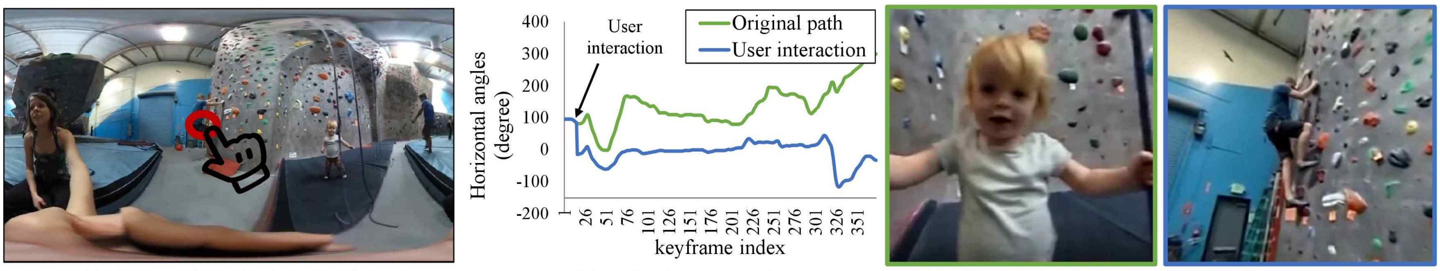
360° 비디오 감상을 위한 사용자 인터랙티브 자동 재생 시스템*

강경국[0], 조성현
대구경북과학기술원
{kkang, scho}@dgist.ac.kr

Interactive and Automatic Navigation for 360° Video Playback

Kyoungkook Kang[0], Sunghyun Cho
DGIST



(a) Change of the viewing direction (b) Optimal camera paths (c) Video content in the original path (d) Video content in the user-specified direction

Figure 1: Interactive and automatic navigation of 360° video. Our system computes an optimal camera path that shows salient areas in a 360° video, and plays a NFoV video based on the path in an online manner. A user can interactively change the viewing direction while watching a video, and the system instantly updates the path according to user intention.

Abstract

In this paper, we propose an interactive and automatic navigation system for comfortable 360° video playback. Our system finds a virtual camera path that shows the most salient areas through the video, generates a normal field-of-view (NFoV) video based on the path, and plays it in an online manner. A user can interactively change the viewing direction while watching a video, and the system instantly updates the path reflecting the intention of the user. Our experimental results show that our system provides more pleasant experience of watching 360° videos than existing approaches.

1. Introduction

A common scenario to view a 360° video on a 2D display such as a computer screen or a smartphone is to crop a part of a video, and render it as a normal field-of-view (NFoV) video. While watching a video, a viewer can manually change the viewing direction, e.g., by dragging a mouse. Unfortunately, it is not comfortable because a viewer needs to constantly adjust the viewing direction as the positions of interesting events can constantly change. Moreover, other interesting events may exist in other directions, and it is easy to miss them.

Recently, a few methods [1, 2, 3] have been introduced to solve this problem and to allow viewers to more comfortably watch 360° videos. These methods automatically analyze an input 360° video and find the best path that navigates the most interesting areas through the input video. A NFoV video is then rendered based on the obtained path so that a viewer can view interesting areas in the input 360° video without any manual intervention. However, these methods produce a single NFoV video without any user interaction and users are not allowed to change the viewing direction when watching a video. Thus, other parts of the original 360° video except for the selected path are completely lost and a viewer cannot see them at all.

In this paper, we propose a novel interactive 360° video navigation system (Fig. 1). Our system finds a

* 구두 발표논문
* 본 논문은 요약논문 (Extended Abstract) 으로서, 본 논문의 원본 논문은 SIGGRAPH. 2019 에 발표예정 되었음.
* 이 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 한국연구재단-차세대정보・컴퓨팅기술개발사업 및 기초연구사업의 지원을 받아 수행된 연구임 (NRF-2017M3C4A7066316, NRF-2018R1A5A1060031) 또한, DGIST Start-up Fund Program의 지원을 받아 수행된 연구임 (2019010063)

high-quality camera path that shows salient events in an input 360° video, and displays a NFoV video based on the path. While watching a NFoV video, a user can change the viewing direction by dragging a mouse. Then, the system instantly updates the path reflecting the user-specified direction.

2. Our Approach

To enable user interaction, our system is designed to consist of two steps: an offline pre-processing step, and an online interactive 360° video navigation step. The pre-processing step estimates optical flow and saliency for a given 360° video. The online navigation step splits the input 360° video into temporal windows. Then, it finds a camera path for one temporal window reflecting user interaction based on the optical flow and saliency information using one thread while playing a NFoV video of the previous window using another thread.

For smooth user interaction, we introduce adaptive control of the temporal window size and saliency-aware path update. Specifically, when user changes the viewing direction, using the default temporal window causes latency as much as the time required for calculating camera path. To reduce the latency, we set the temporal window size to minimum, and gradually increase the temporal window size. During the minimum temporal window, we simply let the camera track the optical flow at the current position, which lead to no latency. A camera path reflecting the user intention is obtained considering the following criteria:

(1) If user-specified viewing direction is in a region with high saliency, the updated path should follow the most salient object near the point.

(2) Even if user-specified viewing direction is in a region with low saliency, we assume that the user still wants to see that direction. Thus, the user-specified direction should be shown for a certain amount of time.

In the online navigation step, we find an optimal camera path that passes through salient parts of the input 360° video and is smooth enough for a user to comfortably watch. Camera path planning consists of three steps: initial path planning, FoV-aware path planning, and path smoothing. The initial path planning step first computes a camera path that follows optical flow while maximizing saliency score using dynamic programming. Thanks to this, our method can find a camera path that effectively tracks moving salient objects, and also avoids jumping back and forth between different salient objects. The FoV-aware path planning step finds a path that is

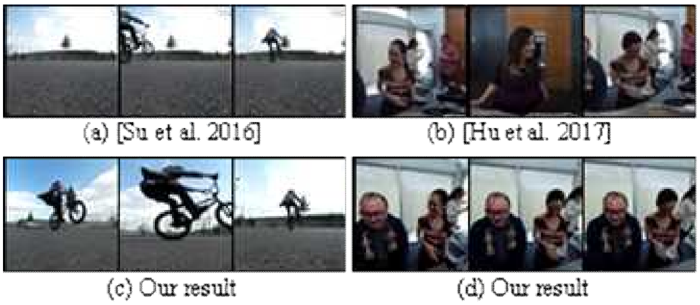


Figure 2: Comparison our results with [1] and [2].

close to the initial path, but more effectively shows the surrounding context as well as the most salient object. Finally, the path smoothing step computes a smooth camera path considering the velocity and acceleration of the virtual camera to suppress remaining jitters and make the resulting camera path smooth enough for a user to comfortably watch.

3. Experimental Results

We tested our system using various videos including datasets of [1, 2] and Youtube videos. Fig. 1 shows an example of our interactive path update. Without user interaction, our system finds a camera path that tracks a kid that is detected to be the most salient. However, as a user changes the viewing direction to a person climbing the wall, our system instantly updates the path to track the person for the rest of the video.

Fig. 2 (a) and (c) show results of Su et al. [2] and our system for video where the content of the video is moving rapidly. Our result tracks the bike more responsively thanks to our effective path planning. Fig. 2 (b) and (d) show results of Hu et al. [1] and our system. Our result shows a more stable camera path without jumping between different people, as we explicitly consider the motions of objects using optical flow. Finally, to compare our high quality path and user interaction system, we conducted a user study including existing approaches [1, 2, 3, 4, 5] and got a favored response from users.

References

[1] Hu, H. N et al. Deep 360 Pilot: Learning a Deep Agent for Piloting through 360 Sports Video, *CVPR*, 2017.

[2] Su, Y. C et al. Pano2Vid: Automatic cinematography for watching 360 Videos. *ACCV*, 2016.

[3] Su, Y. C. & Grauman. K. Making 360 Video Watchable in 2D: Learning Videography for Click Free Viewing. *CVPR*, 2017.

[4] Lin, Y. T et al. Outside-In: Visualizing Out-of-Sight Regions-of-Interest in a 360◦ Video Using Spatial Picture-in-Picture Previews. UIST, 2017.

[5] Pavel, A. et al. Shot Orientation Controls for Interactive Cinematography with 360 Video. UIST, 2017.