Engineering User Interfaces with Beat Gestures

André, Maxime; Bayet, Anthony; JETZEN, Tobias; LUYCX, Pierre; Cauz, Maxime; Dumas, Bruno

Published in:
EICS 2023 Companion - Companion of the 2023 ACM SIGCHI Symposium on Engineering Interactive Computing Systems

DOI:
10.1145/3596454.3597187

Publication date:
2023

Engineering User Interfaces with Beat Gestures

Maxime André
Anthony Bayet
Tobias Jetzen
Pierre Luycx
Maxime Cauz
Bruno Dumas
maxime.andre@unamur.be
anthony.bayet@student.unamur.be
tobias.jetzen@student.unamur.be
pierre.luycx@student.unamur.be
maxime.cauz@unamur.be
bruno.dumas@unamur.be
Univ. of Namur, Namur Digital Institute (NaDI) and Research Center on Information Systems Engineering (PReCISE)
Namur, Belgium

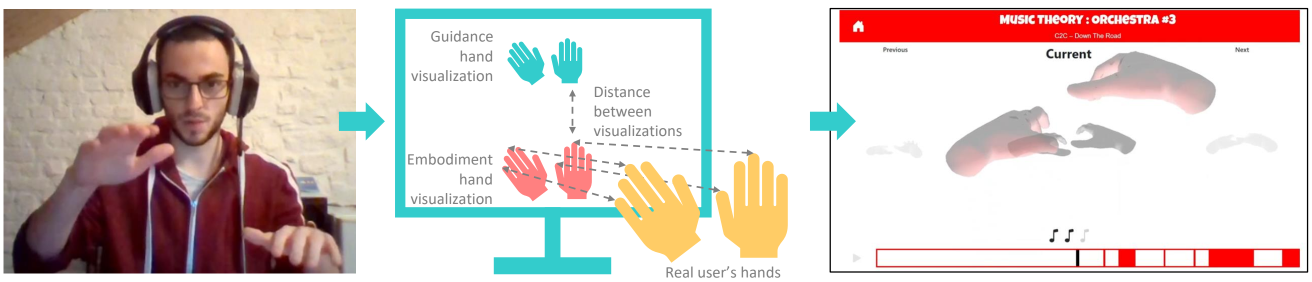

Figure 1: Beat gestures in GesturesHero: a game where the player learns musical rhythm while performing bimanual beat gestures captured by a Leap Motion Controller and rendered by embodiment and guidance virtual hand visualizations.

ABSTRACT

Beat gestures are biphasic up-and-down or back-and-forth movements of hand(s) that are associated with a specific meaning, such as in speech, or without, such as in rhythmic commands. Incorporating beat gesture recognition into user interface engineering involves dynamic recognition of hand pose, identification of movement direction, and calculation of beat number and frequency. We demonstrate a game that uses beat gestures for musical rhythm learning. We aim to understand the impact of real-time embodiment and guidance visualizations synthesizing user hands and gestures and considering the distance between virtual and real worlds.

CCS CONCEPTS

• **Human-centered computing** → **Gestural input**; **User interface programming**; *Graphical user interfaces*; Mixed / augmented reality; • **Hardware** → Tactile and hand-based interfaces.

EICS '23 Companion, June 27–30, 2023, Swansea, United Kingdom

ACM ISBN 979-8-4007-0206-8/23/06.
https://doi.org/10.1145/3596454.3597187

KEYWORDS

Gesture-based user interfaces; gesture recognition

ACM Reference Format:
Maxime André, Anthony Bayet, Tobias Jetzen, Pierre Luycx, Maxime Cauz, and Bruno Dumas. 2023. Engineering User Interfaces with Beat Gestures. In *Companion of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS '23 Companion), June 27–30, 2023, Swansea, United Kingdom.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3596454.3597187

1 MOTIVATIONS

Beat gestures are biphasic up-and-down or back-and-forth movements of the hand(s) [1], with the palm open or closed in any direction, usually towards the ground floor, the ceiling, or towards a certain target such as a surface or an object [12]. Beat gestures can be produced spontaneously without any particular meaning, such as to support the human discourse or communication (*e.g.*, up-and-down hand movement aligned with spoken clauses or when synchronized with some prosodic speech characteristics), or on purpose when a specific meaning is conveyed (*e.g.*, repeat an indication with some rhythm). Gestures are particularly useful when coupled with speech commands [8] and in multimodal interaction [3]. We believe that beat gestures are underestimated in gesture input.

On the one hand, beat gestures are well present in the literature [2, 9]. McNeil [7] structures gestures into iconics, metaphorics, deictic, and beats, which are motorically simple, rhythmic gestures that do not express semantic content but that instead align with the prosody of speech. Kendon [6] structures gestures along a continuum based on speech: gesticulation (beat, cohesives), language-like (iconic), pantomimes, emblems (deictic), and sign language (symbolic). They are among the most frequently encountered co-speech gestures in human communication [1].

On the other hand, engineering user interfaces with beat gestures has received little attention [4]. They could represent an iterative process across a sequence of data [12], an action or a command whose repetition or rhythm matters [5] such as repeating an action in time [4], to select/increase/decrease a value [14], or to indicate a direction [13]. The gesture rhythm indicates the iteration where a function is being applied to some data [12]. Beat gestures could also become deictic when the rhythm is applied to some designated object when they reference an object with a pace that represents a process [11]. McNeil [7] and Theune and Brandhorst [13] found that beats made up 45%, 30% of all gestures in their corpus for narration, for embodied conversational agents, respectively.

2 BEAT GESTURES IN USER INTERFACES

Freeman et al. [5] define *rhythmic gestures* as "touchless gestures that users repeat in a rhythmic manner, as a means of showing their intent to interact. This can help to reduce false-positive gesture recognition, as people are unlikely to perform the exact gesture in a repetitive manner, in time with the input rhythm". They compared five rhythmic gestures, *i.e.*, side-to-side, up-and-down, forward-and-backward, clockwise, and anticlockwise, in various conditions, in different conditions, but these gestures were basically represented through animation, not really implemented.

Incorporating beat gestures into user interface engineering involves identification of both hands (*i.e.*, unimanual and bimanual), dynamic recognition of hand pose (*i.e.*, palm open, palm close, fist, thumbs up), identification of movement direction (*i.e.*, left/right, up/down, backward/forward), and calculation of beat frequency.

To demonstrate this process, we developed GesturesHero, a video game encouraging a single player to perform at particular song milestones a defined gestural dance over a given music rhythm. While seated in front of a computer screen (Figure 1), the player performs beat gestures with one hand or two hands about 10-15 cm above a Leap Motion Controller (LMC, [10]). Based on the Leap Motion development kit (LeapJS), our JavaScript application receives frames data and invokes custom-developed gesture recognizers [10] which check the indicators (*i.e.*, palm normal, hand axes, pinch index, etc.) and validate or invalidate gestures on the basis of empirically defined rules. Checks are effective while the player is making a gesture. Moving hands over the LMC generates new frames to process according to gesture recognizers state machines where the initial state represents the invocation and the final state the gesture recognition result. Each gesture is divided into substates called *gestures part*, such as elementary gesture hand poses and movements. Each hand is processed separately by a distinct instance of gesture part. For example, a gesture part is the trajectory of the hand when it moves up and down. Another gesture part is the position when the hand must stay still. Each gesture part contains logical conditions to be satisfied for transitioning to the next state, the next gesture part. The logical condition becomes true when indicators' value extracted from a frame confirms the expected value of the gesture part (established on trial-error) while respecting an imposed timing. For example, the following rule contains three parts:

$$\neg handOpenedAgain$$
$$\land(currentPinchStrength \geq previousPinchStrength)$$
$$\land(currentPinchStrength < nearHandClosed)$$

To provide the end user with the embodiment, the sense of ownership, and the guidance on beat gestures despite the distance between the real and virtual worlds, two hand representations are used. The *embodiment hand visualization* is a real-time 3D representation of the hands projected in the user interface to match it with the *guidance hand visualization*, a 3D synthesized representation of the beat gesture. The *embodiment hand visualization* is proposed thanks to RiggedHands and Three.js once they are connected to the LMC flow. The *guidance hand visualization* is an animated GIF, pre-recorded with Leap Recorder from its definition. It is played in the background on time when meeting the corresponding song milestones. The application displays the current, next, and previous gestures the player must complete. This carousel helps him to prepare the next gesture, adapt the current one, and understand eventual errors of the previous one. Two layout types of hand visualization are defined: *distant* when maintaining the difference between real and virtual gestures avoids overlapping or *placeholder* when matching both gestures is desired. Both layers are present in the application and chosen by the player according to his feeling of embodiment and his sense of ownership.

3 CONCLUSION

We demonstrate the implementation of a video game intended for musical rhythm learning incorporating beat gestures recognition. Its uniqueness lies in its use of LMC for gesture recognition and visualizations of the hands and gestures. Following the embodiment approach, the player is guided to perform a gestural dance and has the opportunity to learn and adapt beat gestures. Early results are positive and encourage looking forward. Beat gestures and related visualizations could, for instance, be of interest in the field of medical rehabilitation.

OPEN SCIENCE

Our GitHub repository with code is accessible at https://github.com/GesturesHero/GesturesHero. A demonstration video is also visible at https://www.youtube.com/watch?v=QPwSEiEQw2Y

ACKNOWLEDGMENTS

The authors of this paper are very grateful to anonymous EICS reviewers and the associate chair for their insightful and constructive comments on earlier versions of this manuscript. The authors also acknowledge Jean Vanderdonckt for his accurate advice.

REFERENCES

[1] Hans Rutger Bosker and David Peeters. 2021. Beat gestures influence which speech sounds you hear. *Proceedings of the Royal Society B: Biological Sciences* 288, 1943 (jan 2021), 20202419. https://doi.org/10.1098/rspb.2020.2419 arXiv:https://royalsocietypublishing.org/doi/pdf/10.1098/rspb.2020.2419

[2] Paul Bremner and Ute Leonards. 2015. Speech and Gesture Emphasis Effects for Robotic and Human Communicators: A Direct Comparison. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (Portland, Oregon, USA) *(HRI '15)*. Association for Computing Machinery, New York, NY, USA, 255–262. https://doi.org/10.1145/2696454.2696496

[3] Diana Dimitrova, Mingyuan Chu, Lin Wang, Asli Özyürek, and Peter Hagoort. 2016. Beat that Word: How Listeners Integrate Beat Gesture and Focus in Multimodal Speech Discourse. *Journal of Cognitive Neuroscience* 28, 9 (Sept. 2016), 1255–1269. https://doi.org/10.1162/jocn_a_00963

[4] Euan Freeman, Stephen A. Brewster, and Vuokko Lantz. 2016. Do That, There: An Interaction Technique for Addressing In-Air Gesture Systems. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, May 7-12, 2016*, Jofish Kaye, Allison Druin, Cliff Lampe, Dan Morris, and Juan Pablo Hourcade (Eds.). ACM, 2319–2331. https://doi.org/10.1145/2858036.2858308

[5] Euan Freeman, Gareth Griffiths, and Stephen A. Brewster. 2017. Rhythmic micro-gestures: discreet interaction on-the-go. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction, ICMI 2017, Glasgow, United Kingdom, November 13 - 17, 2017*, Edward Lank, Alessandro Vinciarelli, Eve E. Hoggan, Sriram Subramanian, and Stephen A. Brewster (Eds.). ACM, 115–119. https://doi.org/10.1145/3136755.3136815

[6] Adam Kendon. 1988. How gestures can become like words. In *Cross-cultural perspectives in nonverbal communication*, F. Poyatos (Ed.). Hogrefe & Huber Publishers, 131–141. https://psycnet.apa.org/record/1992-98173-004

[7] D. McNeil. 1995. *Hand and Mind: What Gestures Reveal About Thought.* The University of Chicago Press, Chicago, IL, USA.

[8] Vik Parthiban, Pattie Maes, Quentin Sellier, Arthur Sluÿters, and Jean Vanderdonckt. 2022. Gestural-Vocal Coordinated Interaction on Large Displays. In *Companion of the 2022 ACM SIGCHI Symposium on Engineering Interactive Computing Systems* (Sophia Antipolis, France) *(EICS '22 Companion)*. Association for Computing Machinery, New York, NY, USA, 26–32. https://doi.org/10.1145/3531706.3536457

[9] Pilar Prieto, Alice Cravotta, Olga Kushch, Patrick Louis Rohrer, and Ingrid Vilà-Giménez. 2018. Deconstructing beat gestures: a labelling proposal. In *Proceedings of 9th International Conference on Speech Prosody* (Poznań, Poland) *(SpeechProsody '18)*, Katarzyna Klessa, Jolanta Bachan, Agnieszka Wagner, Maciej Karpiński, and Daniel Śledziński (Eds.). Internation Speech Communication Association, New York, NY, USA, 201–205. https://doi.org/10.21437/SpeechProsody.2018-41

[10] Arthur Sluÿters, Mehdi Ousmer, Paolo Roselli, and Jean Vanderdonckt. 2022. QuantumLeap, a Framework for Engineering Gestural User Interfaces Based on the Leap Motion Controller. *Proc. ACM Hum.-Comput. Interact.* 6, EICS, Article 161 (jun 2022), 47 pages. https://doi.org/10.1145/3532211

[11] Arthur Sluÿters, Quentin Sellier, Jean Vanderdonckt, Vik Parthiban, and Pattie Maes. 2022. Consistent, Continuous, and Customizable Mid-Air Gesture Interaction for Browsing Multimedia Objects on Large Displays. *International Journal of Human–Computer Interaction* 0, 0 (2022), 1–32. https://doi.org/10.1080/10447318.2022.2078464 arXiv:https://doi.org/10.1080/10447318.2022.2078464

[12] Amber Solomon, Mark Guzdial, Betsy DiSalvo, and Ben Rydal Shapiro. 2018. Applying a Gesture Taxonomy to Introductory Computing Concepts. In *Proceedings of the 2018 ACM Conference on International Computing Education Research* (Espoo, Finland) *(ICER '18)*. Association for Computing Machinery, New York, NY, USA, 250–257. https://doi.org/10.1145/3230977.3231001

[13] Mariët Theune and Chris J. Brandhorst. 2009. To Beat or Not to Beat: Beat Gestures in Direction Giving. In *Proceedings of 8th International Gesture Workshop, Gesture in Embodied Communication and Human-Computer Interaction, GW 2009* (Bielefeld, Germany) *(Lecture Notes in Computer Science, Vol. 5934)*, Stefan Kopp and Ipke Wachsmuth (Eds.). Springer, 195–206. https://doi.org/10.1007/978-3-642-12553-9_17

[14] Santiago Villarreal-Narvaez, Arthur Sluÿters, Jean Vanderdonckt, and Efrem Mbaki Luzayisu. 2022. Theoretically-Defined vs. User-Defined Squeeze Gestures. *Proceedings of the ACM Human-Computer Interaction* 6, ISS, Article 559 (nov 2022), 30 pages. https://doi.org/10.1145/3567805