3D to display and animate the virtual agent in a 3D scene.

Our tool can generate extra states to improve animation overlap during transition according to the available delay  $t_s - t_r$  between two patterns (see section 3.3).

### 3.2 MIDI pattern recognition

Calls played by real humans have to be recognized to decide if transitions of virtual agent behavior can be achieved. During the performance, information acquisition and processing must be fast, robust and invisible to the spectator (see figure 2). Because of such constraints we decided to avoid sound or visual capture and recognition. Instead, we use MIDI musical instruments such as DRUMPANTS<sup>5</sup>. A drum pants is a wireless electronic drum-set hidden under the clothes. Each time a sensor is struck, a signal is sent to the online calls recognition algorithm. In this way, two-steps recognition is quite simple (see figure 4):

- (1) Identification of the sought-after pattern occurrence (see figure 2);
- (2) Evaluation of the distance between the captured pattern and the theoretical pattern.

<sup>3</sup><https://www.modelio.org>

<sup>4</sup><https://unity3d.com/fr>

<sup>5</sup><http://www.drumpants.com>

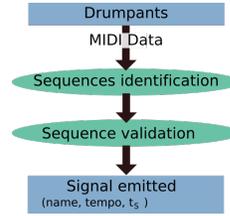


Figure 4: Data flow diagram for MIDI sequences recognition (see figure 1), from body percussionists to virtual human.

notes	states	output
a,12.200s	(a,12.200s)	
c,12.503s	(a,12.200s)	
a,12.706s	(a,12.200s) (a,12.200s)(a,12.706s) (a,12.706s)	
b,12.835s	(a,12.200s) (a,12.200s)(a,12.706s) <b>(a,12.200s)(a,12.706s)(b,12.835s)</b> (a,12.706s)	call1 rejected
b,13.698s	(a,12.200s) (a,12.200s)(a,12.706s) <b>(a,12.200s)(a,12.706s)(b13.698s)</b> (a,12.706s)	call1 accepted <i>tempo = 120bpm</i> $t_s = 14.200s$
b,15.000s		

Figure 5: Calls pattern recognition algorithm.

A call pattern is a sequence of couples (*instrument number, time*), where the time is expressed in number of beats after the beginning of the sequence. The figure 5 describes the calls recognition algorithm. In this example, the algorithm tries to recognize *call1* = (*a, 0beats*), (*a, 1beats*), (*b, 3beats*), ( $t_s$ , *5beats*). The last couple of the sequence corresponds to the instant the other players have to react at, it is not an instrument number to recognize. Each line of the table shown in figure 5 corresponds to the occurrence of a MIDI “note on” events (left column). The *States* column represents the in-building sequences. Each received instrument number is associated to the time it arrived at, expressed in seconds. A sequence tempo will be deduced from these values. A completed sequence may be rejected if its temporal distribution does not match that of the sought-after call. In-building sequences are flushed when a sequence is validated.

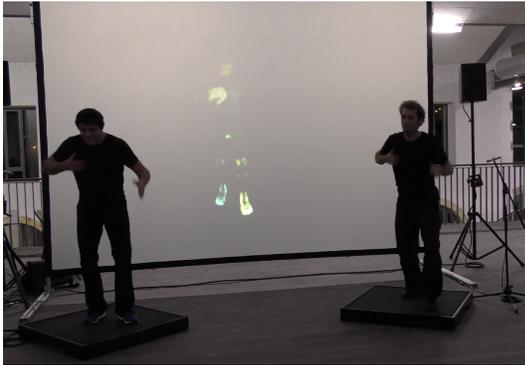
### 3.3 Virtual agent animation

The reception of a call must generate a state change which determines the next pattern that the agent has to play. The easiest solution, which is also the solution we are adopting at the moment, consists in triggering a prerecorded motion captured animation. We recorded all patterns performed at 120bpm by the artists. Each recording corresponded to a single pattern which was repeated several times. Then each repetition was cut and we selected 3 slightly different animations for each pattern: i) the first repetition in which the movement starts from the neutral position<sup>6</sup>; ii) a repetition in

<sup>6</sup>Neutral position: the artist stands still, his arms parallel to his body.

the middle; iii) the last repetition which ends in the neutral position. The first and the last animations are longer than the second one and they are played when the delay  $t_s - t_r$  between two patterns is long enough to allow the *Animator Controller* of Unity3D to merge them through a smoother transition.

The adopted solution suffers from obvious limitations: the played animation will be always exactly the same, which is going to be detrimental to the agent behavior believability. In fact, a body percussionist, as well as any other person, does not move always in the same way, their movement expressivity varies during time; moreover, while performing together, artists inevitably exert an influence on each other. So, the same pattern will not be performed always in the same manner. Such a natural behavior cannot be reproduced when prerecorded motion captures are used to animate a virtual agent. We consciously know that adding intra and inter expressivity variability to the agent behavior would improve the artists' perception of the agent while interacting with it [1]. The whole experience would be more believable and satisfying. For such a reason, we are currently working to improve the agent animation.



**Figure 6: The second performance. The body percussionists and the agent play together. The agent is represented as a set of shining particles.**

#### 4 ARTISTIC APPLICATIONS

A first version of the architecture was used in 2015 to produce a theatrical performance in front of an audience. Since that performance Unity 3D has been used as rendering. At that time, the virtual agent was driven by a “musician of OZ”, a real person which selected on a keyboard the agent behavior according to the body percussionists' calls. This pilot show allowed us to verify the feasibility of such a performance and the interest that it could generate in the audience. Even without a real agent-artist coupling, the public perceived the interaction between them and responded positively to the proposed performance. For such a reason, a second theatrical show, see Figure 6, was proposed in 2016 as soon as the agent was able to perceive and to respond autonomously to the artists' calls. The musician of OZ disappeared and, during the whole performance, the agent behavior was totally autonomous. Again, the audience responded positively and we and the body percussionists were quite proud of the result, however we noticed that the interaction between the agent and the artists took place and was much stronger during the

rehearsals than during the final performance. While rehearsing, the artists were shaping the performance together with the agent, interacting with it through calls very often to evaluate its responses. On the other hand, during the performance, the artistic dialogue between the agent and the percussionists was completely defined and they simply followed a script. This is a limitation we noticed while working with artists. While studying the interaction between the agent and the human, we look for autonomy and some time unpredictability, that is the agent should be free to act proactively and to respond in a way that the user does not expect, like in real life among people. From the artists point of view, such an autonomy is not suitable, since the final performance must unfold strictly as decided.

#### 5 CONCLUSIONS

In this paper we present a first approach of an autonomous virtual agent able to play body percussion with real artists. At present, the architecture is still under construction and the agent behavior is not as sophisticated as we would like it to be. It is autonomous in the sense that it can recognize the body percussionists calls and react to them, but it is missing a proactive behavior that would make it generate its own calls to modify the interaction as it pleases. Moreover, the agent behavior animation is still based solely on motion captured data and, as explained in section 3.3, this solution does not allow the agent to adapt the expressivity of its movements to that shown by the human, which does not improve the feeling of coupling between them. However, even with all this work still to be done, we were able, together with the body percussionists, to product two performances which were positively accepted by the audience.

#### REFERENCES

- [1] Elisabetta Bevacqua, Romain Richard, Julien Soler, and Pierre De Loor. 2016. INGREDIENT: A platform for full body interaction between human and virtual agent that improves co-presence. In *Proceedings of the 3rd International Symposium on Movement and Computing - MOCO'16*. Thessaloniki, GA, Greece.
- [2] Frédéric Bevilacqua, Florence Baschet, and Serge Lemouton. 2012. The Augmented String Quartet: Experiments and Gesture Following. *Journal of New Music Research* 1, 41 (2012), 103–119.
- [3] Frédéric Bevilacqua, Bruno Zamborlin, Anthony Sypniewski, Norbert Schnell, Fabrice Guédy, and Nicolas Rasamimanana. 2009. Continuous Realtime Gesture Following and Recognition. In *Proceedings of the 8th International Conference on Gesture in Embodied Communication and Human-Computer Interaction. GW'09*, Stefan Kopp and Ipke Wachsmuth (Eds.). Springer-Verlag Berlin Heidelberg, 73–84.
- [4] Jean Francois Billeter. 2012. *Un Paradigme*. Allia.
- [5] Mihály Csikszentmihályi. 1996. *Creativity: Flow and the Psychology of Discovery and Invention*. New York: Harper Perennial.
- [6] Anton Nijholt, Dennis Reidsma, Herwin van Welbergen, Riëks op den Akker, and Zsófia Ruttkay. 2008. *Mutually Coordinated Anticipatory Multimodal Interaction*. Springer Berlin Heidelberg, Berlin, Heidelberg, 70–89.