First Person Omnidirectional Video: System Design and Implications for Immersive Experience

Shunichi Kasahara
Sony CSL, The University of Tokyo
Tokyo, Japan
kasahara@csl.sony.co.jp

Shohei Nagai
The University of Tokyo
Tokyo, Japan
shohei.nagai14@gmail.com

Jun Rekimoto
Sony CSL, The University of Tokyo
Tokyo, Japan
rekimoto@acm.org

ABSTRACT
Fully recording and sharing an immersive experience is one of the ultimate goals of media technology. As extensive technical evolution, omnidirectional video is one of promising media to capture an immersive experience. First person omnidirectional video provides a unique experience of world through someone elses perspective. However, difficulties in wearable camera design and cybersickness induced by shaky video has been obstacle to explore applications of first person omnidirectional video. In this research, we introduce the design and implementation of "JackIn Head" a system including a wearable omnidirectional camera and image stabilization to alleviate cybersickness. Our evaluation revealed the alleviation of cybersickness. Then we report the series of workshops to explore user experience and applications in actual use cases such as virtual travel and virtual sports. We have compiled design implications about cybersickness and motion, immersive sensation, visualization and behavior data of spectators in experience with first person omnidirectional video.

Author Keywords
First person; Omnidirectional video; Wearable camera; Image processing; immersive experience;

ACM Classification Keywords
H.5.1 Information Interfaces and Presentation (I.7): Multimedia Information Systems- video; H.5.2 Information Interfaces and Presentation (I.7): User Interfaces - interaction styles, evaluation, User-centered design

INTRODUCTION
Recording and sharing own experience completely has been an ultimate goal of media technology. As envisioned by the SF movie Brainstorm [25], recording and sharing immersive sensations will allow us to virtually experience what we could not experience by our own. Small wearable wide-angle cameras such as the GoPro[1] enables to record a video of the ones

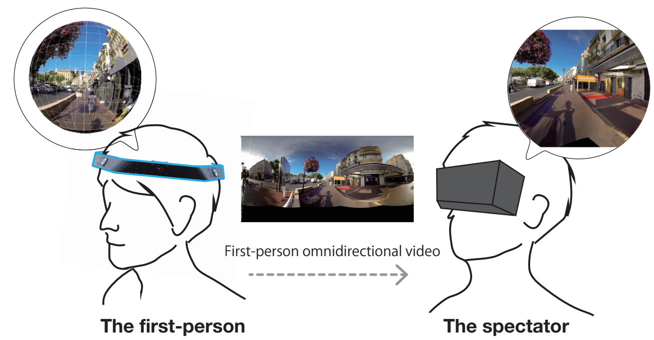

Figure 1. JackIn Head Overview, immersive experience record and playback architecture with a wearable omnidirectional camera.

perspective and then be shared with others. Whereas ordinary cameras capture only a limited field of view in front of them, omnidirectional video cameras are able to capture the entire space. Such a Omni-directional video (ODV) provides spectator experience with impressive and immersive sensations as if they were present in.

There have been active research efforts related to ODV ecosystem in several domains: capturing technology, displaying technology, interaction and control interface technology for ODV, and contents distribution systems for online video workflow. These series of research and development will make ODV more common and bring highly immersive user experience in the context of interactive television and online videos. Therefore, first person omnidirectional video (FODV) will bring a lot of opportunities for ODV applications. Bleumers et al. [6] recently presented a number of interesting findings regarding users expectations of ODV, especially FODV will play important role in various applications such as entertainment, sports viewing, education and simulation training, news casting and therapy.

Although design implications for FODV have been required, it has not been well explored due to two major difficulties. The first is design of capturing device for FODV. A wearable omnidirectional camera attached to the head is one of reasonable design to acquire video from one's perspective. However, there is still no existing solution to satisfy all of robustness, wearability and eye-level position. Another issue is quality of the first-person video from wearable cameras. Video from wearable cameras are often shaky due to the body movement during capturing. Viewers often feel dizzy when

[1]GoPro http://gopro.com/

TVX'15, June 03 - 05, 2015, Brussels, Belgium

http://dx.doi.org/10.1145/2745197.2745202

watching video from wearable cameras, and this gets worse when the visual experience is immersive on a large screen or HMD [11]. This is called as cybersickness. Then this should be investigated to make better user experience.

Our aims are to provide a solution for these existing problems, and to explore design implication of FODV with actual use cases. In this study, we proceed following research steps. We first describe "JackIn Head" system design including a wearable omnidirectional camera and image processing for stabilizing FODV (Fig 1). Then we evaluate how our system alleviates cybersickness and we collect findings from the investigation and interviews. After we understand the risk and prevention of cybersickness with FODV, we perform series of demos and workshops with three activities to prove actual benefit of FODV and collect insight from participants. The results of our inquiries and observations inform design implication and possible applications of FODV.

The main contributions of this paper are: (1) a system design including capturing device and image processing for stabilizing FODV, (2) an investigation of cybersickness and user experience of the system, and (3) design implications for FODV from series of demo and workshop with our proposed FODV system. Out system design and implication will contribute to make FODV a significant interactive video media in ODV eco system.

RELATED WORK

User experience in omnidirectional video

Omnidirectional video is expected as a novel medium that offers immersive and interactive experience. Bleumers et al. [5] reported various users expectations of ODV, and many of them are also supposed to use first person omnidirectional video. Zoric et al. [27] investigated viewing and interaction with panorama video in a TV screen, the result informs benefits of interacting with panoramic content and design challenges including the active - passive viewing and the social aspect. Barkhuus et al. [2] show the possibilities of omnidirectional video streaming for a live performance. FascinatE project developed comprehensive eco-system that enables capture, delivery and reproduction for immersive media including ODV[19]. Decock et al. reported about a case study of an interactive performance called C.A.P.E which uses first person omnidirectional video and HMD, and inform user experience drivers for effects of presence and identification on enjoyment [8].

A series of research and development for immersive displaying technologies provide various options for ODV, such as large screen TV, light weight HMD with mobile screen [2], CAVE like spherical environment [4], HMD with head tracking [3] and projection mapping for the room [12]. In term of researches for user interface, Neng et al. [17] explore about navigation and visualization mechanisms of ODV. Second screen interface [27, 2] and mid air gesture [28] has been studied as an interface for control ODV. Ruiz et al. [21] provide comprehensive interaction design for mid-air gestures to control ODV from individual and collocated usage. These researches effort present many possible scenarios in terms of social viewing with collocated person as well as personal immersive environments and increase the expectation of ODV contents as an interactive media.

Omnidirectional camera

Previous researches of interactive panorama video or ODV has been performed with footages from fixed position. However, in view of various possibilities in ODV, first person omnidirectional video (FODV) should be studied to expand ODV-ecosystem. There have been many approaches that acquire omnidirectional video such as hyperboloid mirror from single camera, multiple cameras, multiple mirror projection.

Recent miniaturization of image capture module and lens technology enables more compact systems and several commercial companies offer omnidirectional cameras from on-vehicle equipment [4] to handheld omnidirectional camera [5]. Another solution could be a holder for multiple wearable camera[6].

However these solutions enable wearable shooting near the first person perspective, the cameras should be located above one's head. This causes the gap between the camera viewpoint and the eye-level position of a wearer. This will lead gaze mismatching [15] and feel of hovering rather than standing at the same position as first person [5]. Furthermore, attaching cameras on top of the head causes shift of the center of gravity toward higher than one's head. This is inappropriate especially for shooting physically dynamic activities such as sports. Kondo et al. [15] proposed the capture devices system that enable eye-level recording of omnidirectional panorama video with uniform resolution with the special designed mirror. They have shot the FODV footage of actual ball sports from the POV of non-player. However to capture the first person ODV form the actual player, the robustness of the head gear and wearability has been should be solved. By considering these advantage and disadvantages in several approaches, reasonable design is required.

Cybersickness

Wearable camera footages contain a lot of shaky scenes, then viewers often become dizzy or nauseous especially with immersive visual environment. Especially in virtual reality applications, it's called "cybersickness"[24]. One of the main theories as to what causes simulator sickness is based on sensory conflict. When the perception of self-motion from visual optical flow patterns is not corroborated by inertial forces transmitted through the vestibular system, humans will experience symptoms of sickness including dizziness, disorientation, and even vomiting[13]. Moreover, an immersive visual setup such as large screen, HMD or CAVE could intensify cybersickness [11]. Inconsistency between visual movement of FODV and physical motion of spectators will cause cybersickness. Then the investigation and prevention technology for cybersickness with FODV is required.

[2] cardboard google.com/get/cardboard/
[3] Oculus rift www.oculus.com
[4] Ladybug http://www.ptgrey.com
[5] THETA theta360.com
[6] 360heros http://www.360heros.com

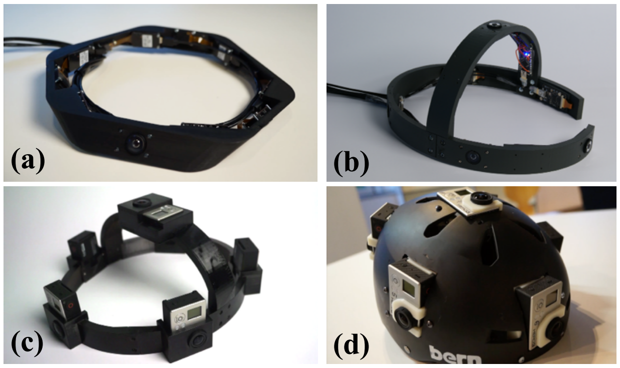

Figure 2. Variation of JackIn Head headgear. (a) and (b) are lighter version with usb cameras. (c) and (d) are HD version with six HD cameras.

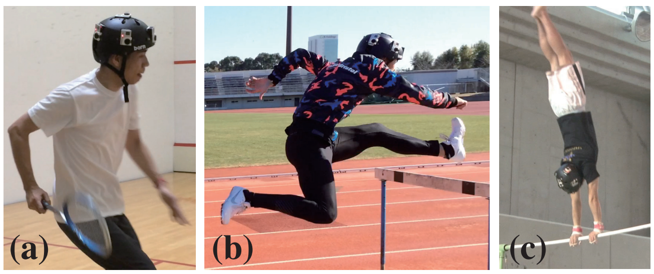

Figure 3. Actual use of the headgear type-d in several situation. (a) playing the squash, (b) track, (c) giant swing on the high bar

Omnidirectional image processing and stabilization

Image stabilization is a fundamental technology to prevent cybersickness, as well as to improve quality of video. Johannes Kopf et al. proposed hyper-lapse generation, which converts first-person video into smooth time-lapse videos [16]. Omnidirectional video has been used in research relating to ego-motion estimation of the camera or self-localization[10, 1, 26], as well as omnidirectional visual immersive simulations[18]. There are various procedure to estimation rotational motion of camera, with only images [10, 26] or with the motion sensors [1]. Bazin et al. proposed a method which can estimation translational movement and rotation [3]. These technology were applied for robot navigation and vehicle control [3, 9]. The estimation of ego-motion from visual information, called as Visual Gyroscope[7], is can be applied for video stabilization. Then, to deploy these technologies for stabilization of omnidirectional video from wearable camera, special requirement of FODV such as intense rotational motion, generation of smoothed movement, and realtime feasibility for practical usage should be also considered.

SYSTEM DESIGN

In this sections, we first give a system design description of JackIn Head we developed: the system for immersive experience sharing with first person omnidirectional video. JackIn Head system includes a wearable headgear(Figure 2) with multiple camerasto capture omnidirectional video, image processing for FODV stabilization, the streaming video data and a playback system with screen or HMD. Video images from these cameras are stitched together into a spherical omnidirectional video and image processing for stabilization is performed. The stabilized omnidirectional video and the head ego-motion information of video are streamed to the viewer device on the spectator side application (Fig. 1). In the spectator side, there are various options for viewing device, example setup is a HMD with head motion tracking (such as Oculus rift [7]) that the spectator uses to look around the first person visual environment.

In the following sections, the design of headgear and image processing are described in detail.

Headgear with omnidirectional camera

The headgear includes six wide-angle cameras with fix position in a rigid body. The head gear was designed with two design considerations. One is lower center of balance, which allows users to move their body and head dynamically. The high center of balance is dangerous for even usual activity. In our preliminary studies, attaching the omnidirectional camera on top of the head affects the physical movement, then weassess that it will be dangerous. Another is that the captured environment should be close to the first person viewpoint. This provides a realistic sensation of the first person perspective. This design of embedding cameras into headgear will produce gap of focal point of each cameras. However this result in noticeable gap between each cameras image especially near by objects are on image seam, we prioritize those design considerations.

We designed several versions of headgear prototype including the light weight version with USB camera (Fig 2 -(a,b)) and Figure 2 -(c,d) shows HD version with 6 HD cameras (Gopro Hero3). Although the light weight version is one of our design targets, existing USB camera module has limitation in terms of image quality and video frame rate. In this study, we employ HD version headgear to record FODV. Figure 3 shows actual use in sports context.

The range of capture is shown in Figure 4, which cover omnidirectional except for the bottom part. Each camera capture 960x1440 pixel, high frame rate (100 fps) with 122.6

[7]Oculus rift DK-2 : http://www.oculus.com/

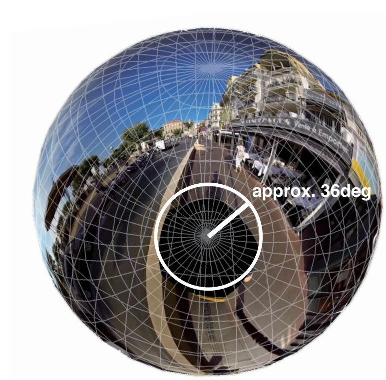

Figure 4. The range of capture area by omnidirectional camera. The headgear can capture spherical omnidirectional video except the bottom 2 x 36deg area.

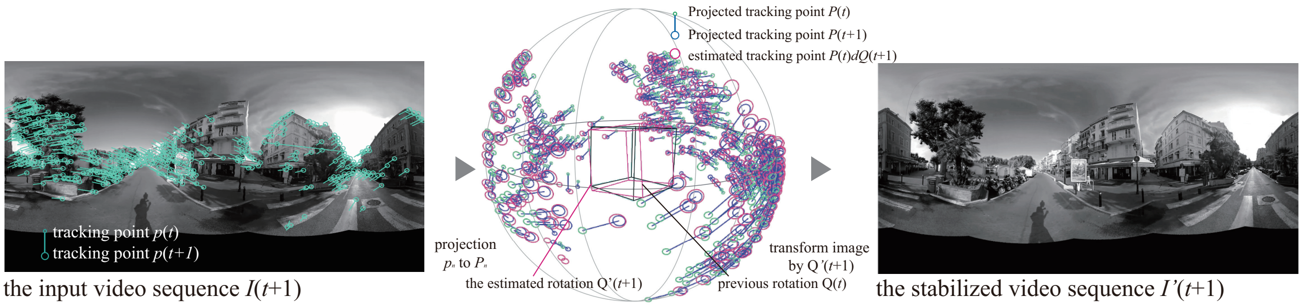

Figure 5. Image processing procedure for estimation of head rotation and stabilization.

deg horizontal and 94.4 deg vertical field of view. Camera calibration for six cameras is performed using the Omnidirectional Camera Calibration Toolbox [22]. In the computer, six video streams are stitched into an omnidirectional video as equirectangular video with a GLSL shader

Design parameter about FOV of each cameras and number of camera are trade off relationship. Wider FOV wearable camera will reduce number of camera, thus will reduce the visual gap in stitched omnidirectional video. However, there is also design trade-off in between FOV and the image quality. Through our initial exploration, longer shutter speed or lower frame rate produce image blur which cause error in stabilization. Another important technical detail is frame synchronization. Stitching multiple lower frame rate videos generates not only visual gap, but also unexpected jitter through the stabilization process. Through our initial exploration, multiple video with 100 fps (i.e. maximum temporal error is 5ms) did not cause this problem. From these conditions, we chose our current configuration for HD version head.

Image stabilization for first person ODV

In our preliminary exploration, we recorded and gather FODVs with our headgear in various activities. We assessed that image motion in FODV is mainly caused by rotation of the wearer. Then we implemented image processing for estimating of rotational motion in FODV, and image stabilization by eliminating rotational motion from FODV (Figure 5).

This stabilizing algorithm has two phases, it first calculates rotation then eliminate its rotation, then it re-produce smoothed rotation to follow the wears head direction with variable parameter as describe later. If this parameter is zero or quite small, the output footage will be **independent view** which are completely decoupled from the head motion of the wearer. By contrast, if the following parameter is larger value, the output will be noise reduced **smoothed view**, which gradually traces the head motion of the wearer.

An omnidirectional video from the headgear is treated as an equirectangular image, which is the standard format for spherical geometry such as global maps.

In each equirectangular video frame $I(t)$, image feature points $p_n(t)(n = 1000max.)$ are extracted by finding the visual corner in the equirectangular image[23]. High latitude and bottom areas are excluded from ROI for this process.

We use the pyramidal KLT method [6] to calculate the optical flow $f_n(t)$ for each $p_n(t)$. Tracked points in the next image sequence $I(t+1)$ are then estimated as $p_n(t+1) = p_n(t) + f_n(t)$.

Next, the 2-D image feature points $p_n(t+1)$ and $p_n(t)$ are converted into 3-D points $P_n(t+1)$ and $P_n(t)$ on spherical geometry with spherical polar coordinates. Here, the radius for conversion does not affect successive processes.

Then, the affine transform matrix $M(t+1)$ to describe the affine transform as $P_n(t+1) = P_n(t)M(t+1)$ is estimated using RANSAC and a differential rotation from $I(t)$ to $I(t+1)$ is acquired as quaternion $dQ(t+1)$.The rationale of using RANSAC is for handling a lot of outlier in the matching space.

Here, the estimation error is calculated. If the error $Err(t{+}1)$ is larger than a threshold, it is considered an estimation failure and the prediction is done using a Kalman filter instead.

$$Err(t+1) = Pn(t+1) - dQ(t+1)Pn(t)$$

By multiplying all differential rotation $dQ(i)(i = s, ...t)$ from the reference start time s to the current time t in every frame, the rotation from the reference start time can be calculated:

$$Q(t) = \prod_{i=s}^{t} dQ(i)$$

The rotation eliminated equirectangular image $I'(t)$ can then be generated by converting $I(t)$ by the inverted rotation $Q(t)^{-1}$:

$$I(t) = I(t)Q(t)^{-1}$$

Note that the sequence of $I'(t)$ is a stabilized video sequence and the sequence of quaternion $Q(t)$ represents the decoupled head ego-motion.

In practice, it often happens that the wearer is walking in a town and turns right and the spectator would like to follow in the same direction the wearer is headed. However, the process described above is just for the elimination of rotational motion, and it forces the spectator to keep heading to the right in order to follow the direction.

By observing the time sequence of quaternion $Q(t)$, in the playback application, the system evaluate whether the movement is noise or intentional motion. If the rotation movement is evaluated as intentional motion, the playback system gradually sift the viewing reference direction Q_{view} of ODV with spherical linear interpolation.

$$Q_{view}(t+1) = slerp(Q_{view}(t), Q(t), k)$$

The parameters value (k) for interpolation varies depends on the context of FODV. A smaller interpolation results **independent view** that suites for video with intense motion such as sports contexts. In contrast, A larger interpolation rate results noise reduced **smoothed view** that suites for human dairy activity.

EVALUATION OF CYBERSICKNESS

In this section we report the evaluation about how our system alleviates cybersickness and collected findings from the investigation and interviews. We conducted experiments to evaluate the cybersickness alleviation and to investigate how individual viewing experience can be achieved.

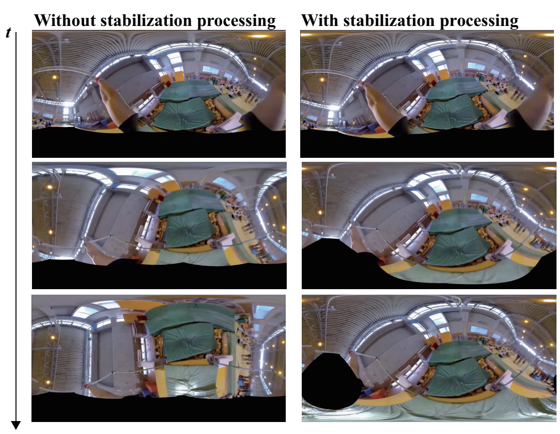

Figure 6. Example of stabilized omnidirectional video sequence.

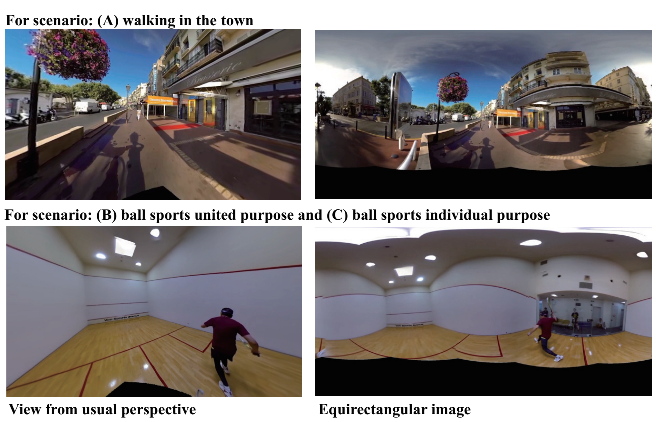

Figure 7. Sample frame from omnidirectional video for test scenario: (A) walking in the town, (B) ball sports, watching to trace the ball, and (C) ball sports, watching to trace another player. Experiment conditions are consists of these video sequences with and without Stabilization.

Experiment Procedure

JackIn Head aims for active viewing experience as well as passive viewing. Thus, to investigate the effect of activity, we designed an experiment with three scenarios: (A) virtually traveling, (B) experiencing a sport, and (C) experiencing a sport with different interest against wearer with intense motion.

Scenario (A): Walking in the town

In this scenario, we assume virtual travel in which the spectator observes an immersive scene of the walking around in an unknown town (Figure7-(A)). Omnidirectional video is captured from JackIn Head gear while walking a street in Cannes at 90 m/min. The estimated average rotational speeds in Euler angle are 12.0(pitch), 19.4(yaw), 13.5(roll) deg/sec. The maximum speeds are 185.5(pitch), 353.0(yaw), 137.5(roll) deg/sec. The participant is asked to find as many potential restaurants for lunch and dinner as possible. In this scenario, The FODV footage was stabilized as **smoothed view**.

Scenario (B): Ball sports, watching to trace the ball

In this scenario, we assume a sports experience in which the spectator observes an immersive scene of playing squash and tries to track the movement of the ball (Figure7-(B)). Omnidirectional video is captured from JackIn Head gear while the squash game in a 6.4 x 9.75 m court with another player also on the same side of the court. The estimated average rotational speeds in Euler angle are 15.7(pitch), 50.7(yaw), 20.3(roll) deg/sec, the maximum speeds are 531.2(pitch), 793.4(yaw), 63.5(roll) deg/sec. The participant is asked to track the ball as much as possible.

Scenario (C): Ball sports, watching to trace another player

In this scenario, we assume a remote assistance situation in which the spectator observes an immersive scene of the squash game and looks at the other player with difference interest. This scenario uses the same omnidirectional video (Figure7-(B)) as scenario (B) so as to compare the two scenarios. The participant acting is asked to track the other player as much as possible. In scenario (B) and (C), The FODV footage was stabilized as **independent view**.

It is assumed that scenario (A) contains a lot of unconscious movement while scenarios (B) and (C) contain intentional movement.For each scenario, the participants perform the tasks with stabilized omnidirectional video and conventional omnidirectional video. The participant uses Oculus DK2 as the HMD with a display refresh rate of 60 Hz; the movie fps is also 60 Hz.

Experiment setup

Participants perform six tasks in total, i.e. Non-stabilization or Stabilization in condition A, B, C, then we describe A-N, A-S, B-N, B-S, C-N, C-B respectively. Each task is two minutes with a recovery time of 30 minutes interval. The order of scenarios is counter-balanced, as is the order of stabilized and conventional vision.

Before and after each task, the participant filled out a simulator sickness questionnaire (SSQ) [14], which is a well-known

Average(SD)	A	B	C
No Stabilization	27.1(26.8)	12.2 (23.2)	30.5 (25.2)
Stabilization	11.8(15.2)	5.6(13.2)	0.6 (11.26)

Table 1. Overview of results for SSQ score change values, average and standard deviation. (A): walking in the town, (B): ball sports, watching to trace the ball, (C): ball sports, watching to trace another player

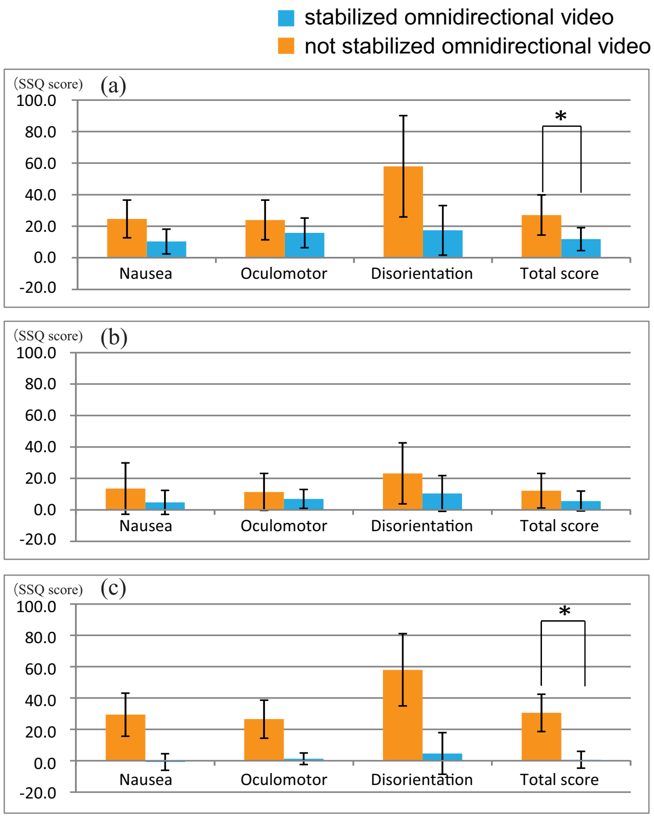

Figure 8. Detail of results for SSQ score change values, about nausea, oculomotor discomfort, disorientation, and general cybersickness. Error bar represent 95 % confidence interval.

measurement tool to evaluate simulator sickness and cybersickness. The SSQ value derived from the questionnaire indicates the degree to which the participant felt nausea, oculomotor discomfort, disorientation, and general cybersickness. Here, the larger the SSQ value, the stronger the feeling of a particular symptom. Twelve volunteers aged 20-38 participated in our study after providing informed consent and were allowed to give up at any time, even in the middle of a task, although the questionnaires had to be filled out in any case. The head motion which can be acquired from the head tracking HMD was recorded during tasks.

Results of Experiment

In general, according to the averaged SSQ values, the stabilization lessened the instances of cybersickness in each scenario. In our analysis, we compare the SSQ score changes before and after each task. Table 1 shows the average and variance value of differential SSQ values for all symptoms and Figure 8 shows the detailed results from the SSQ values for both the total and detailed symptoms. There was a significant difference ($p < 0.05$), between A-N and A-S, also between C-N and C-S. We observed that C-N was the worst condition because half of the participants gave up before task time (2 min) was completed. As for the detailed results, "Disorientation" in scenario (A) and all detailed symptoms in scenario (C) exhibited significant difference between with and without stabilization. Interestingly, in scenario (B), there was no significant difference.

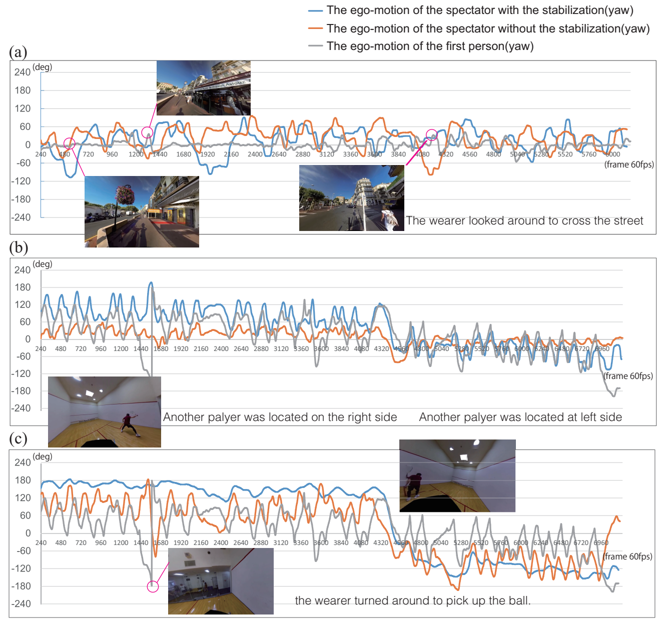

Figure 9. The sequence of yaw rotation of the most representative sample to describe characteristics of the spectator ego-motion.

Analysis and Observation

We also observed the ego-motion sequence of spectators and wearer. Fig 9 shows the sequence of yaw rotation of the most representative sample to describe characteristics. The head motion of spectators can be acquired from head tracking HMD, and the head motion of first-person was acquired from image processing. After completion of all tasks, we conducted interviews about how they felt during viewing FODV, what made them sick and what created good experience.

The analysis of time series data and interviews yields a set of findings that is variable insight and challenges to improve user experience.

Active viewing behavior or not

In scenario (A) without stabilization (A-N) , almost all participants reported that they felt strong disorientation when sudden motion in the direction opposite to their intention occurred. This phenomenon was also evident in C-N in Fig. 9. Interviews with participants who did not feel much disorientation in these conditions revealed that they were able to identify what made them feel disoriented during the task and then refrained from moving their heads as much as possible. Because an active viewing behavior is a significant interaction in ODV viewing experience. In this sense, the stabilization would be a beneficial feature in FODV application. Furthermore, the interview and time sequence data of condition (A)

Figure 10. Overlay indication of wearer's head direction in omnidirectional video. (a) walking in the town, (b) giant swing on the high bar.

reveals that many head motion of spectators stem from the own motivation of spectators in A-S, on the other hands, the head motion in A-N was induced from the motion of FODV as noise.

Synchronization and Asynchronization

In condition (B) this is evident that the wearer was also tracing the ball as well, and the spectator didn't need to trace it actively, in contrast to in condition B-S, the spectator had to trace the ball by their own. This phenomena are also obvious with concurrently synchronized ego-motion sequence in Figure 9. Here note that the condition B-N also shown this synchronous behavior, even though both head motions do not have completely the same value, the direction of the motions was consistent.

Although there is no significant difference between B-N and B-S, seven of twelve participants made comments like *"I had a realistic feeling in B-N conduction"*. Analysis of the head motion sequence (Figure 9) suggests that synchronization in motion of FODV and the spectator will lead lessen cybersickness, and also produce more realistic sensations.

Existence of first person

In interviews after all tasks, several participants reported that condition A-N was worse than C-N, even though both caused conflicted motion. The reason for this can possibly be summed up in the comment that *"in the motion in condition A-N, I could not predict the motion of the video so, even the motion itself was small, I felt stronger sickness"*.

In addition, other participants commented that*"in stabilized FODV condition, I felt it was difficult to understand what the wearer was doing, because I counld not find a cue for movement"*. This indicates the head direction is an important indication of awareness, and was eliminated by stabilization.

This feedback reveals that visualization of the wearer's head direction in the immersive view for the spectator will be key in addressing these problems. We therefore implemented an overlay indication of the wearer's head direction for the viewer based on the rotation information acquired from image processing (Figure 10). We also note that, to reproduce more precise first person video, dedicated eye tracking is required along with JackIn head gear. This should be addressed as technical challenge for future work

Cybersckness induced by translational movement

Some participants commented that even with stabilization, they felt cybersickness temporarily. In depth-interviews about the moment at which cybersickness occurred, it indicated there are two cases in the stabilization footage.

The first case is a sudden change of speed in translational motion, such as lunging forward to pick up the ball in the squash scenario (B) and (C), and suddenly stopping to look around when crossing the street in scenario (A). In general, situations in which the viewer expects to experience acceleration from the visual motion caused temporal cybersickness.

The second case is walking vibrational motion where, participants feel some oculomotor discomfort. This might be caused by the vertical translational motion, which cannot be eliminated by current stabilization process. From those feedbacks, compensating for the translational acceleration change and translational vibration noise should be addressed as technical challenge for future work.

EXPLORATION THROUGH WORKSHOPS

After we found the stabilization works for prevention of cybersickness with the FODV, we performed series of demo and workshop with three activities, a virtual travel, virtual gymnastics and virtual flying experience of paraglider. We selected two sport activities because we found these physical activity will produce a lot of insight and challenges for system and user experience. In series workshops and demos, over 50 people participated our activity. The first person omnidirectional video in each demo was viewed with HMD with head tracking[8] and 15 inch laptop screen. Participants of workshop were also allowed to explore the way of viewing FODV footages, we interactively tested the various viewing experience. The results of our inquiries and observations inform set of finding and possible applications of FODV.

Virtual travel

In the this demo material, FODV that captured walking in street, beach and historical place in Cannes. A part of this material was also used for evaluation of cybersickness. Although the task duration in out experiment for cybersickness was just 2 min, many of participants in the demo enjoyed to watch longer especially with HMD. This virtual travel with beautiful landscape is generally well received. The moment that the wearer was talked by a local person provide realistic feeling such as *"I was a bit upset, because the person in the video talked to me"* commented from a participant. This realistic feeling was provided from the eye-level position of the headgear. However some participant complained about visual gap when watching near object especially a someone's face. Another remarkable finding among them is a possibility of revisiting own experience. The collaborator who recorded this footage commented that *"I felt like I revisited this place from my own perspective. I found an interesting scene behind me that I could not see at that time"*.

Virtual gymnastics

To investigate the possibilities of viewing intense sport activities. We recorded FODV during the giant swing on the high bar and trampoline of gymnastics. Participants both with and without experience of gymnastics watch FODV with our system. In this workshop, we also showed the overlay indication with FODV like Figure 10.

[8]Oculus rift www.oculus.com

Figure 11. (a) : Example view of first person omnidirectional video of the paraglider. (b) HMD playback with a harnessfor practice greatlyincrease realistic feeling and excitement.

In general those contents were well received with both screen and HMD. The person who shot the FODV of the giant swing commented that *"For me, non stabilized version of FODV is similar to own experience. Maybe because I usually focus on the bar during the swing."*. In the other hand, Participants who haven't experience gymnastics commented that *"I could not understand what is going on without the stabilization. The indication graphics helped me understand the movement"*. One of athlete player comments that *"I like to use it for own practice, if it's more light weight"*.

Virtual flying experience of the paraglider

As envisioned by many time from a lot of participants of workshop, we also performed a workshop to examine the virtual flying experience. Participants were persons who usually experience the paraglider as their hobby or as professional. One of participants in the workshop used our headgear to capture FODV during the flight Figure 11-(a). Then we discussed the benefit of FODV and explore the way how to reproduce the experience that they are usually feel.

Many of participants excited about that they can watch the right above and even behind themselves where they could not watch during the flight. One of participant who could not have a flight that day tried our demo and commented that *"This (partially) gave me the satisfaction, I would like to use it for my parents to give virtually experience of mine"*.

Participants found a lot of beneficial use cases in training, education and promoting scenario. One of professional player comments that *"This could be useful fora virtual tutorial, a novice player can understand what will be happened and how they will feel"*. An intermediate player in participants also suggested a possibility ofskill acquisition *"I would like to watch the FODV of the professional perspective, that would be very beneficial instructional material"*. The owner of the paraglider school also mentioned that *"I usually use the GoPro to capture the flight, I would like to capture with this headgear, and use that material for a promotion of the paraglider"*.

Several participant tried our HMD playback system with a harnessfor practice that is hung from the bar, and we found it greatlyincrease realistic feeling and excitement Fig 11-(b). This especially indicates that those kind of haptic equipment to produce an imitational physical sensation with our system will achieve higher immersion.

DISCUSSION AND IMPLICATION FOR DESIGN

Technical challenges

In our studies, one of our aims is to provide technical and design solution for existing problems. Through this exploration, we have also found further technical challenges. From series of workshops with actual sports athletes, we found the headgear should be less weighted and more suitable for wear, and even flexible form factor. The stabilization also should be improved in terms of compensating for the translational acceleration and vibration noise. Due to the form factor of our headgear design, visual gaps in the omnidirectional image also should be addressed. A prior research achieved visually adaptive multi camera image stitching with image feature extraction such as faces and objects [20]. These technology will improve the quality of the first person omnidirectional video. Throughout the series of workshops, we've also got expectation about a realtime version of our system, a realtime processing will be also future technical challenge.

Design Implication

Our JackIn Head systemallow to explore design implication of FODV with actual use cases as our second aim of this research. In what follows, we summarize our findings and insights into design implications.

Cybersickness and FODV

In our evaluation of cybersickness with first person omnidirectional video, we proved that the stabilization process is fundamental feature even for short footages. The findings showed the risk of cybersickness even with footage of walking activities. Especially a conflict between the motion of the ODV and the motion of the spectator will cause an intense sickness, then we need to assess what kind of motion are contained in the video and how the spectator will watch it. Then the stabilization greatly contribute the prevention of cybersickness in immersive viewing environment. Although our investigation is focusing on viewing experience with HMD, these findings can also be useful in other viewing environment for omnidirectional video.

Visualizing motions of the first person

We found that overlaying ego-motion of the wearer in omnidirectional video will help the spectator to understand what the wearer was watching and doing. This feature was especially well received in sports context in the workshop. This design implication is valuable in various applications with a large option of viewing devices. However the current implementation only supports the head ego-motion, adding the gazing information will also open the possibilities of application in the context of professional sports player. musical performance.

Design space for first person omnidirectional video viewing

Throughout our research, we defined design space for viewing application with two dimensions. The first dimension is the activity of wearer that varies from the one physical embodiment matters (ex. sports.) to the one doesnt (ex, sightseeing, travel and shopping). The second dimension is the spectators viewing activity which varies synchronizing, passive and independent. "synchronizing" and "independent" are

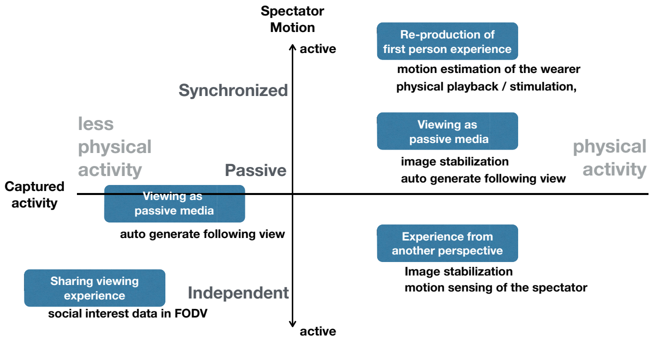

Figure 12. Design space for first person omnidirectional video. This consist of two design dimensions ; What kind of activity of the wearer was captured and how spectators will be expected to see the FODV.

responded to synchronization and asynchronization between the motion of FODV and the spectator respectively. It indicates what kind of user experience is expected, how stabilization should be applied and what other implementation matters. In our early study, the main application region was assumed to be providing another perspective in recorded FODV (Fig 12 - **re-experience from another perspective**), However, other application region has been also found alongside the design space. Other type of experience will be referred in following section.

Designing Synchronization and Asynchronization of motion

The results of our investigation of cybersickness also yield a interesting design possibility about synchronization and asynchronization between the motion of FODV (i.e. the motion of the wearer) and the motion the spectator. Synchronization of motion will lessen the induction of cybersickness, but also more interestingly, will produce an immersive realistic sensation. For instance, adaptive control of stabilization will be one of possible application to increase immersion.

The analysis of the rotational data sequence indicates that the motion synchronization will not need to be the same intensity, the smaller motion with the similar direction would be enough to induce a perceptual cues for virtual sensation. This indicates that we can augment physical sensation and allow users to feel another person's physical motion as their own (Fig 12 - **re-production of first person experience**). In other words, by fine design of FODV application, the spectator can feel the realistic sensation of a physical motion that we cannot perform in real life, such as a triple jump in figure skating.

Viewing angle information of spectators

From an interesting finding about the head motion of the spectator, the stabilization of omnidirectional video will produce meaningful behavior data of the spectator. This is because that the spectator would move their head for their interest in the stabilized ODV without distraction of the head motion of the wearer. The collected date of those information will generate meaningful information about social interest in ODV, and will be significant property in the eco-system of omnidirectional video (Fig 12 - **sharing viewing experience**). For instance, this kind of information will be an important ingredient of automatic contents generation from ODV footage for the conventional passive viewing application (Fig 12 - **viewing as passive media**).

CONCLUSION

Extensive technical developments in recent years allow various immersive experience in the interactive media ecosystem. Omnidirectional video is one of promising media and have capability to capture an immersive first person experience. However, difficulties in capturing device and cybersickness induced by the first-person video have been obstacle to explore applications of first person omnidirectional video.

In this paper we aimed to provide a solution for existing problems and to explore design implications. We first introduced the system design and implementation including capturing device and image stabilization to alleviate cybersickness. Our evaluation revealed the alleviation of cybersickness by our system. Based on our system, we performed the series of workshops to explore user experience and applications. Then we summarized our findings and insight into design implications. These design implications will contribute a further exploration of first person omnidirectional video, but also, more broadly, contribute to expand the eco-system of interactive experience of various media.

We're researching about human to human telepresence" whereby individuals can record and share their own immersive experiences and concurrently experience shared sensations and communicate with others. Toward "human telepresence , this paper presented one tangible topic that focuses recording and sharing ones immersive experience with omnidirectional video.

REFERENCES

1. Albrecht, T., Tan, T., West, G., and Ly, T. Omnidirectional video stabilisation on a virtual camera using sensor fusion. In *Control Automation Robotics Vision (ICARCV), 2010 11th International Conference on* (Dec 2010), 2067–2072.
2. Barkhuus, L., Engström, A., and Zoric, G. Watching the footwork: Second screen interaction at a dance and music performance. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, ACM (New York, NY, USA, 2014), 1305–1314.
3. Bazin, J.-C., Demonceaux, C., Vasseur, P., and Kweon, I. Rotation estimation and vanishing point extraction by omnidirectional vision in urban environment. *Int. J. Rob. Res. 31*, 1 (Jan. 2012), 63–81.
4. Benko, H., and Wilson, A. D. Multi-point interactions with immersive omnidirectional visualizations in a dome. In *ACM International Conference on Interactive Tabletops and Surfaces*, ITS '10, ACM (New York, NY, USA, 2010), 19–28.
5. Bleumers, L., Van den Broeck, W., Lievens, B., and Pierson, J. Seeing the bigger picture: A user perspective on 360° tv. In *Proceedings of the 10th European*

Conference on Interactive Tv and Video, EuroiTV '12, ACM (New York, NY, USA, 2012), 115–124.

6. Bouguet, J.-Y. Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm. *Intel Corporation 5* (2001).
7. Carlon, N., and Menegatti, E. Visual gyroscope for omnidirectional cameras. In *Intelligent Autonomous Systems 12*. Springer, 2013, 335–344.
8. Decock, J., Van Looy, J., Bleumers, L., and Bekaert, P. The pleasure of being (there?): an explorative study into the effects of presence and identification on the enjoyment of an interactive theatrical performance using omnidirectional video. *AI & SOCIETY* (2013), 1–11.
9. Gandhi, T., and Trivedi, M. Parametric ego-motion estimation for vehicle surround analysis using an omnidirectional camera. *Mach. Vision Appl. 16*, 2 (Feb. 2005), 85–95.
10. Gluckman, J., and Nayar, S. Ego-motion and omnidirectional cameras. In *Computer Vision, 1998. Sixth International Conference on* (Jan 1998), 999–1005.
11. Howarth, P., and Costello, P. The occurrence of virtual simulation sickness symptoms when an hmd was used as a personal viewing system. *Displays 18*, 2 (1997), 107–116.
12. Jones, B. R., Benko, H., Ofek, E., and Wilson, A. D. Illumiroom: Peripheral projected illusions for interactive experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, ACM (New York, NY, USA, 2013), 869–878.
13. Kennedy, R. S., Drexler, J., and Kennedy, R. C. Research in visually induced motion sickness. *Applied ergonomics 41*, 4 (2010), 494–503.
14. Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology 3*, 3 (1993), 203–220.
15. Kondo, K., Mukaigawa, Y., and Yagi, Y. Wearable imaging system for capturing omnidirectional movies from a first-person perspective. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, VRST '09, ACM (New York, NY, USA, 2009), 11–18.
16. Kopf, J., Cohen, M. F., and Szeliski, R. First-person hyper-lapse videos. *ACM Trans. Graph. 33*, 4 (July 2014), 78:1–78:10.
17. Neng, L. A. R., and Chambel, T. Get around 360° hypervideo. In *Proceedings of the 14th International Academic MindTrek Conference: Envisioning Future Media Environments*, MindTrek '10, ACM (New York, NY, USA, 2010), 119–122.
18. Neumann, U., Pintaric, T., and Rizzo, A. Immersive panoramic video. In *Proceedings of the Eighth ACM International Conference on Multimedia*, MULTIMEDIA '00, ACM (New York, NY, USA, 2000), 493–494.
19. Niamut, O. A., Kochale, A., Hidalgo, J. R., Kaiser, R., Spille, J., Macq, J.-F., Kienast, G., Schreer, O., and Shirley, B. Towards a format-agnostic approach for production, delivery and rendering of immersive media. In *Proceedings of the 4th ACM Multimedia Systems Conference*, ACM (2013), 249–260.
20. Ozawa, T., Kitani, K. M., and Koike, H. Human-centric panoramic imaging stitching. In *Proceedings of the 3rd Augmented Human International Conference*, AH '12, ACM (New York, NY, USA, 2012), 20:1–20:6.
21. Rovelo Ruiz, G. A., Vanacken, D., Luyten, K., Abad, F., and Camahort, E. Multi-viewer gesture-based interaction for omni-directional video. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, ACM (New York, NY, USA, 2014), 4077–4086.
22. Scaramuzza, D., Martinelli, A., and Siegwart, R. A toolbox for easily calibrating omnidirectional cameras. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, IEEE (2006), 5695–5701.
23. Shi, J., and Tomasi, C. Good features to track. In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on* (Jun 1994), 593–600.
24. Stanney, K. M., Kennedy, R. S., and Drexler, J. M. Cybersickness is not simulator sickness. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 41, SAGE Publications (1997), 1138–1142.
25. Trumbull, D. Brainstorm, 1983.
26. Vassallo, R. F., Santos-Victor, J., and Schneebeli, H. J. A general approach for egomotion estimation with omnidirectional images. In *Proceedings of the Third Workshop on Omnidirectional Vision*, OMNIVIS '02, IEEE Computer Society (Washington, DC, USA, 2002), 97–.
27. Zoric, G., Barkhuus, L., Engström, A., and Önnevall, E. Panoramic video: Design challenges and implications for content interaction. In *Proceedings of the 11th European Conference on Interactive TV and Video*, EuroITV '13, ACM (New York, NY, USA, 2013), 153–162.
28. Zoric, G., Engström, A., Barkhuus, L., Ruiz-Hidalgo, J., and Kochale, A. Gesture interaction with rich tv content in the social setting. In *Exploring and Enhancing the User Experience for Television, Workshop of ACM SIGCHI Conference on Human Factors in Computing Systems, CHI'13* (Paris, France, 04/2013 2013).