

Rap-Style Comment Generation to Entertain Game Live Streaming

Thanat Jumneanbun*, Sunee Sae-Lao*, Pujana Paliyawan[†], Ruck Thawonmas[‡],
Kingkarn Sookhanaphibarn[§], Worawat Choensawat[§]

*Graduate School of Information Science and Engineering

[†]Research Organization of Science and Technology

[‡]College of Information Science and Engineering

[§]Multimedia Intelligent Technology Lab,

*[†][‡]Ritsumeikan University, Japan

[§]School of Information Technology and Innovation, Bangkok University, Thailand
ruck@is.ritsumeikai.ac.jp

Abstract—This paper presents a system for rap-style comment generation for enhancing video game live streaming. As creative storytelling can make live streaming truly entertaining, we investigate music-style comment generation in a video game. Regarding that “rap” is a popular genre of music, which has also been claimed success in promoting the enjoyment and contributing to rehabilitation, we propose a system that generates rap-style comments in real-time for promoting audience experience.

Index Terms—Game live streaming, Speech synthesis, Rap music

I. INTRODUCTION

Live streaming has recently become popular. Twitch, one of the most popular live-streaming platform for video games, had more than 1.7 million streamers and 100 million audiences every month in 2017 [1]. However, a challenge problem is that human streamers cannot stream contents for 24/7, and that the difference in time zone makes some audiences unable to watch streaming content they want. Therefore, recent research has aimed at a use of artificial intelligence to play games (e.g., by Ishii et al. [2]) so audiences can watch live streaming at any time. Nevertheless, one important limitation is the lack of an effective game commentator to entertain audiences.

This paper proposes a rap-style comment generation system for video live-streaming. This idea was inspired by that the rap genre has become very popular—it surpassed rock music, and became the biggest music genre in the U.S [4]—and previous studies (e.g., Tyson et al. [5]) reported a potential use of rap music in stress-relieving. Our system can be adapted to use with any games, but in this study, we implemented it on FightingICE, a fighting game AI competition platform.

II. EXISTING WORK

A. FightingICE AI

Ishii et al. [3] proposed an AI that selects actions using Monte-Carlo Tree Search (MCTS) with highlight cues in the evaluation function (later called HLAI). Their results showed that a gameplay generated by the proposed AI is more

entertaining than that by a typical MCTS-based AI. In this study, we implemented the system by adding a commentary mechanism to this AI by Ishii.

B. Rap Music Generator

Hayashi et al. [6] proposed an application for creating a rap music video with computer animation. They used Aquestalk and TVML (TV program Making Language) to generate computer animation of rap music. Aquestalk is a lightweight speech synthesis engine, used in Japan. The TVML automatically generates the animation from text scripts in Unity, a game engine. In Unity, characters are able to move their mouth according to the prepared scripts as an automatic lip-synch. The system introduced in [6] cannot run scripts in real-time because a user had to prepare a ready script by adjusting the speed of speaking. On the other hand, our system proposes a mechanism for the real-time script reading by adjusting the speech speed of each word to create musical rhythm.

III. PROPOSED SYSTEM

The proposed system¹ consists of two major parts: the Commentary Mechanism embedded in any AIs for comment generation (Fig.1 (A)) and the Rap Synthesizer (Fig.1 (B)).

A. Commentary Mechanism

Commentary Mechanism can be embedded in any AI (HLAI by Ishii et al. [2] is used in this study) for comment generation. During gameplay, this mechanism continuously analyzes game states and generates comments based on rule-based functions. The rule-based functions were built according to six templates, which they were designed to cover all types of game scenes in the FightingICE. The six templates, with numbers of them in (), are as followed:

- **Combo Templates (6):** Comments about the accumulated combo hits and the effects.
- **End Templates (11):** Comments about the ending of the game.

¹<https://tinyurl.com/yadaww5r>

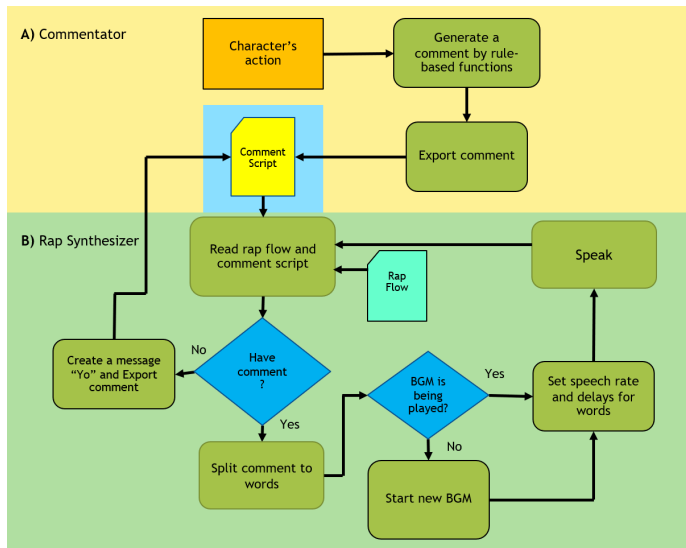


Fig. 1. Process flow diagram: A) Commentary Mechanism embedded into AI for comment generation. B) Rap Synthesizer.

- **General Templates (16):** Comments about the situation not belonging to other types of templates.
- **HP Templates (9):** Comments about which players are in advantage or disadvantage.
- **Open Templates (5):** Introduction of two-player at the beginning of the game.
- **Skill Templates (57):** Comments about skills/ actions of both characters.

A total number of commentary templates in the dataset is 104 . The number of the following template types: Combo, End, General, HP, Open, Skill, and templates are 6, 11, 16, 9, 5, and, 57, respectively. The “Comment Script” (Fig.1 (B)) contains only the most recent comments because whenever new comments are generated, the script will be overwritten.

B. Rap Synthesizer

Script exported by the AI is then processed by the Rap Synthesizer. This synthesizer exploited Microsoft SpeechSynthesizer², an engine capable of converting text strings into an audio stream. Parameters such as voice characteristics, pronunciation, volume, pitch, rate, or speed, emphasis can be set. To process the comments (Comment Script as shown in Fig.1), Rap Synthesizer focus two important actions: setting the speech speed of each word and adding delays between words, in the way that creates rap rhythms.

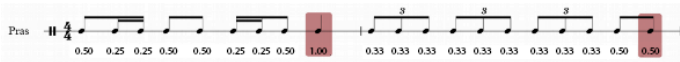


Fig. 2. An example of rap flow diagram representing rap rhythms. Texts below represent the numeric form of notes; these data are stored in Rap Flow.

²<https://docs.microsoft.com/en-us/previous-versions/office/developer/speech-technologies/>

When the system starts, Rap Synthesizer reads and stores the Rap Flow, which is a template of rap rhythm telling at what speed words should be spoken, and how long should be a delay after each word. This Rap Flow is designed based on the rap flow diagram [7] used for hip-hop/rap song’s lyric. The synthesizer converts musical notes in flow diagrams into numeric data based on musical beat, for examples, whole (♩) = 4.00, quarter (♪) = 1.00, and eighth (♫) = 0.50.

Rap Synthesizer keeps playing a preset background beat music with no lyric. To add lyrics, the synthesizer continuously reads comments in the script file written by the Commentary Mechanism (previously mentioned in 3A). The synthesizer will use a comment in the file if that comment is new, or else, create a message “Yo” as a new comment. When there is more than one word in a comment, Rap Synthesizer will first split the whole comment into words. Words are then mapped to notes in the Rap Flow. They are then spoken as lyrics, matching beat with the background music.

IV. CONCLUSIONS

This paper presented a system for rap-style comment generation. The system was tested with the fighting game by integrating in an existing AI a mechanism for generating comments to be spoken as a script. The script was then used by the proposed Rap Synthesizer, which was capable of speaking any English messages in a rap-style. The proposed Rap Synthesizer is universal and able to apply to any games, which it will require only a set of feeding script. We believe that our proposed system will enhance the experience of game live streaming. Our future plan includes changing pitches of words in speeches to make the system able to create comments in other music genres, such as Jazz.

Acknowledgement: This research was supported in part by Grant-in-Aid for Scientific Research (C), Number 19K12291, Japan Society for the Promotion of Science, Japan.

REFERENCES

- [1] J. Deng, G. Tyson, F. Cuadrado, and S. Uhlig, “Internet scale user-generated live video streaming: The Twitch case,” International Conference on Passive and Active Network Measurement. Springer, Cham, pp.60-71, 2017.
- [2] R. Ishii, R. Thawonmas, and T. Harada, “An Analysis of Highlight-Oriented AI Using Fighting-Game Experience,” In 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE), pp. 551-552, 2019.
- [3] R. Ishii, R. Thawonmas, and T. Harada, “A Fighting Game AI Using Highlight Cues for Generation of Entertaining Gameplay,” In 2019 IEEE Conference on Games (CoG), pp. 1-6, August 2019.
- [4] P. Ryan, “Rap overtakes rock as the most popular genre among music fans,” USA Today, 2018. Accessed on: Jan. 31, 2020. [Online]. Available: <https://www.usatoday.com/story/life/music/2018/01/03/rap-overtakes-rock-mostpopular-genre-among-music-fans-heres-why/990873001>
- [5] E. H. Tyson, “Hip hop therapy: An exploratory study of a rap music intervention with at-risk and delinquent youth,” Journal of Poetry Therapy, pp.131-144, 2002.
- [6] M. Hayashi, S. Bachelder, M. Nakajima, and Y. Shishikui, “Rap music video generator: Write a script to make your rap music video with synthesized voice and CG animation,” In: 2017 IEEE 6th Global Conference on Consumer Electronics (GCCE), pp.1-2, 2017.
- [7] R. Komaniecki, “Analyzing collaborative flow in rap music,” Music Theory Online, 2017.