SIGRAD 2016
M. Hayashi (Editors)

Automatic CG Talk Show Generation from the Internet Forum

Masaki Hayashi, Steven Bachelder & Masayuki Nakajima

Uppsala University, Campus Gotland, Visby, Sweden

Abstract

We have developed an application to produce Computer Graphics (CG) animations in TV talk show formats automatically from the Internet forum. First, an actual broadcasted talk show is analyzed to obtain data in regards to camera changes, lighting, studio set up, etc. The result of the analysis is then implemented into the application and a CG animation is created using the TV program Making Language (TVML). The application works in the Unity game engine with CG characters speaking with computer generated voices. We have successfully created a CG generated TV talk show which allows users to "watch" the TV show format generated from the text information coming from the forum on the Internet.

1. Introduction

Due to the advancement of Computer Graphics (CG), and wide spread use of the Internet, many consumer services (e.g. Plotagon [Plo16]) have emerged supporting User Generated Content (UGC). The creation of CG animations is part of this development.

To facilitate UGC further, we have been developing a system which is capable of creating TV-show-like CG animations automatically from text based internet sources [HID*14]. Our key concept is to convert various sources on the Internet into text-based scripts. These then are represented by the CG animations created by the scripts processed in the TV program Making Language (TVML)[www.nhk.or.jp/strl/tvml/english/player2/]. The TVML is the base technology that the author uses to facilitate the development of automatic CG content creation from the text sources.

In this paper, we introduce an attempt to generate a TV-show-like CG Talk Show made by the TVML by converting a HTML script taken from the forum on the Internet. The conversion is done fully automatically so that users can actually "watch" the forum by just clicking a button in the application. The animation is made by real-time CG where CG characters with computer generated voices are discussing with appropriate gestures. The use of camera switching is calculated by the system. Our approach to do this is to first analyze a recorded broadcasted talk show on TV, and then to extract rules and statistical behaviors, and to ultimately implement these as an algorithm in the application in order to generate the CG animation which resembles the original format.

2. Analysis and Algorithm

Figure 1: *Reference talk show on TV (broadcasted on Jan. 1989)*

We picked up a series of Japanese talk shows on TV (see Fig. 1) recorded for our analysis. Through the observation of the reference talk show in terms of who is speaking, camera angles, an elapsed time, we formulated the following process (see Fig. 2).

- Phase-1: To determine whether a camera switching occurs when a speaker changes.
- Phase-2: To determine the next camera angle when the camera switching occurs.

Fig.2 shows the algorithm of Phase-1. There are two factors in Phase-1: "probability of camera switching when a speaker changes" and "duration of certain camera angles".

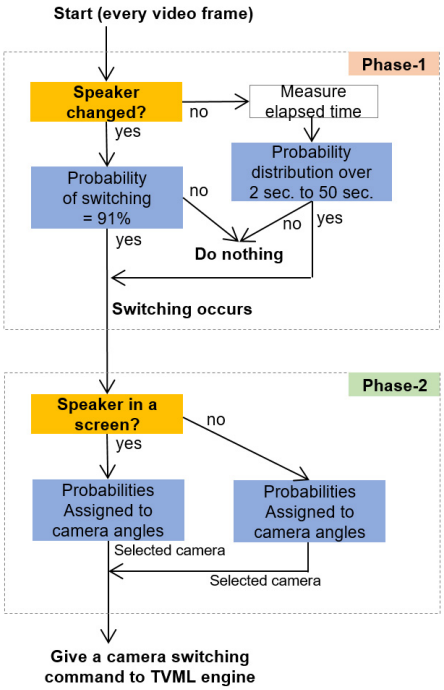

Figure 2: *Algorithm of camera switching.*

When it comes to Phase-2, the camera angles are classified into these six types: "upshot of a speaker", "upshot of a previous speaker", "two-shot of a speaker and a previous speaker", "group shot including a speaker", "shot without including a speaker", "other shot". When the camera switching trigger is sent from Phase-1 to Phase 2, it first checks whether the current speaker is filmed in the current camera angle, then determines the next camera angle based on the possibilities assigned to the next camera angles. We analyzed a video recorded in approximate one hour (number of shot is 264) to determine the factors stated above. The results indicated the probability of switching in Phase-1 is 91The probabilities in Phase-2 are shown in Fig. 3.

	Speaker on a screen	speaker not on a screen
upshot of speaker	51%	73%
upshot of previous speaker	13%	3%
two-shot of speaker and previous speaker	4%	3%
group shot including speaker	4%	6%
shot without including speaker	21%	10%
other shot	7%	5%

Figure 3: *Indicated camera angle possibilities.*

3. Implementation and a Test

The TVML SDK in the Unity3D game engine is used to make the TV-show-like CG animation. It is thus able to create real-time CG animations from a scripts written in TVML. This is done by using CG characters with computer generated voices, superimposing texts, image file displays, movie file playbacks, audio file playbacks, etc.

The algorithm described in Chapter 2 is implemented as a C# program in Unity. Input into the program is the status of "who is speaking" which is obtained from the SDK and an elapsed time sequence measured in the program. The output is a camera switching command given to the TVML engine. The calculation is made in every video frame to make a decision in regards to camera switching by using the probabilities as shown in Fig.3 with random values.

The source for making the talk show animation was the Japanese forum called "2 CHANNERU" [http://www.2ch.sc/]. The reproduced talk show has six CG characters that each post in the forum is assigned to randomly. The necessary camera angles are pre-defined in a TVML script. The camera angles consist of "upshot of each character" (number of the camera is 6), "two-shot of given two characters" (number of the camera is 15) and "other shot" (1 group shot), resulting in a total number of camera angles of 22.

Besides the camera switching described here, character gestures are also added based on a simple heuristic algorithm. The characters change view direction and occasionally nod in real-time to make the animation more realistic.

Fig. 4 shows screenshots of the produced CG talk show. the generated talk show successfully reproduces the desired resemblance to camera switching observed in the original format. The video can be seen at https://youtu.be/twafRtxW1Dk.

Figure 4: *Screenshots of reproduced animation*

4. Conclusion

We have developed a system reproducing a CG talk show with TVML based on the analysis of a real broadcasted TV show. We analyzed the recording of an actual TV talk show

for one hour and extracted the rules and statistical characteristics of camera switching. These were then used to create an algorithm for automatic camera switching. This algorithm was then implemented to TVML SDK in the Unity game engine in order to provide a platform of an automatic talk show generation application. We used the Japanese forum to convert the posts to a TVML script with six CG characters randomly assigned to the different posts. With our implemented algorithm of camera switching and character gestures, the final CG talk show was created.

Our next aim will be to further verify the system, by conducting comprehensive user evaluation test measuring viewer engagement. We plan also to enhance the algorithm for not only camera switching but also camera movement, enabling control of these parameters also in order to further articulate camera movement in TV show formats (e.g. like political debates, stand up comedy, music shows, etc.). We will also focus on the further development of character behavior and movement.

References

[Plo16] Plotagon: http://plotagon.com/ (Accessed Apr. 2016). 1

[HID*14] HAYASHI M., INOUE S., DOUKE M., HAMAGUCHI N., KANEKO H., BACHELDER S., NAKAJIMA M.: T2V: New Technology of Converting Text to CG Animation. *ITE Transactions on Media Technology and Applications 2*, 1 (Dec. 2014), 74–82. 1