

# Design and Implementation of Synchronized Multi-view UCV for second screen service

Bokyung Sung<sup>1</sup>, Green Bang<sup>2</sup>, and Ilju Ko<sup>3</sup>

**Abstract**—Recently, second screen services, in which personal mobile smart devices complement TV watching, are increasingly being presented to users. Camera angle adjustment type services using the multi-view characteristics of live broadcasting produced using multiple cameras for sport, award ceremonies, etc. is viewed as one type of second screen service. In this paper, we show that the multi-view characteristic is present not only in broadcast images but also in User Created Video (UCV) that is produced based on some events occurring in the users life, and propose that Synchronized multi-view UCV through the overlapping part should be applied to Second Screen Services. The proposed synchronization is done using feature similarity matching on the basis of the audio signal. We present an application for Synchronized UCV that should be watched through the multiple camera angle switching method and a redesigned application that is applied to the second screen service structure consisting of DMP, DMS, and M-DMP. We also derive a parallel playing type that is synchronized on the basis of the overlapping parts to determine whether the watching method of the UCV produced under the same event is preferred with user values and usability aspects than the linear playlist playing type.

**Keywords**—User Created Video (UCV), Second screen service, Multiple camera angle, Multi-view UCV

## I. INTRODUCTION

NOWADAYS, various needs for personal video content production and consumption are being satisfied through radical development and popularization of progressive smart devices like Google Glass. Above all, practical use of additional techniques for supplying maximized interactive experiences to audiences from the aspects of content of the video and the watching method are considered in the plan. This means not watching video with a single channel display device but consumption of multimedia contents with a multi-channel structure that combine original videos with related additional multimedia content or control functions. Users have started to actively consume video content using main display devices like TVs with secondary display devices.[1] This change began to accelerate with the launch of the iPad in April 2010. From then on, most smart-pad owners have been using smart devices as a tool for complementary functions while watching TV.[2]

Second Screen Services has been presented as an alternative plan for satisfying the desire of TV viewers for complementary video content. These kinds of services are classified into three types Service based on audio recognition, Social TV Service,

and Camera angle adjustment service. Above all, Camera angle adjustment service focuses on bidirectional interactivity of the content consumption method. It provides multi-view point videos of event genre like award ceremony, performance [3], and sports games [4, 5] through multiple cameras on each location and angle. During the event, audiences can access and watch the broadcast by selecting the camera location and angle they want using second screen device.

Currently, second screen services in the UCV field have been undergoing a rapid increase in production and consumption, unlike the broadcasting and VOD fields. UCV plays a role in the recording real events. This means that UCV recorded videos from cameras in various locations and angles based on a single event have multi-view characteristics. Further, UCV has been produced consistently. These conditions mean that camera angle adjustment service, one of the second screen services, can be applied to the UCV field. Unlike live broadcasting, the camera angle adjustment service for UCV has a time-axis alignment problem that must be dealt with because of the time gap between video production and consumption. In this paper, we overcome the time-axis alignment problem among UCVs using a synchronization process based on audio and implement mobile application for Multi-view UCV. We implement recognition and matching of the overlap part among UCVs based on audio, because we consider that videos of different camera angles are not suitable for stable synchronization. We show that the watching and controlling of Multi-view UCV can be applied to the second screen environment.

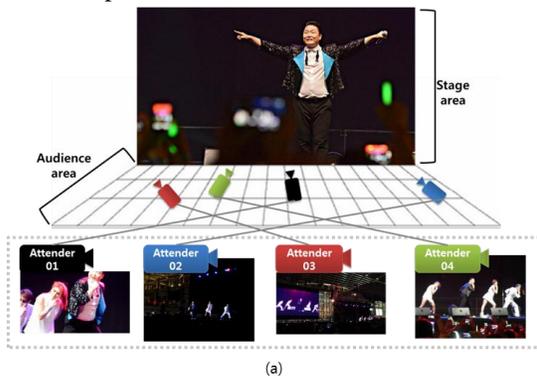
## II. MULTI-VIEW CHARACTERISTIC OF UCV

As previously mentioned UCV recorded by personal users are similar to the multi-view characteristics of videos planned and recorded by video production experts because videos recorded with multiple angles by several personal users are included in the UCV category. Production of UCV with multiple camera angles occurs in the same offline event. People can usually access this kind of event in normal life.

For production of UCV, an event that has multiple camera angles has to also satisfy three conditions: space, time, and the relation in the same time. First, in terms of relation, there has to exist one or more attendees with meaningful judgment of the event and who can record part of the event. Moreover, a relation has to be formed between attendees whether they know each other or not. Second, in terms of space, attendees should be located in a limited same space. The limit on the range of space

Bokyung Sung is a CTO of PDK LIMITED Co., South Korea  
Green Bang<sup>2</sup>, and Ilju Ko<sup>3</sup> are with Soongsil University, South Korea.

can be defined in various sizes, from table to stadium. Third, in terms of time, attendees should exist in a limited same timeline. Events in the same place can occur in different timelines.



However, attendees have to share the same timeline for UCV that has multiple camera angles.

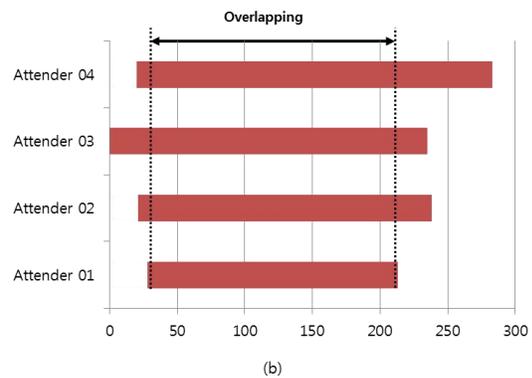


Fig. 1 Overlapping with multiple angles between UCV from the same event: (a) An example of UCV including multiple angle recording situations; (b) Occurrence of overlap part between recorded UCV

Fig. 1 (a) shows a UCV acquisition example that has multiple camera angles. That is, it is reconstructed from the original situation by getting UCV through the retrieved results of keywords that combine concert name and song title of a famous singer. Attendees are freely located in a limited place that is expressed with a mesh in the pictures with each blue, green, red, and black camera. Attendees share same space and timeline and have their own camera angle with Panning and Tilting. Attendees under each condition records part of a performance and upload it to share directly or indirectly with other users. Shared UCV has high probability of including the same contents because it is recorded in a limited space and time. Fig. 1 (b) is the arranged result of each UCV with original event timeline as a criterion. Each UCV has a different recording start time and end time. The recording duration is also different. Each UCV has an overlapping part that share a common timeline. Through the overlapping part, we know that the UCV mentioned is acquired from various camera angles in a common timeline and space.

angles in the same event, it is a resource for observing the angle adjustment method. Because of the existing overlapping part, UCV with the condition mentioned present multi-view characteristic that enable video to be watched with the angle adjustment method. The overlapping part of each users UCV occurs in the same event but can include different images that cannot be identified

### III. SYNCHRONIZATION AND REPRESENTATION AMONG UCVs BASED ON MULTIPLE CAMERA ANGLES

UCVs that share an overlapping part are a record of the same event in terms of content but do not mean they include the same subjects. Generally, attendees record the main common subject but sometimes they are used to record completely different subjects along the event type and the personal perspective of the attendees. Because overlapped UCVs are recorded from various

In this paper, we limit UCV as synchronization is recorded by attendees in the same event. Synchronization of UCVs is carried out to find the overlapping part between various UCVs from event attendees and arrange total UCVs for the reappearing timeline of event that recorded UCV. Fig. 2 schematizes the synchronization process. We determine that audio data are more stable than image data for recognizing the overlapping part of the UCV. For this reason, we use the audio matching method. The matching process consist of each step of the preprocessing audio channel from file, Extracting feature from preprocessed audio, and recognizing overlapping part and arranging by comparing the similarity of the extracted feature data of each UCV. All the UCVs are synchronized on the timeline of attended event through this process.

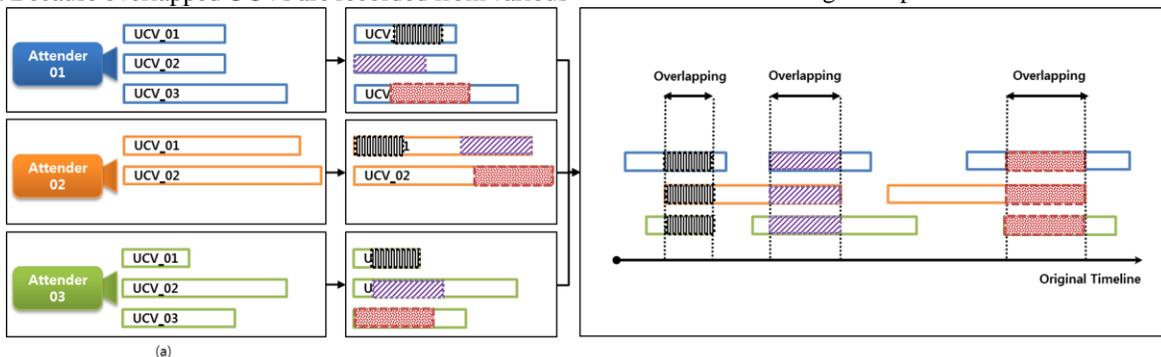


Fig. 2 Synchronizing UCV through the overlapping part: (a) Original UCV; (b) Recognizing overlapping; (c) Synchronizing between UCVs

Preprocessing transforms the audio of the UCV into a suitable form for extracting feature data. Preprocessing involves adjusting different channels, sound quality, format, and wave form of UCV to identical conditions. Each preprocessing step is audio signal processing for making UCV include minimum information for extracting features and have identical conditions. The stereo channel is transformed into a mono channel. Mono channel does not affect how people recognize the content of sound. Sound quality is set at 64 kbps bit rate, 22050 Hz sampling. These conditions are for minimum sound quality of the UCV distributed online. Audio of UCV for extracting features in waveform level is decoded to wave PCM sound format. Then, peak audio normalization is applied for consistent volume level of each UCV audio.

We applied Simplified-Delta MFCCs. Mel Frequency Cepstrum Coefficients (MFCCs)[6] generally used for extracting features from preprocessed audio waves. That is, SD-MFCCs apply Simplified Delta to extracted feature sequences by MFCCs and for simplification of feature data.

MFCCs are generally extracted features from 10~20 ms length frame. That length is in a suitable range for human audible recognition model. However, the proposed system set up was 5 ms and 2.5 ms for each windowing and hop size for minimize loss of information in extracting features. Increase of feature data by shorter windowing and hop size is offset through Simplified-Delta. Using all the extracted features from each frame means that low-level feature form is possible for high accuracy matching. We use reduced data through simplification of feature data. For this, 50 frames of 2.5ms length are simplified to 0.125s length units through feature summation.

Fig. 3 compares the similarity method to recognize the overlapping part through extracted features from UCV.

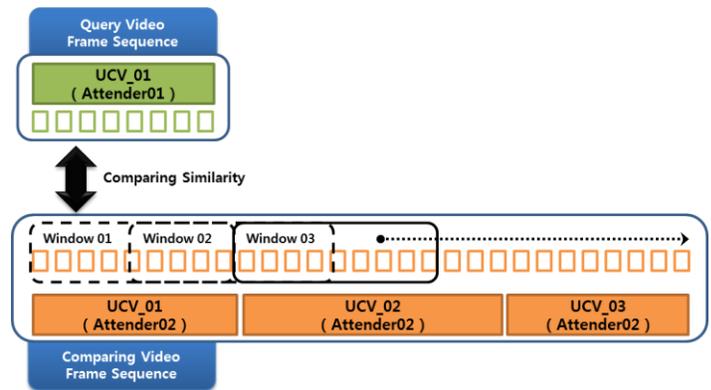


Fig. 3 Example of similarity comparison method

Synchronized UCV from single event can be watched through adjustment of multiple camera angles. Fig. 4 is a mobile application for watching synchronized UCV recorded by each camera angles of multi user in the same event. Two representative pages of this application are playing screen page and timeline control page for controlling angle and timeline.

The page for watching synchronized UCV through adjusting multiple camera angles consist of display area, function area for control, and additional information areas. The direction area supports both methods CROP for dividing one area into multiple areas and Picture in Picture (PIP) for piling up multiple areas. Fig. 4(a) shows that the UCVs of three camera angles displays divided one main area and two sub-areas through the CROP method. The function area for control consists of three functions: play and stop, view switching, and mode switching. The play and stop function signifies playing and stop of synchronized UCV along the original event timeline. The view switching means transforming camera angles between the main area and the sub-area. Fig. 4(b) shows that the User and each users UCV of synchronized UCVs are arranged by one timeline. Users can put the camera angle that user wants into the main area or change the time spot that user wants through selecting certain UCV clip or time spot.

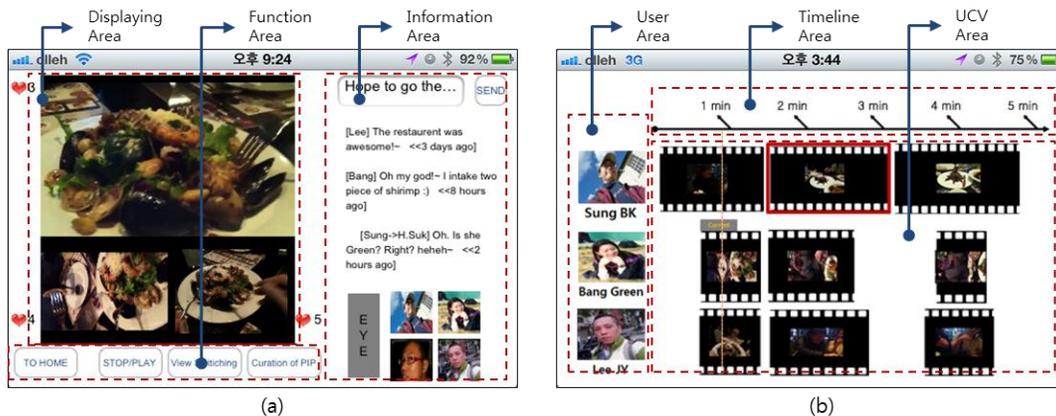


Fig. 4 Mobile application for auto playing synchronized UCV with timeline control function: (a) Playing screen of synchronized UCV; (b) Timeline control for synchronized UCV

#### IV. APPLYING SYNCHRONIZED UCV TO SECOND SCREEN SERVICE STRUCTURE

The implemented mobile app for proposed playing of

synchronized UCV in Section 4 includes control and display of synchronized UCV but does not offer two functions at the same time in one page. The size limitations of smartphone display causes two functions cannot be put in one page UI. The

separated form of video control and display can be applied to Second Screen Service of N-screen display environment like association of smartTV and smartphone.

After various size display device start to possess communication function, it is possible to make up a second screen service for assigning and displaying control functions to each display devices through data transmission among display devices. Synchronized UCV including multiple camera angles can be redesigned as a second screen service, because it also has separated status for control and display functions. Second screen service requests structurally three elements like Fig. 5. Those are DMP for digital media reproduction, DMS for content service (including DMR for digital media rendering), and DMP for digital media upload and control.[7] A smartphone is suitable for DMP that includes UCV recording and uploading functions and synchronized UCVs angles and timeline control functions because a general user usually records UCVs with smartphone and uploads it using the wireless network of the smartphone. Nonetheless, smartphone assigns M-DMP because it is categorized as a mobile device. Next, DMR is requested for synchronizing of the uploaded UCVs.

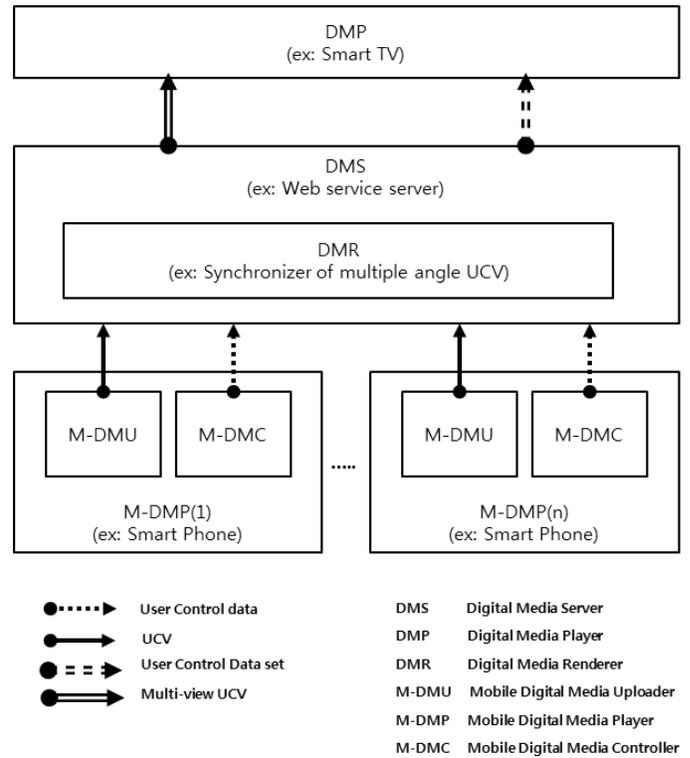
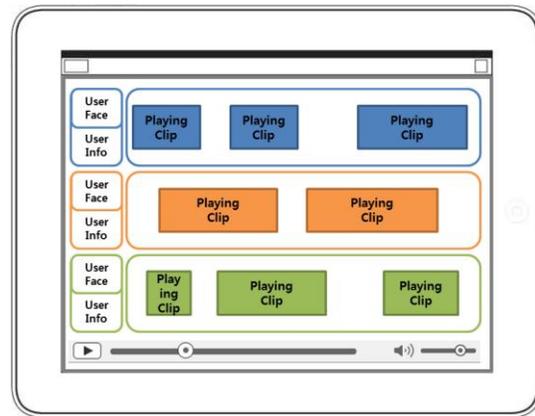


Fig. 5 General system structure of second screen service for UCV

Fig. 5 shows a general system structure of second screen service for UCV. The UCV synchronization method mentioned assigns DMR. DMR receives real-time M-DMC messages of M-DMP and render the UCVs. DMS is requested to transmit synchronized UCV of DMR to DMP. The web service server including DMR plays the role of DMS. DMP is requested for reproducing synchronized UCV rendered along occurring M-DMC message by users playing control and transmitted from DMS. SmartTV is suitable for DMP because SmartTV frequently reproduces video content stored in each device in second screen environment.



(a)



(b)

Fig. 6 Prototype example of applying synchronized UCV to second screen service: (a) DMP (SmartTV) for reproducing synchronized UCV; (b) M-DMP (Smartphone) for controlling timeline and angle adjustment

Fig. 6 shows a second screen service prototype example that watch synchronized UCV including multiple camera angles by

adjusting control of view and timeline. Fig. 6(a) only displays the function of the synchronized UCV applied to Smart TV. Fig. 6 (b) is a synchronized UCV along the original event timeline

structure and control functions are applied to Second Device.

## V. SOME COMMON MISTAKES

In this paper, we proposed multi-view characteristic of UCV via the feature of multiple camera angles. We implemented a mobile application for synchronized UCV by recognizing the overlap part based on audio matching. We then presented a redesigned mobile application as a second screen service prototype. In this prototype, we used smartTV for watching UCV and smartphone for watching control. With this, we proved that the implemented mobile application for displaying UCV can be redesigned then applied to second screen services.

Synchronized UCV is a content form that coincidentally uses videos from users having the same event and displays it. Through the mentioned structural feature and user preference, we can conclude that Synchronized UCV can be used exclusively for second screen services in N-screen environments.

For future work, we have two plans at this point. The first is to get an advanced synchronization method by comparing our proposed synchronization method with other methods. The second plan is to apply it to a real content service.

## ACKNOWLEDGMENT

This research was supported by Next-Generation Information Computing Development Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (No. 2012M3C4A7032783)

## REFERENCES

- [1] Clinch, S., Smartphones and Pervasive Public Displays, IEEE Pervasive Computing, 12(1), 92-95 (2013)  
<http://dx.doi.org/10.1109/MPRV.2013.16>
- [2] Changwoo, Y., Aiwon, U., Hyungwoo, L., Classification of N-Screen Services and its standardization, 14th International Conference on Advanced Communication Technology (ICACT), 597-602 (2012)
- [3] Zihao, Y., Nicholas, D., Mor, N., The Multiplayer: Multi-perspective Social Video Navigation, UIST 10 ACM, 413-414 (2010)
- [4] 4. Stephen, G., Dan, R., Brett, G., Time warp football, EuroITV '09 Proceedings of the seventh european conference on European interactive television conference, 77-86 (2009)
- [5] Frank, R., Michael, G., Meeting the multimedia needs of mobile sports fans, MM '09 Proceedings of the 17th ACM international conference on Multimedia, 471-480 (2009)
- [6] Logan, B., Mel Frequency Cepstral Coefficients for Music Modeling, Proc. Int. Symp. on Music Information Retrieval (2000)
- [7] Scott, S., Digital Home Standards: Choosing and Implementing the Right Ones, IEEE DLNA Presentation (2007)

**Bokyung Sung** is a CTO of PDK LIMITED Co., South Korea  
(e-mail:bksung@pdklimited.com).

**Green Bang** is with Department of Global Media, Soongsil University, Seoul 156-743, Korea(e-mail:greenlbang@gmail.com)

**Ilju Ko**(Corresponding Author) is with Department of Global Media, Soongsil University, Seoul 156-743, Korea(e-mail:andy@ssu.ac.kr)