

# A Research through Design (Rtd) Approach in the Design of a 360-Video Platform Interface

**Brian Michael Poblete**  
De La Salle University  
Manila, Philippines  
brian\_poblete@dlsu.edu.ph

**Emir Christopher Mendoza**  
De La Salle University  
Manila, Philippines  
emir\_mendoza@dlsu.edu.ph

**Julian Paolo De Castro**  
De La Salle University  
Manila, Philippines  
julian\_decastro@dlsu.edu.ph

**Jordan Aiko Deja**  
De La Salle University  
Manila, Philippines  
jordan.deja@dlsu.edu.ph

**Giselle Nodalo**  
De La Salle University  
Manila, Philippines  
giselle\_nodalo@dlsu.edu.ph

## ABSTRACT

Many video interfaces enable multiple sources of input video in displaying and streaming vital information. Most of these setups can be seen in deployed security systems and observer footage that are usually used for surveillance and crisis monitoring. In this study, we consider a crowd-sourced approach to multiple sources of video and aim to design an interface towards multiple possible use-cases. In designing this interface, we performed field studies and on site surveying along with initial user tests to validate our ideas. Research through design was added into the methodology to consider multiple point of views considering varying sources of perspective. Specifically, we catered the design of an initial interface in helping multiple users understand several views from multiples cameras, angles and position. The participants chosen for this study are students who have at least the basic technological ability of using a smartphone and taking a video with it. We intend to extend this study by validating the 360-view and designing an algorithm towards stitching one final view crowd-sourced from multiple cameras and streamers.

## ACM Classification Keywords

H.5.2. Information Interfaces and Presentation (e.g. HCI): User-Centered Design

## Author Keywords

Human Centered Design, User Centered Design, 360 Video, Interface Design

## INTRODUCTION

Videos are widely-used technologies meant for various purposes. They could be useful for surveillance, security, recreating accidents and even capturing memories according to the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*CHIUXID'19*, March 31–April 09, 2019, Bali, Indonesia

© 2016 Copyright held by the owner/author(s). Publication rights licensed to ACM. ISBN 123-4567-24-567/08/06...\$15.00

DOI: [http://dx.doi.org/10.475/123\\_4](http://dx.doi.org/10.475/123_4)

works of [16, 6, 2] With the digital age and the ubiquitous approach to mobile technology, virtually anyone with a mobile phone can capture videos to display a particular object of interest. Mobile phones are also able to gather metadata like location coordinates, duration of video and even lighting scenarios according to the work of [7]. Some video technologies can even detect and capture in-context objects such as smiles, gait, and other humanly-features. This is one consideration we have applied when doing this study.

However these techniques remain underutilized, which could still be implemented in more use-cases. More recently, panoramic technology and other video-stitching technologies have paved the way for the creation of 360-videos, thus improving viewing experiences of users similar to the work of [14]. These conditions and scenarios have not been fully used for actual use-cases that require multiple and crowd-sourced views. Several use-cases involve concert goers and musical events as seen in the work of [11], observing bystanders while recording footage as seen in the work of [12], and even using dash-cams while driving or using the automobiles as seen in the work of [6]. These scenarios are simply natural human events and activities that can be augmented with improved 360-video technology. Having been taken from varying angles and positions, a video that contains unique auditory and visual cues regarding human activity are pieces of information that can be collected when these technologies are used. By analyzing the position and angle at which a video was taken, we can potentially help investigators develop a detailed order of events say in a given incident that has transpired. This along many use-cases can benefit from multi-sourced videos that are stitched together to resemble a 360 video.

In this paper, we will be discussing the process for designing a usable interface that allows multiple live-streams (also known as streams), to be viewed by a user at the same time. This allows a cohesive viewing experience provided that multiple views and angles of a common and single Point-of-Interest (POI). In this prototype, it is proposed that we can provide valuable and significant viewing experiences especially on emergency situations where crowds typically respond by capturing a video. In this paper, we demonstrate how we have



Figure 1. “Screenshot of the Las Vegas Shooting Documentary. In this screenshot we can see how using footage of the event captured from multiple sources paved the way for an easier investigation of the crime scene. A similar but usable approach involving live-streamed video is introduced in this study. Imaged sourced from The New York Times.

used Research-through-Design (RtD) as a methodology in designing and validating an interface before proceeding to the latter parts of the research. The framework follows through the Research-through-Design methodology as seen in the works of [3, 5, 4, 15]. We followed a human-centric approach that had greater emphasis on the users while being able to identify the proper use-cases. Details on the data gathered, instruments used and an analysis of these results are included. We were able to generate three designs for an interface with the best prototype achieving an above average usability score. The goal of this first phase is to be able to understand the needs of the users in our target use case and be able to come up with the appropriate initial design for the prototype.

## RELATED WORKS

We have found several studies relating to the key points of our research: The specific keywords revolve around 360-degree video streaming, POI and videos from multiple angles of said points of interest. We considered understanding Points of Interest as our initial scope. We found numerous studies that combine POI along with several concepts like 360-degree video and multiple angles. In the work of [1], both the area of a user and the areas that a user was previously at was considered in determining the user’s next location. These locations have helped a model determine the prediction of the location-based services. Furthermore, the work of [9] installed a client-server system on mobile phones, integrating GPS, Wifi, GSM and accelerometer sensors to determine POIs. Our proposal combines technologies similar to these and the usage of videos to provide better perspectives on POIs.

In the works of [10, 13, 8] 360-degree videos have been the focus of developing a software prototype. The study of [10] focused on subtle and unobtrusive technologies that gave filmmakers and directors methods to direct attention to certain points of the 360-degree video while maintaining the level of control viewers have for the entire duration of the video. Interestingly, in the work of [13], the same technology was used but this time utilizing an intelligent agent in helping the user find the focus or point of interest. Consequently, in the work of [8], they have enabled technologies of using mini-screens in allowing users find the POI when it is outside the peripheral

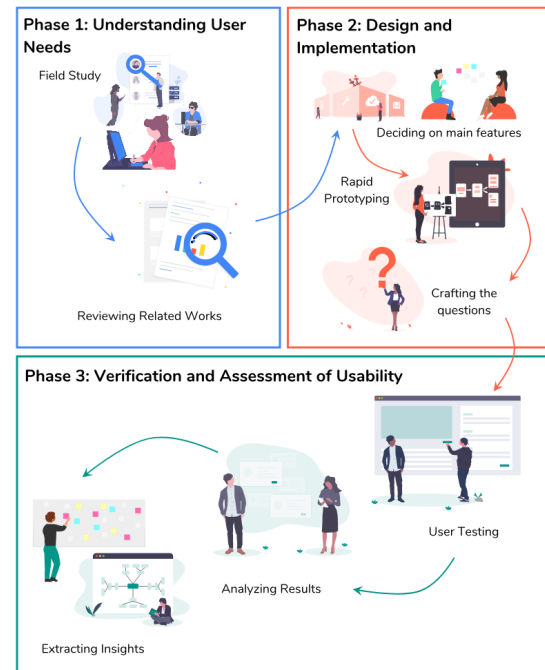


Figure 2. Overview of the three phases in this study. The approach follows a user-centric iterative methodology towards designing an interface. Phase 1 involves a field study and an in-depth analysis of related works in order to devise possible use cases and user needs. Phase 2 includes the development of the prototypes based on the supposed user needs and Phase 3 validates the designs and their features for overall user experience through testing with possible users.

vision called field of view. These mini-screens were referred to as Picture-in-Picture Previews (PIP). These works consider the different approaches and techniques in allowing the user to find and focus on a particular point of interest. It is interesting to note that these studies enabled flexibility and allowing the user to gain more control and freedom in exploring the interface. There are numerous studies that have worked with 360-degree videos in focus alone or with extracting points of interests in videos alone. The mentioned studies took advantage of utilizing 360-videos as the source of data and using these videos to automatically or dynamically detect a point of interest. In our study, we aim to achieve the opposite of this process, specifically capturing multiple angles and being able to produce a specific point of interest and putting them together in a 360-like video interface. We believe that this approach can pave the way for multiple use-cases to easily extract insights that will be usable in decision making, understanding environments and many others.

## METHODS

The first phase of this research includes methods the researchers used that are aimed towards understanding how to create a usable interface for the proposed idea. These include taking videos of a certain area from multiple angles simultaneously, Rapid Prototyping, and User-Testing.

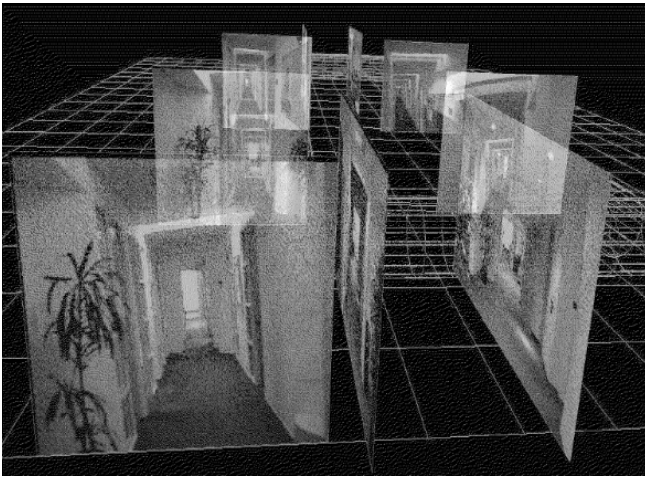


Figure 3. A screenshot of system with multiple cameras used for surveillance systems from the study of [16]

### Video Taking

In order to get an idea of how videos of a certain POI could be played simultaneously, we recorded footage in an open-air space from multiple angles at the same time. Mobile cameras were used to record footage while moving the cameras along a set path. After moving a certain distance in a straight line, the cameras were rotated to face one another before moving them in a straight line in a set direction once again. Once the footage had been taken, we analyzed the footage and chose a specific object from the footage as the POI.

### Rapid Prototyping

After a POI was chosen, this POI was used as a reference point in designing the prototypes for the application interface. Each of us decided to make one prototype for the interface. In order to devise an interface to display the video streams, we took inspiration from video-streaming sites such as Twitch, YouTube, and Netflix. We then created low-fidelity prototypes to give a basic example of how the videos would be displayed. Afterwards, we created high-fidelity prototypes to give a more concrete sample of the interface. We used Gravit Designer to create the prototype screens and Invision to add interactivity to the screens for user testing.

### Participants

Ten (10) undergraduate college students aged 19-20 were recruited through a convenience sampling method. The interviews were scheduled based on the participants' availability and were carried out in a public, indoor space with a high volume of bystanders.

### Study Design

The participants underwent proper research briefing before the actual experiment. They were given consent and waiver forms to ensure proper privacy and handling of data. After the briefing they were free to use the web app as if it was their first time on a brand new website. They were also free to ask the researchers anything they did not understand, provided they tried everything they could to try to understand it without aid. Since there were 3 prototype to be tested, each tester was

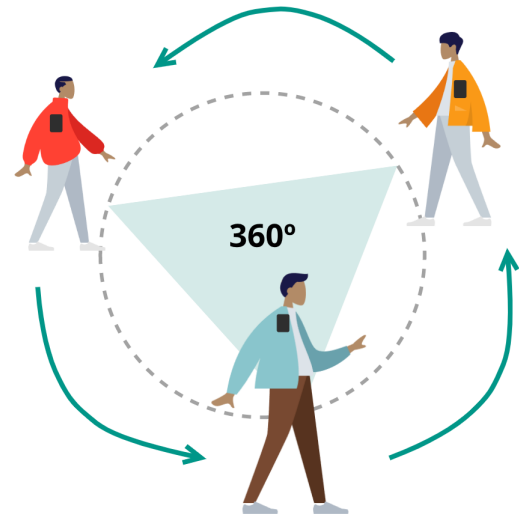


Figure 4. A visual representation of the video taking process and field study. The three people placed cameras on their shirt pockets, started recording, and moved along a set path to get a better perspective on the POI from multiple angles.

given a tester number and selected what prototype number they were testing. We used Google Forms as a platform to collect the answers of our testers. Our setup had 2 laptops with one displaying the Form and the other displaying the prototypes. First, we ask for the testers' consent in answering the form through the starting page of the questionnaire. The starting page states the purpose of our data collection and introduces the researchers. Afterwards, they are asked to tick checkboxes corresponding to which personal information they will allow the researchers to use. These include Age, Profession, Civil Status and Sex. The users then filled up the information they wish to provide and filled in their tester number and the prototype they were currently testing. The participant number was mainly used for the convenience of the researchers so we could easily track their responses as they had to submit the form 3 times, 1 for each prototype. During the interview, each participant was asked three (3) sets of questions with at least two (2) questions per set. The goal of the first set was to identify the current stand of the participant regarding their experiences in learning about ML and the various tools they have used. The goal of the second set was to identify the various problem areas that the participant associated with based on their experience using current ML tools. The third set was used to identify improvements regarding usability of the tool, as pointed out by the participant. The questions that were used are as follows:

- Upon first glance of this website I was able to determine that this page uses multiple points of view.
- With prolonged usage I was able to determine that this page uses multiple points of view.

- Upon first glance of this website I was able to see through the map that there is a P.O.I. in the area.
- With prolonged usage I was able to see through the map that there is a P.O.I. in the area.
- How well can you see the POI from different points of view?
- How well can you see the POI in "Mesh Mode"?
- How did you find navigating through the website? Why?
- How did you find switching from different points of view? Why?
- How well were you able to understand what each of the hotspots do by just looking at them, or clicking on them once? Why?
- How fast were you able to get used to the navigation of this website?
- I was able to understand completely the functions of the system.

The participants answered items 1 to 4 following a 4-point Likert scale with 1 being Strongly Disagree and 4 being Strongly Agree. These focused on determining whether a POI could be easily identified and viewed from multiple angles through the interface or not. The second set of questions (Items 5 and 6) focused on the clarity of the POI. These questions used a 4-point Likert scale with 1 being Very Unclear and 4 being Very Clearly. The last set of questions focused on the usability of the interface. The questions used 4-point Likert scales accompanied by open-ended questions for the participant to explain why they gave that rating. We also asked them to suggest ways to improve functions they did not understand in the event that they find such functions. We utilized descriptive statistical analysis to obtain the mean of the participants' responses to the questions and the mode of their responses. The open-ended follow-up questions enabled us to use the narrative analysis method for the qualitative data.

### Limitations

These methods used for our research include several limitations. Regarding video taking, we simulated the streaming of a POI from multiple angles by systematically recording a POI from multiple angles. However, we could not simulate 360 video streaming as we had no access to a 360 camera and thus could not design for a display that accounts for a 360 video stream. For the study design, the user tests were conducted in a location convenient for the participants. The location was not the most optimal area for conducting user testing due to it being a public area with various distractions present.

### RESULTS

This section contains the results from the user tests and the outcomes of the different prototype versions we prepared. Analysis of the figures are presented in mean values are followed by insights.

**Table 1. Table of average scores for each prototype for each criteria.**

	<b>Focus</b>	<b>Clarity</b>	<b>Usability</b>
<b>P1</b>	3.5	3.1	3.2
<b>P2</b>	3.1	2.9	2.7
<b>P3</b>	3.6	3.6	3.4

**Table 2. Results of the Initial Usability Test done on the three (3) Prototypes. Each area segment concentrated on Focus, Clarity and Usability.**

	<b>Prototype 1</b>	<b>Prototype 2</b>	<b>Prototype 3</b>
<b>Focus</b>			
Q1	3.0	3.1	3.3
Q2	4.0	3.6	4
Q3	3.3	2.6	3.4
Q4	3.7	3.3	3.9
<b>Clarity</b>			
Q5	3.2	3.1	3.7
Q6	3.0	2.7	3.5
<b>Usability</b>			
Q7	3.0	2.3	3.1
Q8	3.5	3.1	3.1
Q9	2.8	2.5	3.4
Q10	3.2	2.8	3.4
Q11	3.5	3.2	3.7
<b>Prototype Average</b>	<b>3.2</b>	<b>2.9</b>	<b>3.5</b>

### User Feedback and Insights

Table 1 indicates that Prototype 3 fared the best in each criterion (Focus, Clarity, Usability). Participants generally liked the easy-to-use design of the prototype. They described the use of icons and images as intuitive and understandable. On the other hand, Prototype 2 had the lowest score for each criterion. This was due mainly to the vagueness of the buttons used and the lack of basic functions such as a back or home button.

It can be observed on the table of averages that the best prototype version is Prototype 3. It garnered a score of 3.5, a few notches higher than Prototype 1 and Prototype 2. In terms of managing focus while using the interface, Prototype 3 scored the highest. However, there are still minor areas for improvement for Prototype 3 in terms of usability, specifically on overall navigation and switching from different points of view. It can also be observed that Prototype 2 scored the lowest especially on the areas of Navigation, understandability of the hotspots upon glances and on determining the POI through a minimap. By analyzing the individual scores for each areas, it is interesting to note that all prototypes scored relatively higher on the criteria of being able to view the POI with multiple angles from prolonged usage. This means that all three versions of the prototype where built for long periods of use.

### Prototype Features

Each prototype features a minimap, a button to activate mesh mode, different screens to view the different points of view and the ability to view in fullscreen the different POV's. Prototype 1 has a navigation bar at the top of the page, the very noticeable "Activate Mesh Mode" button in the middle of the screen and the ability to see what streaming service or social



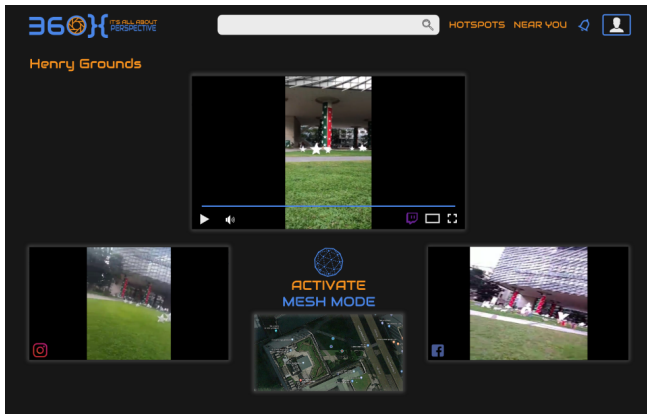


Figure 5. Prototype 1. This prototype shows the different angles, in this case 3 angles, of the streamed POI. In the middle of the screen is a button to activate "Mesh Mode". The navigation bar on the top of the page enables the users to easily move around the website, featuring a "Hotspots" button, a "Near You" button, and a "Home" button which is found on the top left corner. It also shows a minimap on the middle bottom of the screen.

media is being used in each point of view. Prototype 2 has unique labels for each point of view on the top left corner of each screen as well as a label for each point of view on the minimap. Prototype 3 has an interface for Augmented Reality, a scrollable field to browse other events and streams and the ability to click the icons of the other points of view on the main screen. These prototypes are meant for web. Another version will be designed separately for mobile use.

### Comparisons to Previous Studies

We compared the results of this study to previous research findings mentioned by analyzing the main features of prototype 3 and similar features found in the other works. First, the work of [16] details a navigation interface that allows a user to view video streams captured from multiple angles in a 3-D immersive environment, in comparison, our interface allows the user to view these multiple angles through a 360-degree video instead. Another basis of comparison would be the work of [11] where they collect recordings of multiple angles of concerts. Instead of collecting recordings of events, our interface would be able to compile live streams of on-going events and allow the user to switch between them in real time. In the work of [1] they were able to collect the locations and interests from their users' smartphones to generate multiple POI. Our work is able to collect and compile our users' real-time camera feed and location and generate POI from that data.

### CONCLUSION AND FUTURE WORK

We have designed and developed prototypes to provide an interface for the concept of viewing a POI from multiple angles. This was made possible by Research-through-Design and by enabling a user-centric approach in our methodology. This generates a 360-degree video using the data from these angles. The features implemented were inspired from the Las Vegas shooting aforementioned and other related works. To validate the usability of these prototypes, we had these prototypes evaluated and the results analyzed from our data collection and gathered insights from our participants. We also noted

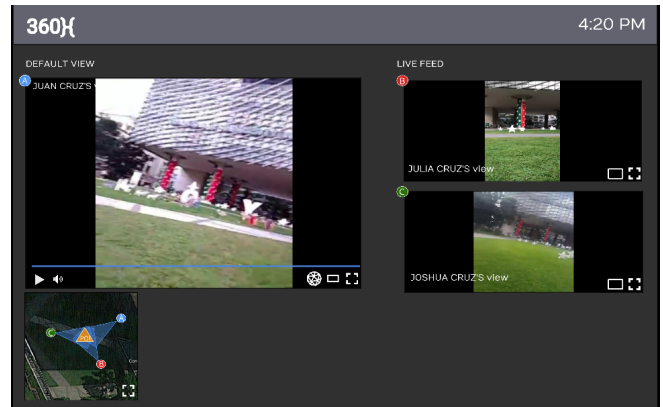


Figure 6. Prototype 2. This prototype shows the different angles on a POI with a main angle being displayed on the left side. The right side shows other angles of the view. Each view angle is indicated on the minimap located on the bottom left of the screen. The minimap indicates the positions of the different view angles (represented by alphabetical letters) as well as the position of the POI (represented by the triangular icon).

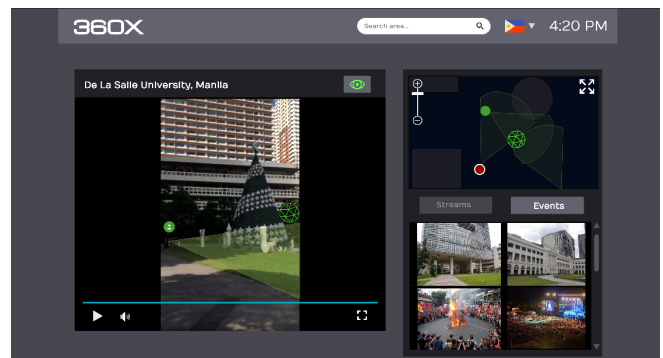


Figure 7. Prototype 3. This prototype has multiple ways to switch between the angles of a POI. In the video player to the left, other people streaming the POI will be highlighted and their field of view will be shown. On the upper right corner, a minimap is shown for an overhead view of the area and POI. The lower right corner will display in a list the different POIs in the area under the events tab and the different people streaming the current POI.

other researches on using location data to generate POIs, improving the user experience of 360-degree videos and their implementations. Our research combines various aspects of these papers to create a flexible crowd-sourced system. For our future work in this project, our second phase will further validate the 360-view following the design of the Prototype 3 as it had the highest average with all of its aspects having the best mean scores. The third phase will involve designing an algorithm that aims to stitch together multiple video streams into one coherent view using a specialized algorithm.

### ACKNOWLEDGMENTS

We would like to thank De La Salle University College of Computer Studies and the Center for Complexity and Emerging Technologies for all the support that they have given us. We would also like to thank the testers and respondents who sacrificed their time for our interaction study. We would also like to thank Prof. Anthony Tang for the inspiration behind this study.

## REFERENCES

1. Abdulrahman Al-Molegi, Izzat Alsmadi, and Antoni Martínez-Ballesté. 2018. Regions-of-interest discovering and predicting in smartphone environments. (May 2018). <https://www.sciencedirect.com/science/article/pii/S1574119217303632?via=ihub>
2. Edward Balassanian and Damon Ganem. 2014. System, method, and product for capturing memories. (July 22 2014). US Patent 8,788,587.
3. Jordan Aiko Deja, Patrick Arceo, Darren Goldwin David, Patrick Lawrence Gan, and Ryan Christopher Roque. 2018a. MyoSL: A Framework for Measuring Usability of Two-Arm Gestural Electromyography for Sign Language. In *International Conference on Universal Access in Human-Computer Interaction*. Springer, 146–159.
4. Jordan Aiko Deja, Rafael Cabredo, and Toni-Jan Keith Monserrat. 2018b. On Building an Emotion-based Music Composition Companion. In *Proceedings of the Asian HCI Symposium '18 on Emerging Research Collection (Asian HCI Symposium '18)*. ACM, New York, NY, USA, 25–28. DOI: <http://dx.doi.org/10.1145/3205851.3205859>
5. Jordan Aiko Deja, Kevin Gray Chan, Migo Andres Dancel, Allen Vincent Gonzales, and John Patrick Tobias. 2018c. Flow: A Musical Composition Tool Using Gesture Interactions and Musical Metacreation. In *International Conference on Human-Computer Interaction*. Springer, 169–176.
6. In-Hwan Han. 2012. Reconstruction analysis of multi-car rear-end collision accidents: empirical/analytical methods, and application of video event data recorder. *Journal of Korean Society of Transportation* 30, 2 (2012), 127–136.
7. James Kai Yu Lau, John P Bruno, and others. 2018. Location and contextual-based mobile application promotion and delivery. (April 3 2018). US Patent 9,936,333.
8. Yung-Ta Lin, Yi-Chi Liao, Shan-Yuan Teng, Yi-Ju Chung, Liwei Chan, and Bing-Yu Chen. 2017. Outside-In: Visualizing Out-of-Sight Regions-of-Interest in a 360 Video Using Spatial Picture-in-Picture Previews. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. ACM, 255–265.
9. Raul Montoliu, Jan Blom, and Daniel Gatica-Perez. 2013. Discovering places of interest in everyday life from smartphone data. *Multimedia tools and applications* 62, 1 (2013), 179–207.
10. Alia Sheikh, Andy Brown, Zillah Watson, and Michael Evans. 2016. Directing attention in 360-degree video. (2016).
11. Prarthana Shrestha, Hans Weda, Mauro Barbieri, Emile HL Aarts, and others. 2010. Automatic mashup generation from multiple-camera concert recordings. In *Proceedings of the 18th ACM international conference on Multimedia*. ACM, 541–550.
12. Samarth Singhal, Carman Neustaedter, Thecla Schiphorst, Anthony Tang, Abhisekh Patra, and Rui Pan. 2016. You are Being Watched: Bystanders' Perspective on the Use of Camera Devices in Public Spaces. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 3197–3203.
13. Yu-Chuan Su and Kristen Grauman. 2017. Making 360 video watchable in 2d: Learning videography for click free viewing. *arXiv preprint* (2017).
14. Nicholas Sula, Peter Seegers, James Lynch, William Gale, and Bradley Kohlmeier. 2010. System and method for stitching of video for routes. (March 30 2010). US Patent 7,688,229.
15. John Edel Tamani, Jan Christian Blaise Cruz, Joshua Raphaelle Cruzada, Jolene Valenzuela, Kevin Gray Chan, and Jordan Aiko Deja. 2018. Building Guitar Strum Models for an Interactive Air Guitar Prototype. In *Proceedings of the 4th International Conference on Human-Computer Interaction and User Experience in Indonesia, CHIUXID '18 (CHIUXID '18)*. ACM, New York, NY, USA, 18–22. DOI: <http://dx.doi.org/10.1145/3205946.3205972>
16. Hanning Zhou, Qiong Liu, Donald Kimber, Lynn Wilcox, and Matthew L Cooper. 2011. System and method for user monitoring interface of 3-D video streams from multiple cameras. (May 17 2011). US Patent 7,944,454.