Face à la mode: A new immersive entertainment system that can train body and facial expressions

Yuria Suzuki

Ayumi Ando

Ryotaro Tsuda

Akihiko Shirai
Kanagawa Institute of
Technology
1030 Shimo Ogino, Atsugi,
Japan
facealamode@shirai.la

ABSTRACT

The Face à la mode application is an immersive entertainment system prototype that can turn a childhood dream of being a fashion model into reality using RealSense and Kinect v2. This application utilizes a virtual studio that can manipulate the facial expressions and poses of players to turn them into models or actors. In addition, a player can become a “superstar” by being on the cover of a brochure. The proposed system can train body and facial expressions in real life.

CCS Concepts

•Human-centered computing → Activity centered design; Gestural input; Interface design prototyping;

Keywords

Kinect v2; RealSense; Selfie; Smile recognition

1. MOTIVATION

This study is based on a proposal for the Laval Virtual ReVolution 2015, with the subject, “Kiddy Dream in Virtual Reality.” Our childhood dream in virtual reality, called Face à la mode, involves becoming both virtual and actual models. Most people put on make-up and maquillage as part of their fashion statement. However, they often forget to focus attention to their smiles and poses. Their smiles are often unnatural or forced, which gives a wrong impression to others. If we can develop an entertainment system that can easily help people develop a natural smile, then they will be given the opportunity to improve their self-confidence and enhance their personality. We consider helping people improve the quality of their life by making them aware of the

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.

VRIC '15, April 08 - 10, 2015, Laval, France

© 2016 Copyright held by the owner/author(s). Publication rights licensed to ACM. ISBN 978-1-4503-3313-9/15/04...\$15.00

DOI: <http://dx.doi.org/10.1145/2806173.2806195>

importance of a nice smile and correct pose.

2. RELATED WORKS

For professional photographers, taking a photo of a nice face is a real-time process. They immediately recognize the expression of a model and present their ideal image by talking to the latter. If photographers take a photo based on their sense of what photogenic means, the model can see a flash and a shutter. This process is a real-time feedback loop that involves the visual, kinematic, and audio senses.

Recently, elements of expressions, such as photogenic timing or a smile, may be obtained using artificial intelligence, particularly smile detection, color space correction, and computational photography.

Manga Generator [1] is an immersive entertainment system that can realize role playing that features comics characters. This system features a real-time cartoon-style graphics and recognizes the posture of a player to fit the frame. The recognition engine called KinEmotion [2] can convert the pose of a player to represent emotional effects. However, being aware of one's smile or other facial expressions is difficult because the players cannot see their face simultaneously.

Face à la mode, which helps people develop a nice smile and pose, is as an application that can improve the quality of life by recognizing those important objectives. Players who use this application can recognize their own facial expressions and poses in real time. Therefore, we do not have to worry that the quality of the output is spoiled because the face of the person overlaps with the magazine design characters. In addition, Face à la mode can invite a natural smile during the experiment through an effect that changes a shutter sound and pose when it became the smile. Face à la mode is an unprecedented new training machine that can improve the quality of life by fulfilling the dream of certain people “to become a fashion model.”

3. SYSTEM CONSTITUTION

This prototype needs to determine the facial expressions or the entire body posture of a player during the real-time experience.

All systems were developed using the Unity3D. To achieve real-time synthesis, called image synthesizer, and to generate a magazine-style image, five layers for the foreground

and players as background are generated. Virtual lighting and color correction are implemented in shaders (GPU programming for real-time).

As a player-centered design, we saved several model poses, which represent professional model techniques, as the ideal postures of kinematics. The kinematics is compared with the current real-time postures of the players using our special technique called KinEmotion. This technique can determine the similarity between the target and current postures as a value between -1 and $+1$. When the value becomes closer to the ideal pose, the sonar frequency becomes higher and its repetition becomes faster.

We used RealSense SDK to recognize the best smiles of the players, which include both adults and children. If the smile value is over a certain threshold, then the system captures the smile with flashing effects. An unconscious laugh is exposed in a significantly short moment. The system supports the conscious and unconscious laugh of players in a short time. A few decorations, such as petals and flashy effects, enhance the laugh of players and the audience with no prior experiences. We assumed that the current RealSense was designed to use a desktop computer or a similar device, such as a notebook PC. A boomer stand used to adapt to the different heights of various players.

The display can realize image multiplexing using the Unity shader of ExPixel [3, 4, 5]. At the end of the experiment, the different images are shown to the next player; the time will start when the player who finished an experience chooses the best shot through a flat panel. This way, it can provide the proper flow of an efficient experience.

4. EXPERIENCE USING FACE À LA MODE

The experience of using Face à la mode is as follows. Visitors can decorate a cover of the official brochure. First of all, a player stands in front of a RealSense camera. A player can see a synthesized image that has a background and magazine-style titles. The appearance of the player will change in real time. When the player changes posture, the sonar and visual effects will change. When the player smiles, the system operates the shutter with continuous speed, which is similar to a burst shooting in a photo studio. Sonar is an audio feedback that can determine the best pose without visual feedback. Particle effects, such as stars, petals, and flashy effects, help the player and the audience to provide a nice smile. After a certain span of time (approximately 60 seconds or less), the player can choose the best photo and print it out. The display can show a secondary image for the next visitor through the ExPixel technology. The printed paper can serve as cover of the official brochure (Negotiated with the Laval Virtual organization beforehand). The motion and smile data are archived in the system, which will be used as reference to explore better experiments.

In this project, we attempted to explore generating the “cover of a magazine” with the best facial expression using a professional body expression for everyone.

This prototype contains the following techniques; (1) Best shot timing with the best smile, (2) Particle effects to keep the best smile, (3) Fashion model-specific standing posture, (4) Magazine logo types, and (5) Color correction.

This real-time image synthesis system could operate using the facial expression and body posture of players only until they have convinced themselves to pose intuitively. The

shooting image is realized using the Kinect v2 with the background removed from the portrait.

The players do not need to face the camera. They can feel and study the ideal pose and facial expression using their knowledge of the practices of professional models and photographers not only by visual feedbacks, but also by audio clues and the audience response. After this experience, a player can test these techniques in real life to obtain better relations with other people. The printed cover will symbolize a childhood dream that was realized using the venue.

5. CONCLUSION

In conclusion, we have proposed to realize our childhood dream of “how to become a model” using a real-time photo machine to make a natural smile and pose. It also proposed a set of techniques on how to be on the cover of the venue brochure in real life.

From a scientific point of view, computer vision can detect a human laugh. However, people do not actually recognize themselves. The proposed system is a type of child application that takes a photo, although visitors will discover the difficulty of consciously smiling naturally. After the experiment, everyone will recognize the importance of smiling real life and not only through a virtual reality technique and computer vision. Visitors can discover with the potential of artificial intelligence to improve their quality of life using Face à la mode.

6. REFERENCES

- [1] Yuto Nara, Genki Kunitomi, Yukua Koide, Wataru Fujimura, and Akihiko Shirai. Manga generator: Immersive posing role playing game in manga world. In *Proceedings of the Virtual Reality International Conference: Laval Virtual, VRIC '13*, pp. 27:1–27:5, New York, NY, USA, 2013. ACM.
- [2] Yuto Nara, Wataru Fujimura, Yukua Koide, Genki Kunitomi, and Akihiko Shirai. Kinemotion: Context controllable emotional motion analysis method for interactive cartoon generator. In *ACM SIGGRAPH 2013 Posters, SIGGRAPH '13*, pp. 75:1–75:1, New York, NY, USA, 2013. ACM.
- [3] Hisataka Suzuki, Rex Hsieh, and Akihiko Shirai. Expixel: Pixelshader for multiplex-image hiding in consumer 3d flat panels. In *ACM SIGGRAPH 2014 Posters, SIGGRAPH '14*, pp. 63:1–63:1, New York, NY, USA, 2014. ACM.
- [4] Koki Nagano, Takeru Utsugi, Kazuhisa Yanaka, Akihiko Shirai, and Masayuki Nakajima. Scritterhdr: Multiplex-hidden imaging on high dynamic range projection. In *SIGGRAPH Asia 2011 Posters, SA '11*, pp. 52:1–52:1, New York, NY, USA, 2011. ACM.
- [5] Wataru Fujimura, Yukua Koide, Robert Songer, Takahiro Hayakawa, Akihiko Shirai, and Kazuhisa Yanaka. 2x3d: Real time shader for simultaneous 2d/3d hybrid theater. In *SIGGRAPH Asia 2012 Emerging Technologies, SA '12*, pp. 1:1–1:2, New York, NY, USA, 2012. ACM.