

Category: Full Paper

Aerial Virtual Reality: Remote Tourism with Drones

A. Fabola*, A. Miller and I. Duncan¹

Affiliations:

¹University of St Andrews, UK.

*Correspondence to: aef6@st-andrews.ac.uk

Abstract: This paper proposes a portable system for group-based exploration of remote landscapes in real-time. The system incorporates a drone for video capturing, a Raspberry Pi for wireless communication, an Android server for streaming and control, and one or more Android clients for rendering the footage in mobile virtual reality headsets. The system has been evaluated from a technical perspective to investigate and optimise resource utilisation, as well as from a user perspective with participants to investigate usability. The findings demonstrate the feasibility of group-based, virtual tours of remote landscapes in real-time using affordable components.

Introduction – Tourism in the New Age

Tourism is a booming industry which contributes to the economic progress of towns and villages, and promotes the heritage and the history of communities. Tourism has taken on different forms over the centuries and in recent decades, the need to explore remote areas have led to the adoption of several forms of remote heritage exploration, usually through the use of multimedia such as audio video and images exploration.

Most forms of remote video tourism are either provided pre-recorded (i.e. where the video is captured offline, stored and then made available on a platform usually through social media) or live (where the video is delivered to users in real-time as events unfold). The emergence of digital technologies, multimedia and social media platforms have given this mode of exploration new light, as users can now immerse themselves in far-away lands through the use of virtual reality and 360° media. Facebook and YouTube platforms now support streaming live 360° images and videos – also known as photospheres and videospheres respectively – which place the user at the centre of the action.

The role of video (and other media) in mediating tourist experiences has been investigated and the findings suggest that such media facilitate imagination (before visiting a place) and reminiscence (after visiting), and also provide access to remote “landscapes and socioscapes”[1]. Furthermore, access to remote landscapes is useful when it is impractical to visit, as some places are hard to reach owing to logistic or financial reasons. Distance, disability and illness may also make some places inaccessible to the public. The need for remote exploration of inaccessible sites provides the motivation for this work and this necessitates a novel kind of system; hence we propose a drone-based, real-time streaming system which provides aerial perspectives on areas that cannot be accessed on land. The system design also supports operation in areas void of power or Internet access, hence it is suitable for deployment in remote locations. Thus, the contributions of this work are the design and implementation of the aforementioned system with emphasis on portability and affordability, as well as an investigation of the value and feasibility of the system for collaborative, real-time heritage exploration.

The remainder of the paper is organized as follows. Section [Related Work] discusses literature in the areas of multimedia, virtual tourism with drones, and live video streaming over wireless

networks. Section [Design] describes the system design and methodology of the work, and Section [Implementation] discusses the system implementation in detail. Section [Evaluation] discusses a user evaluation of the system, Section [Results] provides a discussion of the results and findings from the study and Section [Conclusion] concludes the work.

Related Work – Virtual Tourism, Immersive Media and Drones

The role of video in mediating tourist experiences has been investigated by [1], whose findings suggest that such media facilitate imagination and reminiscence; activities which are usually undertaken before and after touring a remote location. There have also been a proliferation of applications in the form of guides, recommendation systems and content dissemination systems for tourism and cultural heritage. These systems include web-based systems ([2,3]), mobile trail apps ([4]), mobile, immersive systems ([5,6]) and fixed, immersive installations ([7,8]). These work demonstrate the use of 3D and multimedia technologies for remote tourism and cultural heritage dissemination; the incorporation of remote video (as seen in [1]) and virtual reality headsets (as seen in [5] and [6]) are particularly relevant to this work.

User Quality of Experience (QoE) for video streaming and stereoscopic image applications has been investigated and documented in literature. Quality of Experience (QoE) as used in this context, refers to how well a video streaming application performs, as perceived by users [9]. Several techniques have been proposed for QoE estimation and these can be split broadly into subjective [10], objective [11], and hybrid measures [12]. The determinants of QoE while viewing stereoscopic images was investigated in [13] who found that factors such as the baseline of the camera, the contents of the scene, the size of the screen, and the relationship between these factors all significantly affect QoE. The ability of users to distinguish between different levels of stereoscopic video quality has been investigated in [10]; the findings suggest that users generally have low sensitivity to changes in stereoscopic video quality, and depending on scene contents, changes in camera distance may not impact QoE. In addition, the differences between users' perception of stereoscopic and non-stereoscopic images as viewed on a screen, is investigated in [14] and [15] and the findings suggest that stereoscopic images tend to promote more feelings of presence (the sense of being in a virtual environment) than non-stereoscopic images.

Virtual reality scenes that incorporate moving components can be depicted using OmniDirectional Videos (ODVs, also known as 360°, panoramic, spherical or equirectangular videos) as they have the ability to place viewers at the centre of a scene. These videos can be viewed on a screen, in a Computer Automated Virtual Environment (CAVE), virtual reality dome or virtual reality headset. In addition to temporal navigation (rewinding or fast-forwarding) which is typically supported by UniDirectional (i.e. flat) Videos, ODVs also support spatial navigation (looking around the scene), and this is necessary due to their omni-directional nature. Screen-based devices (monitors, phones and tablets) naturally lend themselves to temporal navigation while viewing videos (by clicking or touching the screen), but they are less intuitive for spatial navigation – computer monitors require users to click and drag in order to change the view, and phones/tablets require the viewer to hold up the device and move around. Other (more immersive) means of viewing ODVs such as CAVEs, dome environments and headsets naturally support spatial navigation because the viewer (or part of their body) is physically at the centre of the environment in which the ODV is being viewed, hence by moving/turning/looking around, spatial navigation takes place. They are less intuitive for temporal navigation however, due to the absence of peripherals for point-and-click interaction. Virtual reality headsets naturally support spatial navigation of OmniDirectional Videos (ODVs) and possess the ability to immerse viewers in virtual environments. These features are leveraged by [16], which discusses a system that supports temporal navigation by using gestures that are detected by a headset-mounted motion tracker. ODVs are becoming increasingly more attractive and applicable for various purposes, as [17] discusses the design, implementation and workflow of a system which captures stereo ODVs (i.e. videos in equirectangular projection with left and right pairs of frames for

stereoscopic viewing) for consumption and streaming in virtual reality headsets. In a similar vein, [18] presents a system that enables geographically-dispersed individuals to visually collaborate by immersing users in a headset-based teleconference stream which provides the illusion of presence in remote locations. In addition, social media platforms such as Facebook [19] and Youtube [20] now offer 360° features which support the delivery of immersive ODVs to multiple users over the Internet. Users can interact with these videos on workstations by dragging with point-and-click devices, and also on mobile phones and tablets by moving the device from side-to-side while holding it up in the air. Our proposed system differs from that provided by these social media platforms in the sense that it can provide aerial footage of areas that are inaccessible on land. Also, the proposed system incorporates virtual reality headsets, and although it relays uni-directional video from the drone camera to the headsets, it provides the illusion of omni-directional video and its characteristics (such as an encompassing view and the illusion of presence) by translating users' head movements to the camera gimbal movements thus supporting spatial navigation. Furthermore, given the real-time nature of the video stream, temporal navigation is irrelevant and is hence not supported.

On another note, Unmanned Aerial Vehicles (UAVs, popularly referred to as drones) are gaining interest amongst scholars, industry practitioners and the general public as consumer versions have become available and affordable. This has resulted in use cases such as the use of drones to independently detect and track moving targets in the fields of robotics [21], computer vision [22] and autonomous tracking [23], as well as for thermal imagery and geo-referencing in the environmental sciences [24]. Of pertinence to this work, is the application of drones for virtual tourism, and systems which can stream drone footage to tourists over a network have been proposed. This streaming could be done over the Internet (as seen in [25] and [26]) or a Wi-Fi grid (as seen in [26]), and the clients usually receive the footage in Head Mounted Displays (HMDs) such as the Oculus Rift (used in [25]) or the HTC Vive (used in [26]). Our work is similar to [25] and [26] in the sense that users are imbued with the ability to control the drone's camera using head movements, thus providing the illusion of presence in the remote location. However, our work deviates from the aforementioned approaches in the sense that it focuses on the use of smartphones with affordable headsets (such as the Google Cardboard [27]) so that users can receive the footage without being tethered to a powerful workstation thus ensuring mobility, affordability and minimal use of resources. We also incorporate a social element by enabling the stream from a drone to be received by multiple subscribers. This way, more than one person can enjoy the footage at a time and a control protocol which dictates how (and when) each individual controls the drone has been implemented.

Methodology

An iterative approach was adopted for this work, whereby the project span entails several design-implement-evaluate cycles. The process began with an investigation of the technical capabilities – flight range (distance), multimedia streaming and networking – of drones. This resulted in the formulation of a conceptual model and system architecture (Figure 2). The system components were then implemented, and the components were then combined into a functional system. The system was then probed in order to stretch its limits and optimise it for use in real-world scenarios.

System Design

The system features a drone (Figure 1) which is equipped with a High Definition (HD) camera mounted on a gimbal. The drone is controlled with a joystick-like controller which communicates with the drone via a proprietary wireless infrastructure. The controller also provides an Application Programmable Interface (API) and a Software Development Kit (SDK) which enable the development of apps for Android and iOS smartphones which can be used to send instructions to the drone. The system also features an 802.11n wireless network to facilitate communication between a

server (which is connected to the controller via USB and transmits the raw video from the drone) and one or more clients (which receive and decode the footage and display it in virtual reality mode). The server and clients are both implemented as apps running on Android smartphones. A Raspberry Pi [28] – an inexpensive, portable computer that can be powered with a mobile power bank – is used as a wireless access point to ensure that the system can function in remote areas, independent of electricity and Internet access

After the design decisions were made, a compartmentalised approach was followed in the implementation of the system. The end product was then evaluated from a technical and user perspective. The aim of the technical evaluation was to investigate the capabilities of the system so as to ensure optimal use of resources and improve performance where possible. The user evaluation was carried out to assess the usability of the system and its potential value for group-based, real-time virtual tours of landscapes.

System Implementation

The equipment used and the tasks carried out during the incremental implementation of the system are described in the following sections.

Drone

DJI Phantom 3 Advanced drone [29] was used for development, and the system was also tested with the DJI Inspire 1 drone [30]. The Phantom 3 Advanced has a maximum flight time of 23 minutes and a maximum speed of 16 metres per second, and the remote controller has a maximum transmission distance of 5 kilometres or 3.1 miles unobstructed [29]. The camera can record video at up to 2704x1520 pixels at up to 30 frames per second, but live streaming is capped at 1280x720 pixels at 30 frames per second with a latency of 220 milliseconds. The gimbal can pitch from -90° to $+30^{\circ}$.

Wireless Infrastructure

A Raspberry Pi 3 Model B (Figure 3) running hostapd (Host access point daemon) [31] was used to create an access point which the server and clients connect to. The Pi can be powered using a wall socket charger if available, but in order to support use cases in remote areas, it can also be powered with a mobile power bank. The Pi has a 1.2GHz 64-bit quad-core ARMv8 CPU with an 802.11n Wireless LAN interface [28]. It also has 1GB RAM, 4 USB ports 1 HDMI port and a Micro SD card slot. A 32GB Micro SD card was used as secondary memory for the Pi.

Server

The server was developed as an Android mobile application, compatible with smartphones running Android 4.4 or higher. The server app implements a callback method which receives chunks of the raw video from the drone source and sends the chunks as datagram packets to the subscribing clients over the wireless network. The server also runs a control protocol by which it signals gimbal control (or lack thereof) to each client. This ensures that only one client has control of the gimbal at each point in time.

Client

The client was developed as an Android application, compatible with smartphones running Android 4.4 or higher. The client app runs on a smartphone in a VR headset (see Figure 4) and communicates with the server while connected to the wireless network hosted by the Raspberry Pi. It receives chunks of the raw video from the server and parses the raw video in real-time using ffmpeg, an open-source, multimedia framework [32]. Each parsed frame is rendered on the screen in virtual reality mode. Each client also listens for control packets (sent by the server) to determine whether it has control of the drone gimbal. When a client receives a packet authorising it to control the gimbal, it

continuously sends packets containing sensor readings that describe its orientation in space. This enables the client to control the drone's view, such that the movement of the user's head (and consequently the device) causes the drone gimbal to move, which in turn causes the camera view to change, thus providing the illusion of an immersive first-person experience.

Challenges solved

Decoding raw H.264 stream: As stated in Section [Server], the server application implements a callback (provided by the SDK) which fires when a new chunk of raw video is available from the camera. The video chunks represent a stream of raw H.264 video which must be parsed and decoded in order to be rendered by the client device. Decoding and rendering a raw H.264 video stream on Android is done using the MediaCodec class (Figure 5). The class provides a set of input buffers which must be fed with H.264 Access Units and output buffers which render the decoded frames onto an Android surface [33,34]. Decoding a raw stream therefore necessitates an understanding of the structure of raw H.264 video. A raw H.264 stream is made up of one or more Network Abstraction Layer Units (NALUs or NAL Units) which collectively contain both frame payload and key information needed to decode each frame [35]. The raw stream has delimiters that signal the start and end of the NAL Units, and these NAL Units must be parsed into Access Units (AUs), which can then be fed into an Android MediaCodec decoder that has been initiated and configured. This process was used to implement a parser class that configures a MediaCodec object, continuously parses the chunks of the raw H.264 video into video frames and seamlessly renders them to a virtual reality surface in real-time. The decoding process was designed to run on an independent thread, separate from the thread responsible for receiving the packets so as to ensure that the decoding process is not hampered by the receiving function (which periodically blocks due to the nature of network input/output operations). The decoding process was also optimised to minimise overhead so as to ensure adequate processing power for the other tasks on the client devices.

Portability, mobility and affordability: The system was designed with the requirements of portability – to function with minimal configurations in different environments, mobility – to obviate the use of fixed infrastructure, and affordability – to function with commodity hardware. The system portability is fulfilled by the design of the communication protocol (discussed in Section [Communication and control]), which enables it to function on any wireless network that supports multicast traffic. This implies that the network component of the system can be fulfilled with an off-the-shelf access point, Raspberry Pi, or computer that is capable of hosting a wireless network. Mobility – and by extension portability – is fulfilled by the use of mobile apps for the client and server, which can be downloaded from the Google Play Store, and do not require any pre-configuration owing to the discovery protocol implemented, as discussed in Section [Communication and control]. In addition, the portability and mobility of the system combine to ensure that it can be actualised with minimal costs. While some costs (for example the cost of purchasing or hiring a drone) may be inevitable, a smartphone can be procured or re-purposed to act as a server, many tourists travel with their smartphones and can download the client app for free, and most heritage organisations have wireless infrastructure in place or can install one for the cost of a Raspberry Pi (£30), all of which keep costs to a minimum.

Communication and control: Communication between the server and client takes place via datagram packets over a wireless network. This uses the User Datagram Protocol (UDP) at the transport layer which is characterised by fast, connectionless, albeit unreliable communication [36]. This makes UDP appropriate for video streaming and other applications that are either not sensitive to (minimal) packet loss or do not benefit from retransmission of lost packets. This informed the decision to use UDP for all communication between client and server, and this communication employs both unicast and multicast datagram packets.

When the server is initiated (i.e. the phone is connected to the wireless network, plugged into the drone controller and the server app is launched), it establishes a multicast socket on a port and begins listening for discovery requests from clients. When a client is initiated (i.e. the phone is connected to the wireless network and the client app is launched) it sends a discovery request to the multicast group and port on which the server is listening. Upon receiving this request, the server sends an acknowledgement to the client and adds the client to a list of connected clients. Clients can unsubscribe from the video stream by closing or minimising the app, which generates and sends a packet, notifying the server that it is leaving the stream. Upon receipt of this notification, the server removes the client from the list of connected clients. This ensures that the server can keep track of clients to know how many are subscribed to the stream at any point in time.

By keeping track of connected clients, the server can delegate control of the gimbal to each client in turn and switch control between them. The server implements this control protocol by periodically sending a packet which either authenticates clients to (or instructs them to not) send orientation coordinates. When a client has control of the gimbal, it continuously sends sensor coordinates – a reading about three axes (pitch, roll and yaw) that represents the device's orientation in space – to the server. Upon receipt of sensor readings from the controlling client, the server updates the drone gimbal to reflect the orientation of the client, ensuring that the drone's camera orientation aligns with that of the controlling client.

Only one client can have control of the gimbal at a time, and the transfer of control between clients depends on which of the two control modes – *tour* and *explore* – is selected. When the *tour* mode is selected, the server always assigns control to the same client. This mode is ideal for giving a guided tour of a landscape, where the tour guide's device is the designated client and thus controls the drone's gimbal. When the *explore* mode is selected, the server rotates control between each client, such that one client has control for a fixed duration, then control passes on to the next client in the list, and so on until control gets to last client and goes back to the first client. This mode is ideal for when a group of people wish to explore a landscape at will, such that each individual benefits from controlling the gimbal for a period of time.

Evaluation

A user evaluation was conducted in a real-world setting, where participants in a local community were invited to sign-up to an event involving drone flight over the local landscape. The aim of the user study was to investigate the usability and potential value of the system, as well as how it supports social experiences in the context of heritage exploration. The event took part in the local community park where the participants gathered. The drone was then flown over the local landscape in a slow, circular motion so as to observe the iconic parts of the community. While the drone was being flown, pairs of participants used the virtual reality headsets to receive the drone footage, where one person actively controlled the gimbal while the other passively received the drone footage. The control switched between the pair every minute, so that the active client became passive and the passive client became active. Each pair had 5 to 10 minutes to see the landscape using the headsets, after which they passed on the headsets to the next pair. Data was collected using observations during the event, as well as questionnaires filled by the participants and interviews conducted after the event. There were 8 participants (6 female) aged between 12 and 17, thus there were 4 pairs, and participants in each pair were familiar with each other. The exercise took approximately 1 hour to complete.

Results and Discussion

The findings of the study will be discussed in the following categories: *system usability*, *system control* and *social experience*.

System Usability – How well did the system work?

Overall the feedback from participants was positive. They felt like it was a valuable system for exploring landscapes. This was suggested by comments such as “*it was awesome because you get to see the town from different heights, angles and sides*”, “*it lets you see the world from a different perspective*”, “*...you would not see stuff like that when you are on the ground*”, “*...you can ... see what is around you and high up that you can't see*”, and “*you get to see the town from above*”. Participant's feedback also bordered on feelings of immersion in a virtual environment, as evident from comments such as “*it's like you are in a virtual world*”, “*it's like you are inside a game*” and “*you feel like you are flying*”. Overall, participants thought it was an interesting exercise that they would like to do more often.

System Control – How did users find the illusion of control?

As discussed in Section [Communication and control], a protocol was implemented to periodically switch control of the gimbal between clients, such that one client's movements determine the orientation of the drone's gimbal at any point in time, and that client changes after a set period of time. Given the system set-up which immersed participants in the virtual environment, it was important to investigate how this illusion of control was perceived by each participant, both while in control and otherwise.

Feelings about system control were mixed. Participants were impressed by the ability of the drone's camera to match their head orientation. One participant felt they had sufficient control of the system because “*... you just nod your head and it moves.*” Another participant thought that the illusion of control was unusual at first but after getting used to it, they felt in control. Similarly, two participants felt they were in control some of the time (when the view changed in response to their head movements) but not at other times when their head movements had no effect (presumably because the other individual in their pair had control). One participant felt like they were not in control because “*...it kept moving when I was still...*”, while referring to periods when other individual in their pair had control, and two participants wanted to fly the drone themselves so as to not only control orientation but also location.

Social Experience – How did social interaction unfold?

Participants unanimously felt that the system provided a good social experience. They referred to the ability to “*chat*” to each other while using the headsets and this was important because they could discuss shared views. The provision of shared, common views enabled participants to refer to things in their discussion, as they would have if they were viewing content on a big screen. This improves on the state-of-the-art in headset-based virtual reality systems because these systems often provide personal experiences where a user is totally engaged with a system and is thus cut-off from their real world environment including the people in that environment. However, by providing a shared, immersive view, the system incorporates the social benefits gained when two or more users view content on a screen and the immersive benefits gained when users view content in stereoscopic headsets.

Participants also cited the novelty factor, with one stating that “*...it is something different*”, and with another stating that “*...it is something that everybody likes*”. One participant also stated that “*...we were all talking to each other when we were flying it*”. It is notable that the participant used the phrase “*when we were flying it*”, which suggests that they felt they were in control i.e. they were flying the drone themselves.

Challenges and Limitations

Although a compartmentalised approach has been adopted in designing the system, the server implementation is built using the DJI mobile SDK and is hence limited to DJI drones, which

introduces a vendor lock-in with switching costs. Nonetheless, the design principles and implementation detail can be adapted to develop a similar system using drones from other manufacturers.

Limitations of the technology also imposed challenges on the system functionality. The range of the Raspberry Pi's wireless interface limits the distance that the clients can be away from the server and still receive live stream data. Although this is not an issue in the primary use case, i.e. while using the system outdoors, it poses a significant challenge in settings where the clients need to be at a considerable distance away from the server; for instance in use cases when the clients wish to receive footage in an indoor setting while the server is located outdoors (in a safe flying zone) with the drone pilot. In the same vein, most personal-grade drones have a maximum flying time of around 25 minutes or less as well as a maximum transmission distance beyond which the aircraft cannot communicate with the controller or base station. This imposes further limitations on the duration of a session and the size of the area which can be explored respectively, although the former can be remedied with the use of multiple batteries.

Future Work

A system architecture that expands the reach of the system further (by making the stream available over a Wide Area Networks such as the Internet) is being designed. This will offer the added benefit that users do not have to visit a local heritage centre or attend an event to explore a remote location, but will be able to do so from the comfort of their homes. In addition, more user evaluations will be conducted to investigate how the system can effectively support social and collaborative exploration of heritage sites.

Conclusion

The motivation and methodology for designing, implementing and deploying a system for real-time, remote exploration of heritage sites have been discussed in this work. A system evaluation was conducted which provided input into the re-design process, and a user study was carried out to investigate the potential value and usability of the system. The findings suggest that the system is a valuable and affordable tool for collaborative heritage exploration, as it enables users to observe a landscape from a different perspective and it provides an immersive, yet socially-inclusive experience. The system contributes to the state-of-the-art in remote tourism as it combines the engaging benefits of personal, immersive systems with the social benefits of large displays, hence it offers a new dimension for heritage exploration.

References

- [1] Tussyadiah, I. P., and Fesenmaier, D. R. (2009). Mediating Tourist Experiences. Access to Places via Shared Videos. *Annals of Tourism Research*, 36(1), 2440. <http://doi.org/10.1016/j.annals.2008.10.001>
- [2] Albanese, M., d'Acerno, A., Moscato, V., Persia, F., and Picariello, A. (2011). "A Multimedia Semantic Recommender System for Cultural Heritage Applications," *Semantic Computing (ICSC)*, Fifth IEEE International Conference on, Palo Alto, CA, 2011, pp. 403-410.
- [3] Guarnieri, A., Pirotti, F., and Vettore, A. (2010). Cultural heritage interactive 3D models on the web: An approach using open source and free software. *Journal of Cultural Heritage*, 11(3), 350-353.
- [4] Fabola, A., Davies, C., Kennedy, S., Miller, A., and Allison, C. (2014.). *Mobile Exploration of Medieval St Andrews*, 110.

- [5] Davies, C., Miller, A., and Fawcett, R. (2015). Mobile onsite exploration of parallel realities with Oculus Rift. *2015 Digital Heritage*, 1, 301304.
- [6] Fabola, A. Miller, A. and Fawcett, R. (2015). Exploring the past with Google Cardboard 2015 Digital Heritage International Congress.
- [7] Lo, W., Tsai, Y., Chen, C., and Hung, Y. (2004). Stereoscopic Kiosk for Virtual Museum. *Ics*, 10841090.
- [8] Tredinnick, T. and Richens, P. (2015) "A fulldome interactive visitor experience a novel approach to delivering interactive virtual heritage experiences to group audiences in fulldome projection spaces, evaluated through spatial awareness and emotional response,". *Digital Heritage, Granada, 2015*, pp. 413-414.
- [9] International Telecommunication Union (ITU). (2007, February). Definition of Quality of Experience (QoE) [Online]. Available: <https://www.itu.int/md/T05-FG.IPTV-IL-0050/en>. [Accessed: 12 Sept 2017].
- [10] Goldmann, L., De Simone, F., & Ebrahimi, T. (2010, February). A comprehensive database and subjective evaluation methodology for quality of experience in stereoscopic video. In *IS&T/SPIE Electronic Imaging* (pp. 75260S-75260S). International Society for Optics and Photonics
- [11] Xing, L., You, J., Ebrahimi, T., & Perkis, A. (2010, December). Estimating quality of experience on stereoscopic images. In *Intelligent Signal Processing and Communication Systems (ISPACS), 2010 International Symposium on* (pp. 1-4). IEEE.
- [12] Piamrat, K., Viho, C., Bonnin, J. M., & Ksentini, A. (2009, April). Quality of experience measurements for video streaming over wireless networks. In *Information Technology: New Generations, 2009. ITNG'09. Sixth International Conference on* (pp. 1184-1189). IEEE.
- [13] Xing, L., You, J., Ebrahimi, T., & Perkis, A. (2011, February). Factors impacting quality of experience in stereoscopic images. In *IST/SPIE Electronic Imaging* (pp. 786304-786304). International Society for Optics and Photonics.
- [14] Hkkinen, J., Kawai, T., Takatalo, J., Leisti, T., Radun, J., Hirsaho, A., & Nyman, G. (2008, January). Measuring stereoscopic image quality experience with interpretation based quality methodology. In *Electronic Imaging 2008* (pp. 68081B-68081B). International Society for Optics and Photonics.
- [15] IJsselsteijn, W. A., Bouwhuis, D. G., Freeman, J., & Ridder, H. (2002, May). P15: Presence as an Experiential Metric for 3D Display Evaluation. In *SID Symposium Digest of Technical Papers* (Vol. 33, No. 1, pp. 252-255). Blackwell Publishing Ltd.
- [16] Petry, B., & Huber, J. (2015, March). Towards effective interaction with omnidirectional videos using immersive virtual reality headsets. In *Proceedings of the 6th Augmented Human International Conference* (pp. 217-218). ACM. Chicago
- [17] Anderson, R., Gallup, D., Barron, J. T., Kontkanen, J., Snavely, N., Hernandez, C., ... & Seitz, S. M. (2016). Jump: virtual reality video. *ACM Transactions on Graphics (TOG)*, 35(6), 198.
- [18] Wendrich, R., Chambers, K. H., Al-Halabi, W., Seibel, E. J., Grevenstuck, O., Ullman, D., & Hoffman, H. G. (2016). Hybrid design tools in a social virtual reality using networked oculus rift: a feasibility study in remote real-time interaction. In *Proceedings of the International Design Engineering Technical Conference & Computers and Information in Engineering Conference* (Charlotte, NC). Chicago
- [19] Facebook. Facebook 360 [Online]. Available: <https://www.facebook.com/facebookmedia/get-started/filming-stitching-editing>. [Accessed: 12 Sept 2017]. Youtube Creator Blog. A new way to see and share your world with 360-degree video [Online]. Available: <https://youtube-creators.googleblog.com/2015/03/a-new-way-to-see-and-share-your-world.html>. [Accessed: 12 Sept 2017].
- [21] Boudjit, K. and Larbes, C., 2015, July. Detection and implementation autonomous target tracking with a Quadrotor AR. Drone. In *Informatics in Control, Automation and Robotics (ICINCO), 2015 12th International Conference on* (Vol. 2, pp. 223-230). IEEE. Vancouver

- [22] Choi, H., Geeves, M., Alsalam, B., and Gonzalez, F. (2016). Open source computer-vision based guidance system for UAVs on-board decision making. IEEE Aerospace Conference Proceedings, 2016June. <http://doi.org/10.1109/AERO.2016.7500600>
- [23] Vasconcelos, F., and Vasconcelos, N. (2016). Person-following UAVs. 2016 IEEE Winter Conference on Applications of Computer Vision, WACV 2016. <http://doi.org/10.1109/WACV.2016.7477660>
- [24] Harvey, M. C., Rowland, J. V., and Luketina, K. M. (2016). Drone with thermal infrared camera provides high resolution georeferenced imagery of the Waikite geothermal area, New Zealand. Journal of Volcanology and Geothermal Research, 325, 6169. <http://doi.org/10.1016/j.jvolgeores.2016.06.014>
- [25] Mirk, D., and Hlavacs, H. (2015). Virtual tourism with drones: Experiments and lag compensation. DroNet 2015 – Proceedings of the 2015 Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use, 4550. <http://doi.org/10.1145/2750675.2750681>
- [26] Wu, C., Wang, Z., and Yang, S. (n.d.). Drone Streaming with Wi-Fi Grid Aggregation for Virtual Tour, 01.
- [27] Google. Google Cardboard [Online]. Available: <https://www.google.com/get/cardboard/>. [Accessed: 12 Sept 2017].
- [28] Raspberry Pi Foundation. Raspberry Pi 3 Model B [Online]. Available: <https://www.raspberrypi.org/products/raspberry-pi-3-model-b/>. [Accessed: 12 Sept 2017].
- [29] DJI. Phantom 3 Advanced Specs [Online]. Available: <http://www.dji.com/phantom-3-adv/info>. [Accessed: 12 Sept 2017].
- [30] DJI. Inspire 1 Specs [Online]. Available: <http://www.dji.com/inspire-1/info>. [Accessed: 12 Sept 2017].
- [31] Hostapd [Online]. Available: <https://wiki.gentoo.org/wiki/Hostapd>. [Accessed: 12 Sept 2017].
- [32] FFmpeg. About FFmpeg [Online]. Available: <https://ffmpeg.org/about.html>. [Accessed: 12 Sept 2017].
- [33] Bigflake. Android MediaCodec stuff [Online]. Available: <http://bigflake.com/mediacodec/>. [Accessed: 12 Sept 2017].
- [34] Android Documentation. MediaCodec [Online]. Available: <https://developer.android.com/reference/android/media/MediaCodec.html>. [Accessed: 12 Sept 2017].
- [35] Wiegand, T. (2003). Overview of the H. 264/AVC video coding standard. and Systems for Video , 13(7), 560576. <http://doi.org/10.1109/TCSVT.2003.815165>
- [36] Postel. J. RFC 768 – User Datagram Protocol [Online]. Available: <https://www.ietf.org/rfc/rfc768.txt>. [Accessed: 12 Sept 2017].
- [37] DJI. Inspire Specs [Online]. Available: <http://www.dji.com/products/drones>. [Accessed: 12 Sept 2017]

Display Items



Figure 1. DJI Phantom 3 Advanced flown over the Strath of Kildonan

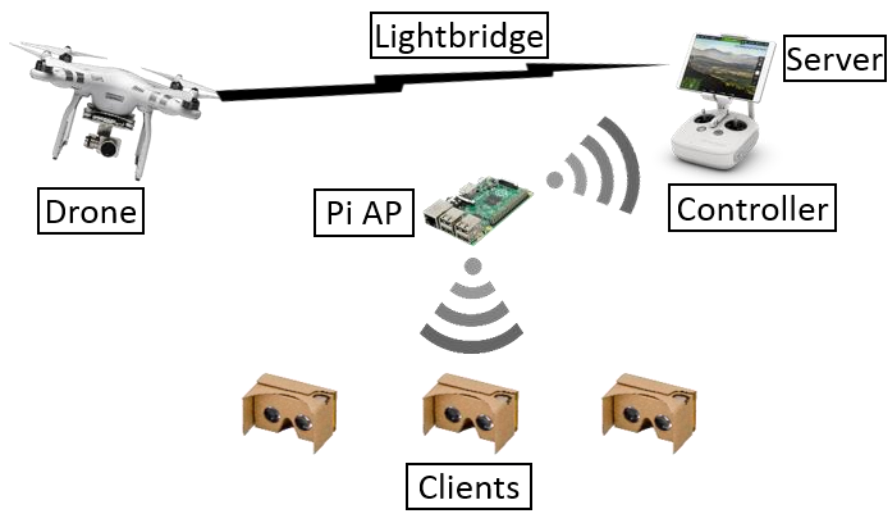


Figure 2. System architecture



Figure 3. Raspberry Pi in a transparent plastic case



Figure 4. Client phone in a Google Cardboard VR headset

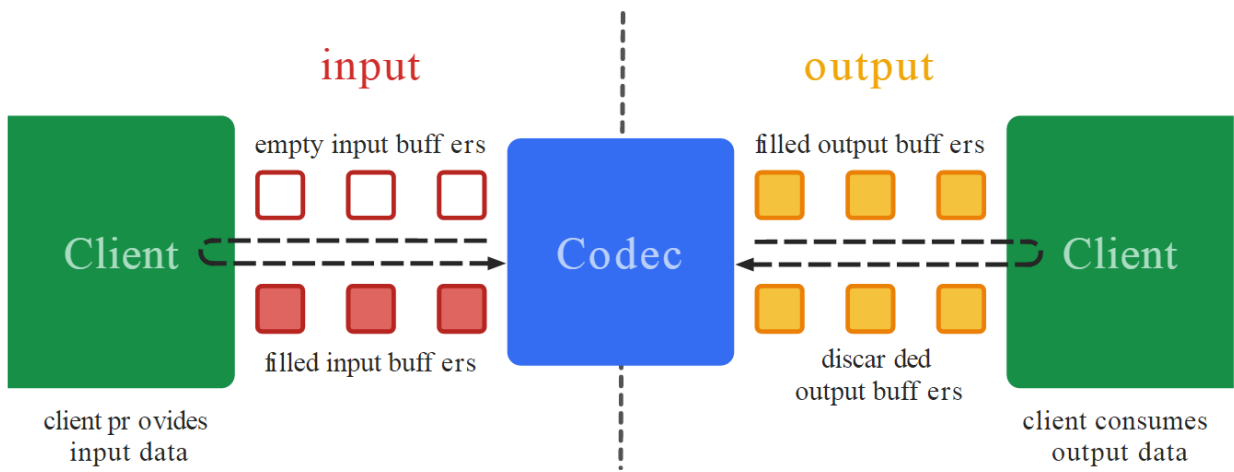


Figure 5. Android MediaCodec decoding process. (Source:[34])