

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.



Exploring travel film beyond the screen.



# [ V I E W F I N D E R S ]

Exploring travel film beyond the screen.





## Acknowledgements

My thanks go to Max Schleser and Gerda Cammaer for getting this project off the ground and keeping it in the air. Your efforts in securing funding, writing proposals, and keeping Viewfinders in the spotlight at every conference that would have us, are greatly appreciated.

Thank you to Tristan Bunn and Lucas Haley for your digital expertise, patience, and wisdom.

To The Enormous Crocodile Company - long may your gleaming smiles roam Oriental Bay.

Thanks also to Hayden and Lachlan for keeping the conversation going over many a flat dinner.

To my partner Hannah for reminding me that humans are supposed to eat and sleep.

And lastly to Havana Coffee. Here's to you guys.



SHANGHAI AIRLINES

上海航空 1112

hudsons  
coffee  
a little love in every drop

www.hudsonscoffee.com

## Abstract

Viewfinders is a design-led research project exploring emerging opportunities for i-docs (interactive documentary) and smartphone filmmaking. The i-doc serving as the practical component is part of a larger collaborative research project between Dr. Gerda Cammaer and Dr. Max Schleser, and focuses on defining new Creative Mobile Media Practices. Responding to recent developments in collaborative media, augmented reality, and computer vision technologies, a mobile web app is designed that proposes new active modes of viewing for the travel film genre.

The Viewfinders app asks the question; what happens when a travel documentary can itself travel? Turning conventional AR on its head, Viewfinders does not seek to augment the live environment but rather positions the real world as the augmentation, harnessing ambient phenomena to augment the i-doc.

To progress the narrative viewer-users are asked to explore their physical surroundings, with geolocation and computer vision algorithms interpreting evolving environmental factors to evoke corresponding sequences of travel footage. The non-linear narrative comprises a series of 60 second clips of peer-generated video that are re-ordered dynamically, with an aim of aligning experiential factors across both the viewer's world and that of the documentary.

Viewfinders has no play button. Taking its cues from the viewer's own journey, the world is its play button.



## Contents

<b>Mise en Scène</b>	<b>12</b>
<b>Framework</b>	<b>16</b>
1.0 Intent	17
1.1 Scope	18
1.2 Research Questions	19
1.3 Hypothesis	20
1.4 Glossary of Key Terms	21
1.5 Methodology	23
1.5.1 Methods	26
<b>Literature Review</b>	<b>28</b>
2.0 The Ascension of Mobile Media	30
2.0.1 Museum of London - <i>Street Museum</i>	31
2.1 Ubiquitous Media	32
2.2 Mobile Media Innovation	34
2.3 Augmented Reality	36
2.4 Mobile Augmented Reality and Locativity	37
2.5 Virtual Reality Cinema	39
2.5.1 Lynette Wallworth - <i>Collisions</i>	41
2.6 Mobile Filmmaking	44
2.7 Database Cinema	45
2.8 Summary of Literature Review Insights	46
<b>Creative Practice</b>	<b>48</b>
3.0 Viewfinders.gallery	50
3.0.1 Early Access Map Interface	54
3.0.2 Google Maps API	54
3.0.3 YouTube and Google Maps Mashup	55
3.1 Virtual Reality	58
3.2 Augmented Reality	59
3.2.1 Summary of Augmented Reality Insights	69





3.3 Computer Vision	70
3.3.1 Video Recognition	71
3.3.2 Clarifai	71
3.3.3 Microsoft Cognitive Services	73
3.3.4 Google Cloud Vision	74
3.3.5 Computer Vision API Performance Comparison	78
3.3.6 Summary of Computer Vision Insights	79
<b>User Interface Design</b>	<b>80</b>
4.0 Wireframing	83
4.1 Digital Mock-Ups	88
4.1.1 Significance of Playback Controls	91
4.2 Interactive Prototyping	92
4.3 Visual Language	100
4.3.1 Pictogram Set	102
4.4 Final User Interface Design	104
<b>Mobile Video Engagement Exercise</b>	<b>114</b>
<b>Conclusion</b>	<b>119</b>
References	122
Figures	130



# MISE EN SCÈNE

The rapid proliferation of smartphones has had far-reaching implications for many fields. To borrow words from prominent scholars Hjorth, Burgess, and Richardson: “As the epitome of ubiquity, mobile media in an age of smartphones requires that we not only attempt to articulate the field but also to more systematically understand its various dimensions - the technical, cultural, social, political, and economic” (2012, p. 1).

Viewfinders responds to this call for investigation, setting its sights squarely on the cultural repercussions of ubiquitous mobile media. Specifically, it examines the new creative avenues opening for amateur travel filmmakers.

Professor of Film Theory Devin Orgeron asserts that “few activities are as ‘collectible’ or as cinematographically inspiring as travel” (2006, p. 77), a contributing factor no doubt to the popularity of the travel film among amateur enthusiasts. This applies even in the earliest days of the medium, when filmmaking was what many would consider these days to be a prohibitively expensive and technically involved process.

A noticeable uptick in amateur travel film production coincided with advancing technology in the 1960s, when it became a common tool for tourists and travellers. The medium was further bolstered by the introduction of the Super8 camera which featured a revolutionary cartridge loading system. In her 2015 publication *Sharing Mobile, Portable, and Personal Travel Stories*, Gerda Cammaer observes that this new user-friendliness really opened travel filmmaking up to the public, in much the same way that the contemporary smartphone allows anyone to try their hand at videography.

Richard Chalfen, an authority on home-movie aesthetics, describes a coordinated effort by the film establishment at this time to stamp out improper practices emerging from a new class of filmmaker (1986). Guidebooks were produced belittling the naïve habits of the untrained eye, bearing warnings such as to avoid any ‘panoram’ on stationary landscapes. To shoot from a moving vehicle at this time was unthinkable. However the format of amateur travel film proved to be as free-spirited as its content, striding fearlessly into new territory. Orgeron believes that the consistency of this dissent indicates a conscious opposition to domestication attempts by the industry. Travel film forged for itself a definitive aesthetic, combining the technical style of home movies with



**Fig 1** Konica Super 8-6TL camera advertisement (c. 1966).



a taste for dynamism, spontaneity, and adventure, while defiantly flaunting practices that were actively discouraged by the film establishment. From our vantage point in 2017 we can now see that the student has become the master so to speak, with the mainstream media now increasingly taking cues from amateur film. The convention of personal intimacy, once considered reproachable, is now a romance mainstay; hand-held camera has found a home in horror; and a subjective first-person perspective features on television shows of all kinds.



▲ Fig 2 The contemporary smartphone user as travel filmmaker.

The arrival of the smartphone has had a decisive impact on the conventions of travel film. The same sense of excitement and exploration prevails, but both format and content have been altered. Early camera phones featured a low-res camera, creating a world characterised by harsh pixelation, distorted colours, and the occasional frame rate variation (Schleser et al., 2009). Technological advancements have moved us away from this early aesthetic. Conventions that set contemporary smartphone travel films apart from their analogue predecessors include the addition of non-diegetic elements like

music and titles, often produced on the phone itself; the birth of the portrait or 'vertical' video; and an increasing tendency towards self-portraiture, particularly evident after the introduction of a forward-facing camera on models like the iPhone 4 (Cammaer, 2015).

With the latest generation of iPhones comes a recording capability of 1080p at 60fps, slow motion at 240fps and apps like Filmic Pro enabling 3K video at 50 megabits per second. Ongoing technological advancements have extended the appeal of smartphone filmmaking, and there are now smartphone film festivals materialising all around the world. I recently attended a screening in Melbourne featuring productions from smartphone filmmakers in countries as distant and varied as South Korea, The United States, and Macedonia.

For all the changes brought about by the arrival of mobile media devices, one thing remains the same. Travel film, even of the smartphone variety, still directs itself toward what may now constitute a dated conception of 'audience'. The last 100 years of filmmaking have seen screens appear in our own homes, on our desks, and now in our hands. However filmmakers maintain an expectation that their audience will be seated, preferably somewhere dark, and even more so somewhere quiet, in order to get the most out of a production. While possible, watching a film on the go brings negligible benefits to the viewer. I believe that the primary reason conditions while on-the-go are unfavourable is because no film has yet been produced specifically with these conditions in mind. Viewfinders investigates the creative challenge of converting distractions like movement and ambient sound into the main attraction.



# FRAMEWORK

## 1.0 Intent

MINA, originally the *Mobile Innovation Network of Aotearoa*, is an international network of mobile media academics who work in association with various practitioners and smartphone film festivals to promote “*cultural and research activities to expand the emerging possibilities of mobile media*” (Berry and Schleser, 2014, p. 4). The objective of this research project is to extend the research of MINA associates Drs. Schleser and Cammaer, by developing an imaginative mobile interface for digital storytelling.

Prominent scholars Goggin and Hjorth make a point of highlighting the unique opportunities for expression and agency that are innate to mobile media (2014), a sentiment echoed by Berry and Schleser who encourage research that explores new technologies, practices, and formats (2014). The Viewfinders project focuses specifically on new mobile technologies, and the implications they have for mobile filmmaking. Media academics and experimental filmmakers form the target demographic for the practical research output, along with amateur travel filmmakers in general.

“We learn most effectively by doing – by active experience, and reflection on that experience. We learn through practice, through research, and through reflection on both.”

- Gray & Malins, 2004, pp. 1-2.

The appearance of new technologies is a driving factor behind a growing demand for practitioners with diverse skill sets who can achieve cross-disciplinary innovation (Design Council, 2010). Throughout this project I refine my technical skills in working with various mobile technologies, culminating in a practical output that offers one example of how nascent technologies could impact future creative practices.

Success is defined not in producing a platform that achieves mainstream adoption or commercial viability, but rather in focusing the lens so to speak on ways of realising the creative potential that mobile technologies hold for amateur and travel film.



### 1.1 Scope

This research project limits its scope to the study of how contemporary mobile devices enable the creative expression of personal travel narratives. The ubiquity and intensely personal nature of these devices have seen them rapidly adopted as the tool of choice for a new generation of amateur travel filmmakers (Cammaer, 2015a), and the consequent shifts in the genre warrant academic scrutiny (Burgess et al., 2012). By focusing on mobile travel film in particular, the research inhabits a space between the disciplines of mobile media and film studies, offering a bridge between the two and continuing the interdisciplinary discussion central to the earlier works of Drs. Cammaer and Schleser.

As a creative practitioner this project allows me to familiarise myself with contemporary mobile media theory; to gain experience in designing for user experiences; to sharpen my understanding of emerging mobile technologies; to advance my technical proficiency working with front-end web technologies; and to propose a new mode of active viewing that challenges the status quo in film.

Though it encounters elements of software development, ethnography, and wider media studies, the heart of this project is mobile filmmaking. Existing in a digital space Viewfinders inevitably demands some degree of front-end development – it should be understood that this development is a means to an end. My objective is not to obtain a full suite of multilingual development skills within a 12-month research project, but rather to use front-end development as a tool to test and refine the Viewfinders proposition. The ultimate product therefore is not a fully-realised piece of software but rather a resolved concept, built on the insights of actual user testing and backed by practical evidence.

**HOW DOES  
EMERGING  
MEDIA SHAPE  
CONTEMPORARY  
TRAVEL  
FILMMAKING?**

**CAN MOBILE  
TECHNOLOGIES  
CREATE NOVEL  
VIEWING  
EXPERIENCES?**



### 1.3 Hypothesis

Film media is not immune to the effects of mobile technology's recent extraordinary acceleration. To quote the godfather of modern media theory himself, the influence of any medium is *"the change of scale or pace or pattern that it introduces into human affairs"* (McLuhan, 1964, p. 8). We can expect then that a new generation of filmmaking instruments, namely smartphones, will introduce a great number of new practices to contemporary travel film. The i-doc and database cinema are both a testament to our hyperlinked world. Mobile technologies take this connectivity and further expand its reach: where the Internet and affordable pocket-cameras democratised film production, the smartphone will diversify modes of consumption.

Pocket-cameras and social networks have already made travel film more intimate, collaborative, and interactive. The remote connectivity of today's smart devices however enables film viewing to become active and experiential, as well as bringing web technologies like Augmented Reality and computer vision right into the modes of production.

**"In every subject area - including design - there is a constant need for new knowledge and there are problems that wait to be resolved..."**

– Muratovski, 2016, pp. 45–46.

## 1.4 GLOSSARY OF KEY TERMS

### COLLABORATIVE MEDIA

Productions that are co-created by many authors. Collaborative film often sources content from its audience, inspiring a sense of ownership and shared identification.

### AUGMENTED REALITY

This term is commonly used to denote information that is visually superimposed on top of the physical world. This study recognises an extended meaning, encompassing non-visual augmentations as well. Therefore AR is any technology – whether visual, aural, tactile, or otherwise – that enhance attributes already present in the environment.

### MOBILE AUGMENTED REALITY

More than simply Augmented Reality facilitated by a mobile device, mAR as its own distinct medium combines AR with the other technologies found on contemporary smart-devices, such as GPS, accelerometer data, and Internet connectivity (David & Schleser, 2013).

### I-DOC

Interactive documentary, also known as ‘web documentary’ or ‘i-doc’ is a recent media form intersecting web multimedia tools and traditional documentary. Unlike the linear progression of a traditional documentary the content of an i-doc is presented in segments that can be navigated in an order of the viewer’s choosing. An i-doc typically has no set running time, and sequences can be replayed, skipped, and re-ordered.

### UBIQUITY

The quality of being always present, and with regard to mobile media, of being always connected to a larger mediascape.

### LOCATIVITY

Locativity allows media to adopt certain location-specific qualities that change the relationship between viewer and subject; *“The capacity of the mobile phone to be out in space, unbound from a fixed location, unlike the television and cinematic experience, and make a relevant association between media content and place for the user”* (Rushton, 2012).

### NON-LINEAR NARRATIVE

In film this term typically refers to a narrative that defies a conventional chronological procession or disrupts the usual pattern of causality. In this study I use it in reference to the logic of the i-doc; its capacity to have no predetermined course; and to have no explicit beginning or end point.

### COMPUTER VISION

A term often used interchangeably with '*image recognition*', computer vision is a new science that trains computer algorithms to automatically extract, analyse, and understand digital imagery (BMVA, 2017). Common computer vision services include text recognition, motion tracking, and general feature recognition.

### EXPERIENTIAL VIEWING

A holistic mode of observation where the needs of all five senses are met to immerse the viewer within the content. Media that endeavour to produce an experiential viewing mode must replicate or otherwise facilitate factors not recorded by the lens.

### MOBILE-MENTARY

A documentary filmed, sometimes also produced, on a mobile device. The qualities of a mobile-mentary are informed by those of its production, often reflecting a sense of intimacy and spontaneity. '*Mobile-mentary 2.0*', a term coined by Max Schleser, alludes to mobile-mentary that embraces global collaboration as facilitated by a contemporary mass-distribution of mobile devices worldwide (Schleser, 2012).

### API

An Application Programming Interface is a set of protocols that allow two or more softwares developed by separate authors to interact. APIs launched by leading web service providers typically grant limited access to third-party developers, so that they can leverage proprietary resources to meet their own needs. One such example is the *Google Maps API*.

## 1.5 Methodology

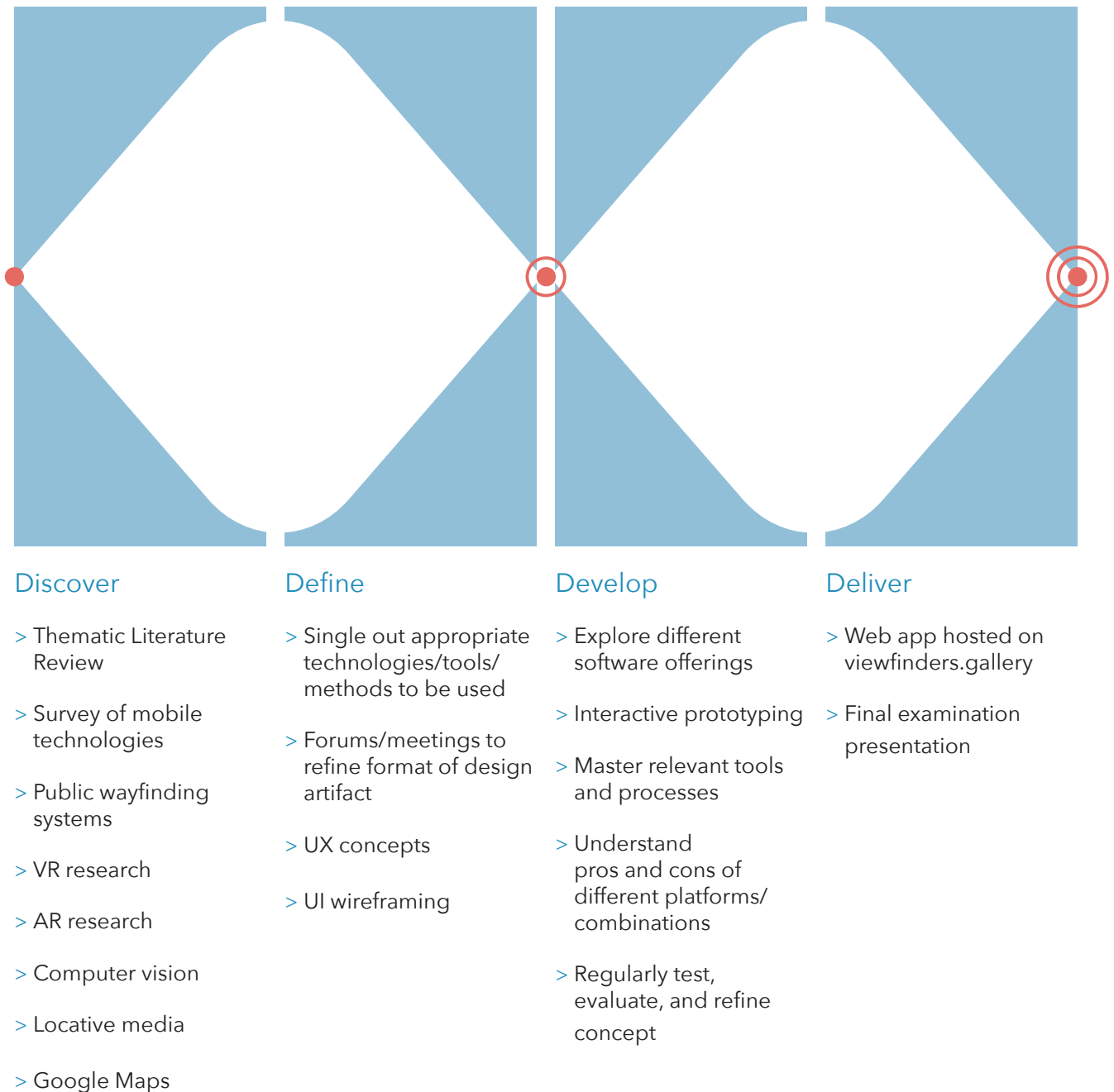
The research methods under the Viewfinders bonnet are fuelled by creative practice. There is little need to quantify any statistical benefits of the final practical output, given the project's aim is to explore, understand, and reconceptualise new mobile technologies as a tool for expression within travel film, which is something of a qualitative endeavour. Furthermore the design artifact that is produced does not presume to identify any one ultimate or 'proper' application for the practices it proposes, but rather seeks to offer one example of how these technologies can be reconceptualised within a filmmaking framework.

**"Today, designers design services, processes and organizations; craft skills no longer suffice... We need Design Research, but as a special kind of research, with methods appropriate to the applied, constructive nature of design."**

– Norman, as quoted in Muratovski, 2016, p. xvii.

Design researcher Linda Candy defines practice-based research as an *"original investigation undertaken in order to gain new knowledge partly by means of practice and the outcomes of that practice"* (2006a, p. 1). The appropriate methodology as described by Candy comprises two parts: a design artifact which should demonstrate originality and contribution to knowledge; and an accompanying text that outlines the significance and context of this claim in words (2006a). As a creative practitioner my method focuses primarily on the development of a web app, here serving the purpose of design artifact; the latter requirement of contextualisation is fulfilled by this body of written work, the exegesis.

Some experts argue that the design process itself sufficiently meets the requirements to be considered applied research (see Swann, 2002; Crouch and Pearce, 2012). My methodology disregards this position. Although practice-based, my research is more akin to the definition of applied research as explained by Muratovski, which is to say systematic and deliberate, something not necessarily required in the methods of a 'pure practitioner' (Muratovski, 2016). *"Applied research aims to generate culturally novel apprehensions that are not just novel to the designer or to the client, but are novel to the field of design as well - the emphasis here is on the process"* (Muratovski, 2016, p. 191).

*The Design Council Double Diamond*

▲ Fig 3 The Design Council Double Diamond.



My method is based on the British Design Council's *'Double Diamond'* framework.

**"In the first diamond many ideas are generated and considered before they are narrowed down. In the second diamond key ideas are selected, developed and tested before being refined again."**

- Design Council, 2015, p. 14.

At the beginning is the **Discover** phase, using the divergent half of the first diamond to conduct an unencumbered survey both of current mobile media theory and of recent technological developments pertaining to mobile media. The survey is opened to the greatest possible number of discoveries by avoiding any blinkering; there are no preconceptions of what the design artifact will be, and an intentional disregard for practicality or budget - true 'blue sky' thinking.

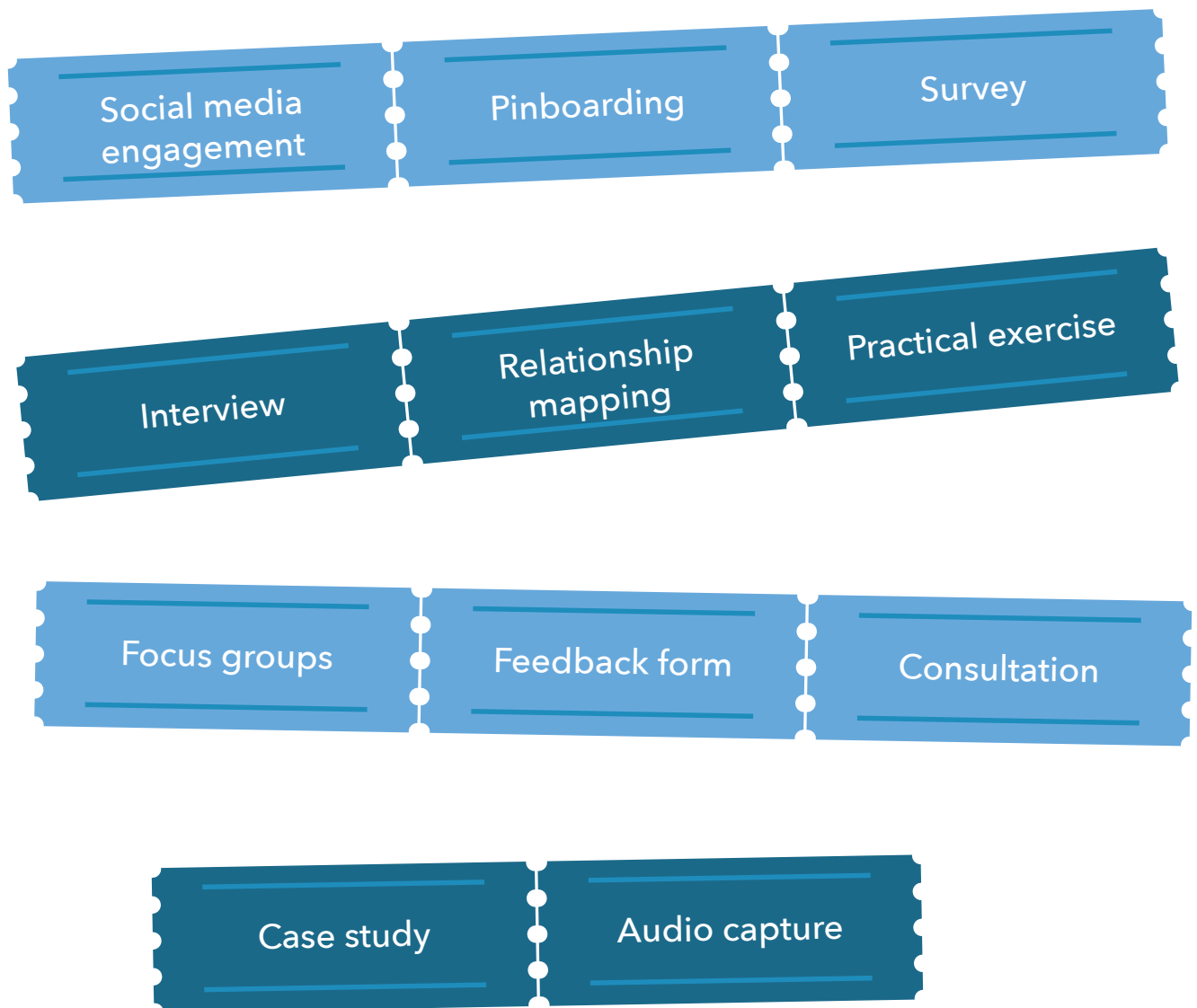
The second phase then converges to **Define** what form the design artifact will take, a process that is informed by the many potential pathways identified in the first phase. Those insights with the most substantial benefits for contemporary mobile filmmaking practice are amalgamated into a single concept.

The second diamond leans more towards application, concerning itself with developing and delivering a practical output that deploys the lessons learnt in the first diamond. The divergence during the **Develop** phase is to explore the many technical tools and methods available, to compare and contrast the pros and cons of each, and to generate myriad prototypes. Design is rarely if ever a linear process and the development stage observes principles of Agile development: incremental delivery; regular testing; responding to change over following a plan; and attention to detail (Fowler and Highsmith, 2001).

Last comes the **Deliver** phase, which here entails publishing the practical research output for the MINA community and delivering the final examination presentation to meet the requirements of the Massey University Master of Design Programme.

### 1.5.1 Methods









# LITERATURE REVIEW



I have conducted a thematic literature review (Muratovski, 2016) to give me a high-level appreciation of contemporary academic theory in the field of mobile media. I have also researched in greater detail the disciplines of mobile and interactive filmmaking. Grounding my research with a literature review allows me to avoid repeating existing research; develop informed research questions; reference the research methods of my peers; ensure the relevance of my study; and adopt a critical stance towards existing ideas.

Particular attention is given to the work of Drs. Schleser and Cammaer, to identify where I might be able to expand on their existing studies of mobile and participatory filmmaking.

## 2.0 The Ascension of Mobile Media

In his 2002 book *Smart Mobs*, media theorist Howard Rheingold notes that mobile phone use took an unexpected turn in the lead up to the new millennium, becoming more a visual than an aural phenomenon seemingly overnight. *“That was when I began to notice people on the streets of Tokyo staring at their mobile phones instead of talking to them”* (Rheingold, 2002, p. 11).

At this early stage phones were not yet capable of sending images, much less video or browsing the Internet, but the inception of text messaging services led to the eruption of subcultures around the world. Rheingold cites shifts in political activism, social interaction, and corporate management as some of the many consequences of taking mobile phone functionality beyond mere voice transmission.



Fig 4 Museum of London's 'Street Museum' app restores historical photos to their original context.

### 2.0.1 Museum of London - Street Museum

As mobile phones gained Internet access, they altered *"the way we spatially interact with each other and with cultural objects"* (Farman, 2012, p. 19). One such example is The Museum of London's *Street Museum* (2010), an app that moves the museum collection onto the street by assigning historic photos and paintings with coordinate metadata corresponding to their narrative. The app then uses AR technology to superimpose the artifact onto the physical location when the user's device is detected at that coordinate. The viewer's relationship with that object cannot help but be affected, achieving a much more immediate connection to the material.

One can hardly fail to appreciate the gravity of a WWI photograph when standing in the middle of an entire city block razed by the Blitz. By comparison the experience of seeing the same photo removed from context on the wall of some distant museum is hardly captivating. Bringing together content and location offers the user a much deeper understanding of the material.

*"[Media] bring together spaces and times that would not normally meet... they transport sounds, thoughts, images and other forms of information between realms that, although maybe entirely dissimilar on a physical level, thereby gain access to a certain shared conceptual or cognitive stratum"*

– de Vries, 2012, p. 177.



## 2.1 Ubiquitous Media

Media theorist Adrian Mackenzie highlights the ‘*naturalisation*’ of Wi-Fi as indicative of a shift in our socio-cultural model (as cited in Han, 2011). The world of the early 21st century features a wide-reaching infrastructure of wireless hotspots and charge points, and its citizens possess a staggering array of personal electronic devices. Connectivity has become simply part of the environment.

**“Wi-fi... has begun to ‘naturalise’ itself in buildings, cities, parks, transport systems, and towns.”**

– Mackenzie, 2005, p. 269

In an Anti-Po-Des journal article David and Schleser advise that our everyday use of smart-device enabled GPS technology has led to location itself becoming a commodity (2013, p. 37), able to be collected and recorded in one’s social media profile. This is certainly the case with Google Maps, with ‘*points of interest*’ added in 2009 (Sterling, 2009) and put to use as a game mechanic by AR apps such as Pokémon Go and Ingress. Goggin and Hjorth describe the hyper-saturation of mobile devices as a “*fertile garden for culture*” (2007, p. 181). We might reasonably speculate that location-aware narratives and dynamic content could become a dominant form of media in the near future.

A standard concern whenever a new communicative media form reaches the mainstream is that it will erode the quality of our human interactions (Farman, 2012). Turkle (2011) warns against the potential adverse effects of too much time dedicated to smartphone use, basing her argument on the idea that live encounters are richer than anything offered by networked mediation. While it is true that mobile telephony can supplant a need to meet face-to-face, it is also true that it enables social interactions that would otherwise be impossible in person. Video calls, once the stuff of science fiction, have become an everyday occurrence, and smartphone apps like Snapchat connect friends asynchronously in a way that avoids scheduling conflicts. Mizuko Ito recognises the importance of co-presence as a social tool in her 2005 ethnographic study Japanese youth. Here we see teenagers using mobile phones to include friends in face-to-face gatherings despite physical absence. From this position, mobile media can be appreciated as a vehicle for shared social experiences.

**“Unlike any other medium before, they let users act as both senders and receivers of messages, wherever they are, immersing them in a vast and virtually ubiquitous network of interconnectedness.”**

– de Vries, 2012, pp. 181–182.

Jason Farman points out that although our mobile devices typically “*resist group involvement... crowds frequently use their own mobile phones to connect in common experiences*” (2012, p. 18). As more mAR applications enter the mainstream the prevalence of communal experiences facilitated by personal devices must also increase. This is convincingly demonstrated by the Pokémon Go global sensation of July 2016. This mobile game broke records when, within two weeks of launching, it amassed over 30 million Daily Active Users (Verto Analytics, 2016). DAUs have since stabilised to a far lower count, but the mAR game’s early mass-popularity meant that its social impact was clearly discernable. Pokémon Go created shared experiences, bringing people into physical contact and altering how they related to one another. An atmosphere of camaraderie emerged, with complete strangers holding conversations in the street: ‘*there’s a Bulbasaur over there*’. It changed how people engaged with space, and with each other. Players were offered a way to revisit familiar surroundings with new eyes, overlaying the mundane with a new significance. Instead of just walking through their neighbourhood, players were traversing the fantasy world that lay beneath, unveiled by their smartphone.

As far back as 2009 mobile data traffic was outstripping voice traffic (Sharma, 2010). Since then the display has remained the most important piece of hardware on what we still counter-intuitively refer to as a ‘*phone*’, with recent reports have even finding that the majority of Internet traffic now comes from mobile devices (Sterling, 2016). Put simply, we no longer use our phones as phones.

## 2.2 Mobile Media Innovation

Arceneaux & Kavoori, writing in 2006, foresaw a new participatory medium in the cell phone, prophetically describing the fledgling device as *“an artifact of popular culture... There are similarities to Internet news groups and e-mail: for example there is no content control; content comes from users”* (2006, p. 51). Six years later they were compelled to revisit this theory, recognising that the technology continues to advance at a blistering pace.

*“These devices are no longer limited to voice and text communication and are quite literally portable, multi-media centers that fit neatly in one’s pocket”*

– Arceneaux & Kavoori 2012, p. 2.

When it was launched in 2007 the iPhone was positioned to meet the needs of an eagerly awaiting telephone market. A decade later, the anniversary edition of the iPhone will be held to an entirely different set of expectations. *“The word “phone” itself feels quaint and outdated”* (Arceneaux & Kavoori, 2012). Apple calls their product the ‘iPhone’, but in truth this name is more an exercise in branding than description. The iPhone is no more a phone than an Aston Martin is a cup holder. It can perform that function, but to pigeonhole it in such crude terms overlooks its real capabilities.

David and Schleser paraphrase the earlier work of Myerson (2001) to surmise that *“users contribute to define the meaning of technology and innovate new ways to deal with technology... mobilisation is not only a technological development, but also a cultural process”* (2013, p. 37). Once a technology is diffused among a population its application inevitably diversifies, regularly filling a role far removed from its intended function.

Howard Rheingold also predicted the cultural impact of mobile media, writing of computer-mediated communication in 1993 that *“[something] big is afoot, and the final shape has not been determined”* (p. xxvi). Bolter and Grusin apply this general observation to storytelling in the early days of mobile phones, noting that mobile media adapts old media to form new customs, reshapes narratives, and generates new forms (1999). This position is reiterated by Berry and Schleser in their later work *Mobile Media Making in the Age of Smartphones*, which opens with the assertion that *“how citizen users and creative professionals represent, experience and share the everyday is changing”* (2014, p. 2). Note that this statement is in the present tense; this remains an ongoing process. The full implications of putting hardware once controlled by the professional sphere into the pockets of the general public are still unfolding.

Professionals and citizen users alike are able to mix media in creative and unprecedented ways. *"With the overlay of location-based services, these experiences and representations are providing new social, creative, and emotional cartographies, which in turn, have created new opportunities for creative practice over the past decade"* (Berry and Schleser, 2013, p. 2).

The unique advantages that mobile media offer to creative practitioners can be summed up neatly in five qualities, as identified by Scott Rushton (2012): ubiquity, portability, personality, connectivity, and locativity.

## 2.3 Augmented Reality

In his classic 1997 survey of the field of AR, Ronald Azuma notes that combining real and virtual spaces is useful because AR *“enhances a user’s perception of and interaction with the real world”* (p. 357), by displaying objects not directly perceptible with the user’s own senses. Imar de Vries paraphrases an assertion from anthropologist Tish Shute, that we have arrived at *“the age of ‘World 2.0’, in which digital layers of information are wrapped around so-called World 1.0 and are made accessible through all-powerful and pervasive mobile media”* (2012, p. 182). In fact, de Vries goes on to declare that new media technologies such as AR can be used to reveal information that in fact was *“already there, hidden from view until an appropriate and well-calibrated looking glass could make it appear right before our eyes”* (2012, p. 183).

AR merges co-existent yet separate realms by *“forging symbiotic relationships between physical reality and virtual dataspace”* (de Vries, 2012, p. 183) to make information that is there-but-not-yet-there accessible. The ways in which we can now interact with virtual environments has completely metamorphosed in two short decades. Whereas in the 1990s one really had only a mouse or keyboard as to interact with the digital world (Snow, 2011), in 2017 touch interfaces, gestural control, and voice commands are all readily available. Given the array of hardware available on today’s devices, perhaps we will soon see the use of location, speed, weather and environment as inputs as well. Not merely as variables to be measured, as is the case currently, but as a means of computer interfacing.

Evidently the dimensions of time and location are not helpful in a purely digital environment, and likewise live global tracking and universal accessibility are unmanageable in an analogue environment. However contemporary mobile media devices are capable of accessing both spaces simultaneously, and the result is a kind of mixed space where the attributes of physical and digital cooperate seamlessly to optimise experiences. This is something that computers simply were not useful for until they could fit into one’s pocket.

## 2.4 Mobile Augmented Reality and Locativity

Mobile AR ('mAR') offers literally a whole new world of cultural opportunities by mobilising distribution channels; *"mAR can create augmented cultural experiences, which can lead to a reinterpretation in the understanding of the current places in the city and simultaneously in a mobile mediascape"* (David & Schleser, 2013, p. 35). When cultural messages are transposed onto different locations, new layers of meaning are added – to both the cultural artifact and the physical space itself. David and Schleser highlight the importance of this development, declaring that while video art surfaced within the gallery, mAR moves into a new realm not confined to any one space. They are critical of the *"commodified brand engagement as seen in the majority of mAR applications that have appeared in Apple's App Store, Nokia's Ovi or the Android Market in recent years"* (2013, p. 35), explaining that these early forays into mAR are naive and not representative of the medium's as yet under-appreciated cultural value.

David and Schleser point to the full suite of hardware fitted to contemporary media devices and the near-permanent state of mobile interconnection as reason within itself to record and evaluate metadata. Contemporary mobile media devices are sophisticated instruments that can measure many variables at once, all the while exchanging information with the wider network. Therefore, David and Schleser advise, it is our duty as creative practitioners to collect and use this metadata to produce imaginative spatial engagements. Where in the past a camera focused on the visual and a compass on the spatial, now we have a device capable of both, a shift that must surely be reflected in consequent film compositions.

**"Mobile applications access various datasets such as GPS coordinates or RSS feeds through their constant connectivity. MAR can connect these forms of information and can present new interactions linking everyday surroundings to the digital media environment."**

– David and Schleser, 2013, p. 39.

Formerly dissociated places are increasingly linked by location-aware technologies that allow their users to inscribe locations with digital information such as text, sound, and images, playing *"an important role in helping people imbue locations with new dynamic meanings and construct new types of urban mobilities and narratives"* (de Souza e Silva, 2013, p. 33). To experience these new narratives and interact with location-based information, people must physically roam. *"As such, they 'read' and 'write' city spaces with the aid of mobile technology"* (de Souza e Silva, 2013, p. 33).

Chris Barker defines place itself as an entirely subjective construct - devoid of meaning besides that which it is assigned by its observer. *"Places are discursive constructions, which are targets of emotional identification or investment"* (2008, p. 376). This idea of location as a sentimental vessel fits well with David and Schleser's assertion that the experience of a location can be changed by adjusting the lens through which it is observed. For David and Schleser, mAR is uniquely situated to create a highly immersive experience because of its capacity to alter the viewer's perception of place (2013, p. 38). Opportunities for new sensorial engagements come about through the exchange of meaning between the location and the mAR application, to synergic effect.

**"The challenge of mAR applications is to shift away from the production of information or even from the production of narratives, to the production of experiences."**

- David and Schleser, 2013, p. 38.

Mobile media devices equipped with multifunctional hardware *"are virtually ubiquitous and they seem to carry an aura of endless potential, which makes them excellent catalysts for innovation dreams"* (de Vries, 2012, p. 184). Claire Boonstra, co-founder of AR industry leader Layar, claims that the full implications of combining all of these features are still far from being realised.

**"We're in the 1994 of the Internet, we're at the moment when radio became TV."**

- Boonstra, 2011.



### 2.5 Virtual Reality Cinema

In conventional cinema the director decides where the audience's attention is directed, using different camera shots to determine the viewer's relationship to the subject. VR cinema introduces a new dynamic, allowing viewers to choose what to watch from a full 360 degrees of content. Where traditionally audiences and content are positioned on opposite sides of a two-dimensional screen, VR cinema puts the viewer them at the heart of the action, not just breaking but entirely removing the fourth wall. VR cinema introduces another dimension not just for audiences but for filmmakers too, who must now devise innovative ways to ensure the viewer knows where to look.

**"For decades, we've been watching as observers in a dark theater... Now, we're taking on the role of observer [inside the movie's] world."**

– Stromberg, as cited in Schwartzel, 2016.

A new generation of consumer-grade VR devices and 360 degree cameras are behind the recent surge in professional VR cinema. While headsets like HTC's 'Vive' and Facebook's 'Oculus Rift' can deliver a premium experience, the fact that the common smartphone can also be adapted for VR viewing has greatly increased the reach of this new medium. A technology that might only have reached the few willing to pay for a dedicated VR system is instead catapulted into the homes of millions, offering rich opportunities for innovative filmmakers. The emergent, costly nature of 360 degree recording equipment does however mean that amateur VR filmmaking is less accessible than its smartphone counterpart.



Fig 5 Filming *Collisions* was an expensive and technically challenging enterprise.

### 2.5.1 Lynette Wallworth - *Collisions*

'*Collisions*' directed by Australian artist Lynette Wallworth is an early example of VR cinema, recorded on-location in the remote desert of Western Australia using a combination of 360 degree cameras, drones, and a quadrophonic microphone (ACMI, 2017). The documentary tells the dramatic story of an indigenous Martu man's dramatic first European encounter – witnessing the Maralinga atomic tests without any frame of reference in the 1950s (Dethridge, 2016).

The Australian Centre for the Moving Image (ACMI), an exhibitor of the work, hosted *Collisions* audiences in rooms filled with *Samsung Gear VR* headsets. The phones used for viewing were linked to a synchronising technology that allowed the content to play simultaneously across all devices. According to ACMI: *"Everyone enters and exits the immersive experience at the same time, enabling a shared experience"* (ACMI, 2017). Personally I would question the accuracy of this claim, because although VR cinema forges a very strong connection to the subject, the viewer can hardly turn to the person next to them to point out something within the film. While the traditional big-screen environment brings people together into one room, VR applications block out the real world precisely so that the senses can be fooled into accepting a virtual one. This is a shift that raises real questions both about cinema as a social activity, and about the role of the viewing environment itself.



**Fig 6** An audience watches Wallworth's *Collisions* at the 2016 Sundance Film Festival.

**"Will a future where we strap boxes to our faces in the cinema aisles be too embarrassing?"**

– Page, 2015.

Something else that occurs to me is how VR cinema renders mobile devices... immobile. There is some degree of irony here. Granted, the reason that Samsung Gear VR incorporates a smartphone is purely to leverage an existing market, but in doing so it robs the device of its most fundamental quality. The result is a user who cannot physically move lest they risk colliding with an object not apparent in their virtual world.

While the best of both worlds may eventually arrive in the form of '*Mixed Reality*' headsets, as of early 2017 we are yet to see a film produced for Microsoft's currently developer-only '*Hololens*' platform, and other Mixed Reality devices remain without so much as a release date (Stein, 2017). For the time being it seems mAR wins in terms of media mobility and locativity, both substantial victories when it comes to expressing amateur travel stories.





▲ **Fig 7** Microsoft's Hololens is currently only available for developer use. No film has yet been produced for the Mixed Reality space.

## 2.6 Mobile Filmmaking

Mobile phones and tablets are intensely personal devices, and the videos they capture subsequently share this character. An honest, often amateur quality is distinctly present in mobile film, conveying a sense of eyewitness account. The absence of any high-end post production work can make the content more relatable; the footage is shot as it might be if you were to visit the location yourself. A low-fidelity aesthetic communicates an air of spontaneity, even carelessness, capturing the true character of a space without the interference of a Director of Photography's subjective idea of what a space 'should' look like. *"Mobile devices make the mundane interesting, the everyday confronted, providing a new lens for viewing the world through a new camera vision"* (Schleser et al. 2009, p. 119).

In a conversation with Gerda Cammaer, I inquired as to how she had taken up the cellphone, and later, smartphone, as a filmmaking tool. She explained that she had always been passionate about experimental filmmaking, but had grown disillusioned with the avant-garde movement as a whole after observing that it had drifted from true pioneering towards simply '*alternative filmmaking*', repeating techniques and processes that although underground introduced nothing new to the field. The true avant-garde, she asserted, was in smartphone filmmaking. Here was a community of filmmakers truly exploring the creative fringes, putting mobile media devices to work in new and unexpected ways that were clever and thought provoking.

**"Mobile filmmakers use mobile media to communicate perceptions about locations via 3gp and mpeg4 videos in a way that HDV video cannot account for."**

– David & Schleser, 2013, p. 36.

## 2.7 Database Cinema

Database cinema is a new media form that constructs a narrative by dynamically picking from a library of scenes. Typically there is no strict beginning or end, and the individual scenes themselves contain precious little in the way of a story, but when combined produce a collective narrative. The order of the database footage may be determined randomly; by the viewer themselves or more recently, algorithmically. With every restructuring of the scenes a slightly different message is conveyed.

The value of database cinema, according to pioneer Lev Manovich, is in its provision of so many elements from which the audience may construct a narrative (2001, p. 218). Kristen Daly expands on this assertion, recognising the importance of the database “[*not as*] a random collection but a collection organized by a particular model” (2010, p. 89). It is not hard to imagine that the introduction of algorithms able to trawl masses of user-generated media for appropriate sequences will have a considerable impact on both database cinema and interactive film in general.

The advent of social media and influence of restrictive bandwidths mean contemporary audiences are now accustomed to segmented narratives. This is seen in the popularity of platforms like Snapchat and Periscope, and to a certain extent even YouTube, whose algorithms recommend ‘scenes’ of footage that extend the narrative of clips already viewed.



## LITERATURE REVIEW INSIGHTS

Mobile devices have already altered the way that we spatially interact with cultural objects and each other

Naturalised interconnectivity is the new normal

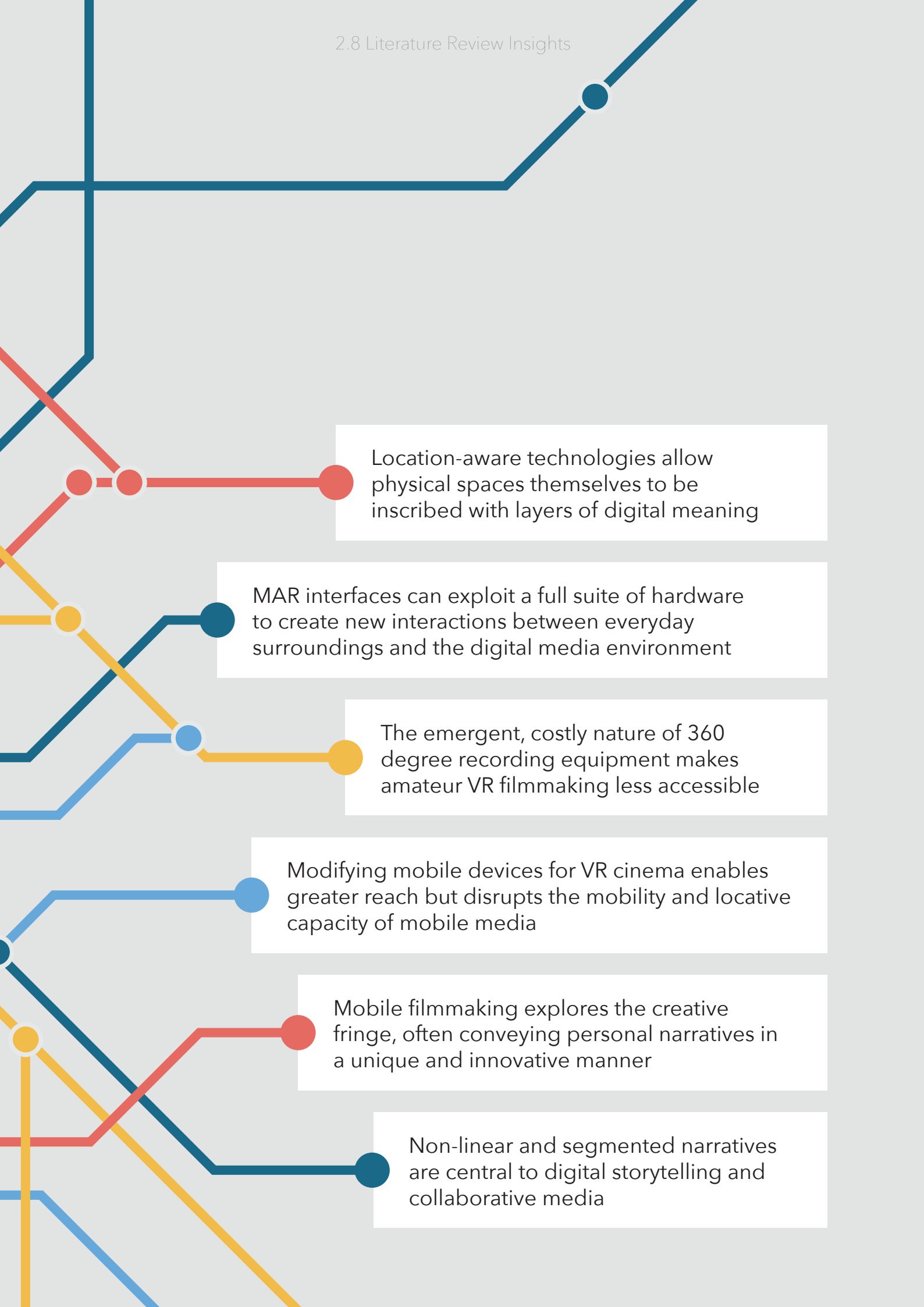
Personal mobile devices regularly support shared social experiences and enable a state of co-presence

As a new technology is democratised its application diversifies correspondingly

Mobile media represent ubiquity, portability, personality, connectivity, and locativity

Digital augmentation allows a perception of already present factors not perceptible with one's own senses



An abstract graphic featuring several overlapping lines in dark blue, red, yellow, and light blue. These lines are composed of straight segments connected by small circles. Some circles are solid colors matching the line, while others are white with a colored outline. The lines and circles are scattered across the page, with some connecting to text boxes and others remaining as standalone elements.

Location-aware technologies allow physical spaces themselves to be inscribed with layers of digital meaning

MAR interfaces can exploit a full suite of hardware to create new interactions between everyday surroundings and the digital media environment

The emergent, costly nature of 360 degree recording equipment makes amateur VR filmmaking less accessible

Modifying mobile devices for VR cinema enables greater reach but disrupts the mobility and locative capacity of mobile media

Mobile filmmaking explores the creative fringe, often conveying personal narratives in a unique and innovative manner

Non-linear and segmented narratives are central to digital storytelling and collaborative media



# CREATIVE PRACTICE

At the beginning of this project I had no preconceived notions of the form or function that the practical design output would take. An untethered approach allows me to avoid settling on any one application too soon, because doing so would mean turning my back on the potential opportunities of another technology left unexplored. Though it meant a less direct path, I made it my objective to explore many technologies; identifying the emerging 'big players' in each field, wherever possible accessing their product offering, teaching myself to use it, and testing the technology myself to gauge its relative value towards mobile filmmaking.

From the outset, Viewfinders aimed to embrace collaborative culture and collate a travel documentary from clips submitted by other hobby filmmakers. Media technologist Tatjana Dzambazova points to the value of doing so, explaining that mobile media heighten the interplay between producer and audience, resulting narratives that "mean so much more to us because we were part of their creation in the first place" (2015). Taking a peer-generated media approach enables Viewfinders to expand its significance for all involved, and by opening itself to a global network of filmmakers the coverage and reach of the documentary is increased exponentially. My first task was to establish a web presence for the project, whose function was threefold:

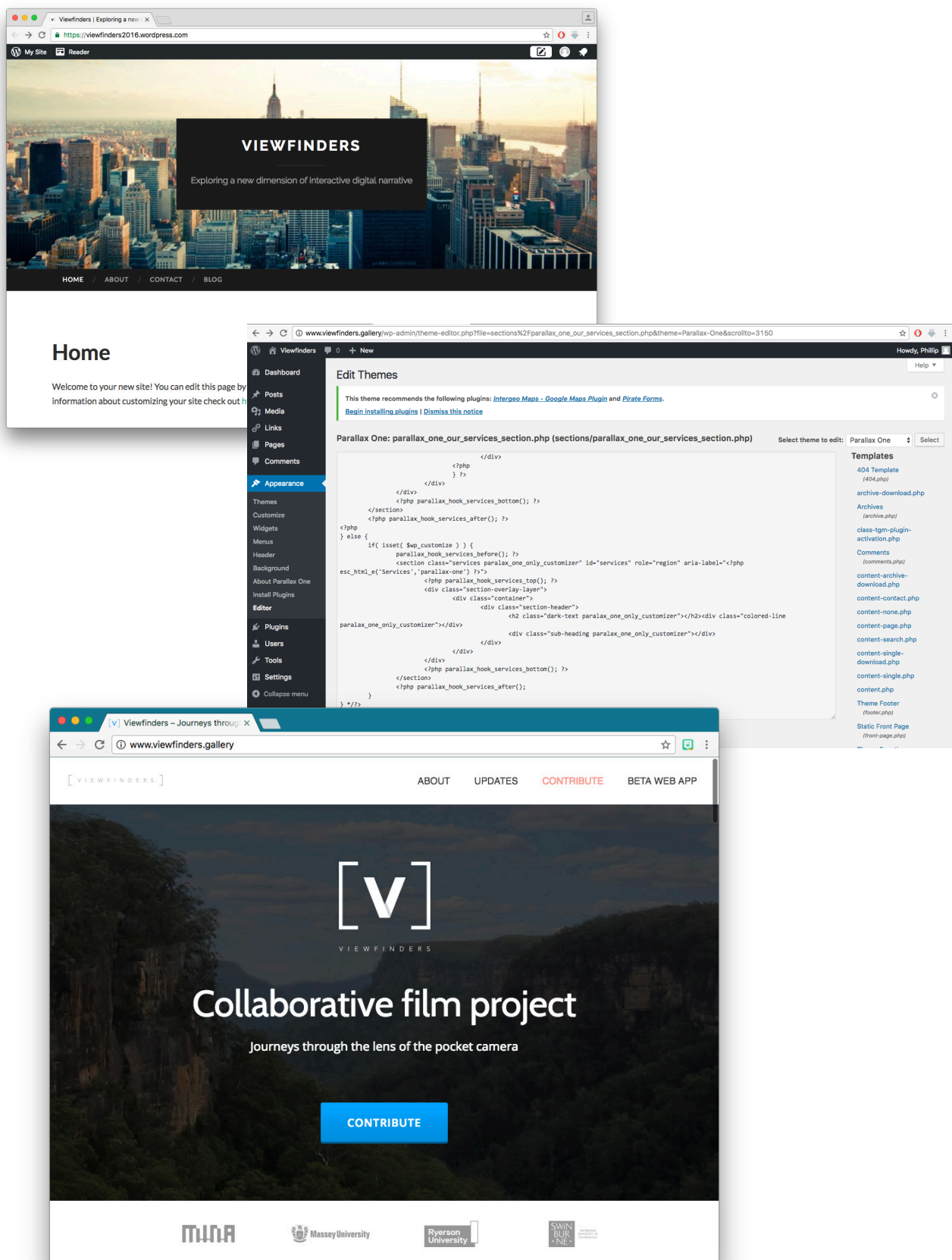
- > To engage members of MINA and the wider amateur filmmaking community.
- > To educate potential contributors on the project intent.
- > To receive digital video files from participants.

### 3.0 Viewfinders.gallery

Producing the website itself was straightforward. I opted for WordPress as my Content Management System ('CMS') because it streamlines the website production process. After installing the setup files onto a server, it was simply a matter of finding a suitable theme and using the WordPress Graphical User Interface ('GUI') to upload all of my content onto the site. I was able to implement a plugin to display a dynamic Twitter feed which I linked to a Viewfinders Twitter account, and another plugin that embedded a YouTube playlist into the landing page of the site. All of this was managed through a front end interface, and exploited an extensive range of existing elements. With the major elements in place I then set about manipulating the template from the back end, accessing the HTML and CSS code to fine-tune the site and ensure it was fit for purpose. This required a few revisions as the Viewfinders brand was refined throughout the project.

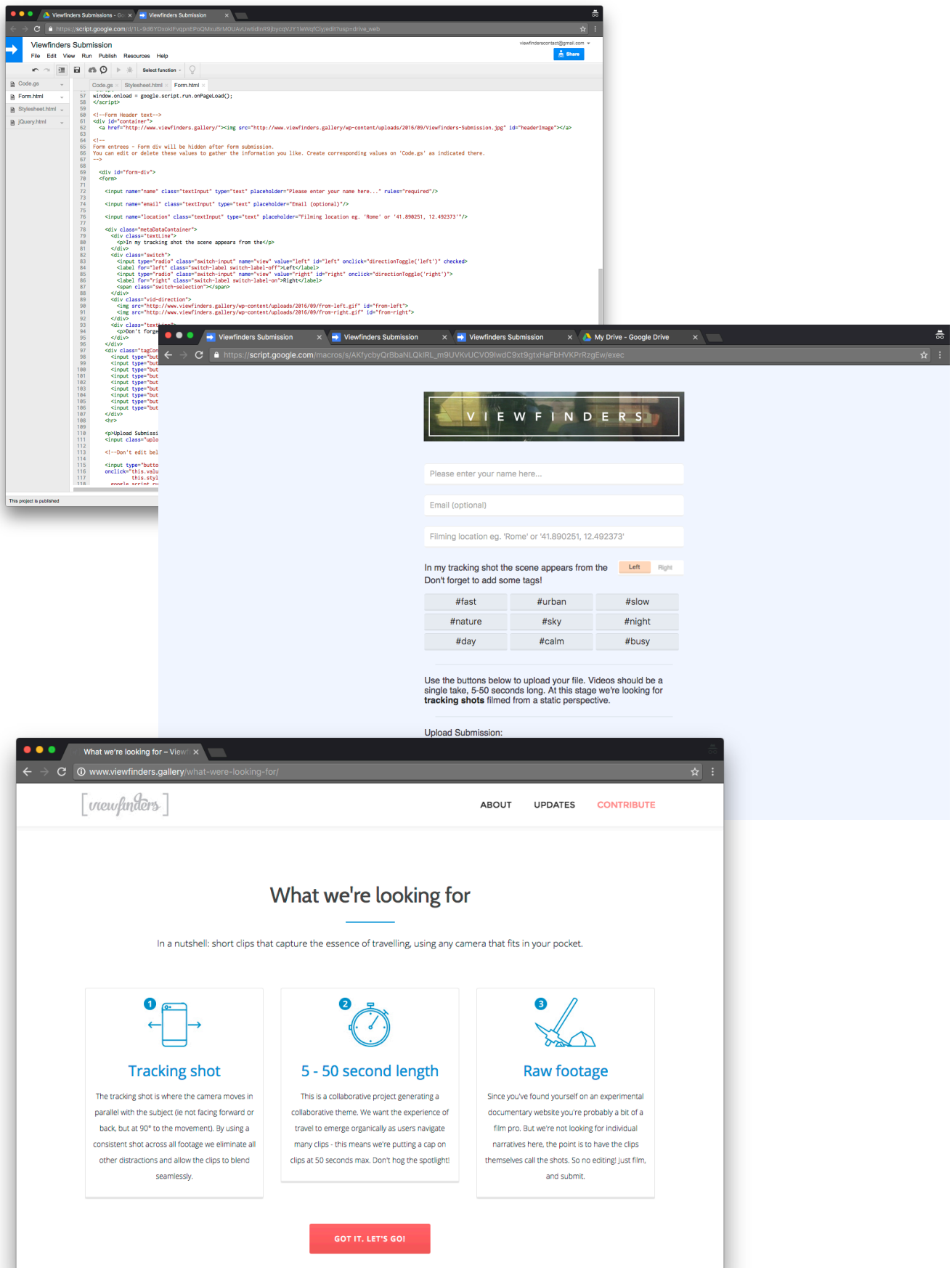
Establishing a system to receive video files proved more complex. Initially I looked at using Google Forms as a service with a similarly streamlined GUI. However I soon found that a significant drawback of the service is that users cannot upload files, since the form only accepts text input. The standard workaround recommended by various online forums is to switch to Google's Drive service, and instead provide contributors a link to a shared folder where they can upload their files. Unfortunately this method requires that the contributor have a Google account, as well as granting universal access to edit or delete the contents of the folder - not to mention it separates the file from its associated metadata input on the form. Other form-building services do exist (*Jotform.com*, *WPForms.com*), however they are overwhelmingly plastered with third-party branding and prohibitively expensive.

After consulting various threads across forum sites *StackOverflow.com*, *Quora*, and *CtrlQ.org*, I determined that the best course of action was to create a web form from scratch written in HTML, and use the HTML service of Google Apps Script to deploy the form as a web app. This rather convoluted workflow relied on some level of familiarity with web development methods, but I saw it as an opportunity to progress my technical proficiency in this area. Using Google Apps Script also had the silver lining of enabling much greater customisation since I would be writing all of the code myself, thus allowing the submission form to reflect the Viewfinders visual identity without any hint of third party branding.



▲ Fig 8 Various stages of Wordpress development.





▲ Fig 9 The original upload form produced using Google Apps Script and subsequent alterations as the Viewfinders brand evolved.



The image displays the 'Viewfinders Submission Form' in a web browser and a Google Sheet titled '\_Submission Data'.

**Viewfinders Submission Form:**

The form is titled 'VIEWFINDERS SUBMISSION' and is at 'Step 1 | Choose your clip'. It includes the following fields and options:

- Please enter your name here...
- What country do you call home?
- Email (optional)
- Click here to work out your video's GPS coordinates (with a globe icon)
- Click here to specify when this clip was filmed (with a calendar icon)
- Submit via: YouTube (selected) or Upload
- Already on Youtube? Great! Paste the URL here.

**\_Submission Data Google Sheet:**

The sheet contains a table with the following columns: Contributed By, Home Country, Email, Timestamp, Terms Status, Coordinates, Place Name, Country, and Date of Filming. The data lists 34 submissions from various countries including New Zealand, The Netherlands, USA, Australia, Canada, Sweden, Romania, and Japan.

	Contributed By	Home Country	Email	Timestamp	Terms Status	Coordinates	Place Name	Country	Date of Filming
10	Phillip Rubery	New Zealand		2016.10.28.23.55	Terms accepted.	-41.234773, 174.367132	Cook Strait	New Zealand	October, 2016
11	Ben	The Netherlands		2016.11.18.17.04	Terms accepted.	51.915921, 4.4900499999999995	Stadsdriehoek	Netherlands	October, 2016
12	Sue Donym	USA		2016.11.19.11.15	Terms accepted.	-27.517715517933613, 153.0499899	Greenslopes	Australia	November, 2012
13	Sue Donym	USA		2016.11.19.11.21	Terms accepted.	-37.8162175, 144.96406820000004	Melbourne	Australia	December, 2015
14	Sue Donym	USA		2016.11.19.11.24	Terms accepted.	45.4408474, 12.315515099999997	Santa Marta	Italy	July, 2014
15	Sue Donym	USA		2016.11.19.11.40	Terms accepted.	-34.9107564, -56.20040460000001	Maldonado	Uruguay	May, 2105
16	Sue Donym	USA		2016.11.19.11.41	Terms accepted.	37.566535, 126.97796919999996	Jung-gu	South Korea	January, 2015
17	Sue Donym	USA		2016.11.19.12.07	Terms accepted.	15.151671572942313, 77.617721557	Andhra Pradesh	India	May, 2014
18	Max	Australia		2017.01.05.15.54	Terms accepted.	22.289254, 114.152331	Hong Kong Island	Hong Kong	September, 2014
19	Max	Australia		2017.01.05.15.56	Terms accepted.	22.249254, 114.102331	Hong Kong Island	Hong Kong	September, 2014
20	Max	Australia		2017.01.05.15.57	Terms accepted.	22.3006017, 114.16741380000008	Kowloon	Hong Kong	September, 2014
21	Terry Cullinan	Canada		2017.01.06.19.45	Terms accepted.	18.2766312, -78.34671879999996	Negril	Jamaica	December, 2016
22	Terry Cullinan	Canada		2017.01.06.20.06	Terms accepted.	18.4073894, -77.10313450000001	Ocho Rios	Jamaica	December, 2016
23	Terry Cullinan	Canada		2017.01.06.20.09	Terms accepted.	18.4329473, -77.1973997	St Ann's Bay	Jamaica	December, 2016
24	Terry Cullinan	Canada		2017.01.07.14.52	Terms accepted.	45.3689234, -75.69258870000003	Ottawa	Canada	January, 2017
25	Terry Cullinan	Canada		2017.01.07.14.57	Terms accepted.	45.3981264, -75.60298870000003	Ottawa	Canada	January, 2017
26	Terry Cullinan	Canada		2017.01.07.15.11	Terms accepted.	45.3582204, -75.62268870000003	Ottawa	Canada	January, 2017
27	Terry Cullinan	Canada		2017.01.07.15.17	Terms accepted.	45.3280234, -75.67208870000003	Ottawa	Canada	January, 2017
28	Catherine L	Canada		2017.01.07.23.28	Terms accepted.	-54.8019121, -68.30295109999997	Ushualia	Argentina	December, 2015
29	Randi Gitz	Sweden		2017.01.08.10.23	Terms accepted.	62.73629159468256, 15.4954862594	Bräcke Ö	Sverige	January, 2017
30	Kim Munro	Australia		2017.02.06.14.54	Terms accepted.	66.071471, -18.653697999999963	Sigljúfjörður	Iceland	January, 2017
31	Toma Peiu	Romania		2017.02.20.00.55	Terms accepted.	33.818828, -118.22582999999997	Los Angeles	USA	February, 2017
32	Luiza Parvu	Romania		2017.02.20.00.58	Terms accepted.	19.393664, -99.174598	Nápoles	Mexico	January, 2014
33	Toma Peiu	USA		2017.03.04.14.33	Terms accepted.	40.737036, -73.99239899999998	New York	Stalete Unite ale Am	February, 2017
34	Elaine	Australia		2017.04.26.11.17	Terms accepted.	35.1236178, 135.77143339999998	Kyoto	Japan	April, 2017

▲ Fig 10 The submission form was launched early to secure as much peer-generated footage as possible for further development.

### 3.0.1 Early Access Map Interface

With the submission process launched there was an immediate need to grant participants access to their material so they could see how their submission contributes to the project. Offering a kind of preview interface encourages repeat submissions, and also worked as a bargaining chip to buy time while the AR interface was developed. At this stage the eventual AR component was still far from a crystallised concept, but I knew that it would respond primarily to the locativity of mobile media. The placeholder application therefore needed to communicate an emphasis on location to give users the sense that they are travelling as they move through the footage.

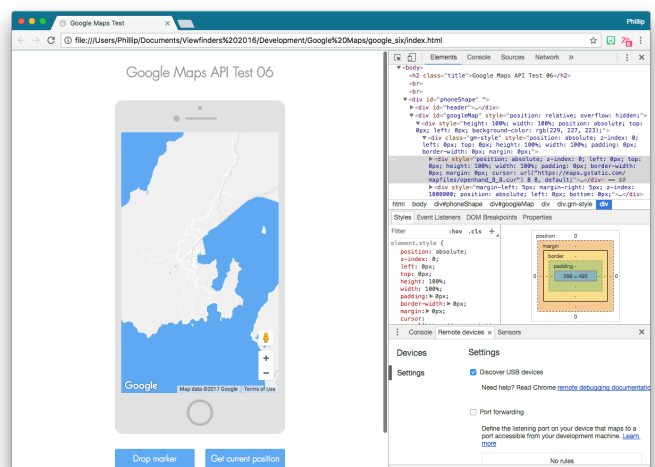
### 3.0.2 Google Maps API

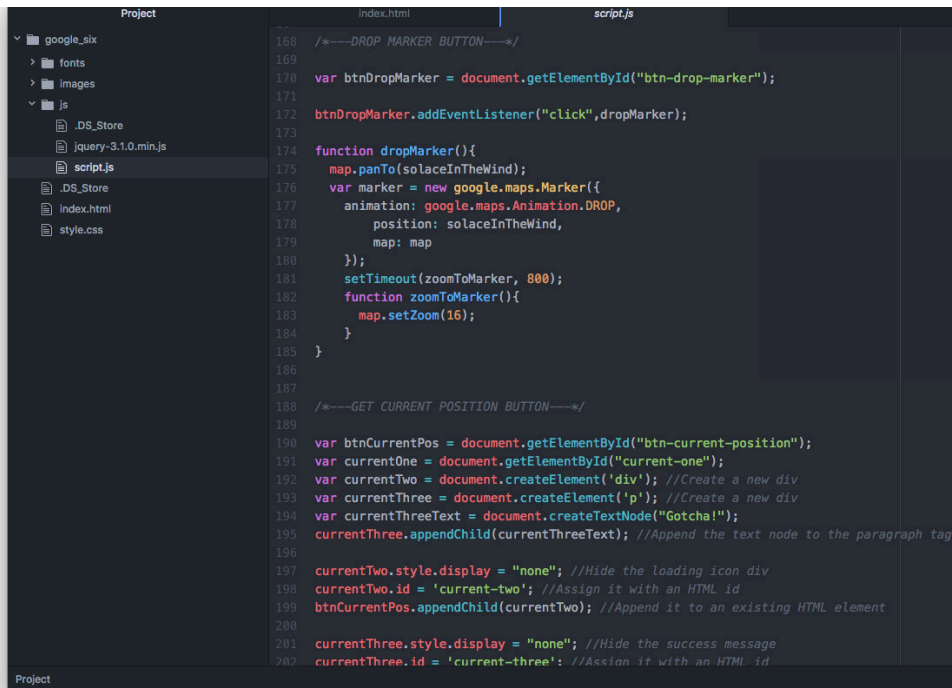
Launched in 2005, Google Maps introduced the world to itself through a revolutionary technology that stitched satellite images together into a seamless composite map of the entire globe (Gannes, 2015). Twelve years later this product is synonymous with wayfinding to the extent that the name itself, like its search engine forebear, has been adopted as a verb in everyday speak. However the impact that Google Maps has had on the modern Internet is even greater than most realise. As an early proponent of the web API protocol, Google Maps was central to ushering in a new trend in web development by exposing the back end services to third-party developers. Through the Google Maps API, the Viewfinders early access interface is able to access select services and data used by Google Maps itself.

**“With this one little API, developers can harness the awesome power of Google Maps to fulfill our wildest geographical dreams”**

– Cater, 2013.

► **Fig 11** One of a series of tests using the Google Maps API: The Graphical User Interface (left), with HTML structure exposed (right).

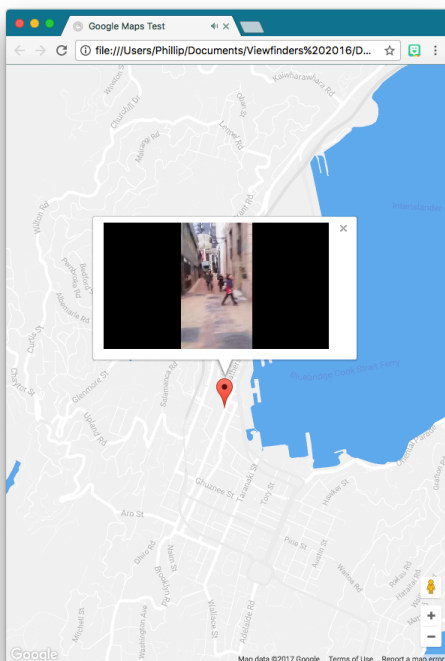




◀ **Fig 12** The Javascript written to interact with Google's 'Maps' API and enact the functions described by the buttons in the interface.

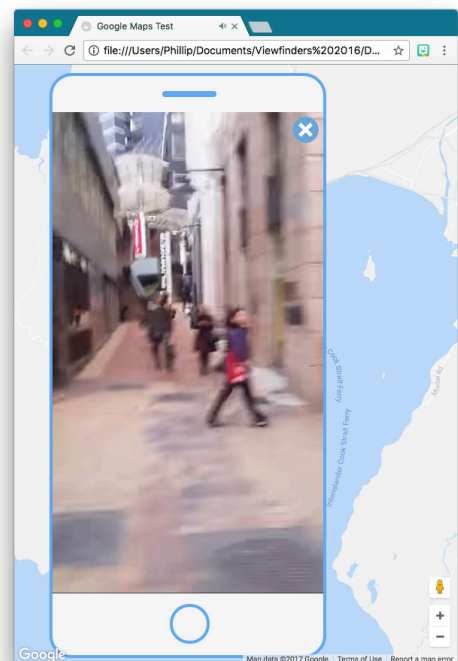
### 3.0.3 YouTube and Google Maps Mashup

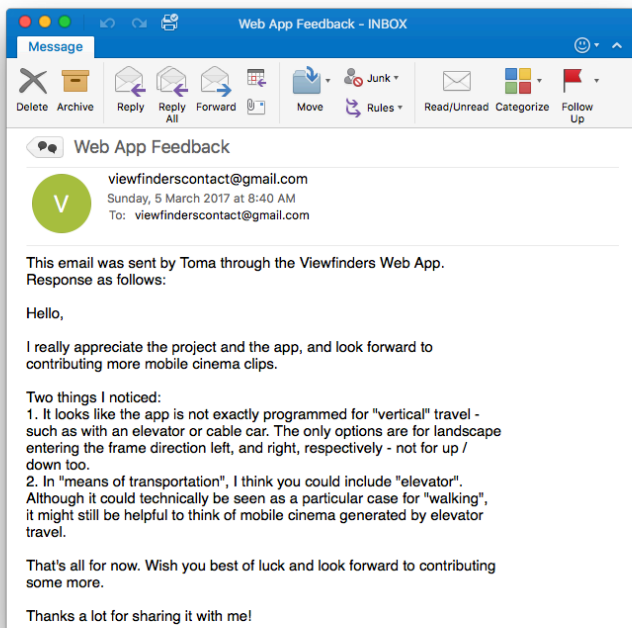
In 2005, Google's release of the Maps API was breaking new ground (Taylor, 2005). Since then however releasing an API has become standard practice and Google has released a whole suite of them, including the Google Drive and Google Sheets APIs that I used to develop the Viewfinders submission form. To ensure that the early access web app retained an inherent association between footage and location, even on desktop machines, I devised a custom 'mashup' (Cater, 2013) of the Google Maps and YouTube APIs to position a video 'iframe' HTML element within a map.



◀ **Fig 13** The default Google Maps API *InfoWindow* cannot be styled, imposing a mandatory margin and close button.

▶ **Fig 14** Using the *infobox.js* plugin for Google Maps incorporates HTML elements directly into the Google Maps canvas - superseding the need for an *InfoWindow* and enabling customised windows such as this phone form.





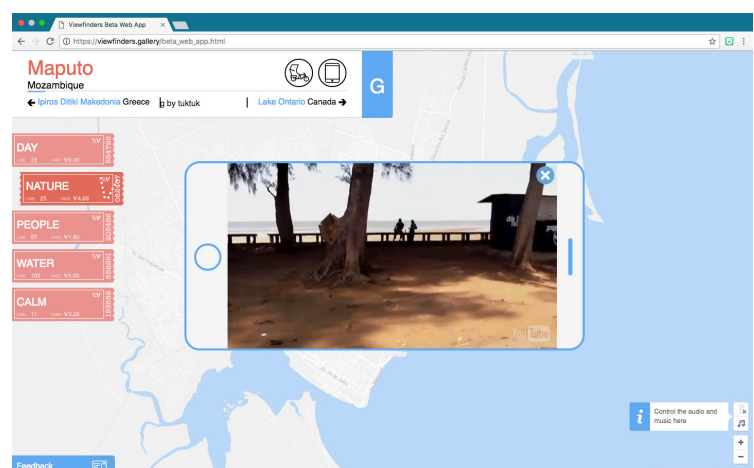
▶ **Fig 15** Embedding a feedback form within the map interface itself promoted the collaborative nature of the project, and brought in valuable insights.

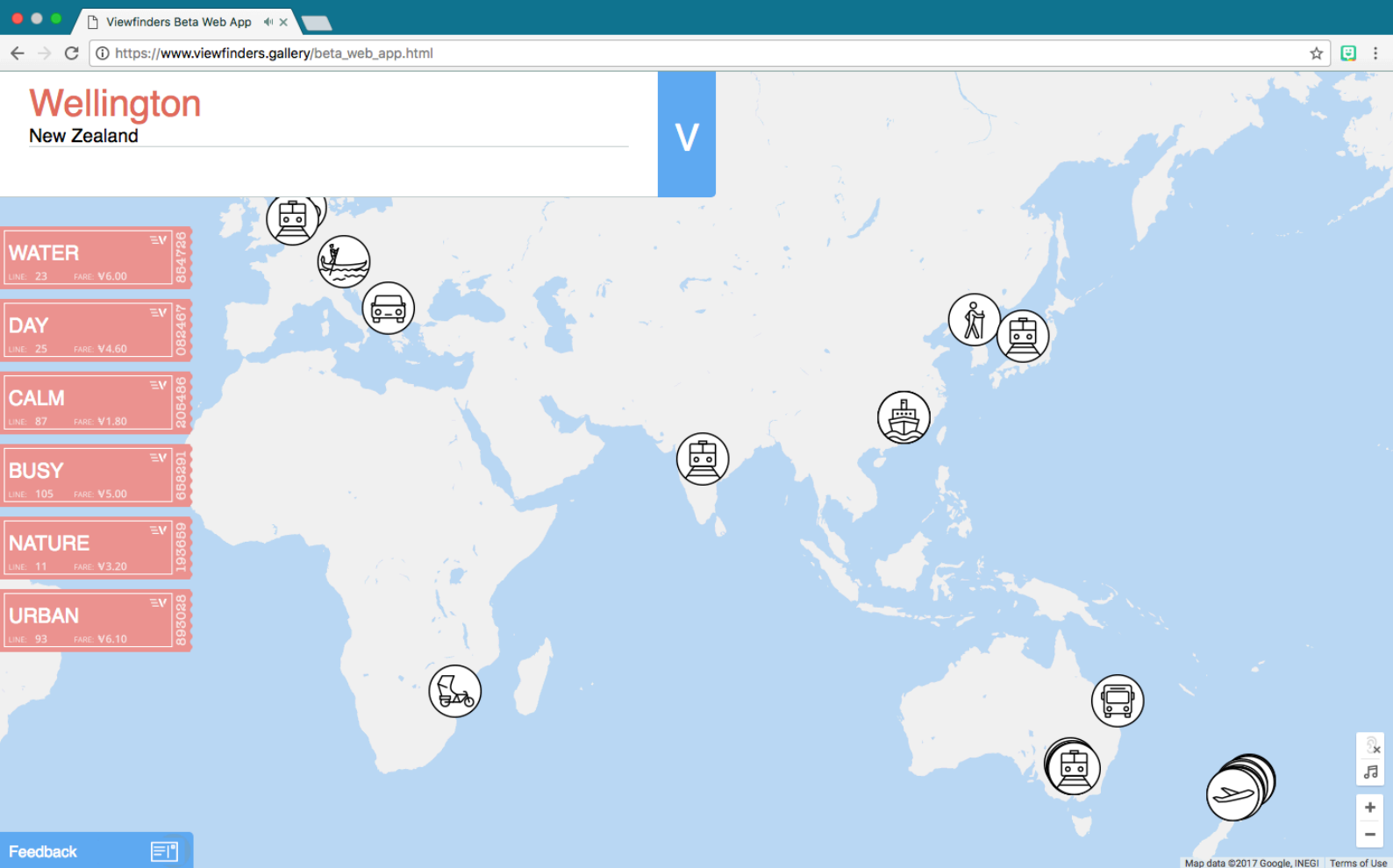
Strictly speaking the Google Maps API does not officially support inserting customised content into the map itself. Users are offered markers and 'InfoWindows' to call out information anchored to a location on the map, but the InfoWindow is styled for use as a text

box and cannot be modified - clearly Google have determined that their maps should be used exclusively for wayfinding. However the popularity of the Maps API means I am far from the first developer to see this as a limitation. Code repository website GitHub hosts an open source initiative called '*Infobox.js*', which is produced and maintained by an informal co-operative of developers. By incorporating the Infobox javascript plugin into my own code I am able to deploy YouTube content natively within a Google Map, which opens the interface up for far more creative interpretations of the service. Given my initial inexperience working with these tools I worked cautiously, producing successive tests independent from one another to avoid corrupting any earlier code, since progress was often hard-won.

The Graphical User Interface was designed using a visual language that echoes airports and bus terminals to reflect the theme of travel, and evolved as each javascript function passed bug-testing and proved viable for use. Through studying the YouTube API documentation I was able to remove the standard playback buttons and hand control of video content over to elements within my own UI. Some prominent changes include removing the ability to pause content, making video footage play automatically when a map marker is selected, and dynamically compiling playlists of peer-generated footage in response to users selecting a thematic '*ticket*'.

▶ **Fig 16** The standard YouTube iFrame behaviour was overridden to prevent pausing the footage - tapping the video makes it play full-screen rather than 'ending the journey'.





▲ **Fig 17** Getting the submission page and early access web app running early in the project allowed time for submissions to accumulate while other research was still ongoing.

With the UI operational the final step was to connect the web app to the live project metadata recorded in a Google Sheets matrix. This was achieved by making a call to the Google Sheets API through Google Apps script, which then returns the metadata in 'JSON' format, a Javascript list structure that holds values to be read by other code.

I deployed the early access interface as a web app and added an access button to the Viewfinders website, as well as adapting the submission form to redirect contributors to the page whenever a submission is completed. In doing so I effected a closed feedback loop - as soon as a person contributes to the Viewfinders project they are able to see how their interaction affected the documentary.

In throwing myself headlong into this world of APIs, plugins, databases, and back-end modification I vastly deepened my understanding of back and front-end web development, learning many tools and methods along the way that enable me to approach the much greater task of developing the eventual mobile interface component of the project.



### 3.1 Virtual Reality

In 2016, when this project began, the mobile technology most noticeably in the spotlight was Virtual Reality. Although many efforts to create a fully immersive virtual environment have been attempted in the past (see Morton Heilig's *'Sensorama'*, 1962), the last few years have finally seen mobile capabilities reach a critical point where the industry is rapidly gaining traction in the commercial space. It seems any technology company worth its salt has some kind of prototype underway, with many now at the point where they can release a fully-fledged consumer product - see Google's *'Daydream'*, Facebook's *'Oculus'*, and Samsung's *'Gear VR'*.

At early meetings with Drs. Cammaer and Schleser we discussed the possibility of creating a VR space for the travel documentary to inhabit where peer-submitted clips would appear in 3D space, with a wayfinding system informed by a kind of airport-signage aesthetic. Early concepts had the viewer-user move around the space by physically stepping forward, with their movement amplified in the virtual space so that a single step might take them across an entire country to arrive at the next film clip. Each piece of footage would be positioned within this virtual world to correspond to the real-world location where they were recorded.

Although VR is expected to bring sweeping changes across the entire mediascape, a VR travel documentary was quickly deemed an unsuitable direction for this project and dismissed early on. When assessed critically through the lens of creative mobile media practice, VR fails to make use of the single most fundamental advantage afforded by mobile devices - their sheer portability. Viewfinders seeks new creative manifestations of travel film, and the prospect of a display that can itself travel simply cannot be overlooked. While it is true that platforms like Samsung's Gear VR use portable smartphones to deliver their offering, this approach is taken purely because leveraging existing smartphone distribution has considerable logistical benefits, and not because the devices offer any technical advantage over a dedicated VR system. The fact is a dedicated machine can easily outperform a smartphone, but consumers are more likely to opt for a relatively cheap smartphone add-on than invest in an expensive purpose-built device. Here we have a curious paradox.

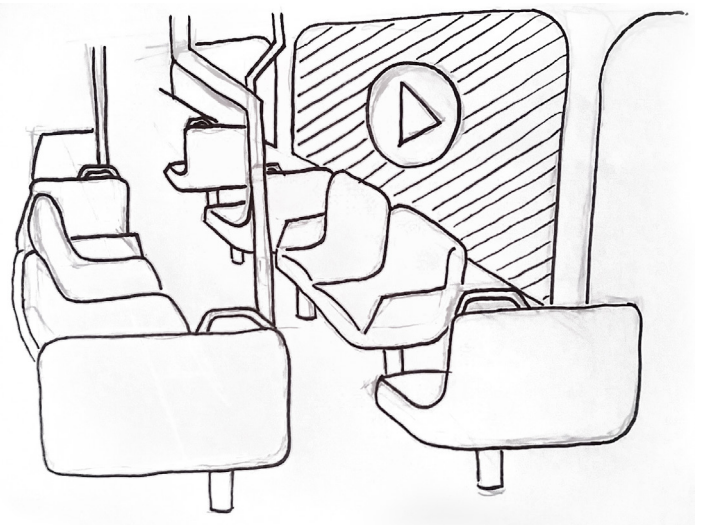
Though VR has been introduced to the consumer market as part of the mobile ecosystem it can only be used while stationary. Additionally, since by definition VR blocks out the real world to supplant it with a virtual one, the location where one uses a VR device is completely inconsequential - all location-specific attributes are erased.



## 3.2 Augmented Reality

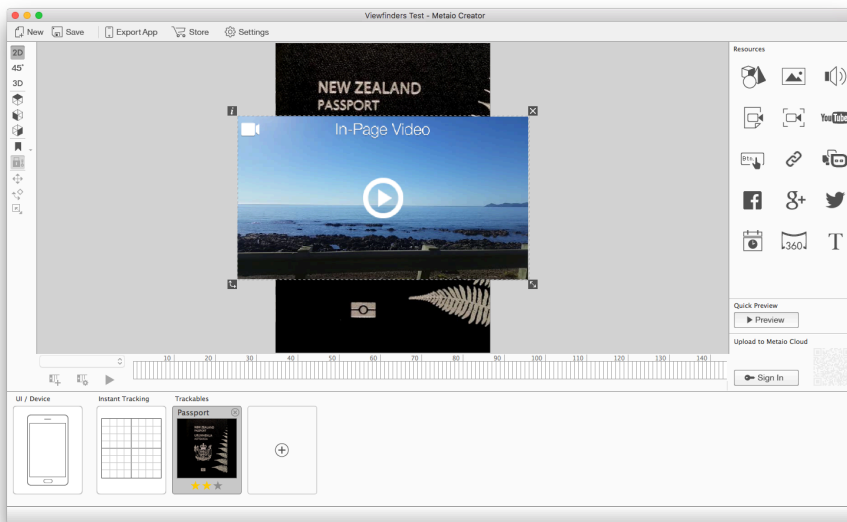
The next logical step in my appraisal of the mobile-mediascape was VR's running mate: Augmented Reality. Unlike VR, AR embraces and actively relies on a perception of the real world. My initial concept using AR was to have a smartphone recognise flat surfaces within the viewer's real-world environment, then have peer-submitted footage overlay a live camera feed on the device display, creating a kind of projection viewable through the smartphone. Ideally, the plane 'projected' onto would be a window, creating the illusion that the viewer is looking out a window to see the documentary footage. Whether the AR app was being used on a bus or in a building, the window would effectively become a portal within the viewer's own physical space and merge the travelling shot into the viewer's live environment.

In producing an application a developer will typically start in an 'Integrated Development Environment' (IDE), a piece of software used to develop and compile applications. The most appropriate IDE for the task would depend on which 'Software Development Kit' (SDK) was deemed best for the task at hand; that is to say I had to begin by finding a package of code that could deliver the functions I wanted, then build my process around how that code works. The lucrative potential of recent AR developments has spawned many new companies, each with a slightly different approach and product offering. My criteria for choosing a suitable SDK was influenced by the 12-month project time-frame, a modest budget, an entry-level experience in software development, and scalability.



▲ **Fig 18** An early idea was to use windows within the viewer's environment to 'project' a tracking shot into, bringing the documentary into the real world.

ARToolkit is a cross-platform tracking library first demonstrated in 1999, and released open source to programmers in 2001. One of the key advantages of ARToolkit is its longevity, spawning a large community of active developers over nearly two decades and giving rise to an extensive and well-maintained development forum. The downside of ARToolkit's seniority to an entry-level developer like myself is that there is no GUI - the SDK expects users to already be proficient using an industry-standard IDE. Throughout my exploration phase I encountered this expectation time and again, with many AR SDKs providing thorough documentation to new users provided they were competent

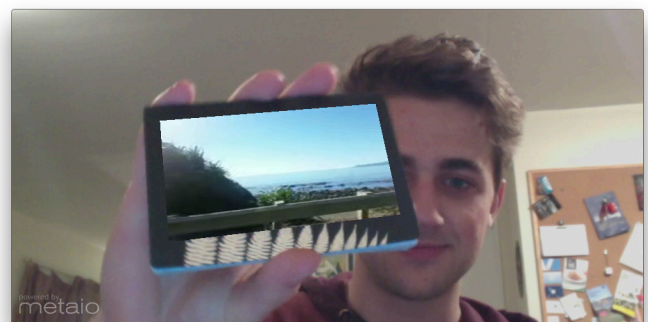


◀ **Fig 19** Inserting video content in the lightweight Metaio Creator, using a static image of a passport as the target.

using any or all of Xcode, Microsoft Visual Studio, or OpenGL. As a practice-based researcher I seek insights through hands-on experience, so although in a commercial setting it would make sense to contract a professional developer already familiar with these tools, doing so in this academic study constitutes a serious compromise. Observing tensions and opportunities as they arise allows me to pivot rapidly and also informs decisions around the form and function of the final design output. In any case, the exploration phase would offer very little in the way of a brief to external parties, since it is only much later in the double-diamond design process that the project settles on a real definition of what it is and what it does.

Equipped with a new appreciation for the complexities of industry level AR SDKs, I next turned to the opposite end of the spectrum: fully-produced AR apps aimed at the consumer market. A standard example of a large corporation dabbling in new technology, Hewlett-Packard's 'Aurasma' app is lightweight, fast, and intuitive to use, claiming to make it *"easy for anyone to create augmented reality experiences"* (Aurasma, 2016). Users are able to define their own target images on which to train their choice of AR content, whether image or video, then test them on their own device or publish them to the Aurasma user community. The pared-back user interface is geared towards maximising reach, making it a first port of call for people with a casual interest in the up-and-coming technology.

▶ **Fig 20** The video plays automatically, with tap to toggle play state enabled by default in Metaio.



The trade-off here is with capability; by removing any option to diversify the AR's application, the result is a kind of one-size-fits-all, with content bound exclusively for use within the Aurasma app. There is no provision for exporting content as a stand-alone entity, and in order to be accessed by other smartphone users, they too must download the Aurasma app, create an account, and view it through this branded portal. There are many other consumer-level apps that are variations on this theme; *BlippAR*,

*Metaio* (recently acquired by Apple), *Geozet*, *Layar*, *Mybrana*, to name a few. Snapchat, originally launched as a network for sharing time-limited image and video content, has in recent years strayed into this space as well, leaning heavily on its ever-expanding library of AR ‘filters’ to keep users engaged. It is clear that while these user-friendly apps have some use for preliminary AR tests, in order to develop customised experiences and break free from proprietary branding I needed to bite the bullet and delve deeper into the world of AR SDKs.



▲ **Fig 21** Wikitude, Augment, and Vuforia are three of the biggest names in AR.

Cautious of the complexities surrounding ARToolkit I next turned to investigate a few of its competitors. Wikitude is one of the bigger players in the nascent AR industry. First launched in 2008, at the time of investigation the latest offering was Wikitude SDK 5, though at the time of writing they have since released SDK 6 - just another example of the rate at which this field is currently advancing. SDK 5 offers different packages for iOS and Android, but also notably extensions for Unity and Cordova, two industry-standard IDEs capable of working across all major smartphone and desktop platforms, as well as having the added advantage of being free to use.

Another major player, Augment takes a hardline commercial approach to AR, providing a start-to-finish service to retailers, manufacturers and educators. Though their paid packages include full design, development and delivery of a custom AR app, Augment also provide a desktop GUI for designers to produce their own targets and AR content, though as with the consumer-level apps investigated previously this content is then only accessible through their Augment smartphone app. Additionally, there is an option to download the Augment SDK, but again this is at cost.

Vuforia proudly claim to be “*The World’s Most Widely Deployed AR Platform*” (Vuforia, 2016). Vuforia software packages are available for a one-off payment, or subscription based if you generate high usage or store app data on their servers. In contrast to the Augment model, Vuforia also offer a development option, which grants free, full-

featured SDK access prior to public release. The real point of difference however is in Vuforia's publicly stated ambition to make AR technology readily available to creative start-ups. A major step towards this came on 1 November 2016 with the announcement that Vuforia would enter into a strategic partnership with Unity, a step that goes beyond simply offering a plugin for the IDE. The frequency of these kinds of developments further exemplify an industry undergoing constant evolution, and a trend towards accessibility.

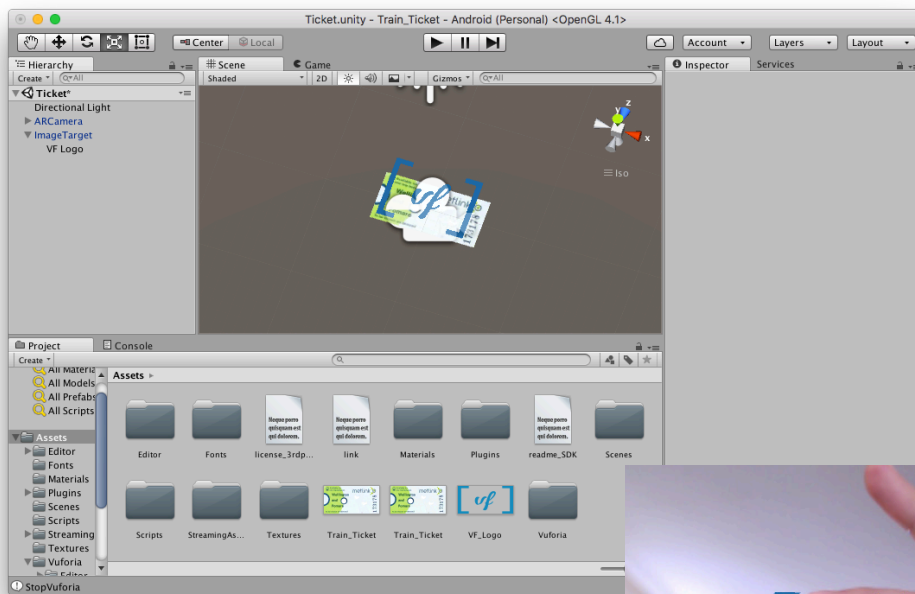
**"Unity and Vuforia share a common mission to democratize AR development... Integrating the offerings will further simplify and enhance the development process to drive greater adoption of AR experiences."**

– Vuforia Product Marketing Director Adam Kerin, 2016.

Though I had no prior experience using Unity I was able to download the IDE for free and teach myself the basics of both the Unity development environment and the Vuforia SDK through online documentation and forums. The ease of using these softwares is actually surprising, and assuming Vuforia do deliver on their pledge to make AR technology more accessible I believe doing so will see AR functionality appear in an ever-more diverse range of applications. My first experiments were very simple, using a two-dimensional photo of a train ticket as the target image to trigger a floating Viewfinders logo. From there I progressed to deploying a video canvas in place of the static logo image, which in Unity is a complex process given the software is intended for three-dimensional game production not interactive documentaries. Several forums advocated using a flat plane object with the video file mapped onto it as a texture, which worked. However it also required developing custom scripts in the C# programming language to control playback, which to take any further would demand that I learn an entirely new programming language. Vuforia clearly anticipate this hurdle for novices however, providing a prefabricated video object for Unity. It seems that Vuforia do genuinely appreciate AR's potential for innovation, and position their products to encourage cross-disciplinary experimentation even from complete beginners.

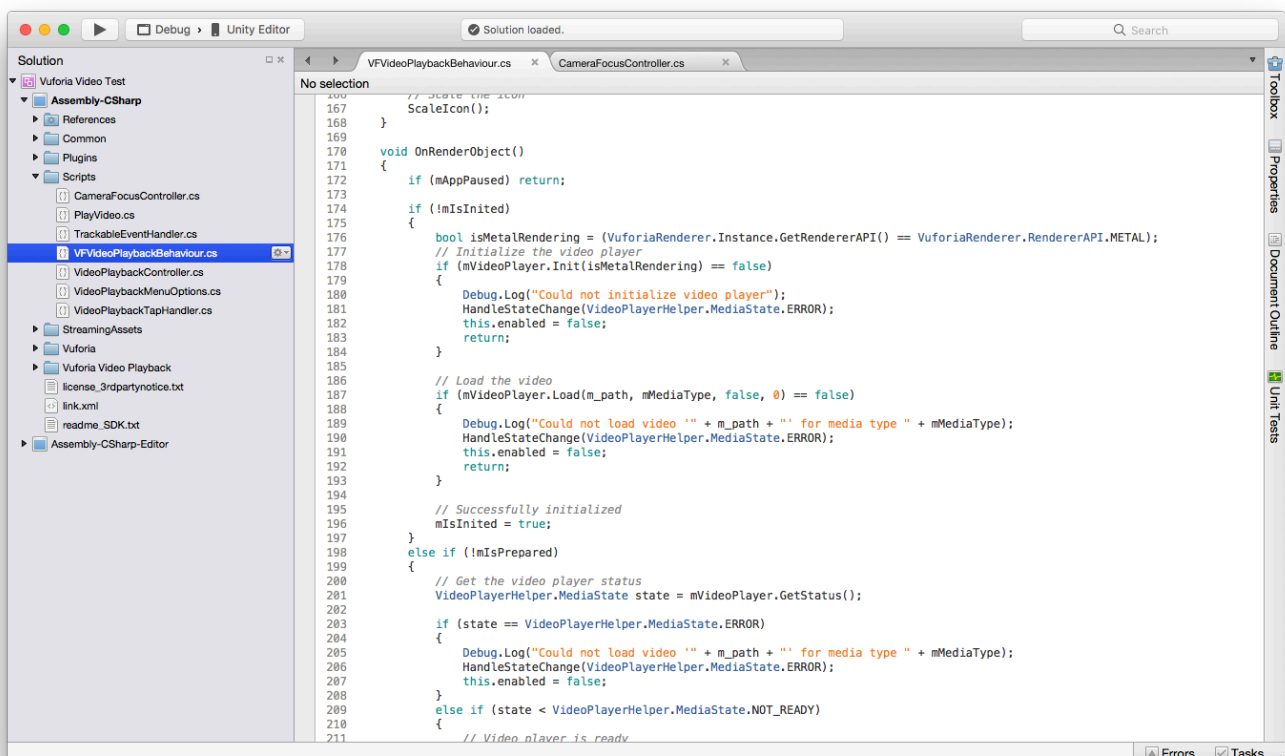
Using Vuforia's video object I was able to produce a canvas that appeared above the target image, and after acquiring some basic C# skills added the ability to tap the video on-screen to toggle pause and play. The video object also included audio, which texture mapping alone in Unity did not.

## 3.2 Augmented Reality



◀ **Fig 22** Using the Vuforia SDK within the Unity IDE to augment a static image to a two-dimensional target - a train ticket.

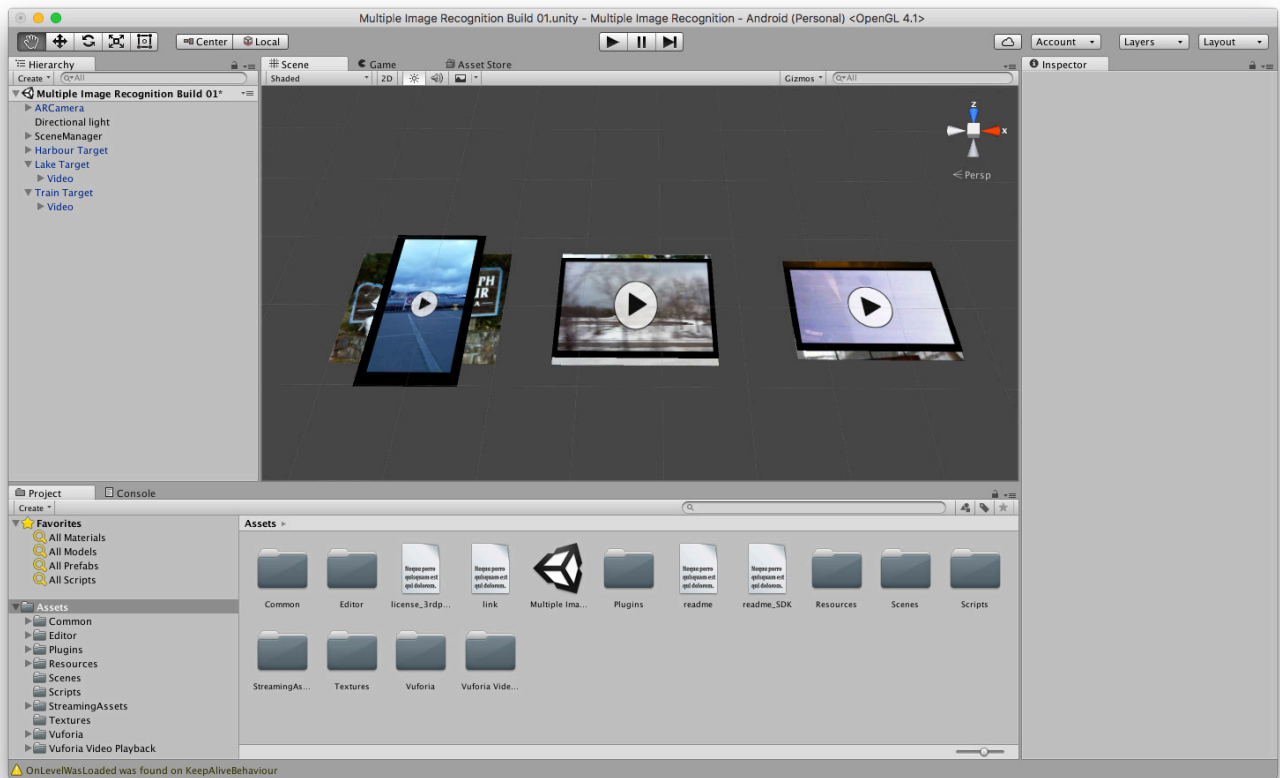
▶ **Fig 23** The resulting AR in action.



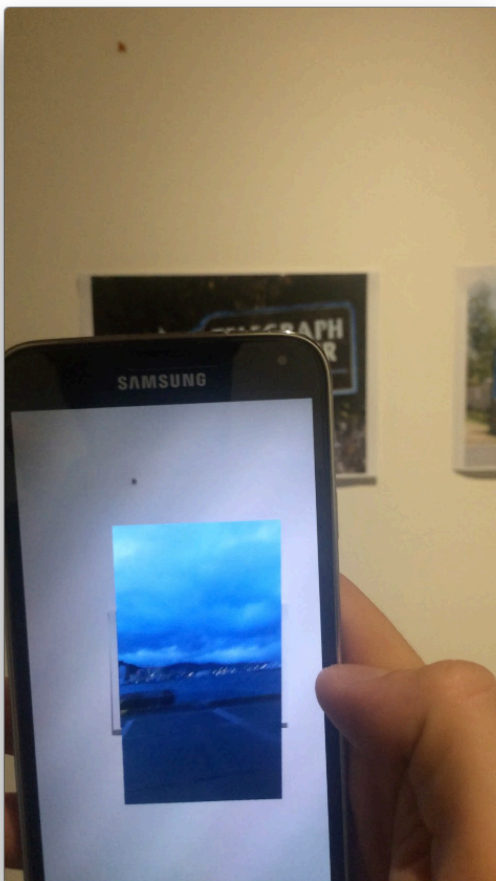
▲ **Fig 24** C# script in Unity's MonoEditor to handle video playback.



## 3.2 Augmented Reality



▲ Fig 25 Assigning video objects to two-dimensional image targets in Unity.



◀ Fig 26 The video object overlays an image target on a flat wall.

▲ Fig 27 When the video object floats out in space, as above the passport used as a target here, it looks strikingly unnatural.



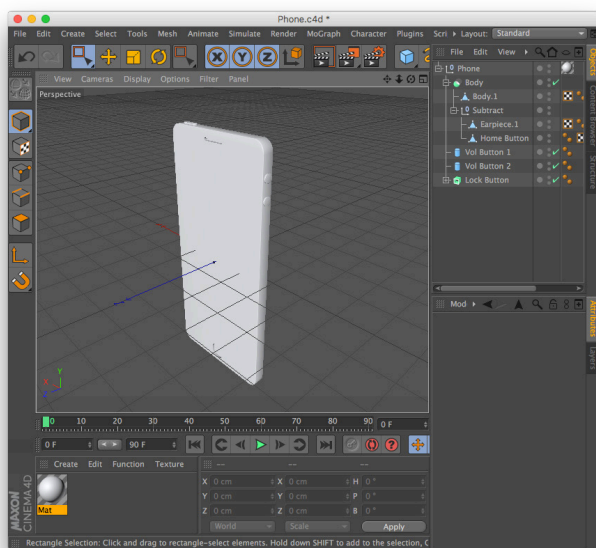
One of the key insights that arose from testing at this point was how unnatural it felt to encounter a piece of footage seemingly floating within the viewer's physical space. In initial tests I mounted the image target on a wall to emulate the use of a window as a target, which had the desired effect: the AR video replaced the target and appeared to co-inhabit the viewer's physical space. However this illusion was shattered when the video canvas was taken away from any existing surface to instead 'float' out in space.

The original concept for the AR documentary would see viewer-users repurpose windows as portals into peer-generated travel footage, but by this stage of the study I had re-evaluated and decided that rather than requiring audiences to seek out windows, the app should evoke content in any environment to encourage viewers to roam. Keeping the audience in a state of constant exploration creates a strong parallel between their live experience and that of the travel documentary. But in taking the video canvas 'off the wall' it appears to float in mid-air, violating the physics that we expect to apply to objects inhabiting our physical space. User testing showed that having a video element unbound from any window already present in the environment causes it to be interpreted as a second screen, thereby achieving the exact opposite of what I intended. Instead of bringing the footage into the world of the viewer, there now seemed to be a screen within a screen, introducing another layer of separation.

My next test attempted to combat this surrealism by introducing a 3D AR element that could act as an intermediary, brokering a relationship between the three-dimensional world of the viewer and the two-dimensional video canvas. Using Cinema 4D I modelled a basic smartphone shape, then combined it with the video object in Unity to produce an object that could more convincingly exist within the viewer's environment. The idea was to position viewers as walking in the footsteps of their peers, encountering smartphones seemingly recording footage in situ.



▲ Fig 28 The AR 'phone' floating above the passport used as an image target.



▲ Fig 29 Modelling the smartphone shape in Maxon Cinema 4D. A video canvas was then overlaid in Unity to provide the phone's 'screen'.

Further testing proved the futility of this approach. Test users responded enthusiastically to the appearance of a 3D AR model, but from a critical standpoint the use of AR was more of a distraction from the video than an aid to engagement. The test audience was captivated by the novelty of the medium, but not the content of the footage. Furthermore, the footage itself again seemed removed from the immediate environment and made poor use of the limited screen space available on contemporary mobile devices.

Having initially opted for an AR smartphone to naturalise the appearance of a video canvas within the viewer-user's own environment, it became apparent that generating a virtual phone was just doubling up on phones unnecessarily. There is no more direct way to incorporate video into the viewer's physical space than simply using the phone in their hand as the canvas, so rather than pushing the agenda for AR I conceded that it is best if video plays full-screen on the mobile device to make use of every available pixel. The challenge then, was to preserve the relationship between physical space and footage that had until now been managed by placing the one within the other. Taking my AR experimentation to its logical end, I trialled the use of markers in place of actual footage floating in the real-world space, indicating recognised objects that could evoke related video sequences. This way the content is understood to be born from the viewer's environment. Tapping one of the markers, shaped like a ticket, then prompts the associated video to play full-screen.

The marker approach worked as hoped, with trials showing that users did indeed perceive a strong link between their own world and that of the documentary. However to produce an object target based prototype required engineering a dirty workaround. The Vuforia SDK provides an option to use 3D objects as targets, but just like their 2D counterparts these target objects must be predetermined. Investigating other AR SDKs reveals the same caveat across the board. There is not yet any provision for dynamic object recognition: in fact any 3D target must be laboriously scanned and hard-coded into the application before it can be recognised. In order to run my test I actually took my laptop down to the coast, took a photo of the surrounding environment, then built the app then and there incorporating the photograph as an image target so the SDK could recognise the scene and display the AR content.



▲ **Fig 30** The interface recognises an island and the sea. Tapping either marker summons a video clip related to the keyword.



► **Fig 31** When users have to switch constantly between an AR camera view and full-screen content it is harder for a narrative to build up across multiple sequences.



Conducting this practical test also showed that it is important to consider emphases in the user experience when AR is used in an i-doc. Visually augmenting the footage within real-world space has been demonstrated to heighten a sense of separation,

but moving the footage behind markers that must be found and tapped turns the experience into more of a treasure hunt than a documentary. Invariably viewers spend more time spent looking for 'tickets' than actually watching the footage. The tension between live camera stream and documentary footage must be carefully managed.

The continuity of narrative was another concern. With viewers required to scan their environment to evoke each clip, the viewing experience was constantly interrupted. It is important that the Viewfinders user interface facilitates a seamless transition between sampling and playback, to promote immersive viewing.

It is interesting to note that since conducting my AR tests in November 2016, Google has announced a new product that looks to overcome limitations around dynamic object recognition. I believe that Google Lens, announced May 2017 (Pichai, 2017) though yet to be released, is just the start of a new wave of experiential media. Mainstream markerless AR technology will exert a considerable influence across the mobile mediascape and beyond. Their present emergent status puts them regrettably beyond the scope of the my own practice for this research project, but the probable launch of a Google Lens API will no doubt bring great opportunities to creative practitioners.

Another major disruption is likely to come from Apple. The technology giant has so far remained uncharacteristically absent from the field of AR, but their 2015 acquisition of Metaio sparked much speculation given Metaio had until that point been one of the leading contenders for AR SDK supremacy (Xu, 2016). The iPhone has already proven itself a reckonable force within the field of travel film (Cammaer, 2015a), and with last year's iPhone 7 Plus confirming Apple's development of dual-rear cameras it is probable that the company is working on camera software that exploits stereoscopic data and depth-mapping (Coldewey, 2016). Once these technologies are made available to Apple's dedicated user-base, shifts in creative practice are inevitable.



## AUGMENTED REALITY INSIGHTS

AR binds visual content to the viewer's physical environment to communicate location-specific meaning

The novelty appeal of AR must be managed to emphasise content over medium

Target-based AR requires predetermination and image-recognition capabilities of AR SDKs are limited

Fullscreen device video is preferable to an AR video canvas: viewers prefer video as a handheld object, not an object within a handheld object

Constantly switching between live AR camera stream and documentary footage perforates the viewing experience

The field of AR is still rapidly evolving and continued disruptions are likely to accompany future commercial releases

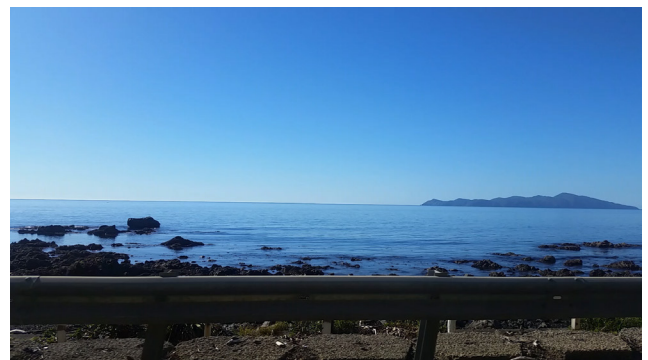
### 3.3 Computer Vision

In 2003 then-photojournalist Evan Nisselson published an article in *The Digital Journalist* that predicted *“The process of visually communicating is in for a drastic shift due to the arrival of cell phone cameras”*. Fourteen years later smartphones are producing films and Nisselson is a digital media technologist. After witnessing his cameraphone revolution play out, he now points to advancements in computer vision technology as heralding *“tremendous new opportunities in analyzing images that... [will impact] every business vertical, from automotive to advertising to augmented reality”* (Nisselson, 2015). Specifically, the improvement of machine learning techniques has made it possible for technology giants like Google and Facebook to train artificial intelligence (henceforth ‘AI’) entities for accurate image analysis. Computer vision takes enormous amounts of data to achieve. For those few companies with the resources to perform such a task, they are able to feed in different datasets to produce image recognition algorithms tailored to certain tasks, then offer access to third parties via an API as a paid service. Image recognition provider Clarifai defines three categories of image recognition: General, domain, and custom (Liu, 2016).

**“The only way computers can accurately identify cats in photos is because they have already learned what cats look like by analyzing millions of pictures tagged with the word ‘cat’”**

– Weiner, 2016.

► **Fig 32** I used a photograph of New Zealand’s coastal State Highway One to test three different computer vision services.



General recognition is useful for broad analyses, recognising high-level features in an image across any number of themes. A photo of a coastal landscape might produce words like ‘water’, ‘scenic’, ‘travel’, and ‘no person’. A domain model specialising in landscapes however would look for keywords relating to that theme, refining results to report ‘ocean’, ‘clear skies’, ‘island’, and ‘sun’, without checking for general purpose features like people or text content. Google, IBM, and Microsoft all have products aligned with these approaches, while Clarifai prides itself on capturing the custom recognition market, producing highly case-specific custom models. *“When you really get down to AI understanding the world in*



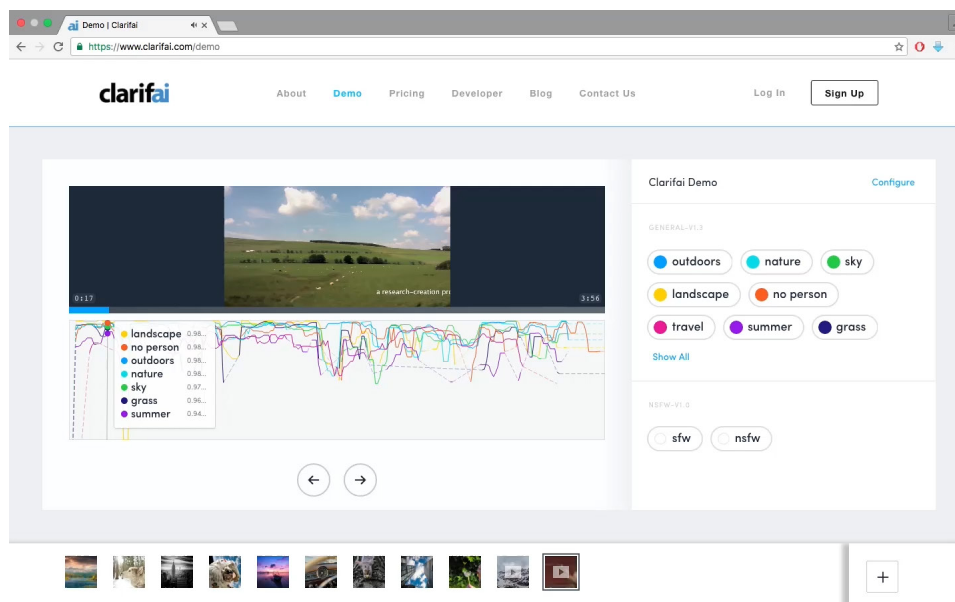
a way that is most useful and relevant to you as an individual and unique human, you'll want a custom model" (Liu, 2016). An example of this might be feeding the same coastal landscape photo into a custom New Zealand tourism image recognition algorithm, which could feasibly report keywords as specific as 'Kapiti Island' and 'State Highway 1'.

### 3.3.1 Video Recognition

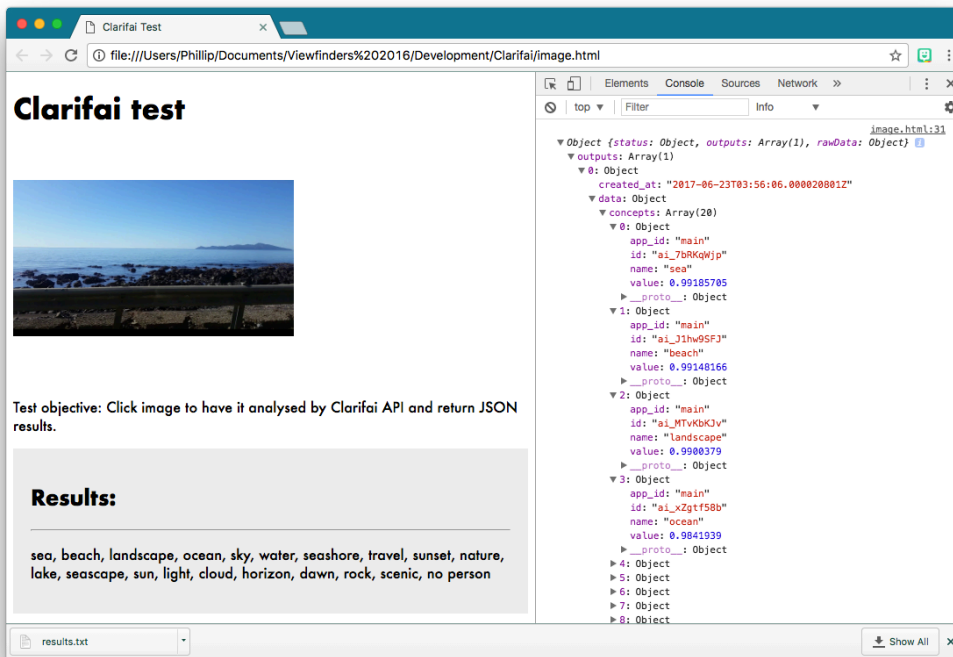
My first computer vision tests looked at how the technology applies to video. Despite having only recently been applied to static images successfully, it has not taken long for computer vision pioneers to adapt it for video applications. It is still early days though: Clarifai launched support for video recognition through their existing API in January 2015 (Clarifai, 2015). Meanwhile Microsoft's dedicated video API released in July 2016 (Yokel, 2016) is currently still in beta, and Google has only just joined the race, releasing a beta in March 2017 (Li, 2017). There are other video recognition APIs as well but I have focused my investigation on these frontrunners. All computer vision algorithms produce a short delay as the image is processed. This is not an issue when handling static images, however it becomes a problem when dealing with a time-sensitive medium like video.

### 3.3.2 Clarifai

The Clarifai website provides a demonstration page where users can upload their own photos and videos.



◀ **Fig 33** A video clip that has been processed by the Clarifai 'general' API on their website. Resulting keywords appear on the right.

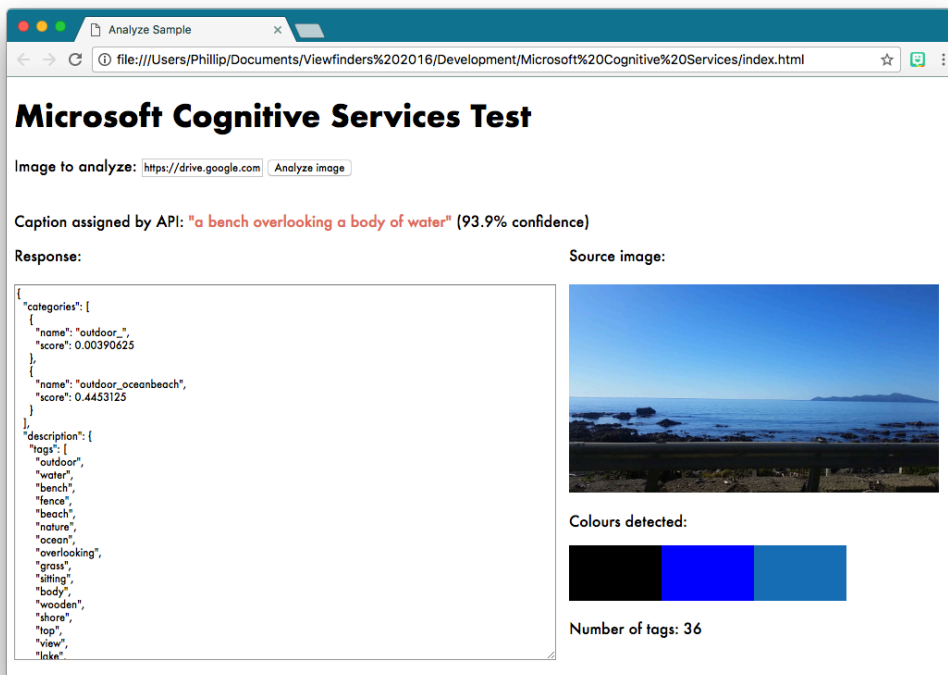


◀ **Fig 34** I produced a basic web interface to send a single frame of footage for Clarifai analysis. The JSON that was returned can be seen on the right in the javascript console, and has also been parsed to extract the keywords, seen at the base of the GUI under 'Results'.

Clarifai handles video through essentially the same process as they do still image: upload to service; process; return resulting keywords in JSON format. However videos are slower to upload, and the time that it takes a video to be processed in full is much longer than that of a single frame. Despite advertising their ability to process videos “at 10x faster than real-time” (Clarifai, 2015), what this actually means is that a 60 second video will take six full seconds to process before any results are returned, and the upload time for a video file further extends the delay. For an i-doc hoping to use this technology the implication is that videos must be stored in an online location that the API can access, to avoid unnecessary upload delays. Secondly, all image recognition functions must occur when the footage is first submitted, not at the time of playback.

Since Viewfinders seeks to interpret a live feed of the audience’s surroundings however, processing video in advance is not an option. I arranged a meeting with a Clarifai representative in New York who told me that a good workaround is to avoid analysing video files altogether and instead design an application that retrieves a single frame each second and submits it as a static image instead. This is usually sufficient to achieve accuracy without demanding too much from the API, the device, or the Internet connection.

Clarifai’s ‘General’ model claims to recognise “over 11,000 different concepts including objects, themes, moods, and more” (Clarifai, 2016), and in my tests proved very accurate, reporting keywords, or ‘concepts’, including ‘sea’, ‘beach’, ‘landscape’, and ‘ocean’ when supplied a frame from a tracking shot I filmed along the Kapiti Coast. A ‘confidence’ score is a metric used by all of the computer vision APIs that I tested to gauge the accuracy of any given keyword: in my tests Clarifai consistently returned 20 keywords per analysis, with confidence averaging above 0.8 meaning that all keywords returned were very accurate.

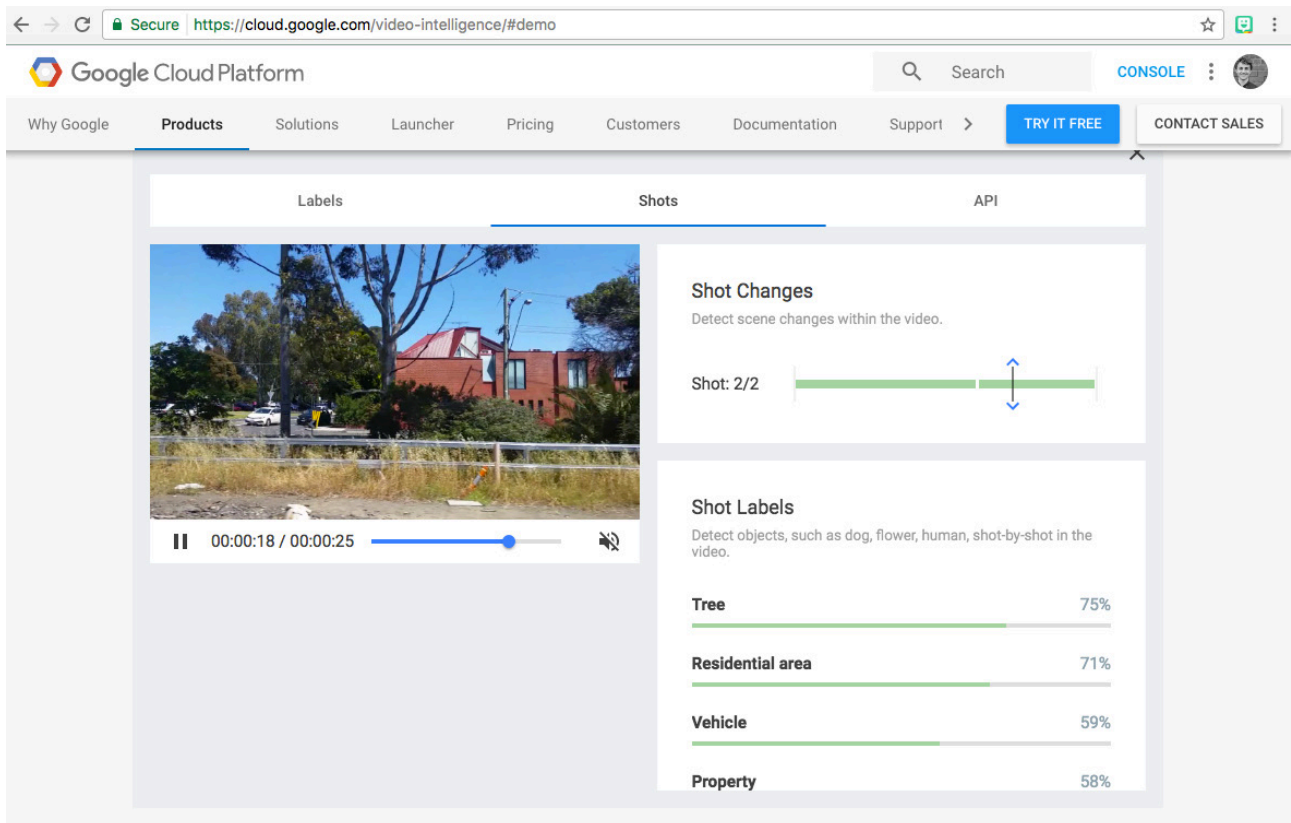


◀ **Fig 35** Microsoft Cognitive Services returns a large number of high-level keywords and produces an image caption based on the overall theme.

### 3.3.3 Microsoft Cognitive Services

Unlike Clarifai Microsoft handles video recognition through a dedicated API. Still in beta, the service offers algorithms for facial tracking, motion detection, video stabilisation, and automated thumbnail generation (Yokel, 2016). Any of these services could bring something new to mobile filmmaking. For example facial tracking may address ethical dilemmas inherent to the tourist gaze (see Cammaer, 2015a) by automatically tracking and blurring faces in public areas. Motion detection could influence the popularity of the close-up 'selfie', allowing filmmakers to move away from the camera and begin filming only when action is detected within the scene. Already apparent in mobile filmmaking, video stabilisation appears set to reach new heights, ushering in a new age for travelling shot montages where footage filmed even on a tuk tuk can match the smooth consistency of a train. The benefits of automated thumbnail generation are less immediately obvious, but will aid the dissemination of amateur film through social media channels, as well as having probably implications for database cinema as artificial intelligence gets better at interpreting footage to isolate important sequences.

The Computer Vision API from Microsoft Cognitive Services (henceforth 'MCV') takes a more general-use approach than Clarifai. Rather than assigning a confidence score to individual keywords, MCV returns a full list, then uses its AI to compose a caption from this list. It is this caption that is then assigned a confidence score. In this test, a frame of the Kapiti Coast prompted the caption *"a bench overlooking a body of water"*, with a confidence of 93.9%. MCV also determined the frame's category (*"outdoor\_oceanbeach"*), and reported the dominant colours (*"black"*, *"blue"*, and *"#1C71AF"*).

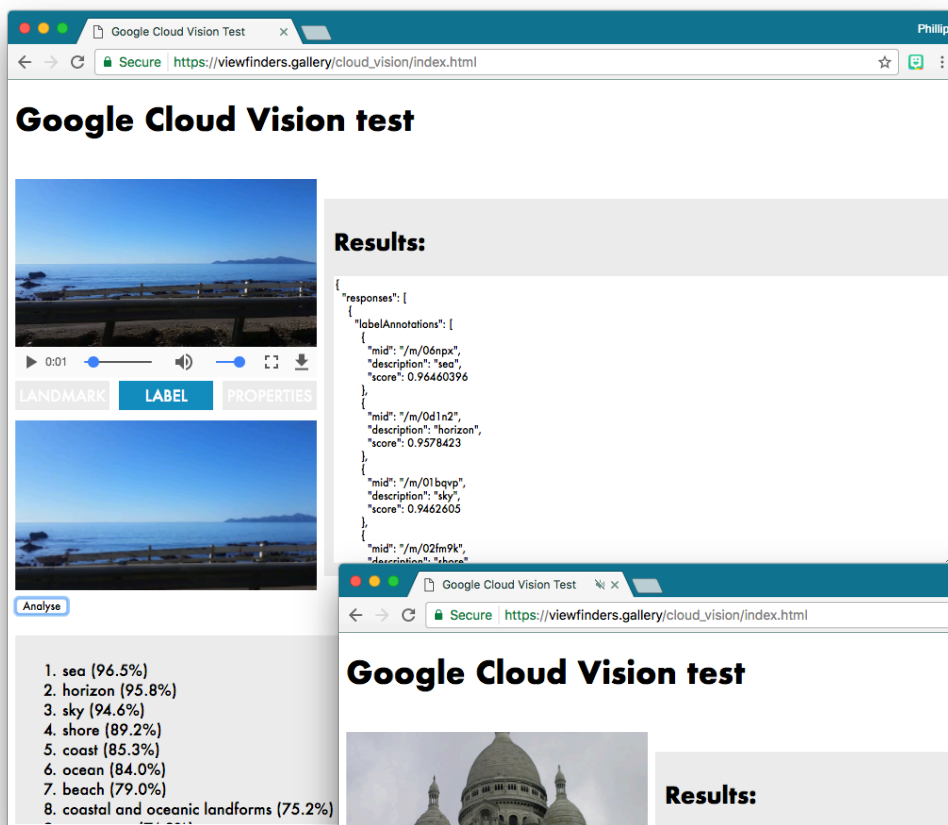


▲ **Fig 36** A travelling shot processed on the Google Cloud Platform website. Note the shot change registered slightly after the halfway point - when the train enters a tunnel.

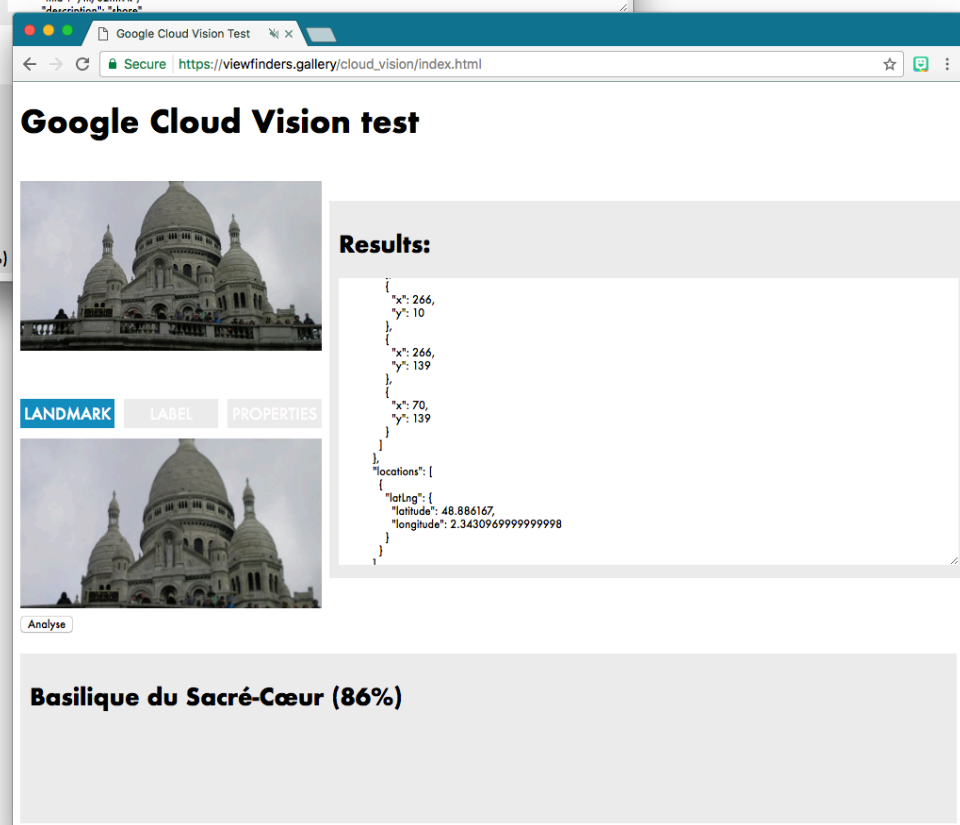
### 3.3.4 Google Cloud Vision

Presumably due to its extremely nascent status, of the computer vision APIs that I tested Google's Cloud Vision was the least accessible. The '*Video Intelligence*' beta API (henceforth 'GVI') is currently only accessible through a proprietary Command Line Interface, requiring prospective users to download dedicated software and execute tasks through typed commands. There is however a working demonstration of the GVI on Google's Cloud Platform website. Registered Cloud Platform members can upload a video to the site for analysis, and results are converted from JSON into a user-friendly visual representation.

Like Clarifai, GVI assigns individual keywords a confidence score, but it also goes a step further and registers when there is a substantial change in visual content, defining a 'shot change' at the appropriate timecode. This is comparable to Microsoft's automated thumbnail generation in that it indicates a shift towards fully automated editing processes. At each change in shot, the new content is submitted for image recognition, associating different keywords with different sections of video. This compromises slightly on accuracy since any keyword identified is only present for part of that shot, but in refraining from analysing content once per second as Clarifai's API does, the analysis is completed far quicker, and the video is also divided into clear thematic sections which could in itself have future benefits for filmmakers. Given GVI's pre-release status and the fact that shot determination is currently their sole service, I suspect Google have a full suite of complimentary functions in development.



► **Fig 38** The Cloud Vision API identifies landmarks and can combine this with coordinate data to locate media, as seen at the bottom of the results. This has obvious benefits for travel filmmakers.

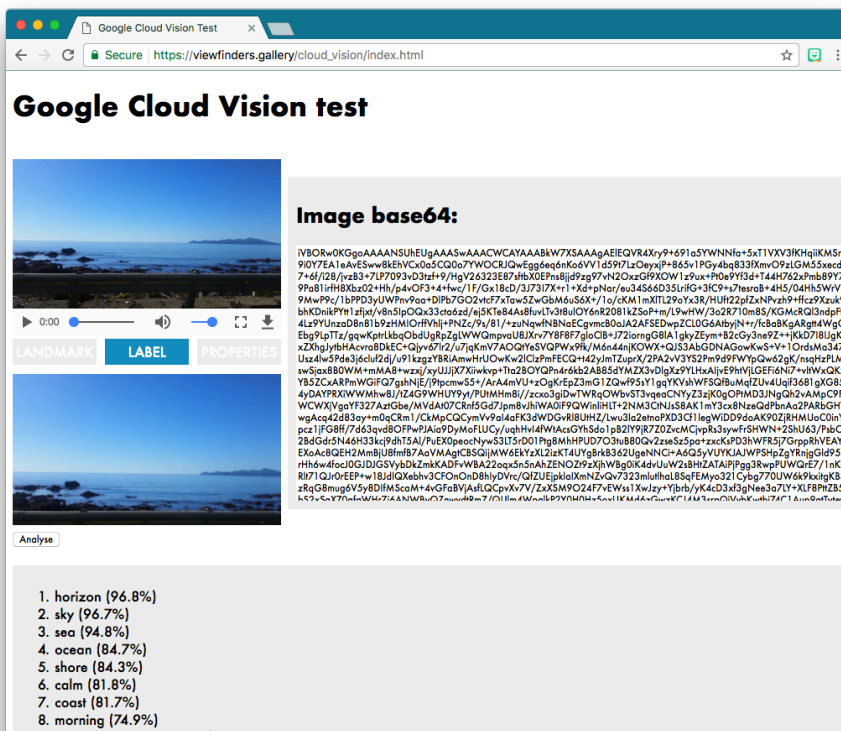


► **Fig 37** The Google Cloud Vision API returns fewer keywords than the other computer vision APIs tested.

I developed another web app interface as before to trial Google's Cloud Vision static image API on individual frames of footage to compare with my earlier Clarifai and Microsoft tests. Google's Cloud Vision services include facial recognition; landmark detection; brand and text recognition; and a unique option to retrieve visually-similar images, an enterprise afforded by their dominance of search engine architecture. An ability to identify landmarks has obvious benefits for travel filmmaking, and like Google Maps is a service enabled by the company's privileged position as gatekeeper to a formidable wealth of data.



The myriad functions made possible through combining Google's services hold real potential for travel filmmakers: by analysing a photograph I took in Paris the Cloud Vision API was able to not only identify the name of the landmark but also retrieve the relevant coordinate data. There is clear potential demonstrated here for a platform that recognises the way location imbues media with meaning, through promoting the intrinsic relationship between context and subject. Paired with some custom code Google's landmark recognition could be used to assemble a playlist of one location through the eyes of many authors, or string together a sequence of shots by a single filmmaker to produce a highlights reel of all the landmarks visited in one journey - the 'collection' of landmarks as proof of visitation is common practice in the social media age (Cammaer, 2015a).



◀ Fig 39 One of the quirks of Google's nascent Cloud Vision API is that data to be analysed must be supplied in raw base64 format, unlike the other computer vision APIs tested which accepted a standard image file. On the right hand side of the GUI here I have exposed the base64 code that was submitted.

► **Fig 40** Below is a sample of the JSON response received from the Google Cloud Vision API when sent the base64 code for the test image (right).

```
{
  "responses": [
    {
      "labelAnnotations": [
        {
          "mid": "/m/06npx",
          "description": "sea",
          "score": 0.9753449
        },
        {
          "mid": "/m/01bqvp",
          "description": "sky",
          "score": 0.96112746
        },
        {
          "mid": "/m/03ktm1",
          "description": "body of water",
          "score": 0.93481535
        },
        {
          "mid": "/m/0d1n2",
          "description": "horizon",
          "score": 0.93224216
        },
        {
          "mid": "/m/01lxd",
          "description": "coast",
          "score": 0.8899417
        },
        {
          "mid": "/m/02fm9k",
          "description": "shore",
          "score": 0.885576
        },
        {
          "mid": "/m/0838f",
          "description": "water",
          "score": 0.84599936
        },
        {
          "mid": "/m/05kq4",
          "description": "ocean",
          "score": 0.8417544
        },
        {
          "mid": "/m/02c9_23",
          "description": "coastal and oceanic landforms",
          "score": 0.7725047
        },
        {
          "mid": "/m/0b3yr",
          "description": "beach",
          "score": 0.7515665
        },
        {
          "mid": "/m/02_ktj",
          "description": "morning",
          "score": 0.73273766
        }
      ]
    }
  ]
}
```



# Computer Vision API Performance Comparison



I analysed the same frame of coastal State Highway One with the general feature recognition models of all three computer vision APIs tested. The following is a summary of their respective performances.

## Keywords Identified

### Clarifai

sea  
beach  
landscape  
ocean  
sky  
water  
seashore  
travel  
sunset  
nature  
lake  
seascape  
sun  
light  
cloud  
horizon  
dawn  
rock  
scenic  
no person

### Microsoft

outdoor  
water  
bench  
fence  
beach  
nature  
ocean  
overlooking  
grass  
sitting  
body  
wooden  
shore  
top  
view  
lake  
park  
standing  
herd  
bird  
field  
people  
eating  
snow  
sandy  
grazing  
table  
hill  
large  
group  
covered  
sheep  
river  
boat  
man  
wave

### Google

sky  
sea  
horizon  
shore  
coast  
phenomenon  
panorama  
ocean  
coastal and oceanic  
landforms  
bay  
inlet  
calm  
beach

## Keyword Accuracy

### Clarifai

85%

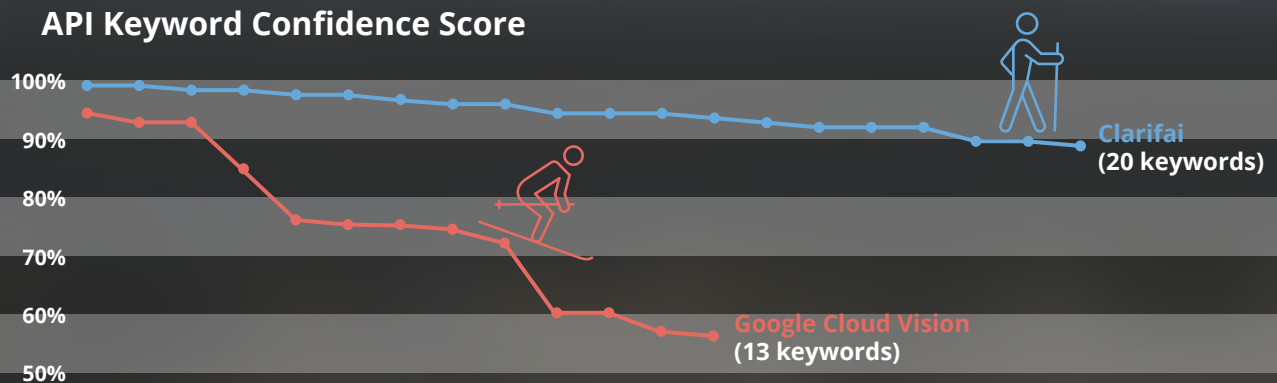
### Microsoft

36%

### Google

92%

## API Keyword Confidence Score



Note: Microsoft does not assign individual keywords a confidence score.



## COMPUTER VISION INSIGHTS

Continuing advances in machine learning are enabling ongoing development of new tools for moving image analysis

Democratic ideologies in many technology industry leaders are driving an API trend that exposes powerful capabilities to third-party developers

As computer vision technology spreads to a diverse array of innovators, the resulting applications will be equally diverse

Static image analysis has proven consistently accurate, though the associated processing delay remains a tension for time-sensitive operations

Dedicated video recognition services are in their infancy, with increasingly specialised functions likely still in development

Regularly sending single frames to a static image recognition API has proven preferable to using a video recognition API in the absence of true real-time processing

Shot isolation, facial tracking, stabilisation and motion detection have implications for algorithmic editing and video database navigation



# USER INTERFACE DESIGN

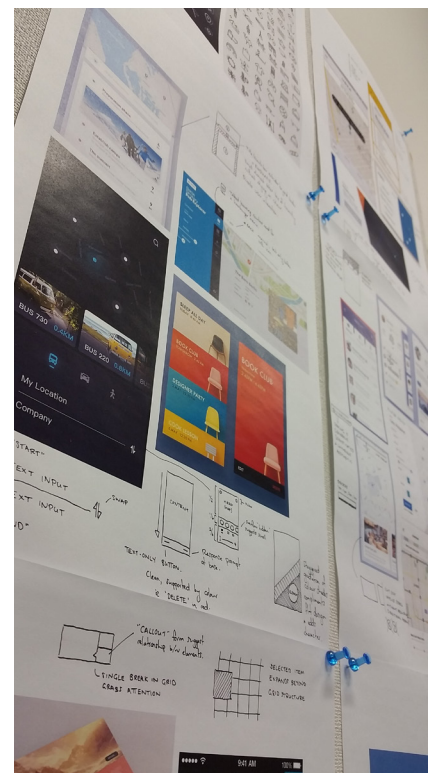


The intention for the Viewfinders practical output is to highlight some of the creative opportunities in store for amateur travel film. The AR app consolidates skills and insights arising from various explorative tests into a single demonstrable product. This app then functions as an access point to the peer-generated footage collected on the Viewfinders website, and facilitates an interactive and experiential mode of viewing that conveys the free-spirited context in which the video clips themselves were produced.

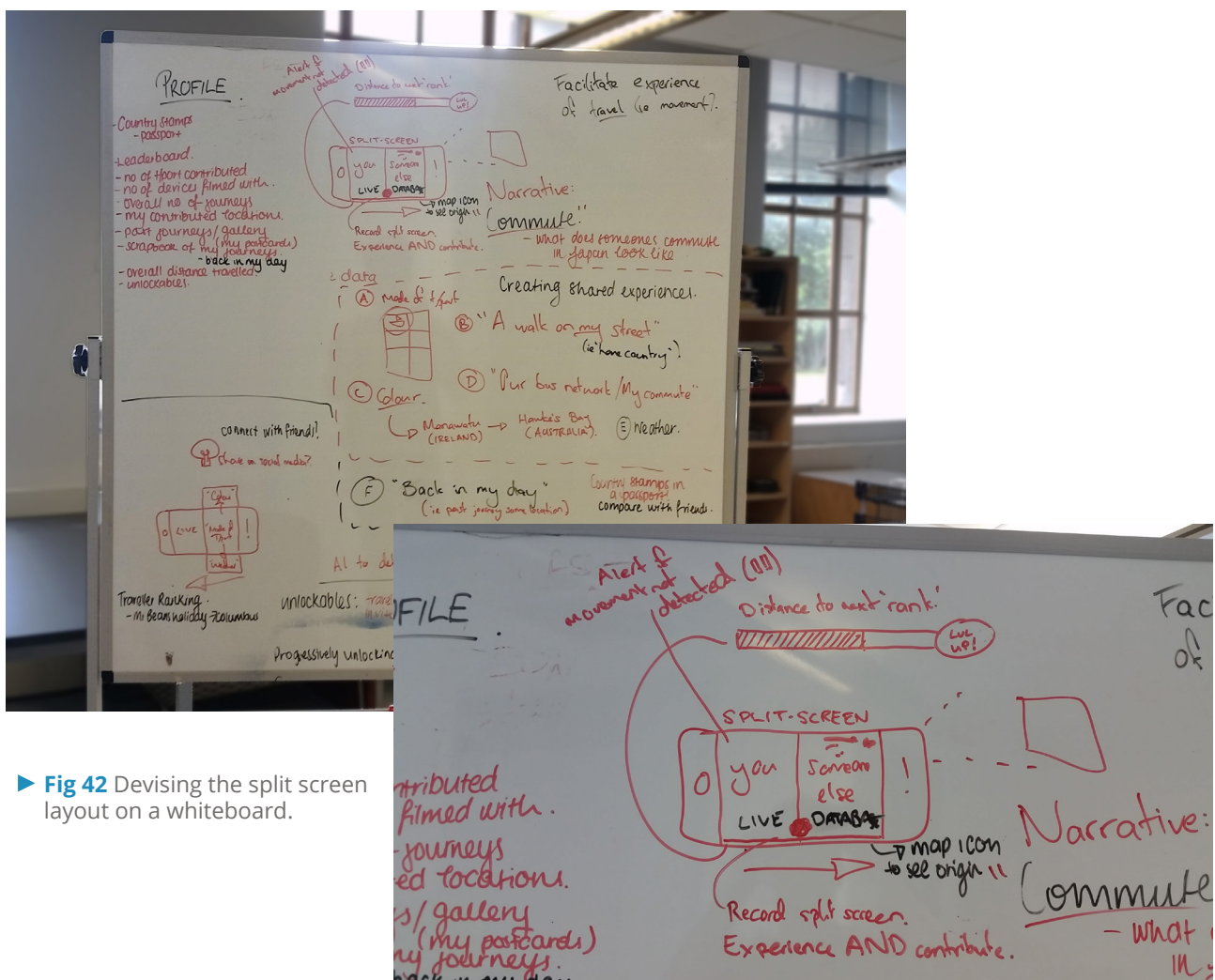
Note that in discussing this interface I use the term '**Scene**' to refer to peer-generated footage submitted to the Viewfinders video database, and '**Scene View**' in reference to a video frame within the interface that hosts these Scenes. The frame hosting the live camera stream on the device is referred to as the '**Live View**'.

My objective was to produce a narrative that relied on the viewer's physical movement in order to progress. By requiring the audience to travel through their own space, experiential factors become mutually apparent between the viewer's own world and that of the travel film, making the viewer feel as though they are inhabiting the footage. Having determined that visually overlaying footage onto a live camera-feed is an inefficient means of drawing the two worlds together, I instead opted to exploit computer vision APIs and engineer a kind of 'reverse AR'. Rather than conjuring supplementary information to amplify the user's experience of the physical world, the interface instead positions the real world as the augmentation, harnessing ambient phenomena to augment the i-doc. I chose computer vision technology because it had proven to be more capable at interpreting its surrounding than any AR APIs I tested; also, by using these interpretations to evoke corresponding Scenes as full-screen content, there is no requirement for any tracking or positioning service, rendering a dedicated AR API redundant. In effect, the interface executes an experiential Augmented Reality rather than the more typical visual AR.

► **Fig 41** Pinboarding is an important tool for exploring design precedents and establishing an overall visual mood for the interface.



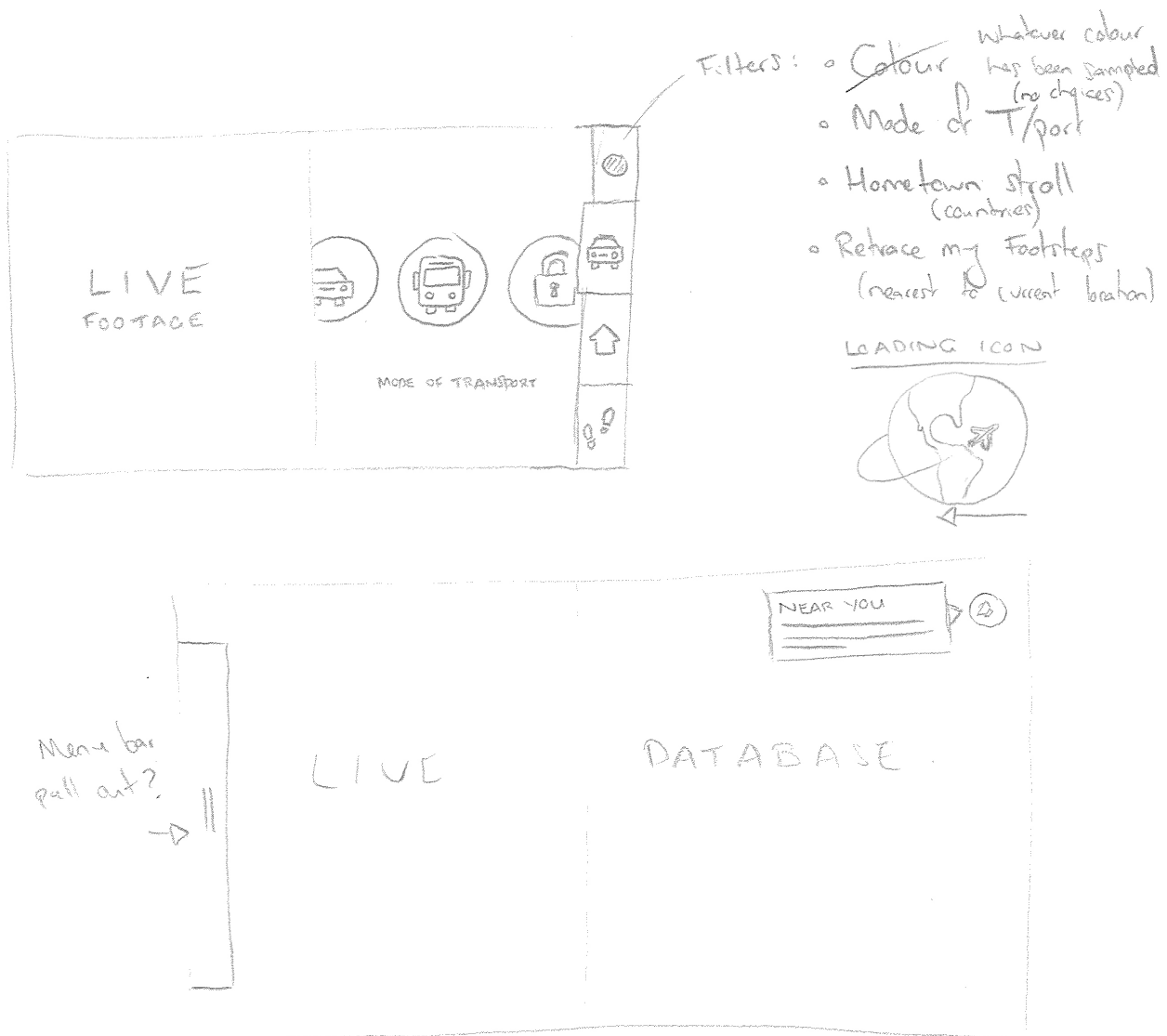
A split-screen layout was devised to combat the constant stop/start dynamic observed in earlier AR tests. Rather than switching constantly from the Live camera feed (used to sample the environment) to a full-screen video canvas playing the resulting Scene, viewer-users can handle both at once. There are additional benefits to this approach: parallels between the Scene and the viewer's own live experience can hardly be missed; further, by presenting production and consumption side-by-side, viewers are compelled to record and submit their own experience for others to encounter. The central division can be dragged left and right to dedicate more screen space to either the Live or Scene view as desired.



► Fig 42 Devising the split screen layout on a whiteboard.

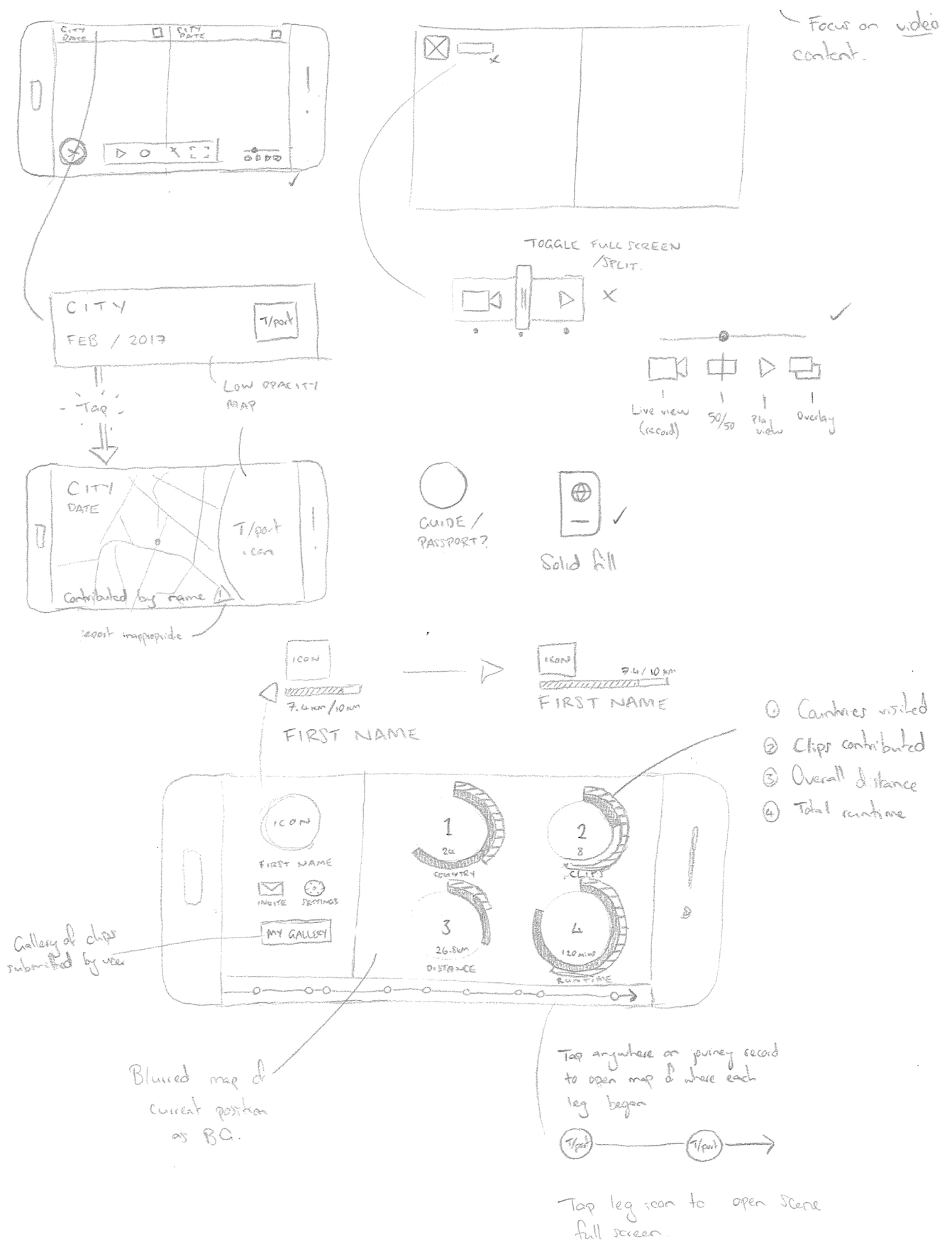
## 4.0 Wireframing

Wireframing is an important tool to crystallise thinking around functionality, usability, and overall aesthetic. The following is a selection of wireframes that illustrate progress from initial ideation through to a full interface.



▲ Fig 43 These rough early wireframes contain the seeds of the concept.

ALL CONTROLS DISAPPEAR AFTER THREE SECONDS.

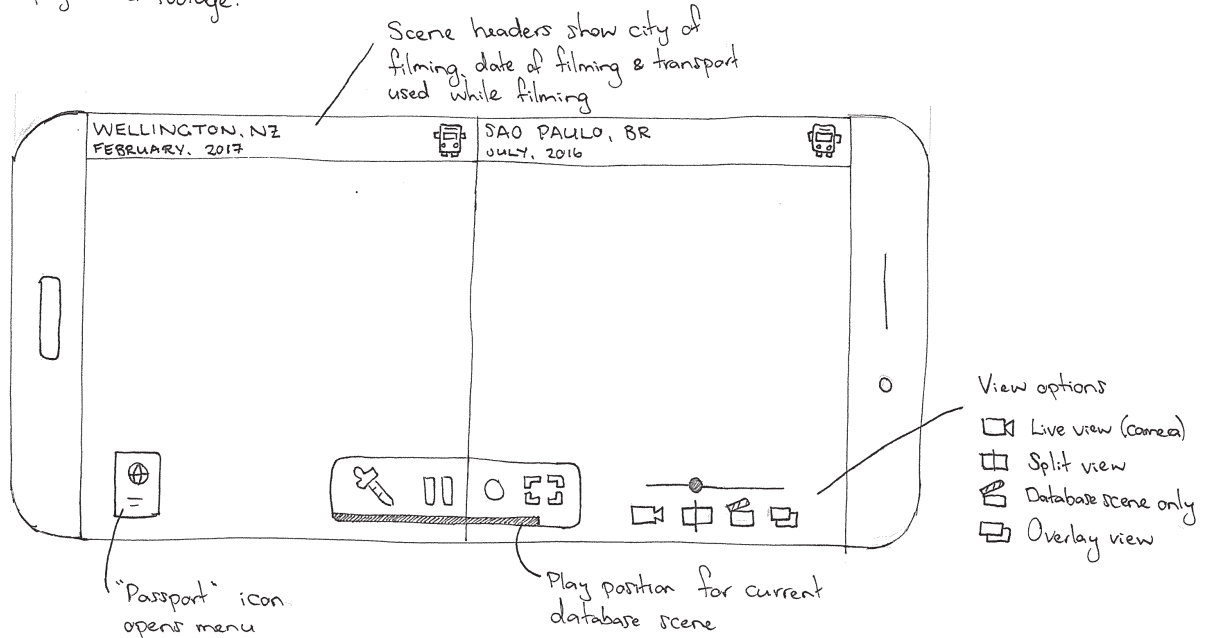


▲ **Fig 44** Elements are restructured to adjust emphases and guide interactions

## MAIN INTERFACE

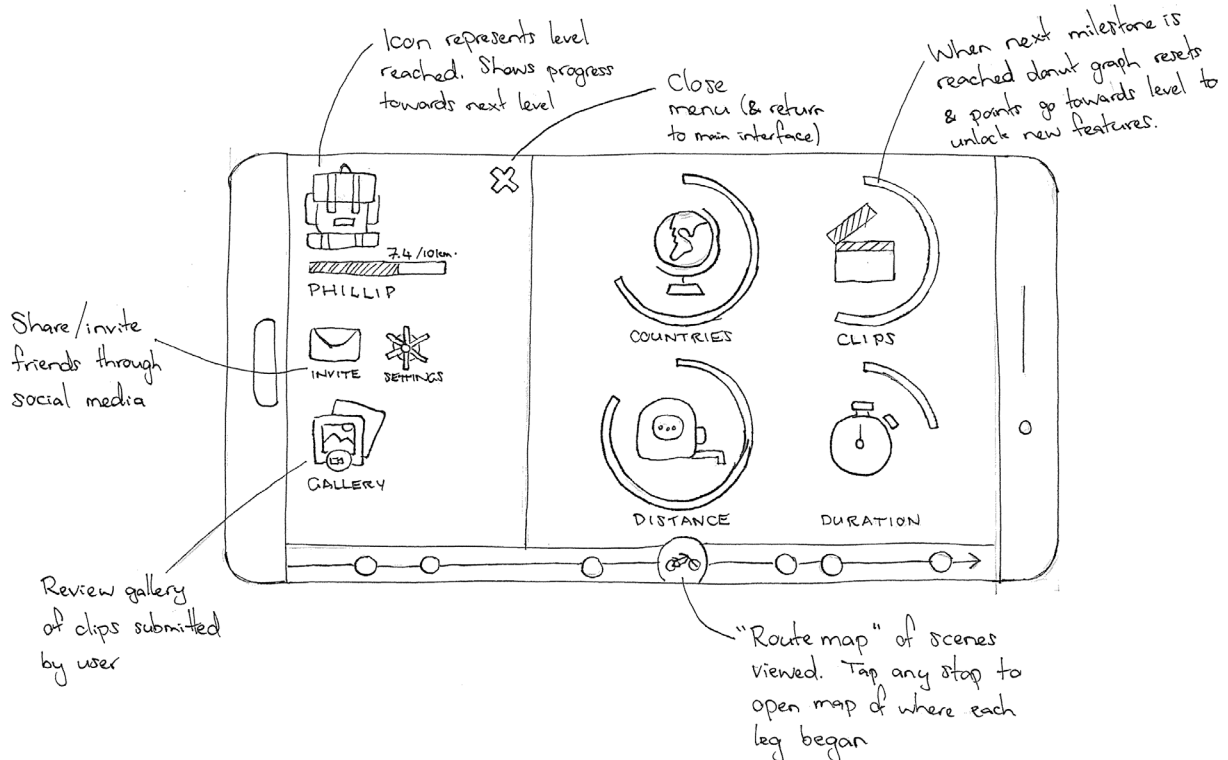
The "eyedropper" control uses the phone's camera to sample the physical environment & play related footage.

Pressing record button saves 50 seconds of footage



Pressing pause takes a still of the live footage to 'pause'.

## MENU

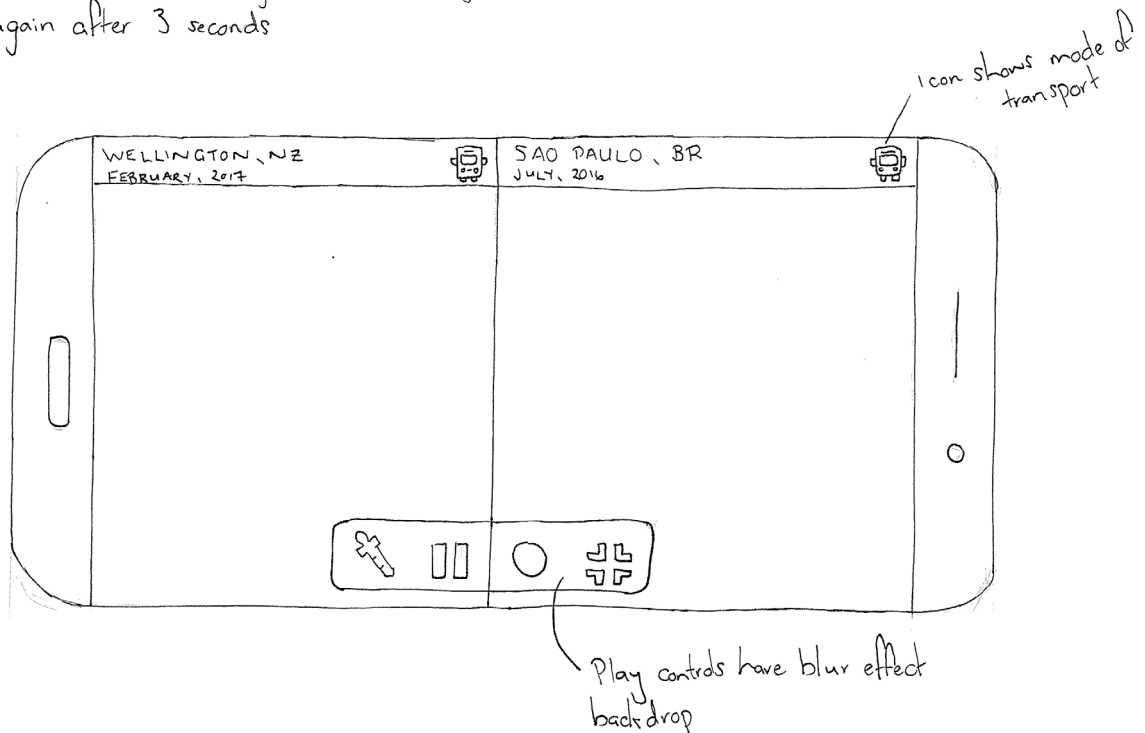


▲ Fig 45 The different pages of the web app bring components together under a common theme or task.

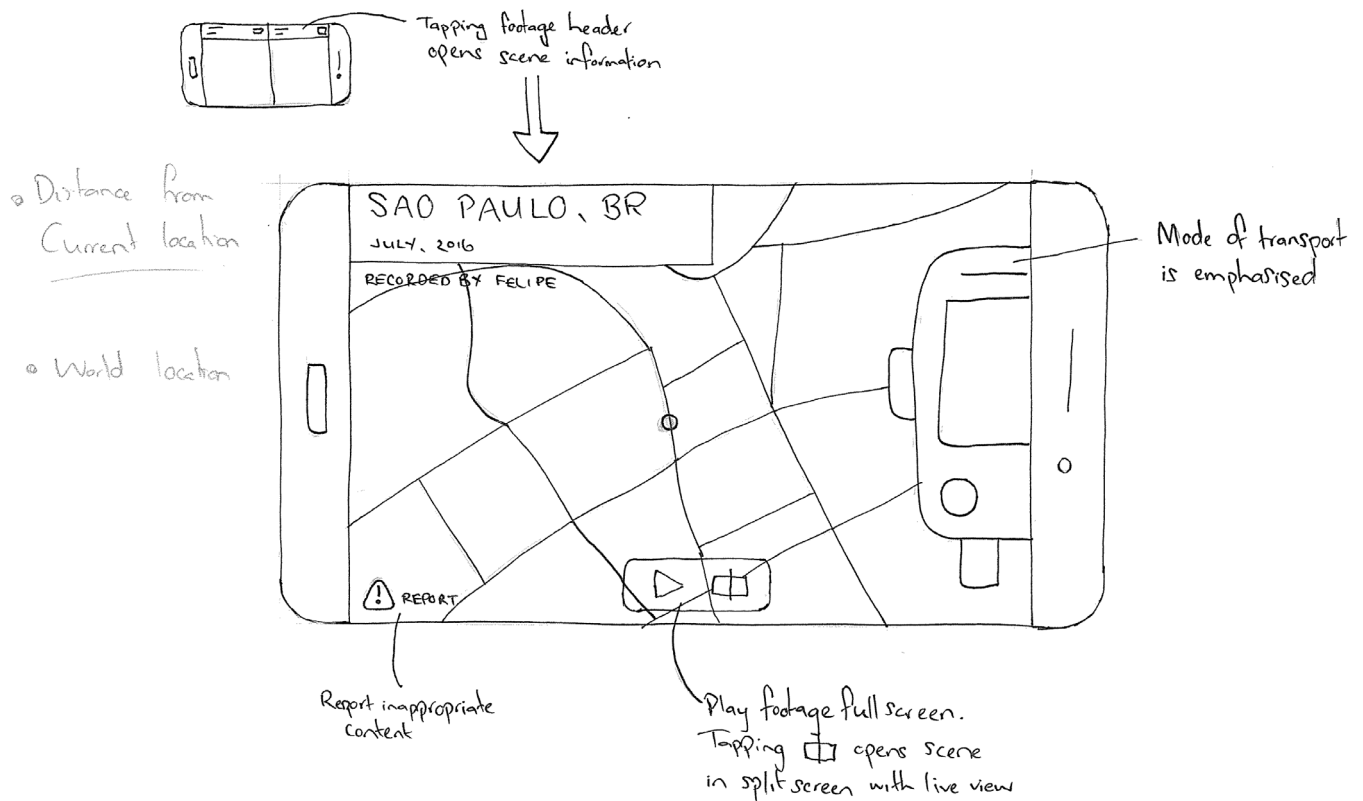


## FULL SCREEN

Tap screen to summon play controls/footage headers  
Fades out again after 3 seconds

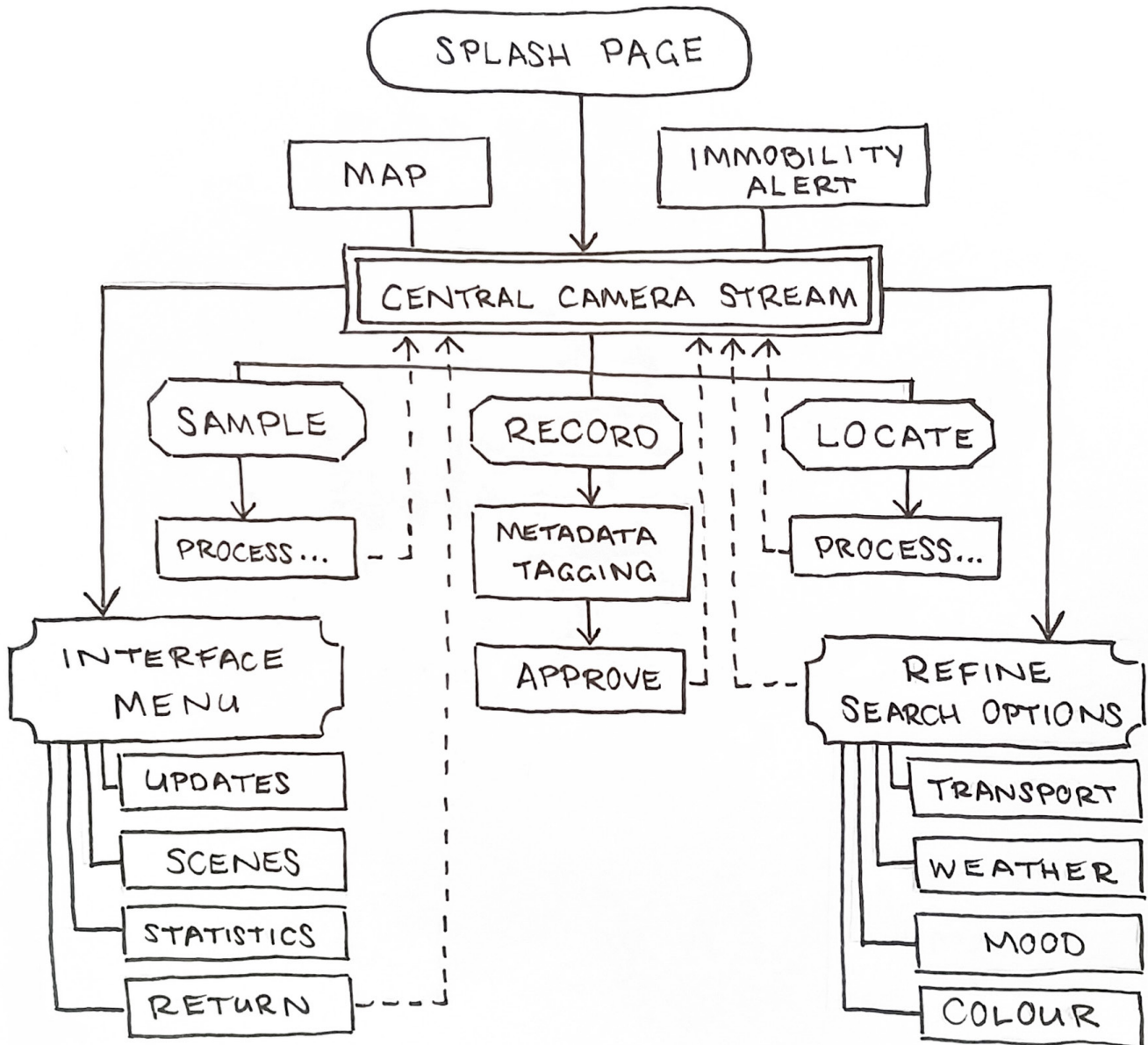


## SCENE INFORMATION



▲ **Fig 46** The travel theme is central to the aesthetic of the interface, and particularly evident in a preference for pan-lingual pictograms rather than text.

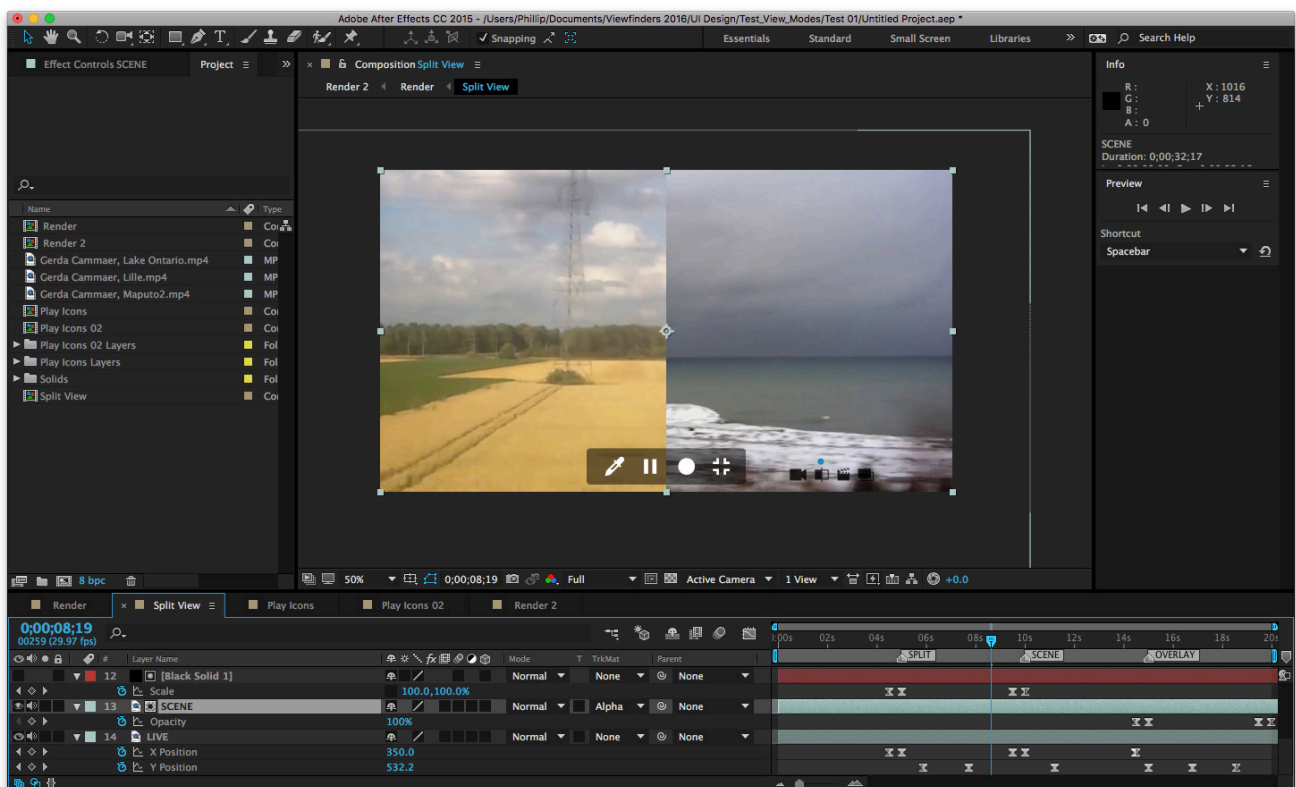
# NAVIGATIONAL STRUCTURE



▲ Fig 47 The web app is structured around the central Live camera stream

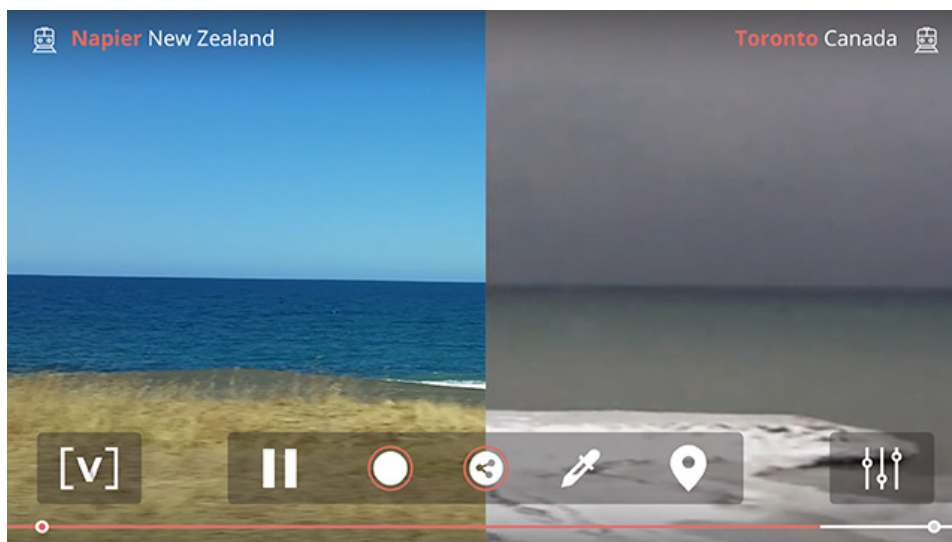
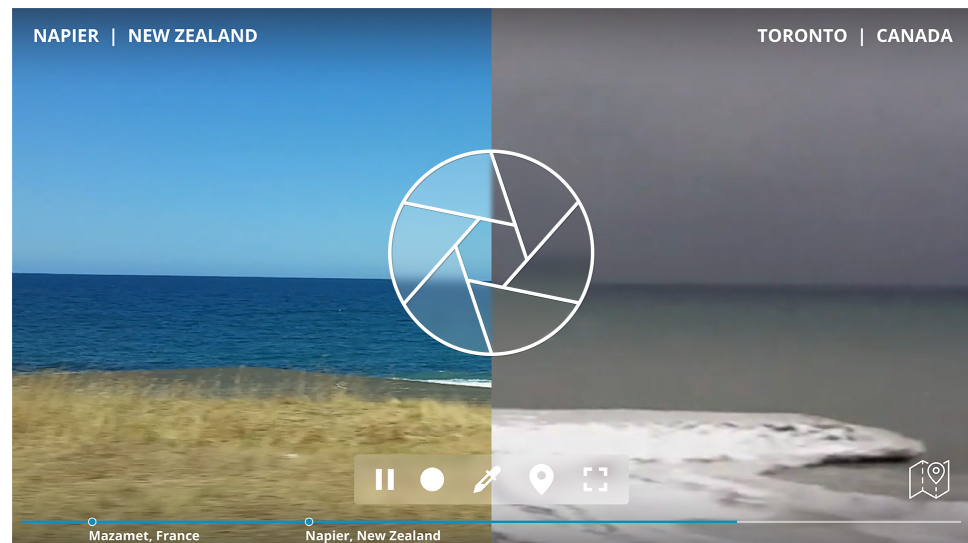
## 4.1 Digital Mockups

Progressing the designs from pencil and paper, these digital UI mock-ups amalgamate and refine the best concepts that emerged during wireframing. Once digitised, interactive elements can be transferred into prototyping software for usability testing; the app's architecture can be assessed; the visual language can be steered in a more cohesive direction; and various design iterations can be rapidly produced and critiqued.



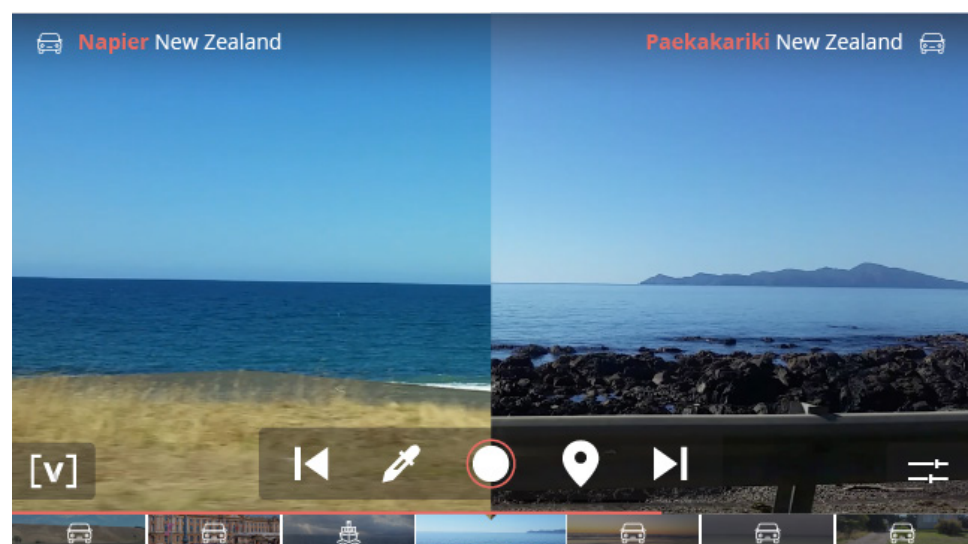
▲ **Fig 48** My first digital mock-up tested the presentation of two videos concurrently, using Adobe After Effects to render a sample where the clips appeared both side-by-side, and overlaid at 50% opacity. To give viewers maximum agency over the material I determined that the best approach was to preserve both options and allow users to change between them as a form of live editing. In later designs the user swipes vertically on the video to alternate between split-screen, overlay, and full-screen.

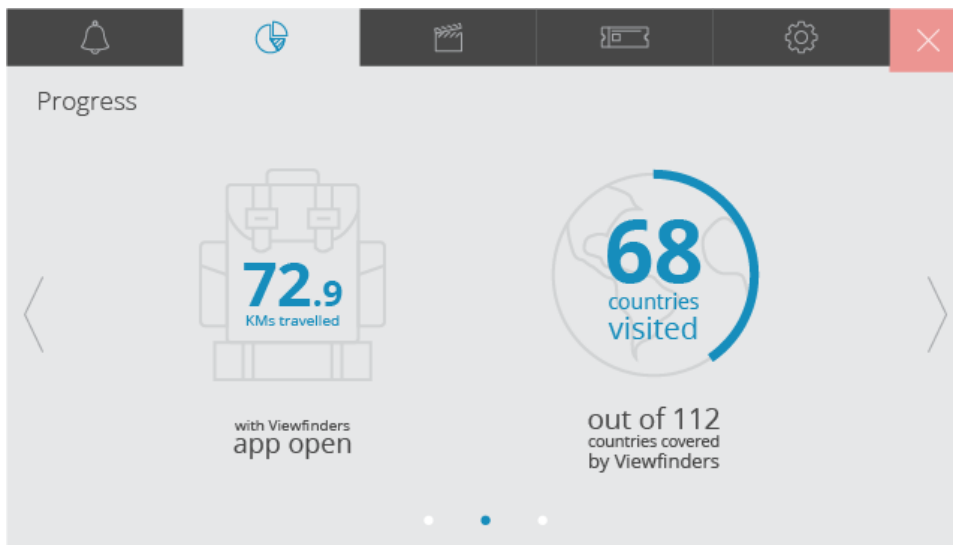
- **Fig 49** Using Adobe Illustrator I created a basic lens aperture form. During wireframing I had envisaged this animating to mimic a snapshot whenever the live camera stream was sampled, but testing proved the animation was too much of a delay.



- ◀ **Fig 50** I trialled a style reminiscent of a route map for the video progress bar, continuing the theme of stops on a journey into the header, announcing the name of the 'stop'.

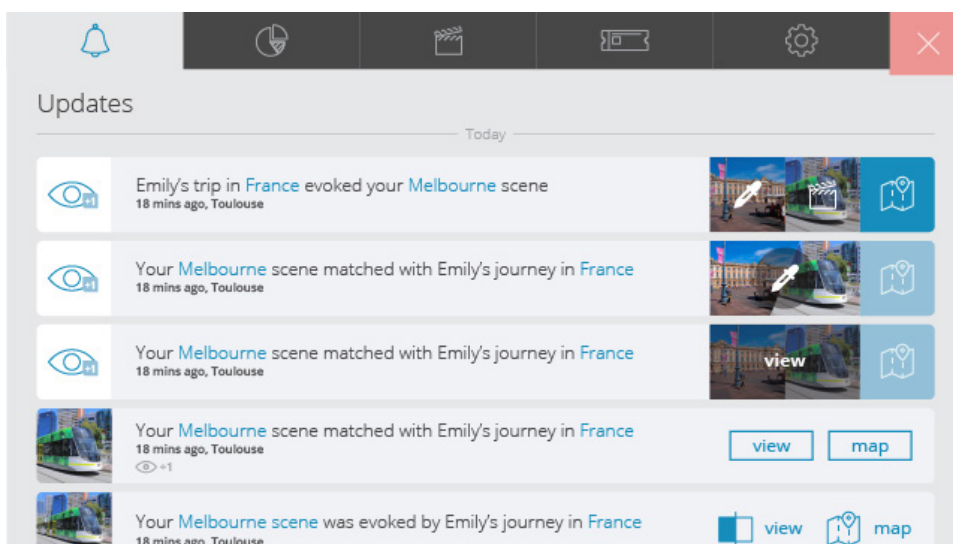
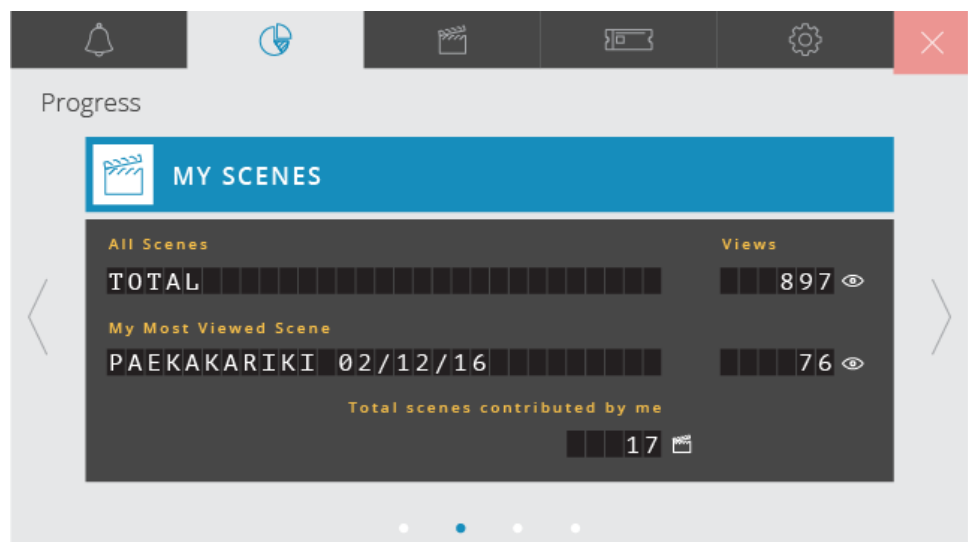
- **Fig 51** Initially I tried to limit the number of UI elements on the screen to conserve room for the footage. After developing an interactive prototype however the importance of visually representing the dynamic playlist became apparent. Without a fixed set of thumbnails on-screen the user is unable to see the impact of their interactions. Rather than limiting my use of space I resolved to simply have all UI elements auto-hide after 3 seconds instead.





◀ **Fig 52** Taking cues from other apps that rely on user generated content (eg. Google Street View), I considered ways to gamify user statistics and incentivise contribution. This aspect was scaled back eventually to ensure the Viewfinders app didn't lose its identity as an i-doc foremost.

▶ **Fig 53** I considered ways that participants could manage their own Scenes within the i-doc. This layout is inspired by arrivals boards to reinforce the theme of travel film.



◀ **Fig 54** Focus groups raised the importance to users of seeing how many times their own Scenes had been viewed by others, prompting the design of an up-to-the-minute log of views and where they had been evoked.



### 4.1.1 Significance of Playback Controls



The inclusion of an image recognition and a geolocation function alongside customary video controls is subtle, but the consequence is more significant than it appears. By housing these functions among the more familiar 'skip' and 'record' buttons, the viewer intuitively understands that although typically used for wayfinding, here these operations serve as a device for playback control. Their role quietly shifts, changing from tools that inform a human user to tools that inform a film narrative. Thus two new mechanisms enter the video lexicon, joining other post-internet inductees like 'share' and 'fullscreen'.

Those familiar with image editing software will know the eyedropper motif represents a 'sample' function, and the inverse-teardrop marker denotes location. When a user taps the sample button a frame is collected from their current live camera stream and analysed through a computer vision API. The resulting keywords inform the assembly of a playlist of related Scenes, queued in order of visual similarity. Note the absence of a play button: the only way for viewers to prompt playback is through physical exploration. The more diversity they encounter in their own journey, the more they will expose to the sample function, and thus the more diverse the resulting Scenes will be. Viewers are cast in an active role, empowered to translate local roaming into a global adventure.

Conversely the geolocation function tells the story of the local area. When tapped, the user's current location is calculated, then used to compile a sequence of Scenes filmed nearby in order of proximity. Hypothetically, the more Scenes contributed to the database, the more likely it becomes to encounter a Scene filmed very near - or even exactly at - the location the viewer currently finds themselves. This shifts the relationship between the viewer and the content to one of juxtaposition. Rather than ambient factors emulating the conditions of a depicted faraway Scene, the viewer is presented with another person's experience of the same location. For example, if the user is a Parisian on their Winter commute past the Arc de Triomphe, the Viewfinders geolocation function might offer a window to see the monument in Summer; or at a different time of day; in different traffic conditions; or simply through the entranced eyes of a tourist.

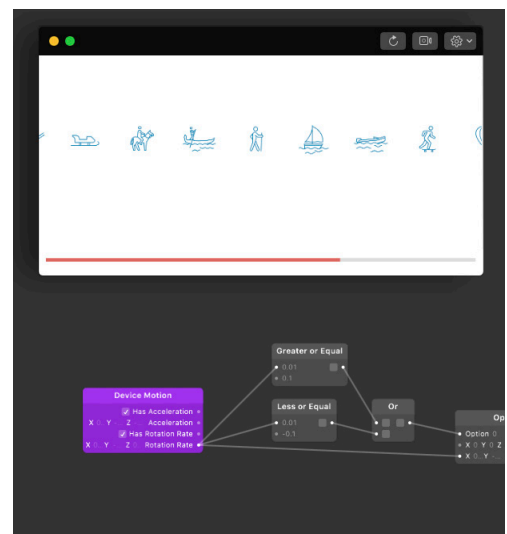
## 4.2 Interactive Prototyping

My UI design coincided with frequent interactive prototyping to enable a regular cycle of drafting, building, testing, and reviewing. Using tools such as Origami, XCode, and PhoneGap, I was able to challenge my own assumptions around how users would interact with the Viewfinders interface. I was also able to gauge how each feature contributed to the overall user experience, preventing the application from straying too far into the realm of social media or gaming. For example I had designed an entire system for users to compare their 'adventure' with that of others – the idea being that users would log-in via Facebook, then compare statistics with friends to claim titles based on distance travelled or Scenes submitted. However interactive prototyping revealed that rather than inspiring competition, the addition of the many pages of profiles, scoreboards, and notifications merely overwhelmed the single screen of actual documentary footage, and confused the message that the application is in fact a tool for experiential cinema.

Graphic elements were brought across from Sketch into Origami, then programmed to react to user interactions using Quartz Composer (hereafter 'QC'), a node based visual programming language. The nodes, or 'patches', contain preset segments of code that can be linked together to rapidly create prototypes. By editing and restructuring the patches a vast array of functions can be compiled. In this example (Fig 55) I produced a rudimentary icon scroller: the code accesses the device accelerometer so that the list of icons scrolls left and right as the user moves their phone left and right.

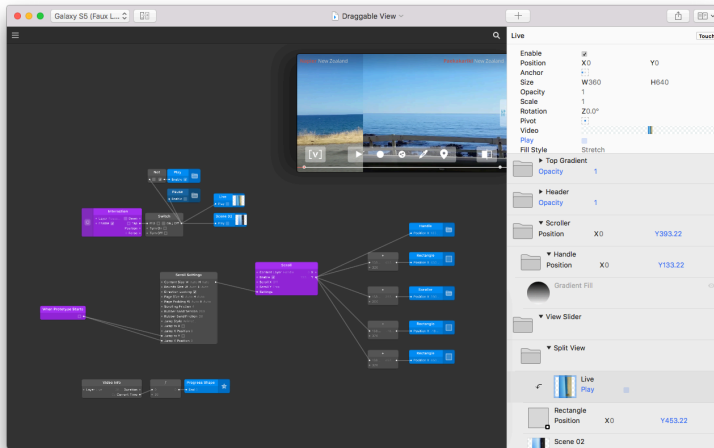
After installing this prototype on my phone for user testing, I found that users habitually try to scroll content with their finger without realising that moving their device has any effect. This was no surprise, but the real lesson was in users' reluctance to use the function at all, even after being instructed in its use. Testers instinctively held their smartphone at arms length to direct it left and right, and described the interaction as awkward and unnecessary. This discouraged me from using the accelerometer as an interface elsewhere in my app, however it could still prove practical in other applications.

For example an experiential mobile-mentary that focuses on a particular locale might direct viewers to look in a certain direction when they are detected at a specific set of coordinates, triggering footage filmed at that exact vantage point – and this data could



▲ Fig 55 Prototyping in Origami Studio.





▲ **Fig 56** In this prototype I discovered that the division between two tracking shots is interpreted more like a natural part of the footage when it can be manually dragged horizontally to keep time with the content.

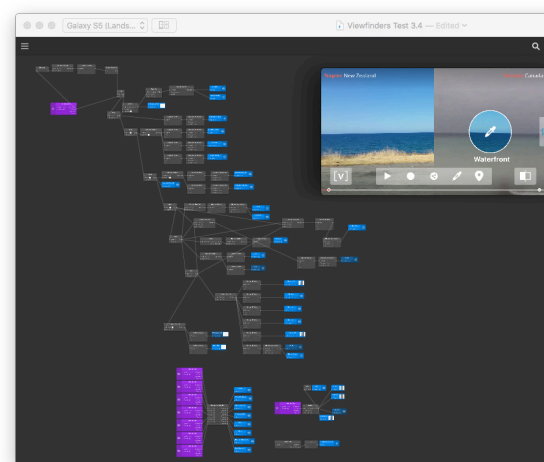
that detects when a user drags the central division between the two video canvases. The code tracks the movement of the user's finger, and resizes both video canvases to make the division appear to move. This feature performed well in user testing. Users were intrigued by the ability to dynamically match a wipe transition to the speed of movement shown in the footage, which was itself tracking horizontally.

By creating a prototype of the animation conceived to illustrate the sample function (Fig 57), I found that it successfully communicated what the function was doing. However in actual practice having a full four-second animation interrupted the viewing experience, and subsequent animations quickly became tedious. Note the number of QC patches required to generate this custom animation: a lens aperture animates open within the Live footage on the left, then 'rolls' across over the Scene footage on the right, taking the sample contents with it. Here a keyword is identified and appears below the circle of sampled footage, before the circle 'flips' to reveal the evoked footage as the new Scene. Later prototypes greatly simplified the visual representation of this function.

Most tests were run on an Android device, but I used Xcode to produce a prototype that could be installed and run natively on iOS. The default behaviour in iOS is for video to

certainly be captured at the time of recording. For a global travel documentary however, the audience cannot be expected to walk in the exact footsteps of another filmmaker.

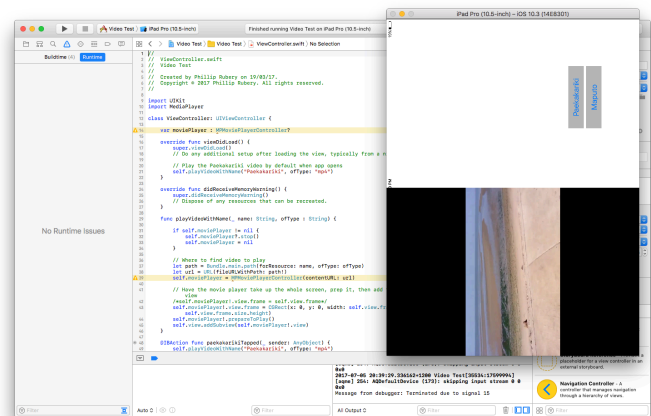
Prototypes produced in Origami Studio on the desktop environment are then run on mobile through the Origami Live companion app, to test performance in the app's native platform. In this screenshot (Fig 56) I have produced a function



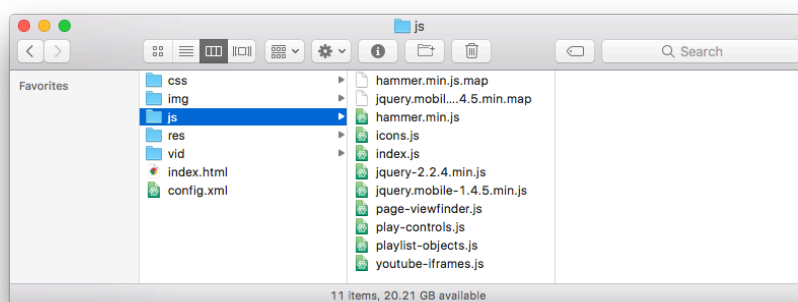
▲ **Fig 57** Some highly specific interactions require many basic QC patches strung together to achieve.

be extracted from its layout and played in a pop-out video player widget, so I had to use the Swift code language to wrangle a custom solution and force iOS to allow the video to play inline. Having no prior experience using Swift this put me at a disadvantage, however after this small success I was satisfied at the proof of concept. Should an iOS version of the Viewfinders app need to be built, a professional developer ought to have at least no technical constraints.

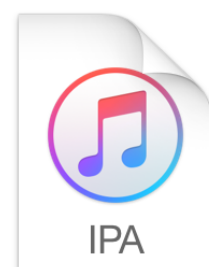
Confronted with the challenge of conducting all further tests twice due to the incompatibility of Android and iOS development, I chose to continue my enquiry by using PhoneGap Build. This service from Adobe is built on top of Apache Cordova, an open-source mobile development framework that packages web development technologies inside a wrapper to run as a native app on any of the major mobile platforms. Taking this approach meant I could use the same code on both Android and iOS, a major benefit alone, but also meant that I could use and continue to advance my existing web development skills across HTML, CSS, and JavaScript. PhoneGap Build allows the developer to create the app as they would a website, then upload the source code to their servers, at which point it is checked for errors, then compiled within an installable native app file.



▲ Fig 58 A basic test run through Xcode to test the differences between local and networked video behaviour.



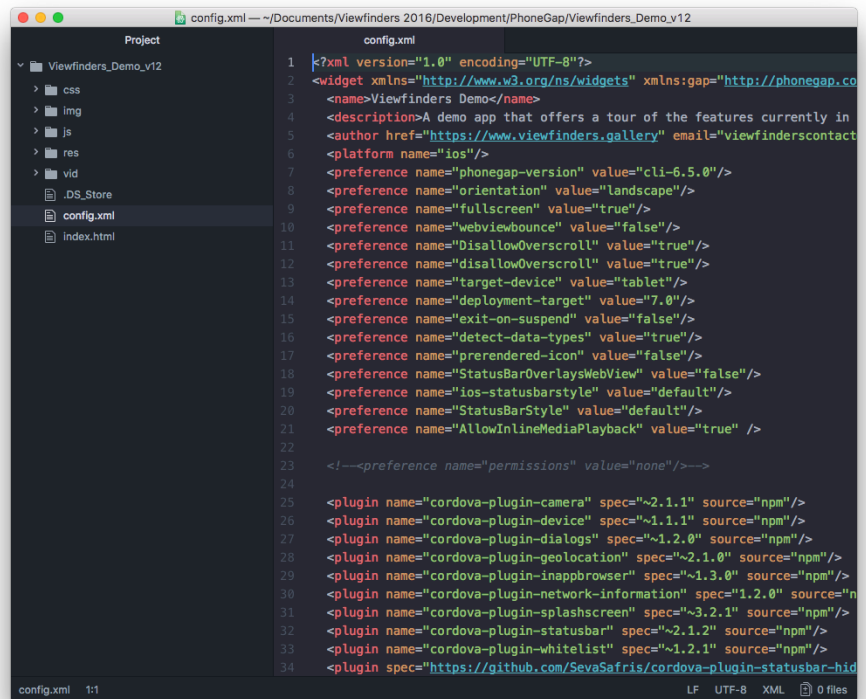
▲ Fig 59 The website folder structure (left) contains all of the standard HTML, CSS, and JavaScript files that make up a website. PhoneGap packages these files within an installable native app file (right).



ViewfindersDemo.ipa

iOS App - 15.2 MB  
 Created Thursday, 30 March 2017 at 9:37 AM  
 Modified Thursday, 30 March 2017 at 9:37 AM  
 Last opened Thursday, 30 March 2017 at 9:38 AM  
 Add Tags...

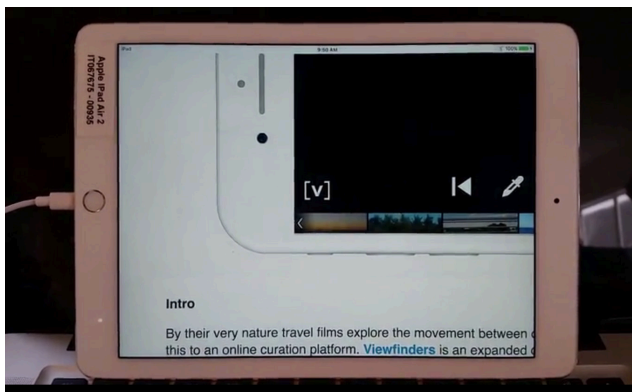
► **Fig 60** The config.xml file sent to PhoneGap along with the regular website files specifies the preferences, plugins, and permissions required for the web app to run natively on each platform.



The first challenge using PhoneGap was to apply the correct configuration to the config.xml file in order for it to pass the error check upon upload. Although the bulk of the app is produced using familiar web development methods, there is still a requirement to lay the

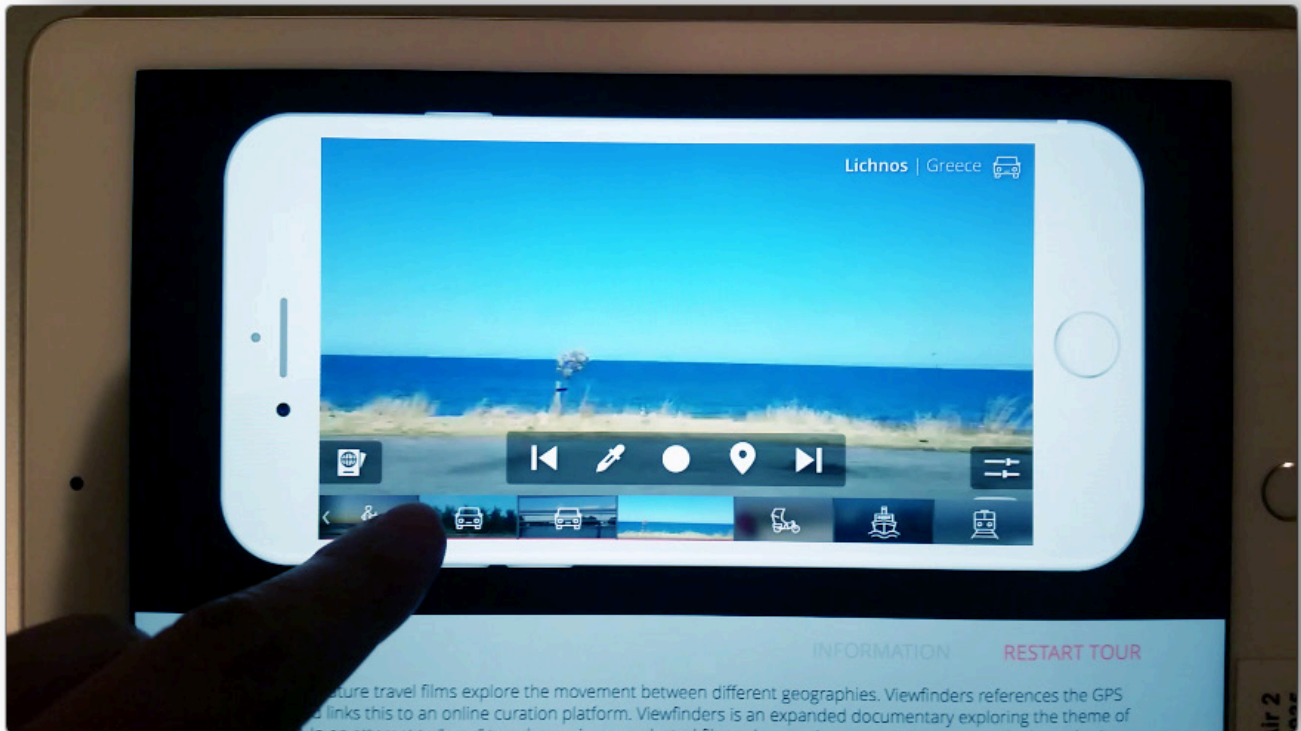
platform-specific foundations that will bridge the divide between the web code and the operating system. While there is sufficient documentation to learn how this is done for a basic app, the configuration must be adjusted to suit the needs of the app in question. Given the innovative nature of Viewfinders, the further I strayed from convention the more challenging it became to identify and configure the preferences and permissions required to support my functions natively across both Android and iOS.

I had anticipated that my own inexperience in native development would be a tension; however I was not interested in producing a flawless rendition of the concept. Instead, my objective was to assess whether my concept was technically feasible. By bringing together disparate methods, plugins, and services – many at the cutting edge of web capabilities – I have demonstrated the feasibility of a web-based mobile-mentary that samples and responds dynamically to its user's environment. Only when the project is complete and I have benefitted from the subsequent interdisciplinary insights can I produce an educated brief, and have an experienced developer produce a more robust product.



► **Fig 61** After several builds where the config.xml file was incorrectly formatted or mismatched the app ID declaration, this was the first successful native installation of a web app on iOS using PhoneGap. Subsequent prototypes corrected the issues apparent here, most noticeably scale.



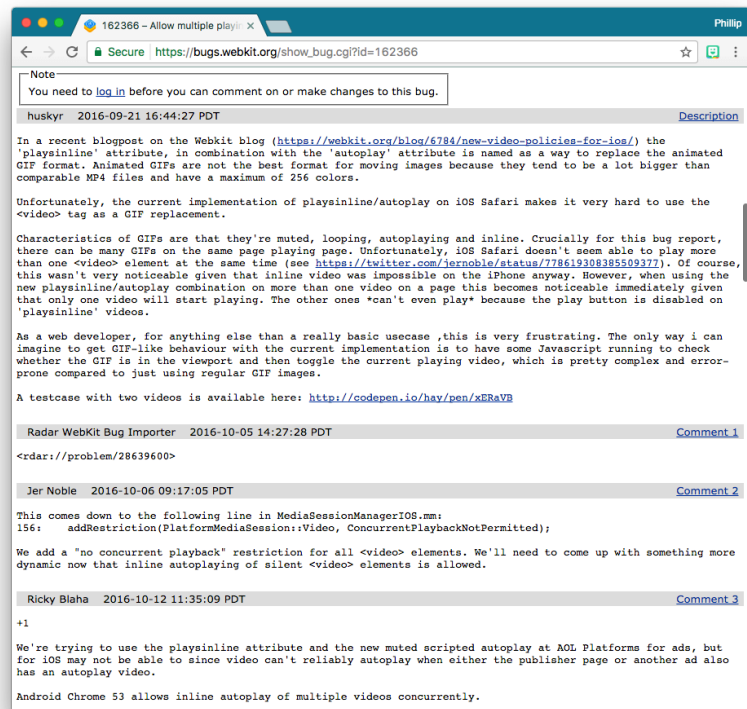


▲ **Fig 62** Javascript animations, a staple web technology, become laggy when installed within a native app.

While testing on an iPad I used a smartphone emulator to ensure usability even at small scale. This test functioned well on desktop, however once installed on the device it became laggy. It turns out that despite pitching itself as a bridge between web and native development, PhoneGap is unable to ‘translate’ some basic web technologies reliably. For example JavaScript web animations lose hardware-acceleration when run as a native app, a consequence of using PhoneGap. The solution was to replace all animations with CSS transitions, and to restrict the total number of moving elements to optimise device performance.

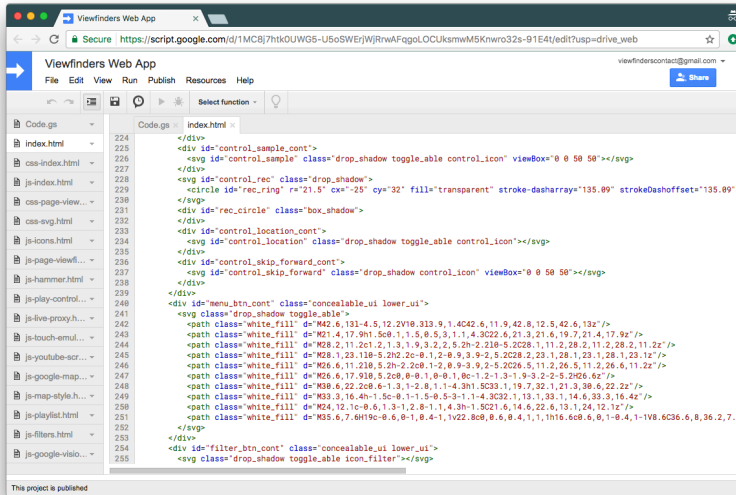
One of my biggest frustrations while developing the Viewfinders web app was Apple’s inflexibility towards handling of web-based video. Although HTML5 canvas and video elements have introduced a world of new possibilities for web developers (Walsh, 2012), outdated platform-specific behaviours in iOS (Kranen, 2016) mean that cross-platform applications like web-apps are limited in how they can actually implement these capabilities. Having overcome the issue of videos popping out of the application layout, I faced the apparently longstanding iOS issue of being limited to one video playing at a time. This limitation simply does not exist on Android, and as it turns out has been a major obstruction for online marketers (Bugzilla, 2016) – hence the animated GIF prevails despite being a relic of the first Internet.

In an interesting development it appears Apple have finally bowed. After months of lobbying (Kranen, 2016), the March 2017 release of iOS 10.3 resolves the issue. For the first time web developers can now take full advantage of HTML5 video inline and autoplay capabilities. This seemingly modest shift is actually the key to a new generation of video-based web-apps, and for those attuned to the happenings of the Internet, likely heralds the long-awaited retirement of the animated GIF.



▲ Fig 63 A comment thread on an online forum discusses the imminent need for consecutive video play behaviour on the iOS platform.

After producing a series of prototypes through PhoneGap I chose to produce the final design output as a standalone web-app. Whereas PhoneGap would allow me to build a native app for each major mobile platform, opting to host the app online circumvents the need for users to install any software on their device. PhoneGap had also struggled to faithfully translate web animations into the native environment: *“since PhoneGap applications are based upon HTML & web views, you are limited by the performance of a web view on a particular platform”* (Trice, 2013). A quick test where I moved the same HTML based code onto the Viewfinders server however proved very effective, optimising both the video elements and the transition animations of the app. Moreover users could simply open the website on their device to use the app, and as evidenced by the popularity of web apps like Google Maps this ease of access can greatly increase usership. Taking cues from my earlier development of a submission form using the Google Scripts web app service, I decided to transfer the latest PhoneGap source code into the Google Scripts environment, then adapt it be operational in its new domain.



▲ **Fig 64** The HTML source code adapted to the Google Apps Script environment. Note the site structure listed on the left-hand side containing the many CSS and JavaScript files that constitute the web app.

The one major compromise of moving to an external web-app was that in doing so the app lost the ability to manipulate how the device interprets the web content, meaning that on iOS most of the default video handling behaviour returned. Though some of these behaviours can be managed through the increasing access that client-side APIs grant to device hardware (Walsh, 2012), Apple remains understandably apprehensive about access to

the device camera. While it is possible for a web app to use the camera, on iOS it must be accessed through an intermediary window, the expectation being that the user will record a video in this widget and before returning to the original interface. As we have seen however, big players are already feeling the pressure to yield to a growing web-app movement, so hardware restrictions like this may change.

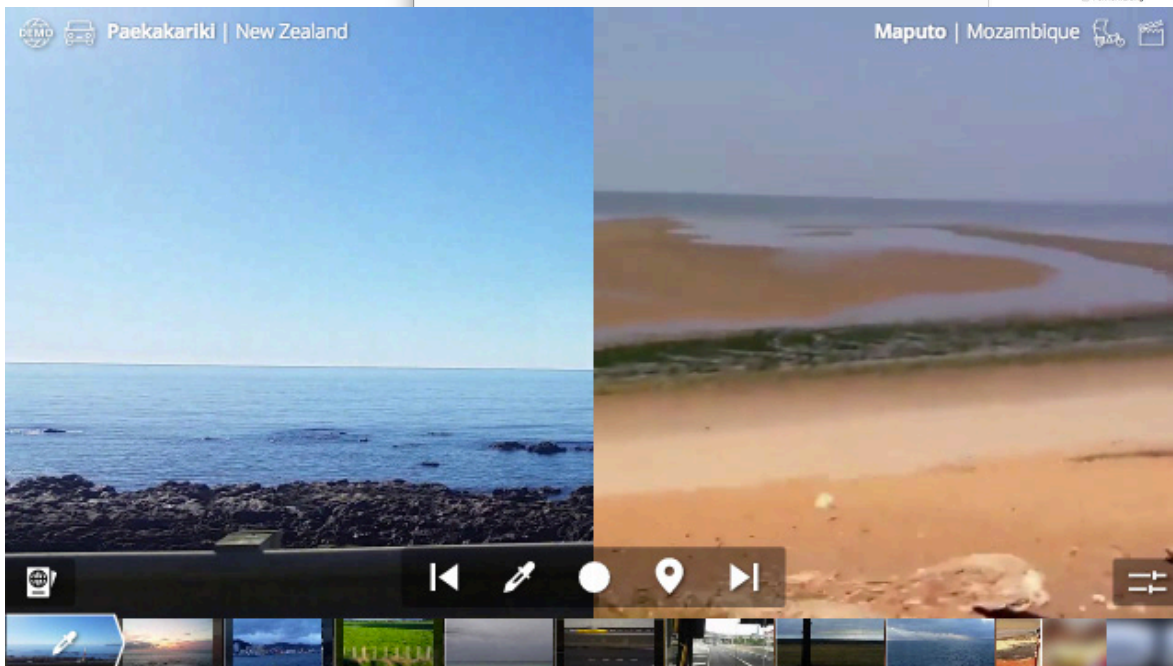
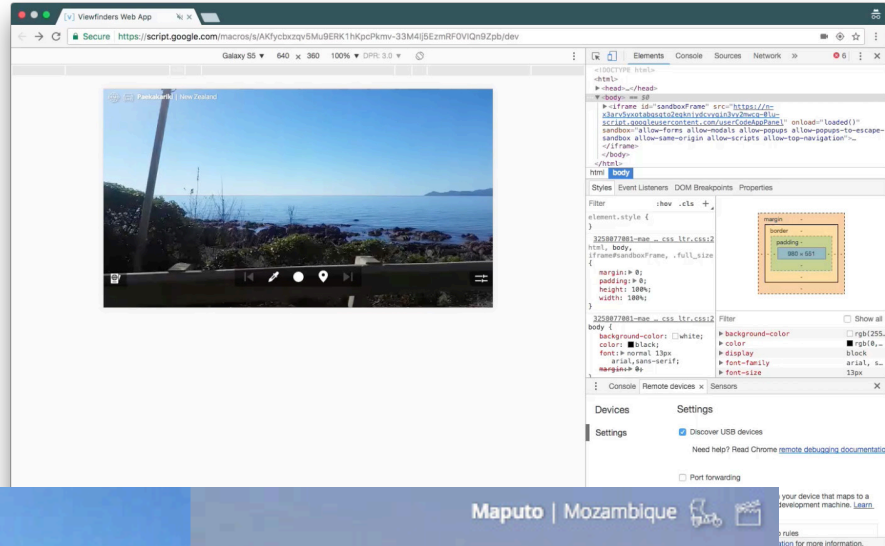
**"I think we'll start to see more applications migrate away from native code to the web as mobile devices and browsers become more capable. Similar to how many desktop applications started off as native apps (e.g., email and documents applications), I think eventually the majority of apps will live on the web rather than on clients with native code."**

– Merrell, as cited in Morony, 2017.

In the interests of progressing the project I resolved to forfeit iOS compatibility in the ensuing prototypes, to focus my efforts on producing a functional web-app proof of concept. Though my practice here is restricted to a 12 month time-frame, my research shows that the Viewfinders web-app concept is at least feasible on iOS, and given more time is certainly a worthwhile pursuit. For the time being, I concentrated on pushing the concept ahead rather than squandering my time on technical issues.

► **Fig 65** The front end of the web-app as hosted on script.google.com, with the Javascript console open to reveal the app's HTML structure.

▼ **Fig 66** The front end of the functioning web app prototype built in Google Apps Script.



The final web-app combines elements of all the technologies tested, including the YouTube API for video storage, retrieval, and playback; Google Cloud Vision API for computer vision capabilities; Hammer.js JavaScript library to handle touch events; Google Sheets for metadata storage and retrieval; Google Maps API for geolocation and custom map generation; and the standard web development toolkit of HTML5, CSS, JavaScript, and jQuery Mobile.





Fig 67 The Viewfinders logo fits into environments associated with travel.

### 4.3 Visual Language

When I developed the Viewfinders brand I wanted it to convey the theme of travel immediately. To achieve this I emulated the ‘airport aesthetic’, mimicking the visual language typically found in terminals and transit lounges. Public wayfinding systems, especially in spaces designed for use by international travellers, employ pictograms so as to be understood across readers that share no common language. Actual text is kept to a minimum, though usually appears in the local language and often is accompanied by a translation into another language appropriate to the region. Professor of visual communication Andreas Uebele advocates the importance of wayfinding systems that reassure. This is crucial when the traveller is confronted with “a multilingual hubbub, a colourful sound, a confusion of voices and a concert of colours” (Uebele, 2007, p. 315). The principle of reassurance is manifest in signage that deploys effective colour coding, clear information hierarchies, and standardised pictograms.

Directly referencing signage system design theory, I produced a visual language featuring minimal tonal variation, block colours, and generous negative space. The Viewfinders logo itself echoes the pictogram aesthetic. Rendered in flat white and overlaying a block colour backdrop, it feels at home within a wayfinding context. The ‘V’ character is subtly accented with a strip of shadow that implies the letterform is layered, a reference to the layering of time and place through AR. The square brackets on either

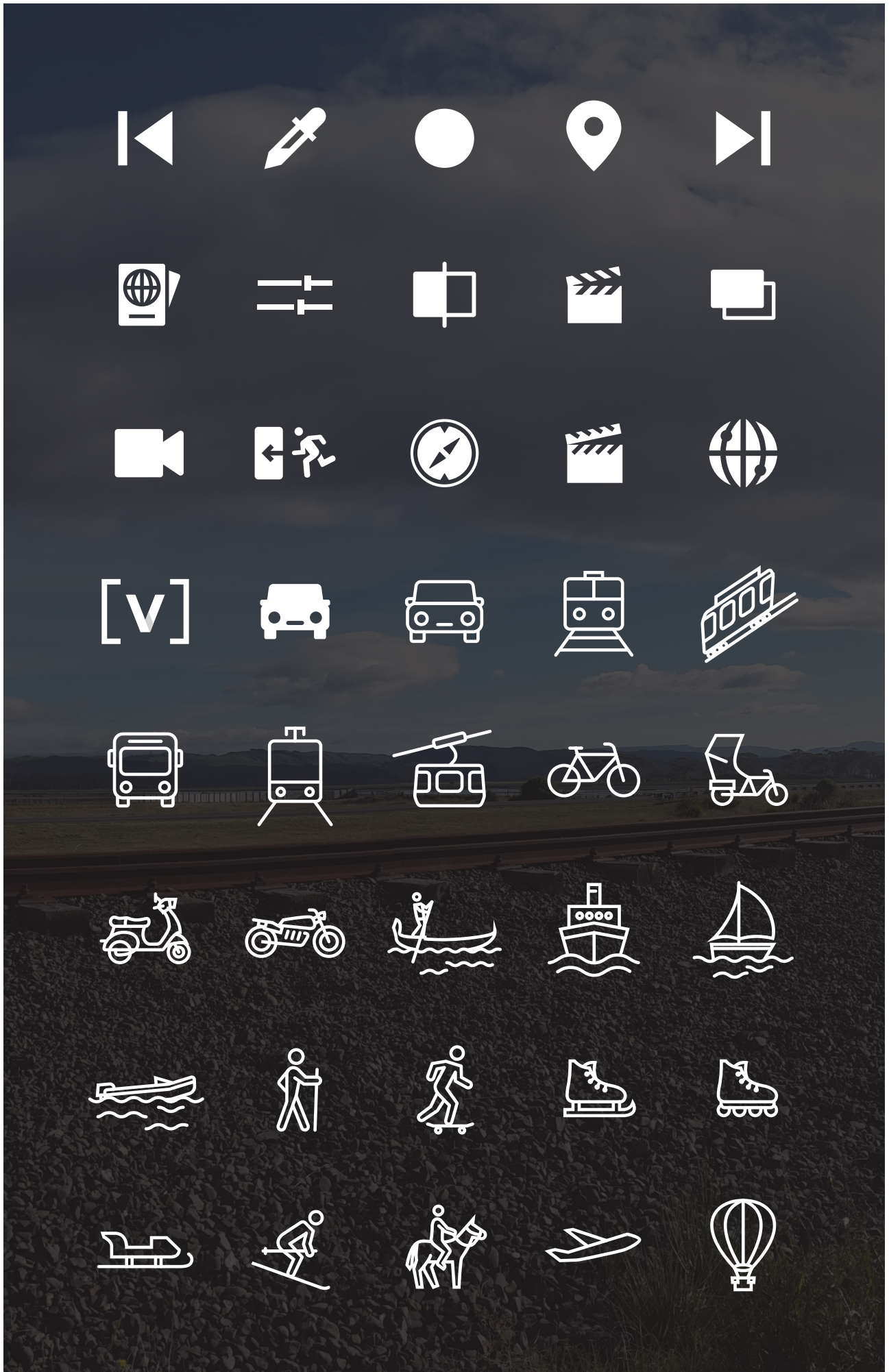


side reference the alignment markings seen in the viewfinder of classic analogue cameras.

**"A signage system can be more than just a sign. It can give a place its identity."**

– Uebele, 2007, p. 5.

All text is presented geometric and sans-serif, and expressed as succinctly as possible: pictograms are preferable so that wherever possible the message is not restricted to speakers of one language. The Viewfinders interface itself embraces travel motifs like tickets, route maps, passports, and split-flap arrival boards.





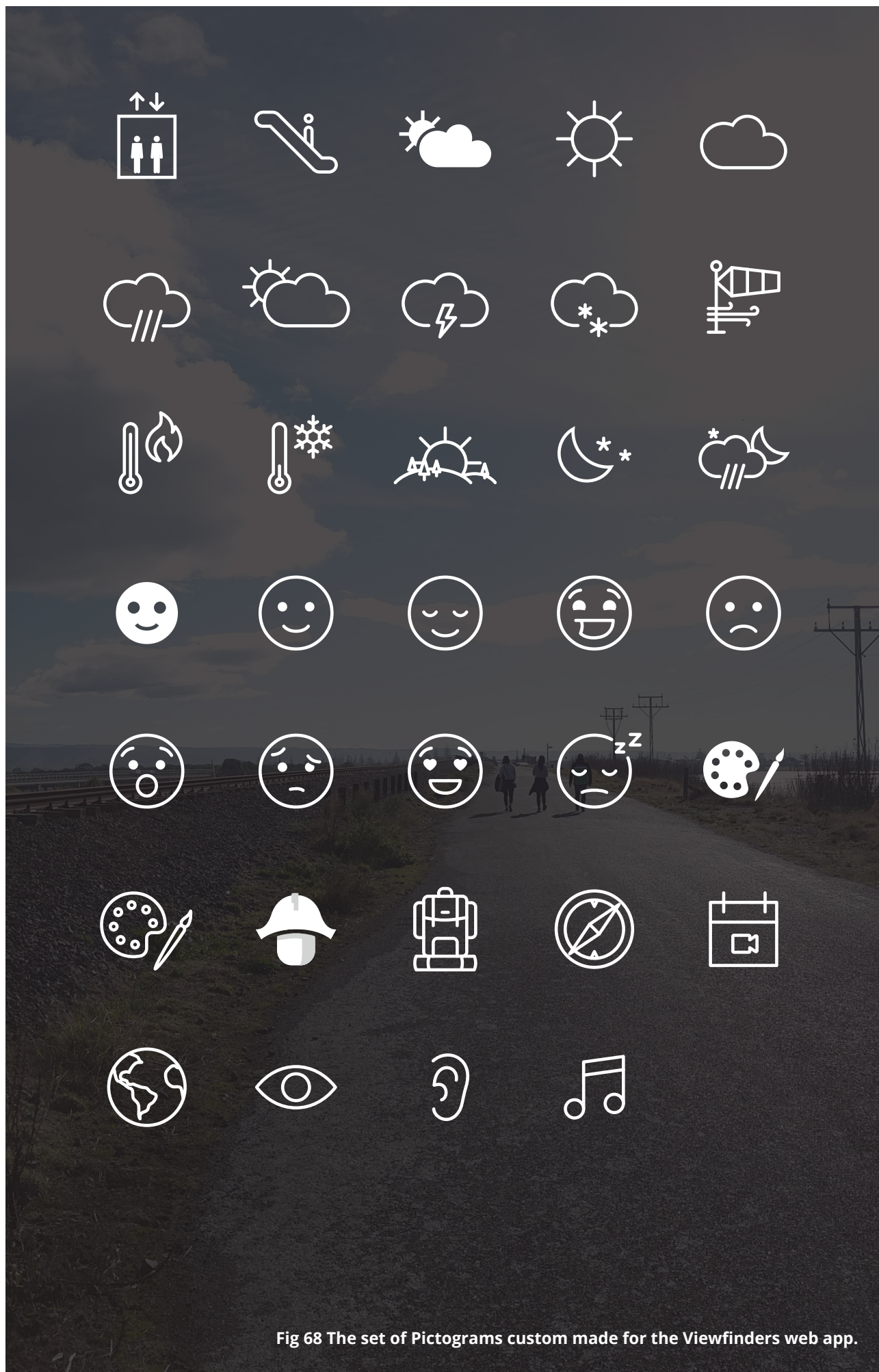
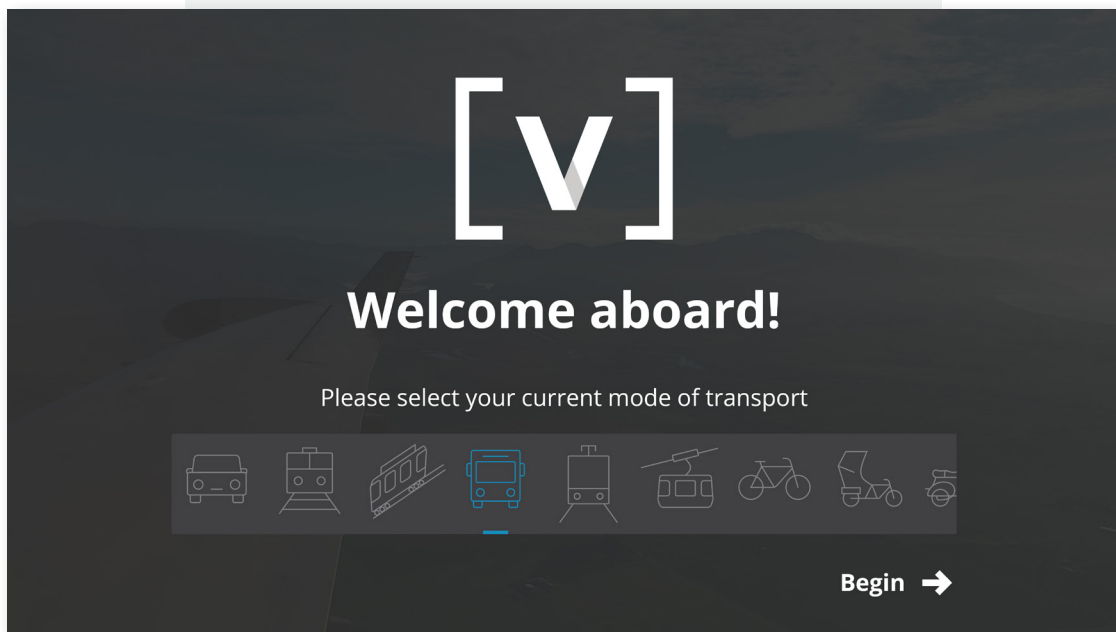


Fig 68 The set of Pictograms custom made for the Viewfinders web app.

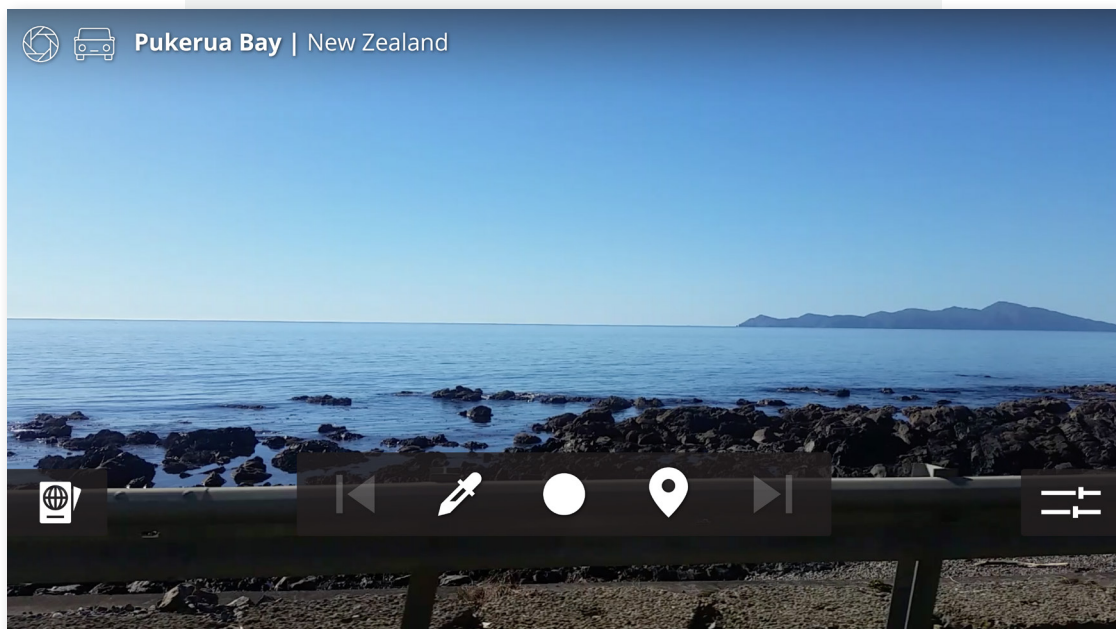


## **4.4 FINAL USER INTERFACE DESIGN**



01

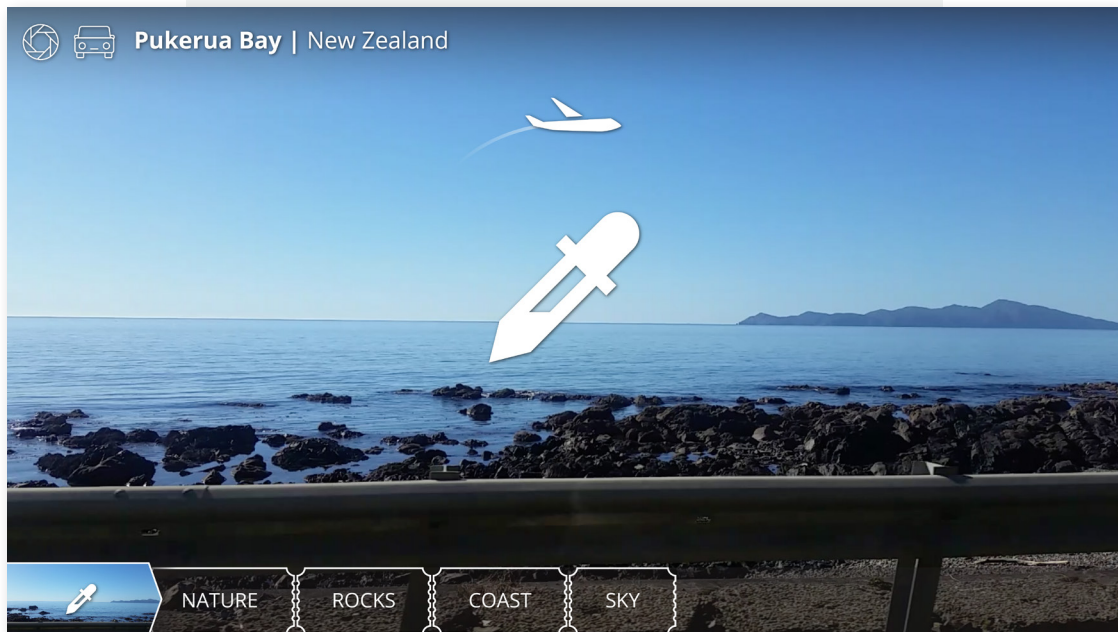
The Viewfinders web app opens on a splash page where viewer-users identify the mode of transport they are currently using. Device camera access is only permitted when triggered by a user interaction, meaning it cannot be activated by a script when the page loads. The 'begin' button serves as both confirmation of the mode of transport, and the prompt to activate the camera stream.



02

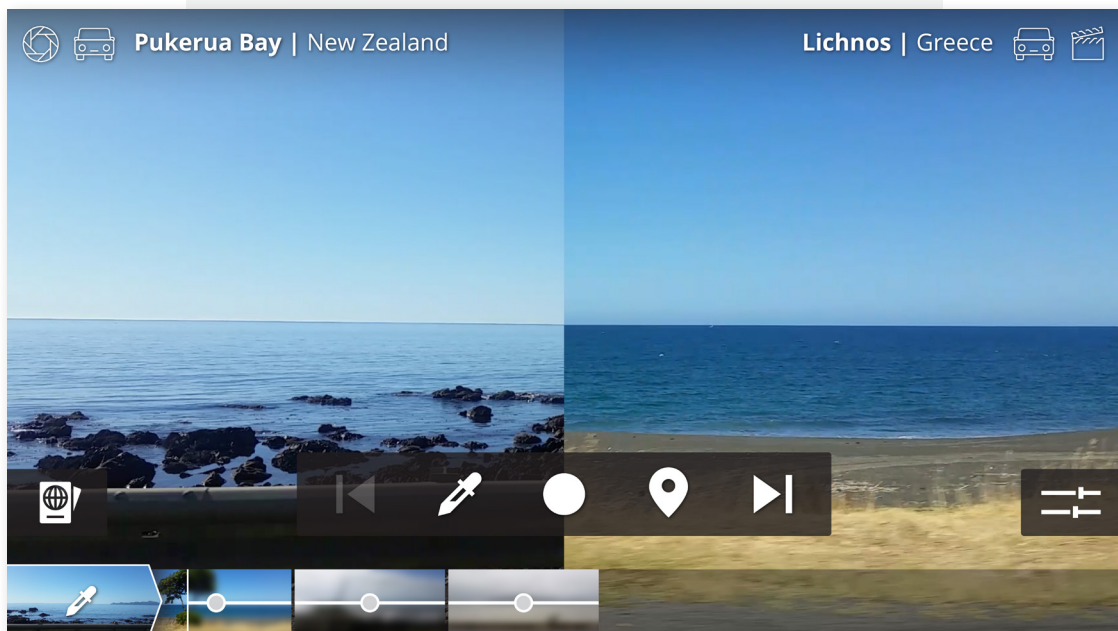
The user interface is centred around this full-screen camera stream. When the page loads Google's reverse geolocation API is used to return the name of the current location, which joins the two pictograms in the header denoting the current mode of transport (car) and that this is the Live view (aperture). At the base are the playback controls (centre), the button to access the menu (passport, left), and the button to refine the narrative direction (sliders, right).





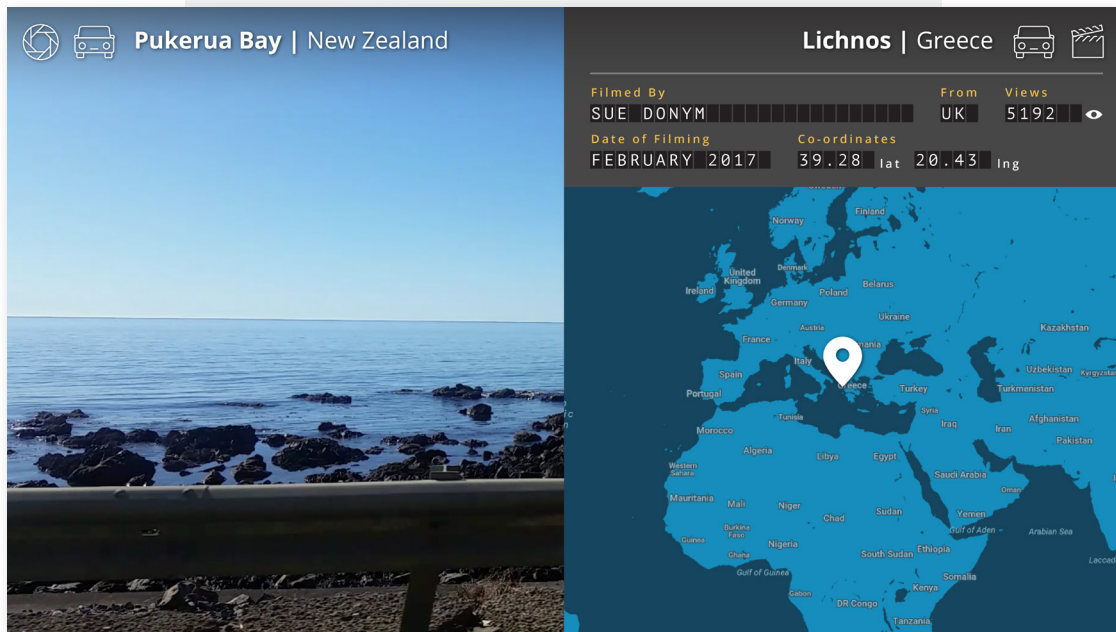
03

After selecting the sample function, the Live video frame freezes, and a loading icon (airplane) appears. Once the computer vision API has processed the frame a series of resulting keywords 'print out' as tickets. These tickets inform the generation of a playlist of Scenes from the Viewfinders database.



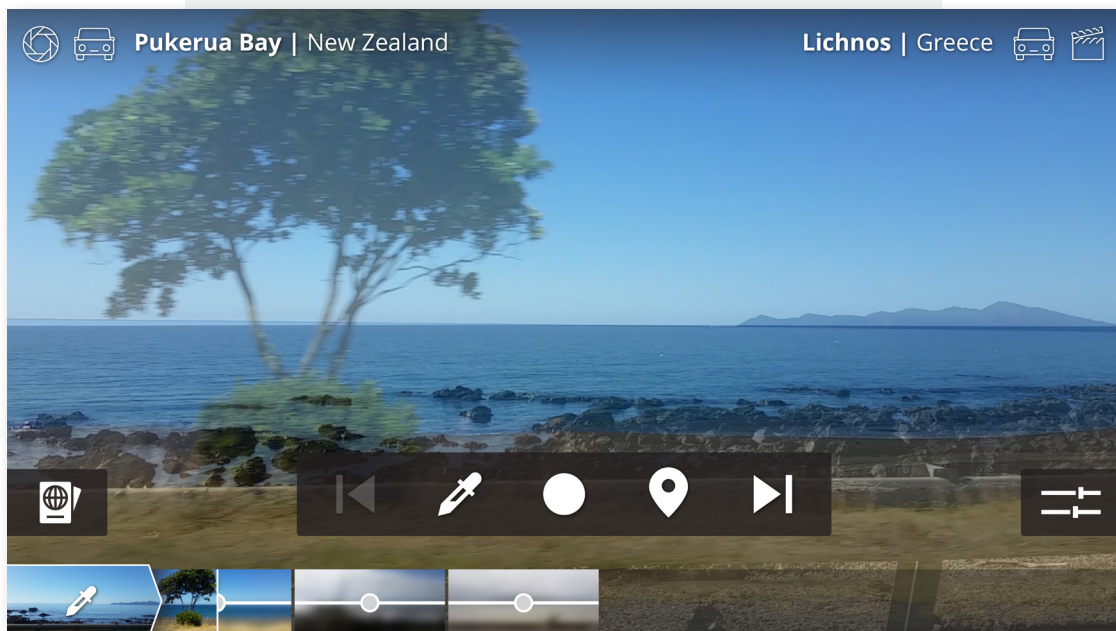
04

After 1.5 seconds on screen the tickets are replaced by thumbnails representing the new playlist. The first Scene begins playing (right), with a header containing the name of the location where it was filmed, and pictograms to show that it is a Scene from the database (clipboard) and what mode of transport it was filmed from (a car). As each Scene plays its thumbnail functions as a playhead, revealing a clear image beneath the blurred original.



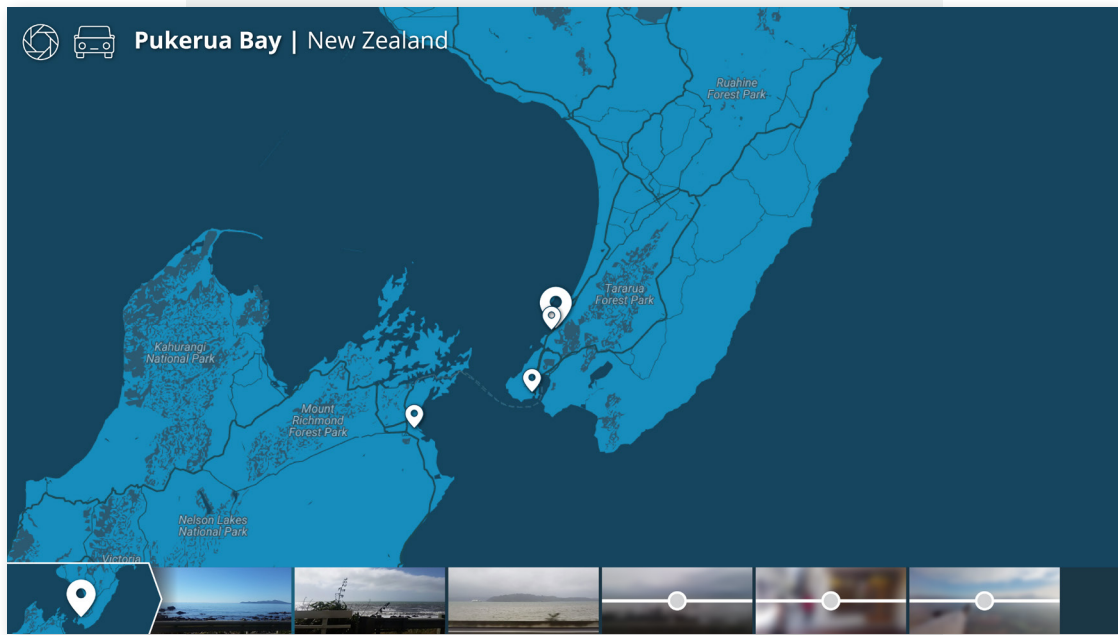
05

To access further information the user can press and hold the video, which summons a world map pinpointing where the Scene was filmed. The name of the filmmaker, the number of times the Scene has been viewed by all users, and the date of filming, are displayed at the top in a panel resembling an airport arrivals board. When the user breaks their contact with the screen the layout reverts to playing the Scene.



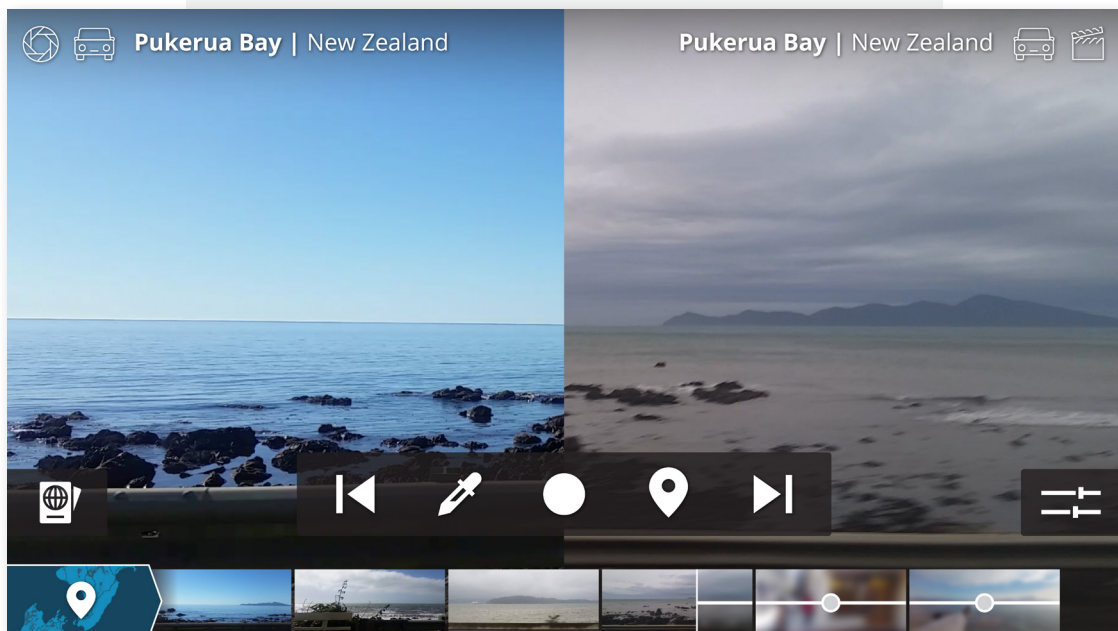
06

Users have the option to alter how the Scenes are displayed. By swiping vertically they can cycle between four viewing states: Overlaid at 50% opacity (as seen here), split view (as seen in the previous figure), full-screen Scene, or full-screen Live camera. The division in the split view can also be dragged to adjust each video's share of screen space.



07

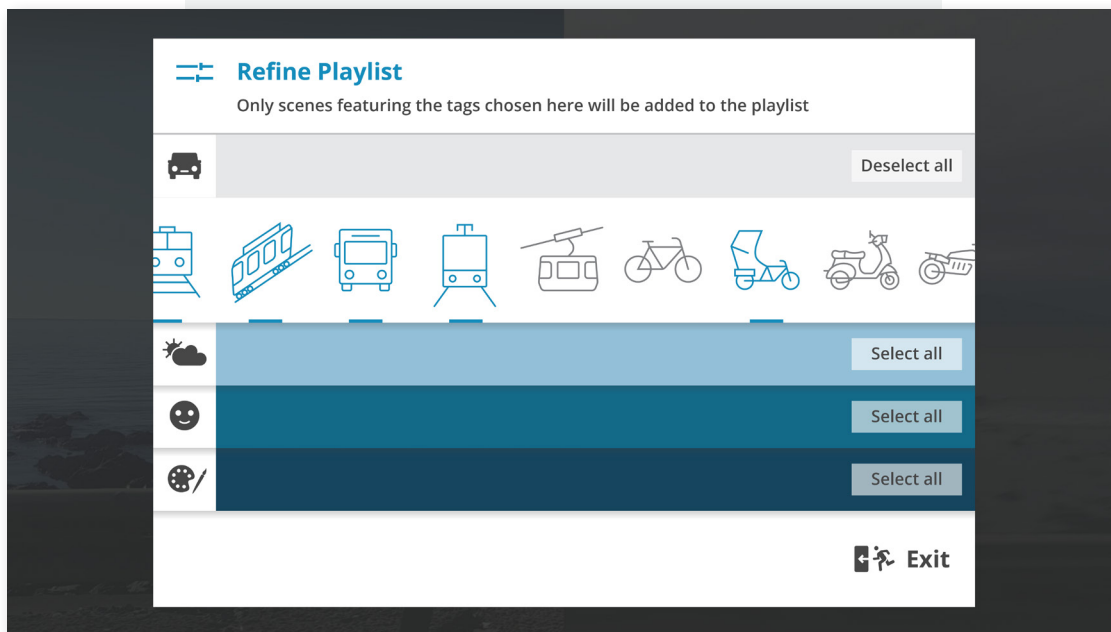
If all queued Scenes play without further input from the user, the same function repeats to continue playback. If the user selects the location function, another three Scenes are queued based on proximity to the current location. This is represented by the appearance of a map: the current location is marked, then the locations of the nearest three Scenes are indicated by smaller markers. Each new Scene is represented by a thumbnail at the bottom.



08

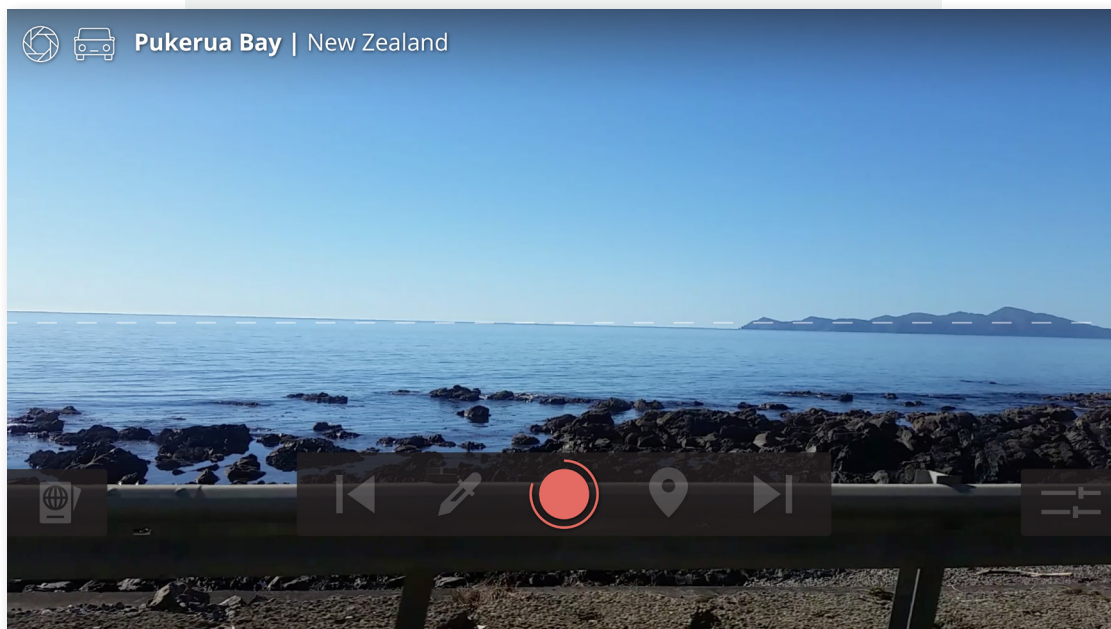
Note that after the location function is used the reference icon at bottom-left changes, announcing the the type of function (visual sample or geolocation), and a corresponding backdrop appears. As before the new Scene begins playing on the right: in this case the nearest Scene was recorded at the user's current location, allowing an insight to Pukerua Bay in two different seasons.





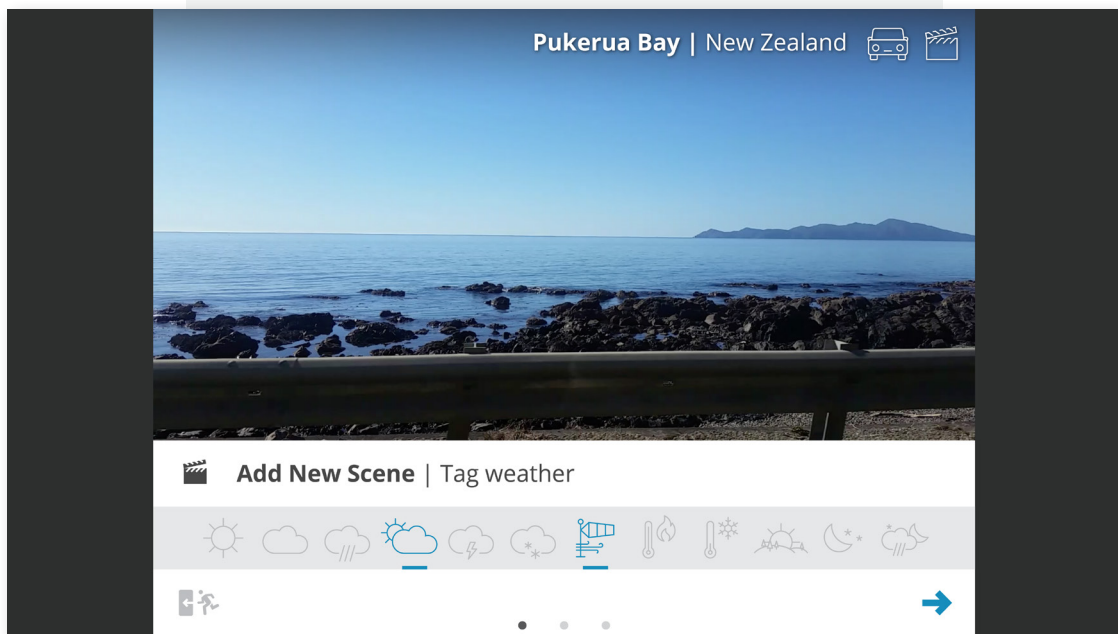
09

When the 'refine playlist' button is selected all other UI overlays disappear and a new window opens on top of the now dimmed Live footage. Users can toggle all icons on and off across four themes: mode of transport, weather, mood, and colour. Only Scenes featuring the active tags will be returned when the next playlist is compiled.

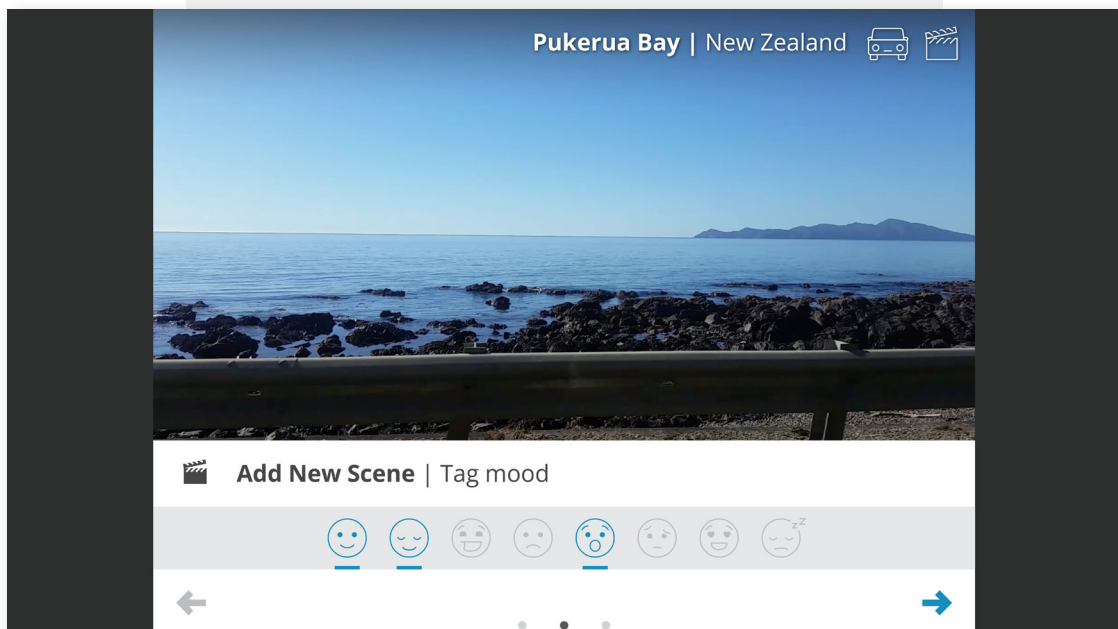


10

Placing a recording button at the centre of the controls prompts viewers to contribute. While recording a horizon guide appears through the centre of the Live view to encourage consistency in alignment across Scenes where the horizon is visible. A ring also progresses around the record button indicating the duration of the recording: Scenes are permitted up to one minute. Tapping the record button again before a minute cuts the recording short.

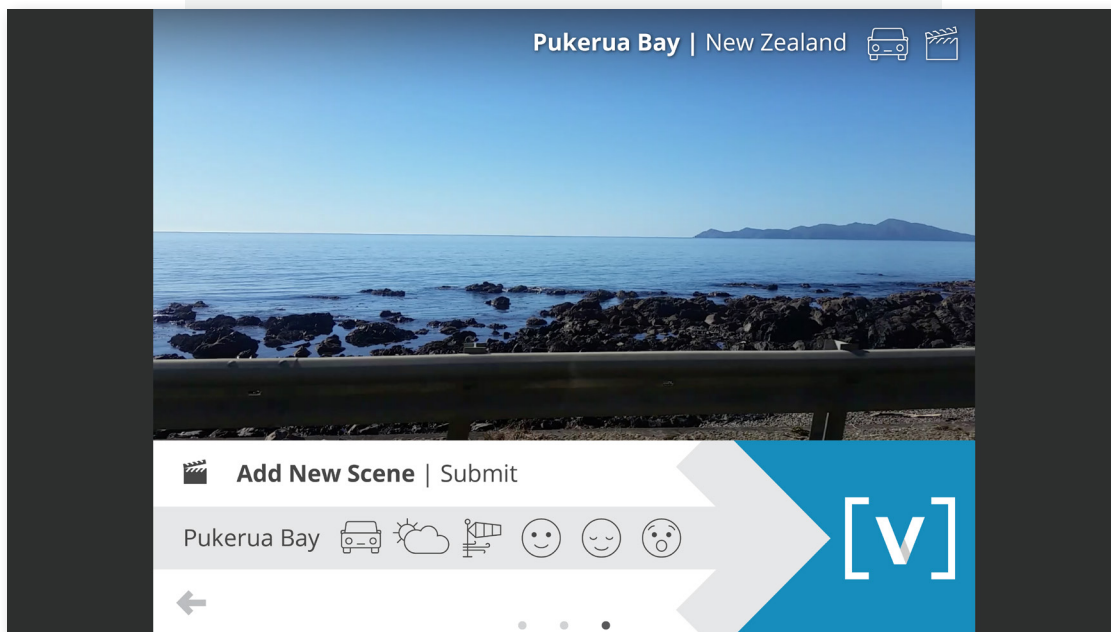


- 11** Once recording has completed the new Scene appears in its own window. While the Scene is uploaded the viewer tags content not able to be identified by computer vision: the weather, and mood. Multiple icons can be selected. To cancel submission, an 'exit' option is provided (bottom-left).



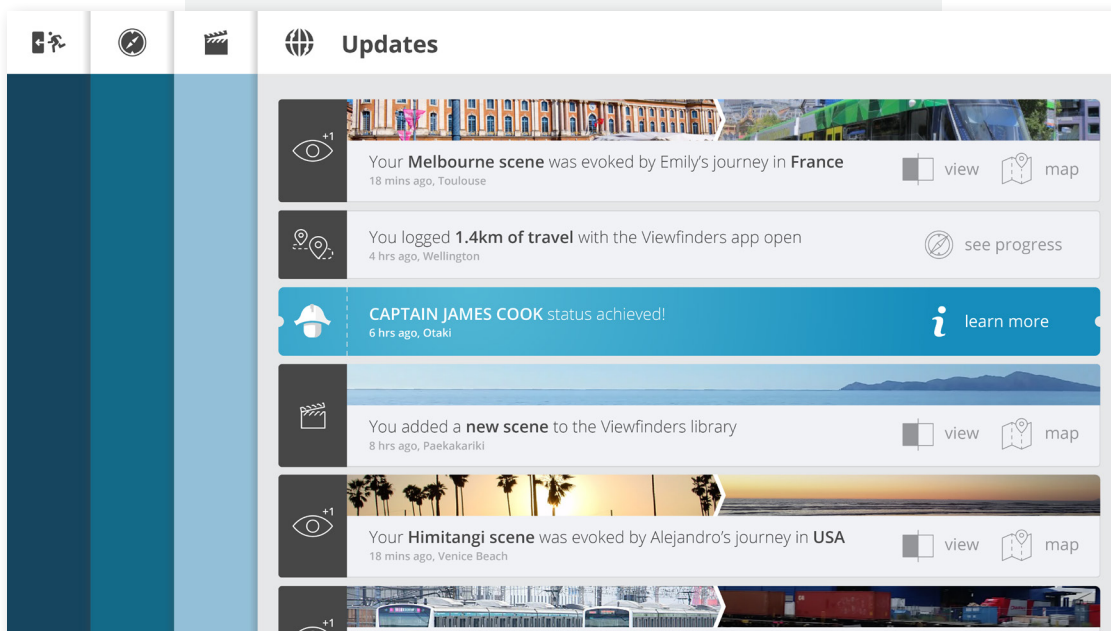
- 12** Tagging the mood of the Scene. Computer vision algorithms cannot yet be relied on to identify these traits, however the dominant colours can of course easily be determined automatically. All other metadata (mode of transport, date and location of filming) are already determined.





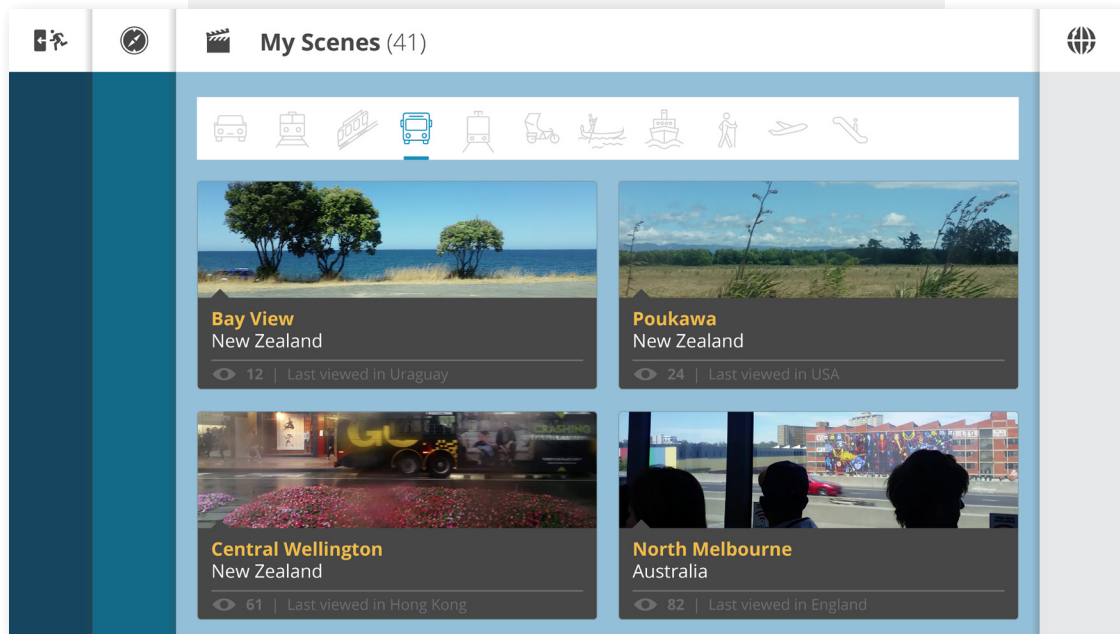
13

After tagging a final review window is displayed. Once satisfied the user taps the Viewfinders icon to submit the new Scene to the database.

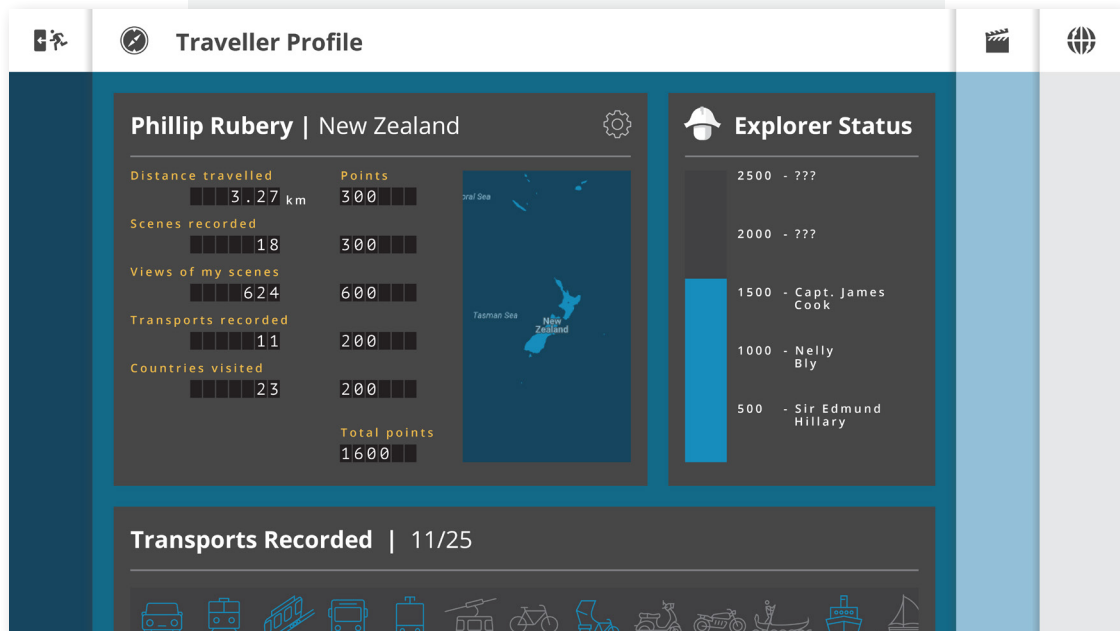


14

Tapping the menu button (passport) from the central video view opens to the Updates panel. The user can see who has viewed their scenes, and from where they were evoked. Tapping one of these notifications reveals a static frame of the environment sampled to evoke their Scene. There are also notifications about statistics, and when a user reaches a new 'explorer status' – a gamified series of levels climbed by contributing scenes and named after famous explorers.



- 15 The Scenes tab of the menu contains a catalogue of Scenes contributed by the user, sorted according to mode of transport (top). From here users can review and delete Scenes.

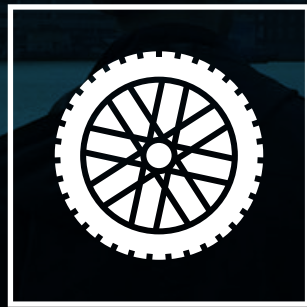


- 16 The last menu tab contains the 'Traveller Profile', which reports statistics associated with the user's account. Tapping the exit pictogram (top-left) returns to the central video view.



17

Since the whole point of Viewfinders is for viewers to take an active role and mirror the movement shown on-screen, the app is programmed to check the user's location at one minute intervals. If their position has not changed, they are not moving, and so are prompted with a message to 'Please continue your journey to resume playback'.



# **MOBILE VIDEO ENGAGEMENT EXERCISE**



To assess whether my proposed active viewing mode really affects the way that viewers experience travel film, I conducted a qualitative exercise on the Wellington waterfront by inviting passers by to ride on a pedalled transport while watching travel footage on an iPad. After a two-minute ride I then stopped the transport and asked the participant a series of ten questions to gauge their experience, recording responses by microphone.

Participation was incentivised by offering \$5 cash to take part, which I considered enough for a coffee on the promenade. As the researcher I took the helm and steered, while participants rode on the back bench so they could safely focus on the iPad footage. The footage shown was a series of traveling shots, starting with a 30 second establishing shot of coastal landscape.



▲ Fig 69 Participants describe their experience.

Questions asked:

1. What is it like to watch travelling footage while moving?
2. Describe the video you just viewed.
3. Describe your own physical journey – consider the sounds, feelings, weather, and other environmental factors.
4. What similarities and differences did you notice between your own journey and the one on-screen?
5. Did watching the video effect your perception of time and place in any way?
6. Did you look away from the video at any point? Why/why not?
7. Did this journey evoke any memories for you, or bring to mind any comparable experiences?
8. If you were to repeat this journey would you unmute the footage, and why/why not?
9. How do you think this experience compares to watching the same clip while sitting in a conventional cinema environment?
10. On a scale from 1 to 10, with 1 being very low, and 10 being very high, how much would you say you felt immersed in the video you just watched?





◀ **Fig 70** The establishing shot featured a section of the Pukerua Bay coastline, intentionally chosen for its parallels with the Wellington waterfront environment.

Responding to question one participants unanimously agreed that the novel viewing mode affected their engagement with on-screen content. Most reported that the surrounding stimuli made them highly attuned to motion, more so than if they were watching the video or riding the cart as independent activities.

Overall there were conflicting reports on how this engagement was affected: where one participant declared it difficult to focus on the footage due to the distractions arising from pedalled transport, another reported total absorption in the footage: *“really I was concentrating on the video so I wasn’t really absorbing a lot of what was happening around me”*. Prompted to elaborate, this latter participant explained that their concentration was in order to “figure out” what was in the footage, likely influenced by the format of the exercise – they knew questions would be asked at the end. The distraction mentioned by the former participant as it happens was due to concern that their phone would slip out of their pocket during the ride, which I have deemed an anomaly unlikely to be a factor for others using the Viewfinders app. It does however illustrate the demands of ‘active viewing’, which is that viewers must remain aware of things happening in 360 degrees. Many participants seemed wary of potential hazards, a factor which would likely be mitigated if the viewer was a passenger on a train or a bus rather than a pedalled crocodile cart. Asked if they had looked away from the iPad, many responded that they had indeed directly compared what was shown on screen with the live environment.

***“I think I looked away when it was the harbour scene, to see if I could recognise what was in the video as being on harbour... to match the footage with what was around me.”***

– Exercise participant.

The effect of peripheral vision was consistently mentioned as a departure from regular viewing experiences. *“Particularly because of my peripheral vision, I’d see something out of the corner of my eye. We’d turn a corner and I’d be curious to see what was there.”* *“You still have your peripheral vision and you feel like everything is moving beside you and things are moving like... landscape-ways, on the screen”*.

Participants also underscored the sense of authenticity provided by ambient conditions. The iPad was muted during the exercise, which I contend would usually counteract engagement, but when asked if they would unmute the audio given the option most pointed out that the presence of sounds in the surrounding area were sufficient: *"I think [not having audio] adds to the atmosphere of what's going on... there are exterior sounds taking place, there's still a particular draw in from hearing train tracks or wheels hitting the ground... it definitely helps you inhabit the experience."*

## Participant Quotes:

*"Because of the bumps and things, it's that shaking motion that you get when you're travelling, when you're on a train, when you're in a car, you get that kind of bumping..."*

*"I didn't need sound. Because... there was sound around us anyway because we're out in the open."*

*"Oh it's much better being in the fresh air and actually moving. I enjoy that much better."*

*"Especially with the movement and feeling the wind on my face it was very much like I was there."*

*"This is definitely more 3D. Actually 4D. In the sense that you're not just sitting. I don't know if it was a coincidence or not that the video is us driving, as you're actually driving, but it gives that additional dimension to it."*

*"I feel like it was very unusual because I know where I am but the video made it feel like I was somewhere different. Like when the plane was taking off and you're going over the bridge it feels like you're taking off, and yeah it made you think 'oh what am I actually doing?'"*



▲ **Fig 71** A participant answers questions about their journey experience.



# CONCLUSION



In recent years new media devices and a hyperlinked global society have exerted considerable influence on filmmaking practices. From the dismantling of linear narrative in interactive documentary and database cinema, to the emergence of modularised video content services like Snapchat, Periscope and Instagram, viewers are now accustomed to video that is participatory, interactive, and delivered in bite-sized segments. In 2017, the smartphone is king. Its capacity to exist beyond the traditional cinema domain means that it is uniquely positioned to offer novel viewing experiences unlike any contained within the four walls of the theatre.

**“Today, traditional media companies face the adaptive challenge posed by the Internet. That challenge is not just the technology itself, but how it has altered people’s habits of media consumption.”**

– Han, 2011, p. 78.

Just as the arrival of the Super 8 camera forever changed the face of amateur travel film, so too the arrival of the smartphone has been a force of disruption. Accessible, practical, and ubiquitous, amateur filmmakers are empowered to produce and distribute first-person experiences more easily than ever before. Unlike the Super 8, these devices untether not just the producer but the audience as well. Emergent practices should acknowledge this new phenomenon: the freedom of the viewer. For travel film especially, the potential for viewers to inhabit spaces with explicit parallels to on-screen content is formidable.

Like Wallworth’s *Collisions*, Viewfinders removes the fourth wall to place its audience centre-stage. Viewers are aware of things moving in their peripheral vision, and completely encapsulated in a diegetic soundscape. However in complete contrast to VR cinema, Viewfinders achieves this not by blocking out the real world but by embracing it, directly sampling it to match content with space. Where *Collisions* viewers are bound to a chair, Viewfinders takes its audience into the world so that their own exploration can recreate the travel experience of the Scene. Museum of London’s *Street Museum* also demonstrates the impact of using location as a contextual device. By positioning a user exactly where a historic photograph originated, *Street Museum* calls on ambient environmental factors to fill in for those not captured by the lens on the day of capture, bringing the scene back to life. The wind in their hair, the smell of the sea: these were the things that Viewfinders test participants enjoyed. But while *Street Museum* is limited to a few predetermined historic sites, Viewfinders has been designed for use anywhere, and in place of an institutional archive Viewfinders features content contributed by the users themselves. Viewfinders demands not just visual engagement but all of the senses combined. “*The future of storytelling is placing the audience at the heart of the experience*” (Barrett, 2016).

Viewfinders pilots an experimental approach to travel film by acknowledging recent innovations across web app development, geolocation, computer vision, augmented reality, and participatory media. It is specifically designed to reject passive viewing, with a narrative that depends entirely on the viewer's interactions with their own environment to progress. The choice of viewing environment directly informs the content evoked, and viewers are positioned to inhabit the same conditions as those on-screen. Furthermore, by bringing the means of production and consumption into a single activity the viewer is compelled to participate, and by sharing their own experience the story continues on.

**"From passive listeners to active participants, we all have the power to take initiative in this frontier of democratised capture and creation.... New stories can be experienced differently, in all parallel ways, across time, across geography, across cultures; and most importantly, these stories will mean so much more to us because we were part of their creation in the first place."**

- Dzambazova, 2015.





## References

- Aagaard-Hansen, J. (2007). The Challenges of Cross-Disciplinary Research. *Social Epistemology*, 21(4), 425–438.
- ACMI. (2017). *Collisions*. Retrieved from <https://guides.acmi.net.au/collisions/>
- Arceneaux, N., & Kavoori, A. (2006). *The Cell Phone Reader: Essays in Social Transformation (Digital Formations)*. New York, NY, Peter Lang Publishing Inc.
- Arceneaux, N., & Kavoori, A. (2012). *The Mobile Media Reader*. New York, NY, Peter Lang Publishing Inc.
- Aurasma. (2016). *Get Started*. Retrieved from <https://www.aurasma.com/get-started/>
- Azuma, R. (1997). A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, 6(4), 355–385.
- Barker, C. (2008). *Cultural studies: Theory and practice*. London, Sage Publications.
- Barrett, F. (2016). *Felix Barrett - Burn the Seats: Audience Immersion in Interactive Theater* [Video file]. Retrieved from <https://vimeo.com/79522231>
- Berry, M., & Schleser, M. (2014). *Mobile Media Making in the Age of Smartphones*. New York, NY, Palgrave Macmillan.
- Benghozi, P., Bureau, S., & Massit-Folléa, F. (2009). Définir l'internet des objets. In Waroquiers, C., & Davidson, S. (Trans.), *L'internet des Objets: Quels Enjeux Pour l'Europe?, The Internet of Things: What Challenges for Europe?* (pp. 15–23). Paris, Éditions de la Maison des Sciences.
- BMVA. (2017). *Vision Overview*. Retrieved from <http://www.bmva.org/visionoverview>
- Bolter, J., & Grusin, R. (1999). *Remediation: Understanding New Media*. Cambridge, MA, The MIT Press.
- Boonstra, C., Lens-FitzGerald, M., & van der Klein, R. (2011). *Layar - Impactful Augmented Reality in Your Everyday Life* [Video file]. Retrieved from [https://www.youtube.com/watch?v=HW9gU\\_4AUCA](https://www.youtube.com/watch?v=HW9gU_4AUCA)



- Bugzilla. (2016). *Allow Multiple Playing Videos On A Page With 'Autoplay' And 'Playsinline' Attributes*. Retrieved from [https://bugs.webkit.org/show\\_bug.cgi?id=162366](https://bugs.webkit.org/show_bug.cgi?id=162366).
- Burgess, J., Hjorth, L., & Richardson, I. (2012). *Studying Mobile Media: Cultural Technologies, Mobile Communication, and the iPhone*. New York, NY, Routledge.
- Cammaer, G. (2015a). Sharing Mobile, Portable and Personal Travel Stories: From "I was there" to "I was there". *Mobile Media Innovation: Proceedings and Occasional Papers Related to the 2014 MINA Conference*, 5. Retrieved from <https://ctechjournal.aut.ac.nz/paper/sharing-mobile-portable-and-personal-travel-stories/>
- Cammaer, G. (2015b). *Travelling Shots on Vimeo* [Video file]. Retrieved from <https://vimeo.com/144995895>
- Candy, L. (2006a). *Practice Based Research: A Guide* (Report No. 1 - November 2006). Retrieved from <https://www.creativityandcognition.com/resources/PBR%20Guide-1.1-2006.pdf>
- Candy, L. (2006b). *Research and Creative Practice* (Dissertation, University of Technology Sydney, Sydney, Australia). Retrieved from <http://research.it.uts.edu.au/creative/linda/CCSBook/Jan%2021%20web%20pdfs/Candy.pdf>
- Cater, M. (2013). *A Brief History of API Based Web Applications*. Retrieved from <https://blog.smartbear.com/api-testing/a-brief-history-of-api-based-web-applications/>
- Chalfen, R. (1986). *Media Myopia And Genre-Centrism: The Case Of Home Movies*. *Journal Of Film And Video*, 38(3/4), 58-62.
- Clarifai. (2015). *Clarifai Introduces Video Recognition*. Retrieved from <http://blog.clarifai.com/portfolio/clarifai-introduces-video-recognition/>
- Clarifai. (2016). *General Model*. Retrieved from <https://developer.clarifai.com/models/general-image-recognition-model/aaa03c23b3724a16a56b629203edc62c>
- Coldewey, D. (2016). *Here's why the iPhone 7 Plus has two cameras*. Retrieved from <https://techcrunch.com/2016/09/07/heres-why-the-iphone-7-has-two-cameras/>
- Crouch, C., & Pearce, J. (2012). *Doing Research in Design*. London, Berg.

- Daly, K. (2010). *Cinema 3.0: The Interactive-Image*. Cinema Journal, 50(1), 81–98.
- David, G., & Schleser, M. (2013). Mobile AR: Creating Augmented Experiences. *Anti-Po-Des Design Journal*, 2, 35–45.
- de Souza e Silva, A., & Frith, J. (2012). *Mobile Interfaces in Public Spaces: Locational Privacy, Control, and Urban Sociability*. London, Taylor & Francis Group.
- de Souza e Silva, A. (2013). Mobile Narratives; Reading and Writing Urban Space with Location-Based Technologies. In Hayles, N., & Pressman, J. (Eds.), *Comparative Textual Media: Transforming the Humanities in the Postprint Era* (pp. 33–52). Minneapolis, MN, University of Minnesota Press.
- de Vries, I. (2012). Through the looking cell phone screen: dreams of omniscience in the age of mobile augmented reality. In Arceneaux, N., & Kavoori, A. (Eds.), *The Mobile Media Reader* (pp. X). New York, Peter Lang Publishing Inc.
- Dethridge, L. (2016). *Virtual Reality Film Collisions Is Part Disaster Movie, Part Travelogue And Completely Immersive*. Retrieved from <http://theconversation.com/virtual-reality-film-collisions-is-part-disaster-movie-part-travelogue-and-completely-immersive-66563>
- Design Council. (2010). *Multi-Disciplinary Design Education in the UK: Report and Recommendations from the Multi-Disciplinary Design Network*. Retrieved from <http://www.designcouncil.org.uk/sites/default/files/asset/document/multi-disciplinary-design-education.pdf>
- Design Council. (2015). *Innovation by Design: How design enables science and technology research to achieve greater impact*. Retrieved from <http://www.designcouncil.org.uk/sites/default/files/asset/document/innovation-by-design.pdf>
- Documentary Network. (2011). *Documentary and New Digital Platforms, portrait of an ecosystem in transition – A Documentary network's Study*. Retrieved from <http://obsdoc.ca/wp-content/uploads/ancien/recherches.e/portrait2011/index.html>
- Dzambazova, T. (2015). *Tatjana Dzambazova - Recreating Reality* [Video file]. Retrieved from <https://vimeo.com/139386007>

- Farman, J. (2011). *Mobile Interface Theory: Embodied Space and Locative Media*. New York, NY, Routledge.
- Farman, J. (2012). Historicizing Mobile Media: Locating the Transformations of Embodied Space. In Arceneaux, N., & Kavoori, A. (Eds.), *The Mobile Media Reader* (pp. 9–22). New York, NY, Peter Lang Publishing Inc.
- Fowler, M., & Highsmith, J. (2001). *The Agile Manifesto*. Retrieved from [http://andrey.hristov.com/fht-stuttgart/The\\_Agile\\_Manifesto\\_SDMagazine.pdf](http://andrey.hristov.com/fht-stuttgart/The_Agile_Manifesto_SDMagazine.pdf)
- Gannes, L. (2015). *Ten Years of Google Maps, From Slashdot to Ground Truth*. Retrieved from <https://www.recode.net/2015/2/8/11558788/ten-years-of-google-maps-from-slashdot-to-ground-truth>
- Gauntlett, D. (2011). *Making is Connecting: The Social Meaning of Creativity, from DIY and Knitting to Youtube and Web 2.0*. Cambridge, UK, Polity Press.
- Given, L. (2008). *The SAGE Encyclopedia of Qualitative Research Methods*. Thousand Oaks, CA, Sage.
- Goggin, G. & Hjorth, L. (2007). *Mobile Technologies: From Telecommunications to Media*. London, Taylor & Francis.
- Goggin, G. (2012). *New Technologies and the Media: Key Concerns in Media Studies*. London, Palgrave Macmillan.
- Goggin, G., & Hjorth, L. (2014). *The Routledge Companion to Mobile Media*. London, Routledge.
- Gray, C., & Malins, J. (2004). *Visualizing Research: A Guide To The Research Process In Art And Design*. Aldershot, Ashgate Publishing Limited.
- Han, S. (2011). *Web 2.0*. New York, NY, Routledge.
- Hayles, N., & Pressman, J. (2013). *Comparative Textual Media: Transforming the Humanities in the Postprint Era*. Minneapolis, MN, University of Minnesota Press.
- Ito, M. (2005). Introduction: Personal, Portable, Pedestrian. In Ito, M., Okabe, D., & Matsuda, M. (Eds.), *Personal, Portable, Pedestrian: Mobile Phones in Japanese Life* (pp. 1–16). Cambridge, MA, MIT Press.



- Katz, J., LaBar, W., & Lynch, E. (2011). *Creativity and Technology: Social Media, Mobiles and Museums*. New York, NY, MuseumsEtc.
- Kerin, A. (2016, November 1). *Unity Partnership Announced* [Web log message]. Retrieved from <https://www.vuforia.com/Blog/Unity-Partnership>
- Kranen, H. (2016). *iOS 10 Promises Better Video Handling In Safari, But Doesn't Deliver*. Retrieved from <https://www.haykranen.nl/2016/09/22/ios-10-video-playsinline/>
- Laverty, S. (2003). Hermeneutic Phenomenology and Phenomenology: A Comparison of Historical and Methodological Considerations. *International Journal of Qualitative Methods*, 2(3), 1-29.
- Leedy, P. & Ormrod, J. (2010). *Practical Research: Planning and Design*. Boston, MA, Pearson.
- Li, F. (2017). *Announcing Google Cloud Video Intelligence API, and more Cloud Machine Learning updates*. Retrieved from <https://cloud.google.com/blog/big-data/2017/03/announcing-google-cloud-video-intelligence-api-and-more-cloud-machine-learning-updates>
- Ling, R., & Donner, J. (2009). *Mobile Communication*. Cambridge, UK, Polity.
- Liu, A. (2016). *Train Your Own Visual Recognition Model And Search Any Image With Custom Training & Visual Search*. Retrieved from <http://blog.clarifai.com/train-your-own-visual-recognition-model-and-search-any-image-with-custom-training-visual-search/>
- Mackenzie, A. (2005). Untangling the Unwired: Wi-Fi and the Cultural Inversion of Infrastructure. *Space and Culture*, 8(3), 269-285.
- Manovich, L. (2001). *The Language of New Media*. Cambridge, MA, MIT Press.
- McLuhan, M. (1964). *Understanding Media: The Extensions of Man*. New York, NY, McGraw-Hill.
- Morony, J. (2017). *Could HTML5 Web Apps Mean The End Of Native Mobile Apps?* Retrieved from <https://www.joshmorony.com/could-html5-web-apps-mean-the-end-of-native-mobile-apps/>

- Muratovski, G. (2016). *Research for Designers: A guide to Methods and Practice*. London, Sage Publications Ltd.
- Myerson, G. (2001). *Heidegger, Habermas and the Mobile Phone*. London, Icon.
- Nisselson, E. (2003). Why will wireless camera phones revolutionize the photography industry? *The Digital Journalist*, 3(5). Retrieved from <http://digitaljournalist.org/issue0305/cameracorner.html>
- Nisselson, E. (2015). *Advances In Computer Vision & AI Giving Rise To Billion Dollar Opportunities* [Video file]. Retrieved from <https://vimeo.com/130817268>
- Orgeron, D. (2006). Mobile Home Movies: Travel And Le Politique Des Amateurs. *The Moving Image*, 6(2), 74-100.
- Pichai, S. (2017). *Google Keynote* [Video file]. Retrieved from [https://events.google.com/io/schedule/?section=may-17&sid=\\_\\_keynote\\_\\_](https://events.google.com/io/schedule/?section=may-17&sid=__keynote__)
- Rheingold, H. (1993). *The Virtual Community: Homesteading on the Electronic Frontier*. Reading, MA, Addison-Wesley.
- Rheingold, H. (2002). *Smart Mobs: The Next Social Revolution*. New York, NY, Perseus Books.
- Rieser, M. (2012). *The Mobile Audience: Media Art and Mobile Technologies*. Amsterdam, Rodopi.
- Rushton, S. (2012). Calling Ahead: Cinematic Imaginations of Mobile Media's Critical Affordances. In Arceneaux, N., & Kavoori, A. (Eds.), *The Mobile Media Reader* (pp. X). New York, NY, Peter Lang Publishing Inc.
- Schleser, M., Baker, C. and Molga, K. (2009). Aesthetics of Mobile Media Art. *Journal of Media Practice*, 10(2), 101-122.
- Schleser, M. (2011). *Mobile-mentary: Mobile Documentaries in the Contemporary Mediascape*. Sarbruecken, Lambert Academic Publishing.
- Schleser, M. (2012). *Mobile-mentary (Mobile Documentary) 2.0* [Video file]. Retrieved from <https://vimeo.com/35411132>

- Schwartzel, E. (2016). *Virtual-Reality Movies: Get Ready for the VR Revolution*. Retrieved from <https://www.wsj.com/articles/virtual-reality-movies-get-ready-for-the-vr-revolution-1457030357>
- Sharma, C. (2010). *Global Mobile Data Market Update 2009*. Retrieved from <http://chetansharma.com/globalmarketupdate2009.htm>
- Simon, H. (1982). *The Sciences of the Artificial*. Cambridge, MA, MIT Press.
- Snickars, P., & Vonderau, P. (2012). *Moving Data: The iPhone and the Future of Media*. New York, NY, Columbia University Press.
- Stein, S. (2017). *Magic Leap's Mixed Reality Future Is Blurrier Than Ever*. Retrieved from <https://www.cnet.com/news/magic-leaps-mixed-reality-future-is-blurrier-than-ever/>
- Sterling, G. (2009). *Google Maps Plots Attractions, Points Of Interest*. Retrieved from <http://searchengineland.com/google-maps-plots-attractions-points-of-interest-23592>
- Sterling, G. (2016). *Mobile Devices Now Driving 56 Percent Of Traffic To Top Sites – Report*. Retrieved from <http://marketingland.com/mobile-top-sites-165725>
- Snow, B. (2011). *Look at the Numbers: The Mouse and Keyboard Aren't Dying*. Retrieved from <https://www.theatlantic.com/technology/archive/2011/08/look-at-the-numbers-the-mouse-and-keyboard-arent-dying/242939/>
- Swann, C. (2002). *Action Research and the Practice of Design*. *Design Issues*, (18)2, 49–61.
- Tarbert, K., & Wanis, B. (2016, August 1). *How Pokémon Go Changed Our Lives*. *Woman's Day* (Australia Edition), 18–19.
- Taylor, B. (2005). *The World is Your JavaScript-Enabled Oyster*. Retrieved from [https://googleblog.blogspot.co.nz/2005/06/world-is-your-javascript-enabled\\_29.html](https://googleblog.blogspot.co.nz/2005/06/world-is-your-javascript-enabled_29.html)
- Trice, A. (2013). *Performance & UX Considerations For Successful PhoneGap Apps*. Retrieved from <http://www.tricedesigns.com/2013/03/11/performance-ux-considerations-for-successful-phonegap-apps/>

- Uebele, A. (2007). *Signage Systems & Information Graphics: A Professional Sourcebook*. London, Thames & Hudson.
- Verto Analytics. (2016). *Pokémon GO: Thirty Days Later, is the Game Still Living Up to the Hype?* Retrieved from <http://www.vertoanalytics.com/pokemon-go-thirty-days-later/>
- Vuforia. (2016). *Vuforia.com*. Retrieved from <https://www.vuforia.com/>
- Walsh, D. (2012). *Camera And Video Control With HTML5*. Retrieved from <https://davidwalsh.name/browser-camera>
- Weight, J. (2014). Participatory Mobile Media Projects as Academic Practice. *Journal of Creative Technologies (MINA Special Issue)*, 4, 97-111.
- Weiner, K. (2016). *Why Image Recognition is About to Transform Business*. Retrieved from <https://techcrunch.com/2016/04/30/why-image-recognition-is-about-to-transform-business/>
- Xu, T. (2016). *A New Contender For Top Dog In The Augmented Reality SDK Market Has Appeared* [Web log message]. Retrieved from <https://www.linkedin.com/pulse/new-contender-top-dog-augmented-reality-sdk-market-has-tao-xu-%E5%BE%90%E6%B6%9B->
- Yokel, C. (2016). *Video API Documentation Overview*. Retrieved from <https://docs.microsoft.com/en-us/azure/cognitive-services/video/home>

## Figures

Fig 1. Artist unknown. (c. 1966). *Konica Super 8-6TL Camera Advertisement* [Print advertisement]. Retrieved from <https://www.thrillist.com/vice/best-1960s-classic-print-advertisements-cigarettes-alcohol-cars>

Fig 2. Rubery, P. (2017). *Project work*.

Fig 3. Design Council. (2017). *The Design Council Double Diamond* [Diagram]. Retrieved from <http://www.designcouncil.org.uk/sites/default/files/asset/document/innovation-by-design.pdf>

Fig 4. Independent. (2016). *Street Museum App* [Digital photograph manipulation]. Retrieved from <http://www.independent.co.uk/news/pictures/step-into-the-past-ghostly-hybrid-images-that-reveal-the-sights-of-london-then-and-now-9154557.html#gallery>

Fig 5. Brundle, P. (2016). *Collisions* [Digital photograph]. Retrieved from <http://press.sundance.org/press/category/new-frontier?page=11>

Fig 6. Robertson, A. (2016). *Collisions Screening VR Sundance* [Digital photograph]. Retrieved from <https://www.theverge.com/2016/2/9/10940236/sundance-virtual-reality-residency-lynette-wallworth-collisions>

Fig 7. Microsoft. (2016). *Microsoft Hololens* [Digital photograph manipulation]. Retrieved from <https://i.ytimg.com/vi/pLd9WPlAMpY/maxresdefault.jpg>

Figs 8–20. Rubery, P. (2017). *Project work*.

Fig 21. Wikitude. (2017). Wikitude Logo [Digital image]. Retrieved from [https://upload.wikimedia.org/wikipedia/en/a/ae/Wikitude\\_logo.png](https://upload.wikimedia.org/wikipedia/en/a/ae/Wikitude_logo.png). Augment. (2017). Augment Logo [Digital image]. Retrieved from [http://3.bp.blogspot.com/-uINRDtnt0kA/VZ8xpnUH7-I/AAAAAAAAABCs/kJ\\_xvvGvZks/s1600/AugmentDevLogo.png](http://3.bp.blogspot.com/-uINRDtnt0kA/VZ8xpnUH7-I/AAAAAAAAABCs/kJ_xvvGvZks/s1600/AugmentDevLogo.png). Vuforia. (2017). Vuforia Logo [Digital image]. Retrieved from <https://media.licdn.com/media/AAEAAQAAAAAAAAAdIAAAAJDJhMjlmMjRhLWNkZTUtNDU1NC1hOTk3LTMzYzQ1M2IxOTVIOA.png>

Figs 22–71. Rubery, P. (2017). *Project work*.