Interactive Narrative Facial Expression Animation Generation by Intuitive Curve Drawing

Yanxiang Zhang, Yan Ling
University of Science and Technology of China

ABSTRACT

This paper presents a type of interactive facial expressions animation generation system based on traditional montage techniques and generative narrative concepts. It could allow users to produce narrative facial expression animation by interactively and intuitively drawing a plot curve based on pre-recorded facial expression animation clips, thus provide high flexibility and creativity to users with interesting interaction.

Keywords: Facial expressions animation, narrative, interactive, curve drawing.

Index Terms: facial animation; facial expression database; integrated narrative generation system; narrative structure; narrative technique; control mechanism;

1 INTRODUCTION

Faces are one of the primary visual channels used to convey emotional information in humans. Facial expressions may be an intrinsic part of emotional responses and more developed social interactions [1]. These movements convey the emotional state of an individual to observers [2]. Facial expressions are a form of nonverbal communication that could convey social information between humans [3]. Edward T. Hall, a renowned social anthropologist, believed that more than 65% of social messages are shared through non-verbal communications, such as body language and facial expressions [4].

The current study shows that the same facial muscular movement is associated with the same emotion in all peoples through inheritance [5]. So facial expression is a universal communication language of emotion. Some artists are also famous for their facial expressions, such as Mr. Bean.

Some selected images of special facial expression from photos, especially those from famous people, have been widely used in social media as a cultural symbol or social idea. Such as in Facebook, QQ, Wechat, Line, and so on, meme (a picture or drawing of unknown origin that's taken on a shared, iconic quality, coupled with a resonating message [6]) or sticker is a very special and high-efficiency communication symbol. Some users also use facial expression photos of themselves as memes to communicate with others.

Facial expression animation is a vital aspect of making facial expression alive. In order to generate a facial expression animation, some researchers focus on synthesizing realistic facial expressions from photographs [7] and 3D models [8], or real-time face tracking [9]. Yi Z et al. proposed a framework that consists of two modules, a novel two-stage neural-network generator and a novel post-processing module known as Animating Through Warping (ATW). This framework successfully animates a 4K facial image [10]. Zhao proposed a descriptive emotion command model to bind a triggering action, the semantic meaning, psychology measurements. This model uses emotion as input to produce animations out of facial expressions from drawn characters [11].

Animating facial expression could also be a great improvement on meme or sticker. Currently, most of the memes are still images. To make memes or stickers alive, Apple has released an App on iPhone X named Animoji, allowing users to mirror their expressions to custom animated characters for use on messages app. Still, it only works with a depth camera on iPhone X.

The existing facial expression animation technologies work in a monotonous mode and result in only one animation in a process. It is difficult to generate a narrative facial expression animation, which is usually much more impressive than a tedious facial expression animation.

The authors present a kind of interactive facial expressions animation generation system based on traditional montage techniques and generative narrative concepts. This system could allow users to produce narrative facial expression animation by interactively drawing a plot curve based on pre-recorded facial expression animation clips, thus provide high flexibility and creativity to users with interesting interaction. Interactive narrative is a form of digital interactive experience in which users can create or influence dramatic storylines through their actions. An interactive narrative system aims to immerse users in a virtual world to convince them that they are an integral part of the unfolding story and that their actions can significantly change the direction and/or outcome of the story [12]. It is a new method of creating rich expression animations to meet users' needs who expect to use animated expressions intuitively and playfully.

2 IMPLEMENTATION METHOD

At first, we prepared several different facial expression animation clips that can be used to generate facial expression animations. The clips could be recorded or produced by facial expression animation software. We can qualify a facial expression animation clip to be "GOOD" or "BAD" and assign a score. For example, the clip of smiling or laughing will be classified as "GOOD" and will have a high score, and the clip of sweeping or crying will be a "BAD" clip and will have a low score (Fig. 1). The prototype technique proposed in this study can be combined with AI and other similar technologies. This Ingenious approach demonstrated by the authors can bring new possibilities for the development of AI technologies. Since the authors do not have a technical basis for AI development, this study used a flash environment that the authors' team has a technical basis for and supports multimedia development. Although flash has become obsolete, the ideas and ingenious methods presented in this paper can still be seamlessly integrated into animations based on AI technology.

petrel@ustc.edu.cn
lingyan77@mail.ustc.edu.cn

978-0-7381-1367-8/21/$31.00 ©2021 IEEE
DOI 10.1109/VRW52623.2021.00088

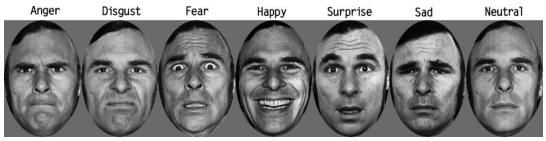

Figure 1: Clip of happy or surprise will be classified as "GOOD" and will have a high score, and the clip of sad or fear or anger or disgust will be a "BAD" clip and will have a low score. (Examples were taken from the Ekman and Friesen (1976) and Matsumoto and Ekman (1988) pictures of facial affect)

The following table shows a rough score assignment for six kinds of basic facial expressions. Scores may vary for a certain type of facial expression to a different extend.

happy	surprise	neutral	disgust	anger	fear	sad
8-10	**6-7**	**5**	**4**	**3**	**2**	**1**

The facial expression animation clips' score varies from 0 to 10, 0 is the lowest score, and 10 is the highest score.

Each clip lasts about 2 seconds, and to achieve a seamless connection between different clips, the first 0.3 seconds and the last 0.3 seconds of each clip are currently processed as neutral expressions. If the system can be combined with AI technology in the future, the transition between two expressions can also be realized directly.

Fig. 2 shows the composition of two different animations based on the same group of prepared clips. With each clip corresponding to a different ending, it is possible to create various scenarios by interconnecting the clips after user inputs.

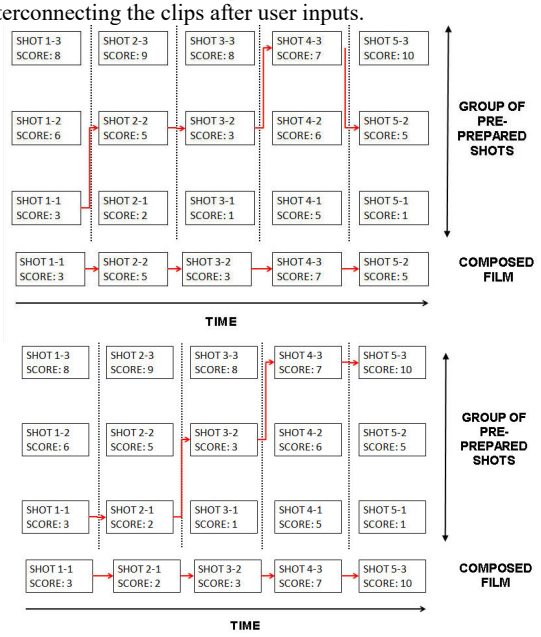

Figure 2: Generation of two different facial expression animation based on the same group of pre-prepared facial expression animation clips

If there are 10 clips of different score, and the generated animation also have 10 different clips, then there will be 10×9×8 ×7×6×5×4×3×2×1 kind of possible facial expression animation could be generated. It is really enough to form a narrative animation.

During the editing process, clips are grouped into libraries and each clip that forms one animation can only be retrieved from one library. Besides, the system is designed in such way that each clip can only appear only one time in each facial expression animation. Database of Clips will be recorded by the below format:

```
<?xml version="1.0" encoding="utf-8" ?>
<data>
        <movie src="01.flv" name="flv1" score="1"
description="description01" />
        <movie src="02.flv" name="flv2" score="2"
description="description02" />
        <movie src="03.flv" name="flv3" score="3"
description="description03" />
        <movie src="04.flv" name="flv4" score="4"
description="description04" />
        <movie src="05.flv" name="flv5" score="5"
description="description05" />
</data>
```

To allow the user to generate a facial expression animation using those clips, we create an interactive user interface that allows user inputs, defining how the animation will be composed. The current version presents the user with three input options: a random generation button that will generate a facial expression animation by randomly associating different clips, show in Fig. 3.

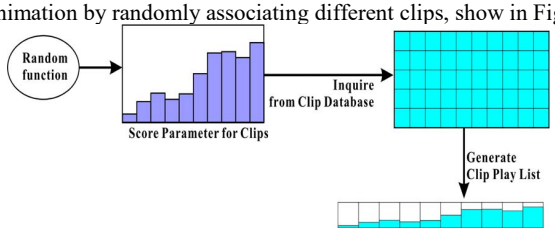

Figure 3: Generation of narrative facial expression animation based on Random function

A curve-based input that will generate a narrative facial expression animation based on a curve drawn by the user and refers to the "Score" attribute shows in Fig. 4. Finally, a "Length" option refers to the number of clips that will form the animation. In most cases, the random generation option results in animation with no logic, while an animation generated based on user input curve will present more logic.

The algorithm used in this design is to sample the curve drawn by the users equidistantly and correspond to the extraction of specific segments.

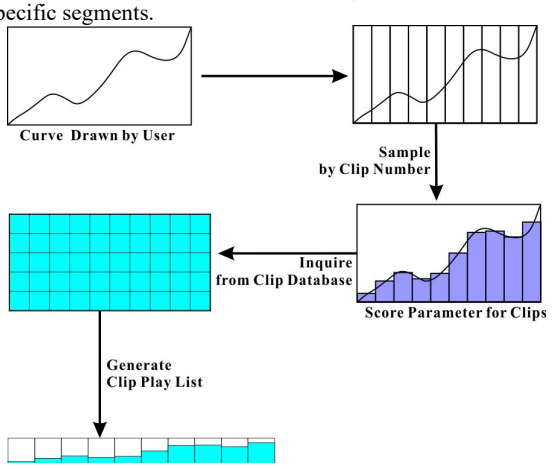

Figure 4: Generation of narrative facial expression animation based on the pre-prepared clips by user input

3 RESULT

The system could offer users the possibility to assemble a certain number of clips that will form a narrative facial expression animation. The user interface of the system is presented in Fig. 5.

The interface is composed of four main windows. The "Library" window refers to a specific folder where the clips are saved. All clips are grouped into libraries to avoid the clips that do not have any attribute appearing in the same movie. The "Playlist" window shows the list of clips that formed the generated animation.

Figure 5: System interface

The "Generate" window offers the user the generation options. The first option "Chart" gives the user the possibility to draw a curve line that will base on the "Score" attribute of the clip, and that will define how good or how bad each animation ends (Fig. 6). Users could draw a curve intuitively to represent their emotions and get a narrative facial expression animation. By clicking the Minimize button at the upper left corner, users can minimize the system interface, and then the user-generated expression animation can be played on the background interface.

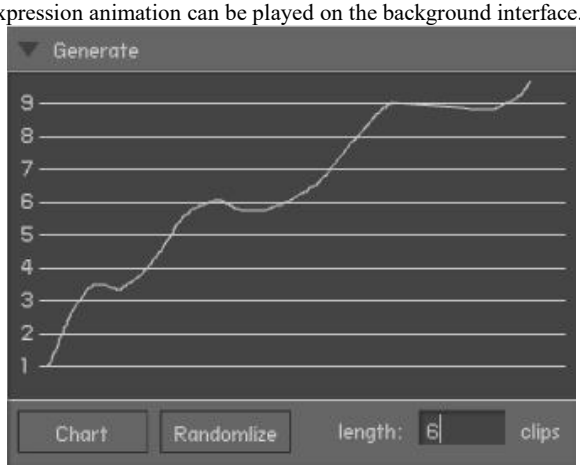

Figure 6: Curves lines defining the "Score" of the animation over time

As shown in Fig.6, the higher ordinate means the better emotion. The curve is drawn manually, which is a more intuitive visual representation of emotion. The "Randomize" option can be used to generate an animation composed of random clips. Animations generated using this option, in most of the case lack of logic in their scenario. The user also has the possibility to limit the number of clips that will form an animation by defining the length of the animation. The "Import/Export" window gives the user the possibility to save and open a sequence of clips used to compose an animation. With the "SaveScript" button, users can save the script of the animation. A flash media player is situated in the background of the interface with the options of playing, pausing, or stopping the playback. Right-click anywhere on the interface, it will appear (Fig.7).

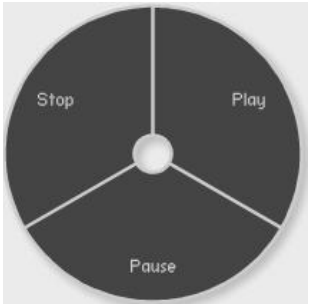

Figure 7: Player playback options

The final user-generated expression animation is shown in Fig.8. Each clip lasts about 2 seconds, and there are 0.3 seconds of neutral expressions at the start and end part of each clip. When the clip starts to play, it first morphs from the neutral expression to the current expression and finally returns to the neutral expression. Thus the transition of the whole video clip is very natural.

Figure 8: Generation of a narrative facial expression animation

4 User Feedback

Fifteen college users (7 males and 8 females) were recruited for this study to test the system designed by the authors. They were recruited on the condition that they like to use memes or emojis to communicate on social media. The final feedback is organized as follows:

The system enables a richer expression form of the generated expression animation, and the process of implementation is straightforward. It would be troublesome to use the Animoji feature of the iPhone and other devices to generate abundant animation driven by a person directly. In contrast, this system can provide an interesting and very intuitive way for people to operate. However, this system can only generate facial expression animations by using preset clips. Some users hope it can be used in combination with different animation apps.

5 Conclusion and Discussion

This system is a useful tool for diversifying the narrations that a facial expression animation can tell. With some predefined clips, the user can choose to watch and experience various facial expression animations with different endings in-between or at the end of the animation.

The flexibility of the system is based on the number and the qualities of the clips. A higher number of clips will increase the possibility to diversify the story of each animation, and the way the attributes of each shot are assigned will define the efficiency of the system.

Currently, the system is designed with flash. In the future, we plan to make it an App that could work with; thus, the composed facial expression animation can be a very interesting dynamic meme or sticker that users could use on message apps.

The mechanism described in this paper could also be applied to existing facial expression animation systems, thus allowing narrative design ability in facial expression animation production.In the later stage, the expression library can be enriched through different length expression clips to achieve non-linear mapping of expression animation and curves, and the AI technology can be integrated so that expressions of different scores can be directly morphed to obtain more natural and diverse expression animations.

REFERENCES

[1] Correia-Caeiro, Catia, Kun Guo, and Daniel S. Mills. Perception of dynamic facial expressions of emotion between dogs and humans. *Animal cognition,* 23.3 (2020): 465-476, 2020.

[2] Fridlund, Alan J. Human facial expression: An evolutionary view. *Academic Press,* 2014.

[3] Russell, James A., and José Miguel Fernández Dols, eds. The psychology of facial expression. *Cambridge: Cambridge university press*, Vol. 10, 1997.

[4] LiveseySolar. Cultural differences in body language and universal facial expressions, https://liveseysolar.com/cultural-differences-in-body-language-and-universal-facial-expressions/, January 2014.

[5] Eckman, Paul. Universal and cultural differences in facial expression of emotion. Nebraska symposium on motivation. *University of Nebraska Press,* Vol. 19, 1972.

[6] Zittrain, Jonathan L. Reflections on internet culture. *Journal of Visual Culture,* 13.3 (2014): 388-394, 2014.

[7] Pighin, Frédéric, et al. Synthesizing realistic facial expressions from photographs. *ACM SIGGRAPH 2006 Courses,* 19-es, 2006.

[8] Waters, Keith. A muscle model for animation three-dimensional facial expression. *Acm siggraph computer graphics,* 21.4 (1987): 17-24, 1987.

[9] Cao, Chen, Qiming Hou, and Kun Zhou. Displaced dynamic expression regression for real-time facial tracking and animation. *ACM Transactions on graphics (TOG)*, 33.4 (2014): 1-10, 2014

[10] Yi, Zili, et al. Animating Through Warping: An Efficient Method for High-Quality Facial Expression Animation. *Proceedings of the 28th ACM International Conference on Multimedia,* 2020.

[11] Zhao, Z. Live Emoji: Semantic Emotional Expressiveness of 2D Live Animation. arXiv preprint arXiv:1902.03529, 2019

[12] Riedl, Mark Owen, and Vadim Bulitko. Interactive narrative: An intelligent systems approach. *Ai Magazine,* 34.1 (2013): 67-67, 2013.