

## HKUST SPD - INSTITUTIONAL REPOSITORY

---

Title

Authors

Source

Version

DOI

Publisher

Copyright

This version is available at HKUST SPD - Institutional Repository (<https://repository.ust.hk/ir>)

If it is the author's pre-published version, changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published version.

# Live Emoji: A Live Storytelling VR System with Programmable Cartoon-style Emotion Embodiment

Zhenjie Zhao    Feng Han    Xiaojuan Ma  
Department of Computer Science and Engineering  
Hong Kong University of Science and Technology  
Hong Kong, China  
{zzhaoao, fhanab, mxj}@cse.ust.hk

**Abstract**—We introduce a novel cartoon-style hybrid emotion embodiment model for live storytelling in virtual reality (VR). It contains an avatar with six basic emotions and an auxiliary multi-modal display to enhance the expression of emotions. We further design and implement a system to teleoperate the embodiment model in VR for live storytelling. Specifically, 1) we design a novel visual programming tool that allows users to customize emotional effects based on the emotion embodiment model; 2) we design a novel face tracking module to map presenters’ emotional states to the avatar in VR. Our web-based implementation makes the application easy to use. This is an accompanying paper extracted from [1] for the demo session in IEEE AIVR 2019.

**Index Terms**—Emotion embodiment; social presence; avatar; storytelling; education; visual programming

## I. INTRODUCTION

Virtual reality (VR), with its ability to immerse audiences in a scene, is a promising new medium for engaging storytelling [2]. To improve an audience’s experience, there is a growing interest in the role of a presenter in engaging users [3], [4], *i.e.*, telling stories to others. In conventional face-to-face storytelling, the physical presence of a presenter [5], especially the sharing of emotion, bounds the narrative to experience and renders it unforgettable [6]. Similarly, in the digital field, embodying the presence and emotion of a presenter plays a crucial role for engaging storytelling [7].

Existing artists’ practice show that it is possible to create immersive VR experience through composing 2-dimensional (2D) images [8]. To lower the barrier to perform engaging live storytelling in VR, we take a 2D approach to create emotion embodiment. Specifically, taking inspiration from common emotion expressing methods in comics [9] and cartoon animation [10], we propose a novel method to embody a presenter’s emotional states by integrating two components: 1) a simple cartoon-style animated avatar to represent emotions directly and 2) an auxiliary display of multimodal effects to enhance emotional intensity. We then adopt a design process [11], and design and implement a system, *Live Emoji*, to create and express the emotion embodiment.

## II. HYBRID EMOTION EMBODIMENT IN STORYTELLING

Following three key design requirements: clarity, iconic, and multi-modalities from comics [9] and cartoon animation [10], we propose to use a hybrid model to embody a presenter’s emotions in the virtual counterpart, which combines a central

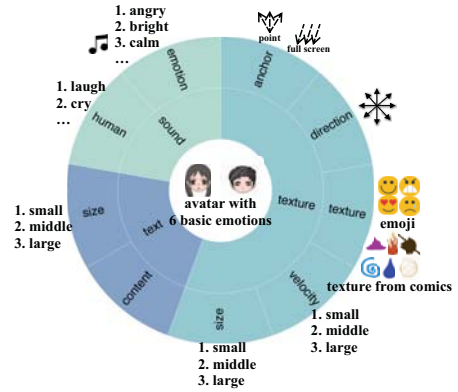


Fig.1: Illustration of basic design space of our hybrid emotion embodiment.

expressive avatar and auxiliary displays of special effects. The basic design space is illustrated in Fig.1. We adopt the design guidelines from Ma *et al.* [12] and invite two graphics designers to create two virtual characters, a boy and a girl, with the ability to animate Ekman’s six basic emotions (Fig.1-center). For special effects, we learn from motion lines in comics [9] to animate textures and support textual commentary and non-speech sound to enhance emotion expressing.

## III. SYSTEM FOR SUPPORTING EMOTION EMBODIMENT

To enable ordinary users to compose and perform narrative in VR, we design and implement a live storytelling support system called *Live Emoji*. As illustrated in Fig.2, a live VR story coming out of our system consists of three components: 1) a VR scene, 2) a verbal story told by a presenter, and 3) our emotion embodiment module. We design and integrate a block-based visual programming tool into our system so that users can easily create and customize such effects. During storytelling, a face tracking module analyzes the presenter’s facial information in real-time and directly maps the detected emotion, eye state, and head position to the avatar. The presenter can also actively trigger the pre-edited multimodal effects through keyboard and mouse. Meanwhile, the audience wears a HMD to explore the VR scene and watch the performance.

The interface used by the presenter during the live performing is shown in Fig.3-A. In particular, a) is the menu

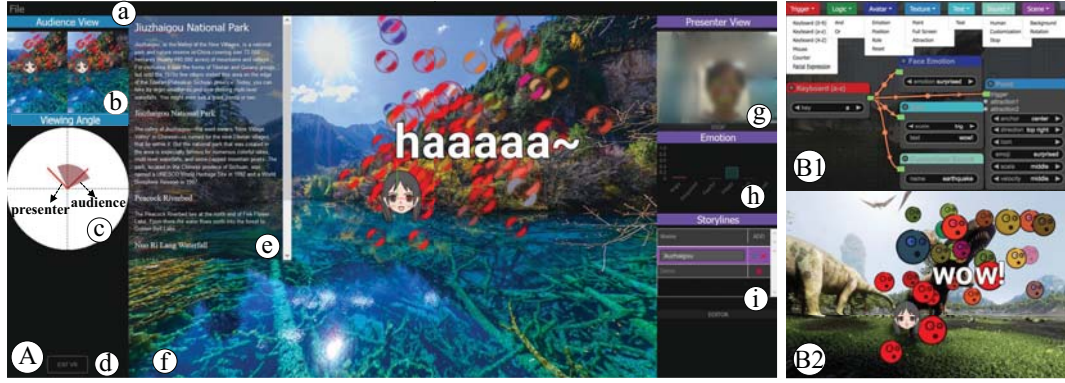


Fig. 3: A: The system interface for the presenter, including A-a) menu, A-b) audience view, A-c) viewing angle, A-d) VR button, A-e) script editor, A-f) main canvas, A-g) camera view, A-h) recognized emotion distribution, A-i) storyline view. B: Block-based visual programming tool, including B1) the editing interface and B2) a multimodal effect example.

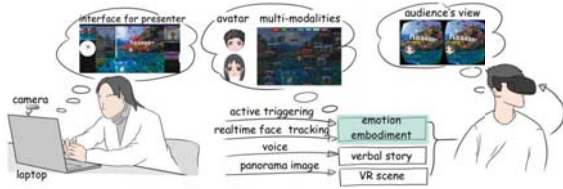


Fig. 2: Illustration of the proposed live storytelling VR system *Live Emoji*.

for importing and exporting edited effects; b) is the real-time VR view mirroring the two screens in the HMD; c) is an overhead view to show the different viewing angles of the presenter and the audience as marked in the figure; d) is the button for starting and stopping displaying in the HMD; e) is a script editor for users to add, modify, and delete performance scripts; f) is the main canvas where users can dragging the scene to change the presenter's view angle; g) is the camera view that captures the presenter's face; h) is an emotion recognition visualization view that displays the real-time recognition results across six basic emotions; and i) is the *storyline* view that records all the default and user-edited effects, where we call *storyline* as a group of edited effects. Users can add, delete, and (re-)name a storyline. Pressing the button **EDITOR** takes users to a visual editing tool (Fig.3-B1) to edit effects. Fig.3-B2 shows the corresponding edited effects of Fig.3-B1. An audience with a HMD can look around 360-degree, and the avatar and auxiliary visual effects are always in a front 2D plane to attract his/her attention.

We implement the system as a web-based application using Javascript, and run it on a laptop with an Intel CPU i7 2.8GHz, 16GB RAM, and Nvidia GeForce 1070 GPU. An Oculus Rift [13] is used for audiences to view the story scenes. The current implementation only supports one-to-one storytelling.

#### IV. CONCLUSION AND FUTURE WORK

We propose a new form of emotion embodiment by combining facial expression and auxiliary multi-modalities for live storytelling in VR. We design a proof-of-concept system called

*Live Emoji* to support authoring and performing with an emotionally embodied avatar. In the future, we plan to extend the system to support multiple audiences and remote storytelling, which can better empower real-world applications.

#### V. ACKNOWLEDGEMENTS

This work is supported by the Research Grants Council of the Hong Kong Special Administrative Region, China under Grant No.: T44-70716-N.

#### REFERENCES

- [1] Z. Zhao, F. Han, and X. Ma, "A live storytelling virtual reality system with programmable cartoon-style emotion embodiment," in *IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, Dec 2019.
- [2] R. Pausch, J. Snoddy, R. Taylor, S. Watson, and E. Haseltine, "Disney's Aladdin: first steps toward storytelling in virtual reality," in *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*. ACM, 1996, pp. 193–203.
- [3] A. Alohal, K. Kunze, and R. Earle, "Run with me: designing storytelling tools for runners," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016, pp. 5–8.
- [4] Z. Deng, P. Benckendorff, and J. Wang, "Blended tourism experiencescape: A conceptualisation of live-streaming tourism," in *Information and Communication Technologies in Tourism*, 2019, pp. 212–222.
- [5] D. Lipman, "Improving your storytelling: beyond the basics for all who tell stories in work or play," in *American storytelling*. August, 1999.
- [6] P. Guber, "The four truths of the storyteller," *Harvard Business Review*, vol. 85, pp. 52–9, 142, 01 2008.
- [7] R. M. Baños, C. Botella, M. Alcañiz, V. Liaño, B. Guerrero, and B. Rey, "Immersion and emotion: Their impact on the sense of presence," *Cyberpsychology & Behavior*, vol. 7, no. 6, pp. 734–741, 2004.
- [8] S. Hayden, "Facebook artist makes a case for VR manga in homage to iconic 'dragon ball' scene," Retrieved October, 2019. [Online]. Available: <https://www.roadtovr.com/facebook-artist-makes-case-vr-manga-homage-iconic-dragon-ball-scene>
- [9] S. McCloud, "Understanding comics: The invisible art," *Northampton, Mass*, 1993.
- [10] F. Thomas, O. Johnston, and F. Thomas, *The illusion of life: Disney animation*. Hyperion New York, 1995.
- [11] M. W. Easterday, D. G. R. Lewis, and E. M. Gerber, "The logic of design research," *Learning: Research and Practice*, vol. 4, no. 2, pp. 131–160, 2018.
- [12] X. Ma, J. Forlizzi, and S. Dow, "Guidelines for depicting emotions in storyboard scenarios," in *8th International Design and Emotion Conference*, 2012.
- [13] Oculus Rift, "Oculus Rift: get into the best of VR gaming," Retrieved March, 2019. [Online]. Available: <https://www.oculus.com/rift>