GIM3D: a Pilot Study of Intuitive 3D Motion Control HMI

Marine Capallera, Antoine Lestrade, Dimitri Kohler, Elena Mugellini, Benoît Le Callennec, Benoît Le

GIM3D: a Pilot Study of Intuitive 3D Motion Control HMI
GIM3D : une Etude Pilote d'IHM de Contrôle de Mouvements 3D Intuitifs

Marine Capallera
marine.capallera@hefr.ch
HumanTech Institute, University of Applied Sciences and Arts of Western Switzerland // HES-SO
Fribourg, Switzerland

Antoine Lestrade
HE-Arc, University of Applied Sciences of Western Switzerland // HES-SO
Saint Imier, Switzerland
antoine.lestrade@he-arc.ch

Dimitri Kohler
HE-Arc, University of Applied Sciences of Western Switzerland // HES-SO
Saint Imier, Switzerland
dimitri.kohler@he-arc.ch

Elena Mugellini
HumanTech Institute, University of Applied Sciences and Arts of Western Switzerland // HES-SO
Fribourg, Switzerland
elena.mugellini@hefr.ch

Benoît Le Callennec
HE-Arc, University of Applied Sciences of Western Switzerland // HES-SO
Saint Imier, Switzerland
benoit.lecallennec@he-arc.ch

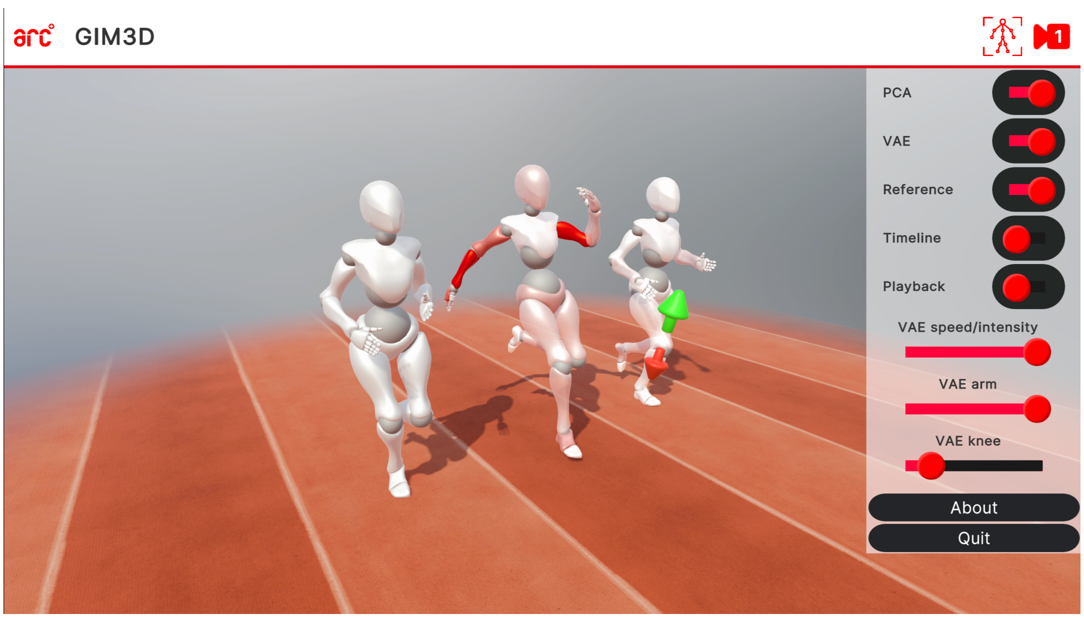

Figure 1: Interface and controllers of GIM3D application - final prototype

Abstract

In many fields such as sports training, rehabilitation, and virtual prototyping, the ability to communicate movement information rapidly, intuitively, and unambiguously is essential. In sports, a key challenge for coaches is conveying precise motion adjustments to athletes. This short paper introduces **GIM3D**, a user interface designed to enable fast and intuitive manipulation of 3D running movements generated. A pilot study with five participants shows a strong willingness to use the tool (F-SUS score = 85.5). The UEQ-S resulted an overall average of 2, with higher pragmatic quality (2.35) than hedonic (1.65). Users found the tool supportive, clear, and efficient. Initial results demonstrate the efficiency and plausibility of the generated movements, highlighting the potential of **GIM3D** as a tool for both practitioners and researchers in human motion science.

IHM '25, Toulouse, France
2025. ISBN XXX-X-XXXX-XXXX-X/XX/XX

Résumé

Dans de nombreux domaines tels que l'entraînement sportif, la rééducation et le prototypage virtuel, il est essentiel de pouvoir communiquer des informations sur les mouvements rapidement, intuitivement et sans ambiguïté. Dans le domaine du sport, l'un des principaux défis pour les entraîneurs consiste à transmettre aux athlètes des ajustements précis de leurs mouvements. Cet article présente **GIM3D**, une interface utilisateur conçue pour permettre la manipulation rapide et intuitive de mouvements générés de course en 3D. Une étude pilote menée auprès de cinq participant·es montre une forte volonté d'utiliser l'outil (score F-SUS = 85,5). L'UEQ-S a donné une moyenne globale de 2, avec une qualité pragmatique (2,35) supérieure à la qualité hédonique (1,65). Les utilisateurs·trices ont trouvé l'outil utile, clair et efficace. Les premiers résultats démontrent l'efficacité et la plausibilité des mouvements générés, soulignant le potentiel de **GIM3D** en tant qu'outil pour les praticien·es et les chercheurs·euses en science du mouvement humain.

CCS Concepts

• **Human-centered computing** → **User studies**; *Visualization techniques*; • **Computing methodologies** → Machine learning approaches.

Keywords

3D animation, 3D movement, Machine Learning, User Interface, Pilot Study

Mots clés

Animation 3D, Mouvement 3D, Machine learning, Interface Utilisateur, Etude Pilote

Reference Format:

Marine Capallera, Antoine Lestrade, Dimitri Kohler, Elena Mugellini, and Benoît Le Callennec. 2025. GIM3D: a Pilot Study of Intuitive 3D Motion Control HMI. In *IHM '25 : Actes étendus de la 36e conférence Francophone sur l'Interaction Humain-Machine, Novembre 4–7, 2025, Toulouse, France*. 5 pages.

1 Introduction

Efficient and accurate communication of movement-related information is crucial in many fields, such as sports training, rehabilitation and virtual design. In the sports context in particular, one of the major challenges for trainers is to convey subtle technical adjustments to athletes. To address this challenge effectively, it's important to understand how athletes acquire and refine their motor skills. As described by Taylor et. al [16], motor learning progresses through three stages:

(1) **Cognitive**: learner understands what to do
(2) **Motor**: learner performs the movement and improves how to do it
(3) **Autonomous**: the movement has become a reflex

In the first two stages, the way in which information about the movement is transmitted is crucial. In particular, we need to be able to communicate this information intuitively, quickly and unambiguously to maximize progress and ensure rapid, continuous learning. However, current tools rely mainly on indirect modes of communication - speech, text or video - which frequently introduce an element of ambiguity, particularly when it comes to fine-tuning gesture corrections.

In response to this limitation, we present **GIM3D**, a system and user interface (UI) designed to enable the rapid and intuitive generation of 3D racing movements. Drawing on advanced statistical models such as Principal Component Analysis (PCA) and Conditional Variational Auto-Encoders (CVAE), **GIM3D** can be used to synthesize customizable movement sequences.

This work in progress represents a first step toward a larger vision: enabling direct, unambiguous communication of 3D movement in real time, as seamlessly as spoken or written language. Future development will focus on integrating **GIM3D** into real-world coaching scenarios, expanding the range of supported movement types, and refining the interactivity and user control of the system.

2 Related Work

Several tools already exist to help sports coaches improve their athletes' performance. The most common approach is to use video analysis softwares such as Dartfish[1] and Kinovea[2]. The first step is to film the athlete in action. These programs can then slow down the video, advance frame-by-frame, and draw angles, trajectories, and annotations on the screen. In this way, any flaws in the movement can be highlighted and communicated to the athlete. Another method is to use motion capture tools to create a 3D representation of the athlete. This enables movement to be viewed from all angles, and kinematic data such as speed, acceleration, and joint angles to be extracted more accurately [14]. However, these tools require specific equipment and a controlled environment. These two techniques focus solely on the kinematic parameters of the movement. In addition, although the second approach creates a 3D model of the athlete, it does not allow interaction with the latter. Such interaction would enable movements to be generated in accordance with semantic parameters, such as "stand up", "speed up" or "rise the knees".

Regarding intuitive controls for 3D motion generation, Guay et al. generate 3D motion by drawing action lines [5] and then spatio-temporal curves [6]. All these techniques are based on the kinematic parameters of 3D motion, not on style or general form. Similarly, Harish [7] and Mahmudi [8] projects use kinematic trajectories drawn by users for 3D movement generation for artists. Moreover, regarding the visualization of important 3D movements parameters, Assa et al. [1] present a new technique to abstract the important information in a 3D movement and represent it with a limited number of images. Goldman et al. [3] summarize an entire video in a single image. To the best of our knowledge, there is no work available that visualizes variations in the general shape of a 3D movement (style) in a relevant and meaningful way.

On another hand, generative motion synthesis has been an active field of research. Traditional approaches are based on kinematic and dynamic models, to which deterministic algorithms such as linear interpolation or principal component analysis are applied. With the development of deep learning, motion synthesis methods have evolved. Mourot et al. [9] publish a literature review on the use of neural networks in skeletal animation. Among more recent approaches, Tessler et al. [17] propose a generative antagonistic networks, to create 3D character animations by supplying a sequence

[1] https://www.dartfish.com/
[2] https://www.kinovea.org/

of semantic instructions to the model. Tang et al. [15] present an approach capable of generating motion transitions between two poses in real time, based in particular on a Conditional Variational Auto-Encoder (CVAE) [13]. Also, based on CVAEs, Chen et al. [2] introduce a solution that generates various potential motion sequences in real time from a user's actions.

3 GIM3D design and functionalities

GIM3D (for *Génération Intuitive de Mouvements 3D*) lays the technological foundations for using 3D motion as a means of direct communication. More specifically, we have created a touch-screen UI (Android and Windows) application that enables:

- intuitively generate 3D running movements: using advanced statistical models such as Principal Component Analysis (PCA) and Variational Autoencoders (VAE);
- display generated 3D movements with relevant visual cues: **GIM3D** uses visual cues superimposed on 3D movements to ensure efficient transmission of information between runners and trainers.

Interface design and development follow a user-centered approach. A first meeting with a running coach took place at the Swiss BioMotion Lab in CHUV (Lausanne) to observe and note the advice given to an athlete running on a treadmill. The main dimensions mentioned by the coach were foot placement, pelvis position (too low, too far back...), knee height, torso (straightened, forward), arm spread (angle, distance) and finally head (gaze). Of course, the dimensions can be linked together. For example, to lift the knees, you can straighten the torso. Moreover, they depend on the individual and his/her muscular strength. In agreement with the coach, only three parameters will be modeled first in order to evaluate the relevance of the visualization and controllers in the first instance.

The choice of touchscreen interfaces is twofold. Firstly, they offer several advantages in terms of ease of use, thanks to direct control [10, 12, 18]. In addition, this choice was also discussed with the coach, who could see an extension with the possibility of drawing on the interface.

To generate real-time 3D movements, a database of running movements from the Mixamo library[3] is used. These data served as the basis for two types of advanced statistical models implemented in Python and easily parameterized by the user. This data is then integrated directly into the UI application for display. **GIM3D** can be used on a touchscreen device as an Android application (tablet) or Windows executable (touch-screen laptop).

The main functionalities of the touch-screen UI developed with Unity are the following (Figure 2):

(1) Open/close menu and show/hide interface
(2) Select manikins (PCA and/or VAE models) and the "ideal" run manikin
(3) Modify the course of VAE and PCA manikins using controllers (3 in this proof-of-concept to simplify handling: elbow distance, knee height, trunk)
(4) Display comparison in movement difference intensity
(5) Select pre-selected point of view camera (front, profile, 3/4 orientation)
(6) Personalize camera view and zoom in/out
(7) Control animation (pause, advance/rewind frame by frame)
(8) Exit the application

This interface was used in a pilot study described in the following section.

4 User tests

The objective of this pilot study is to evaluate the user experience and usability of the prototype. It should be noted that the choice of model between PCA and VAE was not part of the evaluation, although functional testing and model evaluation were carried out in parallel. Similarly, controller semantics were not directly evaluated. Participants were asking to interact freely with the three available controllers and interface components.

Five participants tested the interface (one coach, two athletes, two people who did not consider themselves athletes). On average, runners cover between 9 and 14 km twice a week or 20 km three times a week. The majority were familiar with technology (3/5 on a Likert scale with 1 "not familiar at all" and 5 "very familiar). One person considered themself unfamiliar and one person rather familiar with technologies.

The pilot study was being carried out using several complementary evaluation methods such as think aloud throughout the test period, behavioral observation (use of controllers during pause or animation, ease/difficulty of interacting with controllers, familiarization with the interface, etc.). Participants interacted with **GIM3D** using a touchscreen computer in tablet format (HP EliteBook 14"). The test lasted a maximum of 20 minutes. It proceeded as follows:

- Explanation of project context and purpose of user testing [5']
- Iteraction with the interface and functionalities [5'-10']
- Questionnaires (UEQ-S [11] and F-SUS [4]) and discussion [5']

5 Results

This section summarizes the main results of the pilot study regarding usability, user experience, and subjective feedback.

Usability The average F-SUS score is (M=85.5; std = 8.91). This means that people who have tested the application would be willing to use it. However, it is important to note that few people tested the prototype, but this shows that the functionalities implemented are satisfactory for this proof-of-concept version.

User Experience The user experience (from UEQ-S) has an average of 2 (on a scale between -3 and 3), with an average pragmatic quality of 2.35 and an average hedonic quality of 1.65. From a pragmatic point of view, participants found the application rather easy to use and clear (2.6), supportive (2.2) and rather efficient (2.0). These tests were designed to evaluate the pragmatic aspect of the application in the first instance. From a hedonic point of view, the solution was judged to be inventive (2.0), interesting and rather exciting (1.8), and slightly leading edge (1.0).

Feedback Participants shared some feedback while using the interface and after interacted with:

- If I could, I think I could use it to self-correct

[3] https://www.mixamo.com/

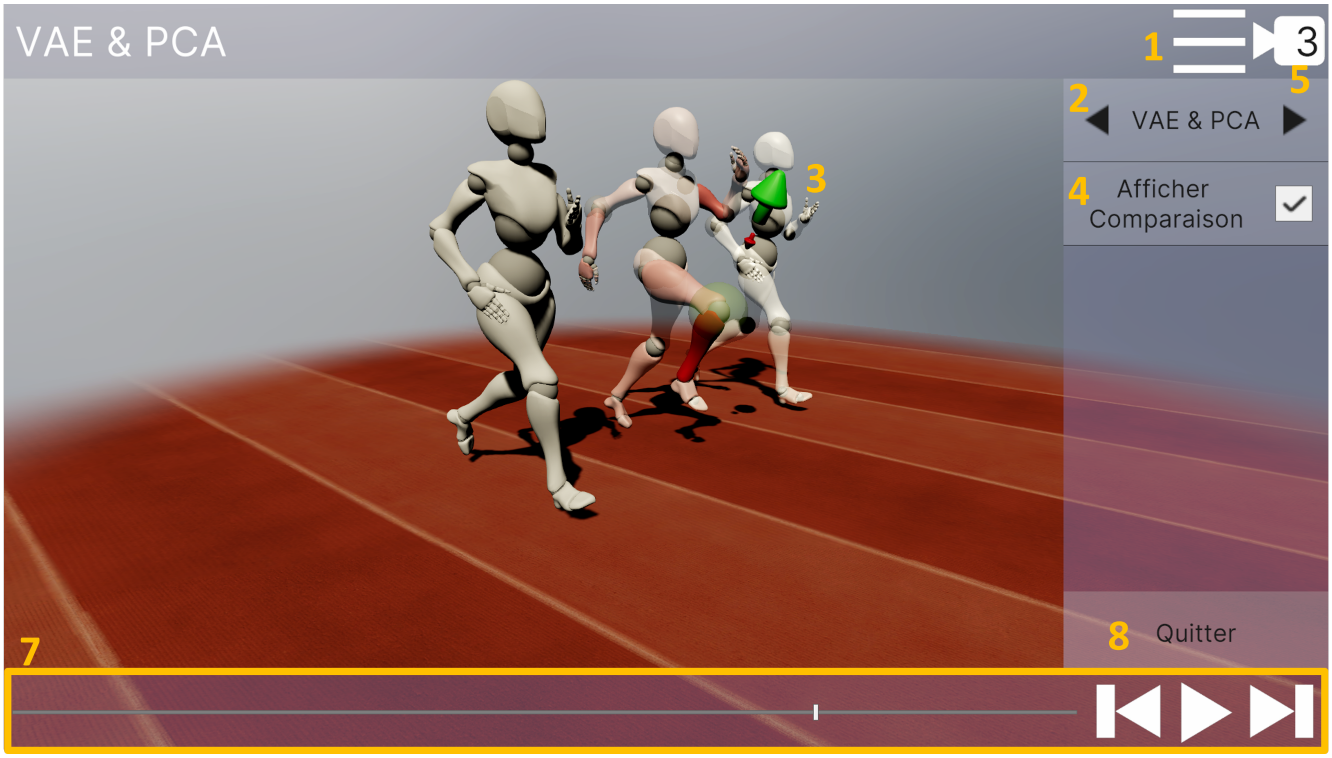

Figure 2: First version of UI used for pilot study

- It would also be interesting to be able to save sessions for learning: know how many times I have to do it to correct myself
- The coach needs to be trained to better understand the behavior of the controllers, and what the manipulations affect.
- Simple to use, it's helpful to have the ideal race next to me, because I learn a lot from models and comparisons - it helps with calibration
- For me, it's not relevant without the coach unless I can have data on efforts (or be able to capture this data)
- I think it would be very interesting for someone just starting out. It's important to have [a coach] at the start, but you could do without the coach afterwards.

Additional observations Participants manipulated the controllers mainly when paused, but also when running to "match" the animation to the ideal run. They used the controllers a lot, a bit "groping" to understand the controllers' behaviors. One participant wanted to model his own run with the configurable manikin, so as to be able to compare himself with the "ideal" run. All participants appreciated the presence of the "ideal" run manikin to be able to compare the two manikins. "The comparison is really useful". Finally, several participants would like to have other metrics, such as being able to tilt the ground, change stride, speed and effort indicators (linked to breathing, heartbeat, etc.), or to have other types of run, such as sprinting.

6 Conclusion

This short paper introduces **GIM3D**, an intuitive 3D motion generation interface. Although the current state of our system does not yet allow us to work on the actual movement performed by the athlete, it does allow us to control a 3D manikin performing a run. Real movement analysis and generation is currently done with video or motion capture tools (e.g., [14]) or, generative approaches using deep learning for realistic motion synthesis [2, 15, 17] that achieve high accuracy or realism, but often lack intuitive interaction mechanisms.

However, research on intuitive control of motion, such as sketch-based trajectories [5–8], has demonstrated the potential of higher-level interactions, but these approaches remain mostly limited to artistic or animation contexts rather than sports and performance analysis. Similarly, works on motion visualization [1, 3] have emphasized abstraction and summarization, but without addressing the real-time generation and control of new movements.

In this context, **GIM3D** aims to bridge these gaps. It extends the idea of using intuitive controls to generate 3D movements in real time [7, 8]. It combines intuitive controls with the real-time generation of 3D motion.

A pilot study shows that both the prototype and the vision are promising. Moreover, this project is part of an ambitious longer-term vision to measure and analyze 3D movements in real time, in order to communicate crucial information automatically and live. It is not limited to sports, but also targets all fields requiring the generation and display of 3D movements. Several parts of the

Python package developed during **GIM3D** would be of interest to the scientific community. The package has been developed in such a way as to make it easy to use and integrate. It will be reused in future research projects. In particular, further work is needed to provide tutorials and relevant and varied examples of use to increase its usability.

Acknowledgements

The authors would like to thank all the people who contributed to this paper, design of the experiment, implementation of the interface. The authors would also like to thank Mr. Baptiste Ulrich and Mr. Julien Favre from the Swiss BioMotion Lab (CHUV) for their collaboration, insights and feedback.

References

[1] Jackie Assa, Yaron Caspi, and Daniel Cohen-Or. 2005. Action synopsis: pose selection and illustration. *ACM Trans. Graph.* 24, 3 (July 2005), 667–676. doi:10.1145/1073204.1073246

[2] Rui Chen, Mingyi Shi, Shaoli Huang, Ping Tan, Taku Komura, and Xuelin Chen. 2024. Taming Diffusion Probabilistic Models for Character Control. In *ACM SIGGRAPH 2024 Conference Papers* (Denver, CO, USA) *(SIGGRAPH '24)*. Association for Computing Machinery, New York, NY, USA, Article 67, 10 pages. doi:10.1145/3641519.3657440

[3] Dan B Goldman, Brian Curless, David Salesin, and Steven M. Seitz. 2006. Schematic storyboarding for video visualization and editing. *ACM Trans. Graph.* 25, 3 (July 2006), 862–871. doi:10.1145/1141911.1141967

[4] G. Gronier and A. Baudet. 2021. Psychometric evaluation of the F-SUS: Creation and validation of the French version of the System Usability Scale. *International Journal of Human-Computer Interaction* (2021). doi:10.1080/10447318.2021.1898828

[5] Martin Guay, Marie-Paule Cani, and Rémi Ronfard. 2013. The line of action: an intuitive interface for expressive character posing. *ACM Trans. Graph.* 32, 6, Article 205 (Nov. 2013), 8 pages. doi:10.1145/2508363.2508397

[6] Martin Guay, Rémi Ronfard, Michael Gleicher, and Marie-Paule Cani. 2015. Space-time sketching of character animation. *ACM Trans. Graph.* 34, 4, Article 118 (July 2015), 10 pages. doi:10.1145/2766893

[7] Pawan Harish, Mentar Mahmudi, Benoît Le Callennec, and Ronan Boulic. 2016. Parallel Inverse Kinematics for Multithreaded Architectures. *ACM Trans. Graph.* 35, 2, Article 19 (Feb. 2016), 13 pages. doi:10.1145/2887740

[8] Mentar Mahmudi, Pawan Harish, Benoît Le Callennec, and Ronan Boulic. 2016. Sketch-based per-frame inverse kinematics. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation: Posters* (Zurich, Switzerland) *(SCA '16)*. Eurographics Association, Goslar, DEU, Article 3, 1 pages.

[9] Lucas Mourot, Ludovic Hoyet, François Le Clerc, François Schnitzler, and Pierre Hellier. 2022. A Survey on Deep Learning for Skeleton-Based Human Animation. *Computer Graphics Forum* 41, 1 (2022), 122–157. doi:10.1111/cgf.14426 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14426

[10] Cuong Nguyen, Yuzhen Niu, and Feng Liu. 2014. Direct manipulation video navigation on touch screens. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services* (Toronto, ON, Canada) *(MobileHCI '14)*. Association for Computing Machinery, New York, NY, USA, 273–282. doi:10.1145/2628363.2628365

[11] M. Schrepp, A. Hinderks, and J. Thomaschewski. 2017. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *IJIMAI 4* 6 (2017), 103–108.

[12] Andrew Sears, Catherine Plaisant, and Ben Shneiderman. 1993. *A new era for high precision touchscreens.* Ablex Publishing Corp., USA, 1–33.

[13] Kihyuk Sohn, Honglak Lee, and Xinchen Yan. 2015. Learning Structured Output Representation using Deep Conditional Generative Models. In *Advances in Neural Information Processing Systems*, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett (Eds.), Vol. 28. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2015/file/8d55a249e6baa5c06772297520da2051-Paper.pdf

[14] STT System. 2024. Motio 3DMA. https://motio.stt-systems.com/

[15] Xiangjun Tang, He Wang, Bo Hu, Xu Gong, Ruifan Yi, Qilong Kou, and Xiaogang Jin. 2022. Real-time controllable motion transition for characters. *ACM Trans. Graph.* 41, 4, Article 137 (July 2022), 10 pages. doi:10.1145/3528223.3530090

[16] Jordan A. Taylor and Richard B. Ivry. 2012. The role of strategies in motor learning. *Annals of the New York Academy of Sciences* 1251, 1 (2012), 1–12. doi:10.1111/j.1749-6632.2011.06430.x arXiv:https://nyaspubs.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1749-6632.2011.06430.x

[17] Chen Tessler, Yoni Kasten, Yunrong Guo, Shie Mannor, Gal Chechik, and Xue Bin Peng. 2023. CALM: Conditional Adversarial Latent Models for Directable Virtual Characters. In *ACM SIGGRAPH 2023 Conference Proceedings* (Los Angeles, CA, USA) *(SIGGRAPH '23)*. Association for Computing Machinery, New York, NY, USA, Article 37, 9 pages. doi:10.1145/3588432.3591541

[18] Fong-Gong Wu, Hsuan Lin, and Manlai You. 2011. Direct-touch vs. mouse input for navigation modes of the web map. *Displays* 32, 5 (2011), 261–267. doi:10.1016/j.displa.2011.05.004

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009