

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.

Hypervideo

Interface Design for Collaborative Documentaries

Exegesis submitted in partial fulfilment of the Master of Design
at Massey University, Wellington, New Zealand

Tim Turnidge

2014

Acknowledgements

Thank you to my supervisors Dr. Max Schleser and Antony Nevin for your guidance and support. Thank you to Julieanna Preston for facilitating the research process and for your feedback. Thanks to Klaus Kremer for always being around to bounce ideas off and for your help with the printed material. Finally, thank you to friends, family and my partner Ashley Kinsey for your love and support.

Abstract

Hypervideo is developed in response to the collaborative mobile-mentary 24 Frames 24 Hours. Through practice-led research an online interface is designed that creates a dynamic remix of user submitted and workshop generated videos. By means of leveraging contemporary web technologies such as APIs, metadata and video databases the interface presents an interactive documentary as a way of exploring innovative possibilities of web 2.0. The interface facilitates a unique viewing experience, which encourages new ways of experiencing and inspires the creation of mobile movies. This interactive documentary film form is influenced by Soft Cinema (Manovich, 2002) and the Korsakow system (Thalhofer, 2000).

The interface is developed through an iterative design process in response to the emerging significance of metadata in online viewing formats and the vertical/horizontal video design problem.

Keywords:

interface design, database cinema, interactive documentary, metadata, web APIs, mobile media

Contents

1.0 Introduction	7
1.1 Research Questions	9
1.2 Metadata	9
1.3 Web APIs	12
2.0 Research Process and Theoretical Framework	14
2.1 Creative Mobile Media	14
2.2 Contemporary Mobile Media	17
2.3 Database Cinema	19
2.3.1 Man With A Movie Camera	19
2.3.2 Soft Cinema	20
2.4 Interactive Documentary	23
2.4.1 Korsakow	24
2.5 Online viewing platforms	25
2.5.1 YouTube	25
2.5.2 The Original 24 Frames 24 Hours Platform	26
2.6 Interface Design for Mobile Media: Portrait, Landscape and Square	27
3.0 Hypervideo: Design Experiments and Dynamic Interfaces	34
3.1 MINA + Snapr App	35
3.2 Popcorn	35
3.3 24 Frames Popcorn	36
3.4 24 Frames 24 Hours Vimeo	36
3.5 YouTube API Experiments	37
3.6 24 Frames 24 Hours YouTube	38
3.7 Metadata Editor	38
3.8 24 Frames 24 Hours Split	39
3.9 Concept	40
3.10 Hypervideo: 24 Frames 24 Hours (previously titled Butter)	41

4.0 Conclusions	48
4.1 Future Developments and Research	49
5.0 References	53
5.1 Video references	57
5.2 Image references	58
Appendix A: Hypervideo Design Experiments and Dynamic Interfaces	60
Appendix B: Raw Vimeo metadata	88
Appendix C: Vimeo response to geolocation feature request	89
Appendix D: Raw YouTube video metadata	90
Appendix E: Notification of Low Risk Research/Evaluation Involving Human Participants	93

1.0 Introduction

Hypervideo is a practice-led research project which explores new web technologies through interface design to produce a dynamically-generated international collaborative mobile-documentary. The project begins by examining historical precedents within the fields of mobile media, locative media, database cinema and interactive documentary to inform the design of the interface. In addition to this, a number of design experiments are conducted to test the possible applications of contemporary web-technologies such as mobile video metadata and APIs within the creation of a collaborative documentary.

Hypervideo is created in response to the 24 Frames 24 Hours international documentary developed by Dr. Schleser. The Hypervideo project engages closely with the documentary as the 'Hypervideo: 24 Frames 24 Hours' interface serves as the platform which 24 Frames 24 Hours is assembled and viewed on.

"24 Frames 24 Hours is an international collaborative mobile documentary capturing the life in 24 hours in 24 different cities. The feature documentary asks participants to contribute personal stories about their cities, which are filmed on mobile devices. The project examines collaborative practices and applies Cinema verite and Kino-Pravda practices in the digital realm." (MINA, 2013)

The Hypervideo interface presents an innovative platform for experiencing the 24 Frames 24 Hours mobile-mentary (Schleser, 2011) through an interface which leverages upon new media concepts and recent developments in web technology. The interface shows how metadata, an often invisible but pervasive aspect of web media, can be used to enhance the content of mobile video by providing context for stories about communities.

The Hypervideo project positions itself within a creative mobile media practice aligned with festivals and conferences such as SXSW¹, MINA² and Power to the Pixel³.

Manovich describes Soft Cinema as generating “new cinema forms using the key technology of the information society - a digital computer” (Kratky & Manovich, 2005). In a similar way, the Hypervideo project aims to generate new video forms through interface design using key technologies of today: mobile devices, metadata and web APIs.

The title Hypervideo draws parallels to the hypertext and hypermedia concepts coined by Ted Nelson in 1965 (Joyce, n.d.). The prefix hyper is derived from the Greek “above, beyond or outside”. In the case of hypertext this refers to extending beyond the linear constraints of written text. The idea of hypertext was conceptualised long before it was technically possible to create. With contemporary technology and innovations in new media, what might modern day hypervideo look like?

1 <http://sxsw.com>
2 <http://mina.pro>
3 <http://powertothepixel.com>

1.1 Research Questions

The research process is initiated by three questions:

- How can interface design respond to the new formats of video created through mobile media?
- How can metadata and web APIs be used in an interactive documentary context to create connections for users and participants?
- How can new web technologies be used in the design of new online video forms?

1.2 Metadata

```
ube.com,2008:playlist:PLLi1FLPbGJ98B3gMynmeW0xuA4gO6Qd0K:PL2KLEVHEHDe97_kx_Vf
: "2013-01-15T02:12:20.000Z"}, "updated": {"$t": "1970-01-01T00:00:00.000Z"}, "c
chemas.google.com/g/2005#kind", "term": "http://gdata.youtube.com/schemas/2007
data.youtube.com/schemas/2007/categories.cat", "term": "Animals", "label": "Pets
", "title": {"$t": "Broccolicat"}, "content": {"type": "application/x-shockwave-fla
C4?version=3&f=playlists&app=youtube_gdata"}, "link": [{"rel": "alternate", "type
watch?v=RHE_msXeaC4&feature=youtube_gdata"}, {"rel": "http://gdata.youtube.com/
on/atom+xml", "href": "http://gdata.youtube.com/feeds/api/videos/RHE_msXeaC4/re
schemas/2007#mobile", "type": "text/html", "href": "http://m.youtube.com/details?
com/schemas/2007#uploader", "type": "application/atom+xml", "href": "http://gdata
xpR5Sw?v=2"}, {"rel": "related", "type": "application/atom+xml", "href": "http://gc
v=2"}, {"rel": "self", "type": "application/atom+xml", "href": "http://gdata.youtub
meW0xuA4gO6Qd0K/PL2KLEVHEHDe97_kx_VfOaD10_aILxwyqhemk4JckpsEY?v=2"}], "author"
data.youtube.com/feeds/api/users/ZealNZ", "yt$userId": {"$t": "VcH04ohsO6qlxFl3
ment", "permission": "allowed"}, {"action": "commentVote", "permission": "allowed"}
```

Fig 1. Example of metadata in JSON format.

Metadata is a central concept in this project because it enables the algorithmic editing⁴ used in Hypervideo to operate. Coined by Philip Bagley in *Extensions of Programming Language Concepts* (1968), metadata can be defined as “data about data”. In a web context, metadata is used to describe digital content (for example, a digital image or video). Generally, metadata provides information about an aspect of the data (content) such as:

4 Algorithmic editing is where editing is performed through pre-defined algorithms, executed by code.

- The creator or author of the data.
- The location where the data was created (geo-locative metadata)
- The time and date the data was created or modified.
- The file format of the data.

For example, the metadata for a digital video file might include: the video dimensions, the bit-rate, the duration of the video, subtitles, who the author is, when it was shot, what format it is saved as, among others.

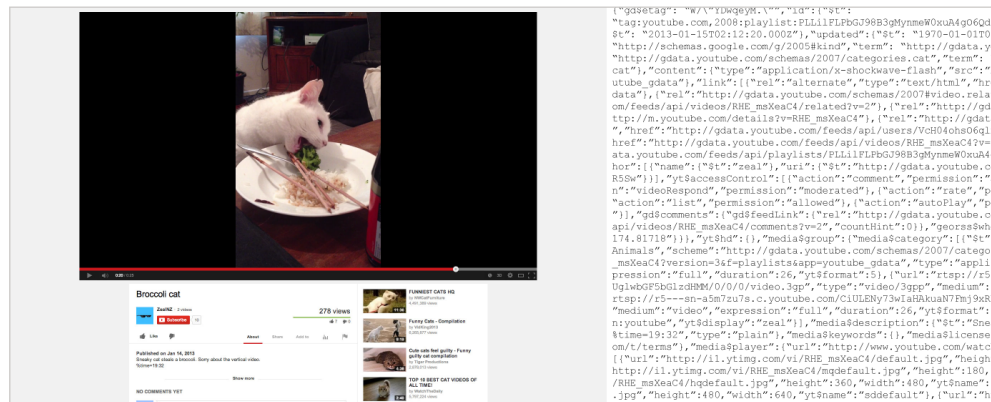


Fig 2. Example of a YouTube Video and metadata in JSON format.

Metadata is an important concept in algorithmic editing contexts such as database cinema or interactive documentaries, where editing is performed through pre-defined algorithms. Metadata is often used to determine the outcome of an algorithmic process, such as those used in the Korsakow System and in Soft Cinema. The Korsakow System, invented by Florian Thalhofer in 2000, is used to create non-linear video narratives, known as Korsakow-films, by using the SNU (smallest narrative unit) concept to apply metadata to media to determine which clips can link to one another (Thalhofer, 2013). Korsakow is discussed further in section 2.4.1 Korsakow.

Lev Manovich's Soft Cinema work "Texas", released in 2002, uses metadata to categorise spatial qualities of video clips and sound files with descriptions such as "city view", "private interior" or "public interior" (Kratky & Manovich, 2005). This metadata is then used in an algorithm to determine which media to display next in the sequence. Soft Cinema is discussed further in section 2.3.2 Soft Cinema.

Metadata is generated in multiple ways and at different points within the process of generating media. Using digital video as an example, some metadata is generated when the video is shot (for example: geolocate metadata - the location the video was shot) while other metadata is generated during editing (for example, the duration of the video). When a digital video file is uploaded to a video sharing platform such as YouTube, yet another layer of metadata is added containing information such as the video title, links to the author's YouTube Channel, whether the video is public or private, etc.

In Hypervideo, an additional layer of metadata is added when a video clip is added to the video database. This extra layer adds metadata that is specific to the 24 Frames 24 Hours documentary, for example, the theme that applies to the video. This metadata is used as the driving force behind algorithms at work behind the scenes of the interfaces which decide which videos to be shown and when.

The Hypervideo interface is designed in response to the types of metadata accessible through the Application Programming Interfaces (or APIs) of contemporary video platforms such as YouTube and Vimeo.

1.3 Web APIs

An API (Application Programming Interface) is a protocol which allows one system to access another system. Web APIs serve a similar function, allowing one website to access data or services hosted by another website. The web API has a fairly short history, with Google being the first large company to launch a JavaScript API in June 2005 with the Google Maps API (Google, 2005). Many contemporary web platforms such as YouTube, Vimeo, Twitter and Facebook offer web APIs allowing third parties to access their data⁵.

The YouTube API describes two levels of access: “Simple” and “Authorized” (Google, 2013). Both APIs allow read access to public data relating to videos uploaded to YouTube, YouTube channels, YouTube users, video playlists, comments posted to public videos, among others. The Authorized API also allows write requests to be made on behalf an authorised user, for example editing the title or description of a video (assuming the authorised user owns the video), liking a video, or subscribing to a channel.

Vimeo has a similar API to YouTube except that the two levels of access are labelled: “Simple” and “Advanced” (Vimeo, n.d.).

The Google Maps API offers “services and data related to mapping” (Google, 2013), such as the generation of interactive maps and markers. The API also provides services for geocoding (converting text addresses to coordinates), traffic information, turn-by-turn navigation, among others.

APIs enable interface designers to leverage the existing infrastructure and communities that online platforms such as YouTube and Vimeo already have in place. APIs also provide new possibilities for visualising metadata through tools such as Google Maps.

5 APIs can be used to both read and write data. Each API provider will have different specifications about what data can be read publicly, what data requires authorisation to be read and what data can be written (which always requires authorisation).

Hypervideo makes use of a number of APIs to explore ideas put forward in Manovich's Soft Cinema. The project has used the Google Maps API, to display the world map and plot the video markers in the correct location; and the Vimeo and YouTube APIs, to access the video databases and to display the videos and associated metadata within the interface.

2.0 Research Process and Theoretical Framework

2.1 Creative Mobile Media

This section begins with an examination of the history of creative mobile media as a way of placing the current state of creative mobile media practice. In the context of this project, the term “mobile media” refers to media created both on mobile devices and for mobile devices.

As 24 Frames 24 Hours uses video assets generated on mobile devices, it is important to understand their limitations and unique attributes to inform the interface design process. These characteristics set a starting point for the Hypervideo design process.

Creative works made on mobile devices are typically characterised by:

- Low pixel resolution (for visual works)
- Short duration (for visual) or short text (for textual)
- Close proximity to the subject (for visual works)

Historically, the low resolution of visual works created on mobile devices can be attributed to physical and technological constraints of the devices themselves. For example, camera sensors had to be scaled down to allow them to fit within the body of the mobile device, resulting in low image quality. Relatively low storage space (again a product of physical size constraints and the

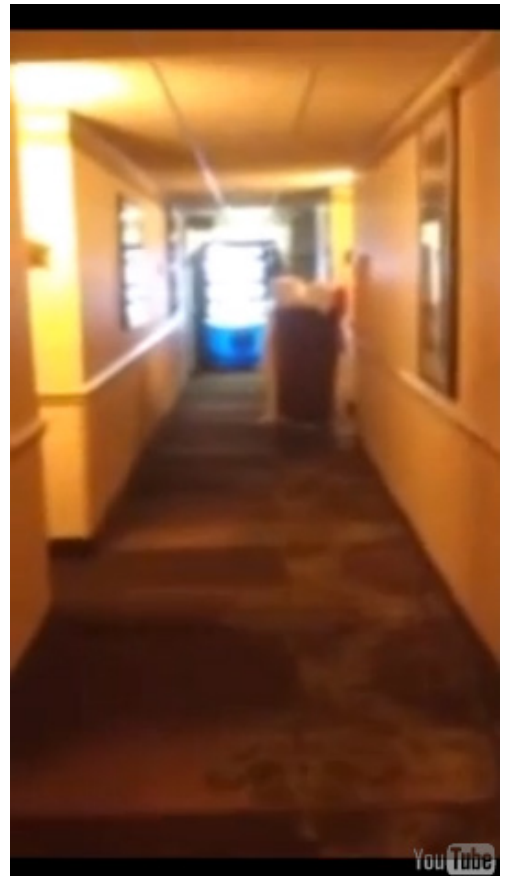


Fig 3. Still image from video produced in the 24 Frames 24 Hours filmmaking workshops.

technology of the time) restricted video to short durations and low resolutions.

The necessity of close proximity to the subject within visual mobile media works is a result of the physical form of the mobile device. Achieving a true optical zoom is not feasible due to the thin dimensions that are desirable in a mobile device, as there is no room for a lens to move toward or away from the subject in relation to the camera sensor. While a digital zoom is often offered as an alternative, this method of zoom simply enlarges and crops the captured image, further degrading quality and resulting in pixelation.



Fig 4. Still image from video produced in the 24 Frames 24 Hours filmmaking workshops.

The limited bandwidth over mobile networks has also contributed to the characteristics of media works created on mobile devices. Generally, mobile networks are slower and more expensive than the landline, copper cable or fibre-optic counterparts which puts a premium on any data that must be transferred in the creation or consumption of media. For this reason, within mobile media, there is an emphasis placed on keeping file sizes low to reduce the amount of data transferred. In the case of audio or visual works, a smaller file size is achieved by reducing resolution or bit-rate, using compression algorithms or by shortening the (temporal) length of the media.

In recent times technological constraints of the past have been lessened as better camera technologies allow for higher quality images; larger, cheaper storage allows for higher resolutions; and faster, cheaper networking allows for larger file sizes to be transmitted. However there still exists a gap in image quality between mobile device

cameras and dedicated cameras. In addition to this, many of the constraints still remain in effect even when the technological barriers no longer remain, for example the 140 character limit in Twitter tweets is a remnant of SMS character limits.

Videos created in the 24 Frames 24 Hours filmmaking workshops are shot on mobile devices and exhibit some of the mobile media characteristics discussed above. As Hypervideo engages with this mobile generated content, it is important that these characteristics are taken into account in the interface design process.

Hypervideo embraces the fact that the resolution of mobile media is lower than that of traditional broadcast media by displaying the videos within smaller video frames that approximate the dimensions of the mobile devices that the videos were shot on. Using smaller frames not only reduces the effect of lower resolutions in video but also allows for creative opportunities to be explored by combining multiple videos at once in spatial montage.

The geolocative abilities embedded into mobile devices is another defining characteristic. From triangulation using cellular network towers through to GPS, mobile devices have the ability to locate a user in space. In media created for mobile devices, this can be used to give the user an experience that relates to the location they are currently in. One example of this is the early mobile work *Stampede* (Raby & Dunne, 1998) which used WAP (wireless application protocol) to geolocate a user's position within the city in relation to cellular network towers. Users are able to follow a stampede of virtual reindeer as they move from cellular tower to cellular tower and could get virtually "trampled" if caught in the same cellular area as the stampede.

Similarly, media created on mobile devices can be tagged with geolocative metadata which is able to identify where that media was created. Hypervideo uses this geolocative metadata to place media within a location, establishing a link between the story and the place.

2.2 Contemporary Mobile Media

The Hypervideo project investigates how contemporary web technologies such as metadata and APIs can be applied to a mobile media project, specifically the 24 Frames 24 Hours documentary. Because of this, the design of Hypervideo is influenced by the constraints and conventions of contemporary mobile media.

What were originally constraints came to define the aesthetic for the emerging mobile medium. The character limit on SMS text messages and the shortening of audio-visual media works to save on file size encourages brevity: saying more with less. Writing in her article *Exploring Mobile Media Performance and Embodying the Network: Visual Communication Re-Mixed*, Camille Baker describes how these constraints can “become an asset, rather than a hindrance to image-making” (Baker, 2013).

The way that pixelation in video now comes to represent authenticity is in part a result of these aesthetic constraints. Baker writes that “the tension, the poor image quality, created by the compression and image artefacts, lends a rawness to the medium and gives it an authenticity, ‘realness’ or ‘liveness’, making it seem more personal” (Baker, 2013). Speaking at the 3rd Mobile Creativity and Mobile Innovation Symposium Dr Miriam Ross outlines qualities that differentiate mobile video from professional video and thereby represent an authentic, undoctored view as:

- Shaky, hand-held footage
- Lack of stage lighting
- Lack of cuts
- Wind noise
- Lack of closure
- Often shot in portrait orientation



Fig 5. Still from New Zealand's 3 News showing user generated content

This can be seen in news reports where user submitted content, with its pixelated aesthetic, is used to suggest a subjective human observer. This is in contrast with the clean, high-definition footage that is filmed by the regular film crew. In 24 Frames 24 Hours this aesthetic adds authenticity to the footage, reminding the audience that the videos are showing personal snapshots of real life from the perspective of everyday people.

In the context of interface design this means that dynamic systems must be created to cater for user submissions. In contrast to industry formats which are standardised, user and workshop generated content require a more flexible design approach.

The ubiquity of mobile devices means that mobile user generated content is more prolific than ever, speaking at the 3rd Mobile Creativity and Mobile Innovation Symposium, Dean Keep states that "The Kodak moment has been replaced by the mobile moment". These mobile devices mean that the ability to generate video content is more accessible than ever, especially in developing

nations which accounted for 5.2 billion of the 6.8 billion mobile cellular subscriptions in 2012 (International Telecommunication Union, 2013).

In addition to this, the mobile devices has some cultural connotations which are important influences on contemporary mobile media, such as the intimate nature of the phone. Susan Kozel describes mobile media devices as deeply personal: “[they] encourage or inhibit human exchanges. They are portable; they accompany us for hours, days and seasons; they span moods and activities, cycles and rhythms of life” (Kozel, 2010)

The proliferation of mobile devices and media generated by them creates opportunities for interfaces to be designed that are able to cater to the unique technical attributes of contemporary mobile media such as geolocate metadata and portrait orientation video. Hypervideo focuses on these unique attributes and explores how they can be applied in the creation of an interactive documentary interface.

2.3 Database Cinema

Database cinema and interactive documentaries are two media genres which share some aspects with Hypervideo. In traditional cinema, the narrative is generally fixed and linear, whereas in database cinema, the narrative is developed dynamically through the sequencing of discrete scenes which are pulled from a database. This sequencing can be random or generated through a series of rules or algorithms.

2.3.1 Man With A Movie Camera

Man with a Movie Camera is a 1929 documentary by Dziga Vertov. The film itself can be considered one of the earliest examples of database cinema as all

the scenes within the film were shot separately without a predefined narrative in mind, the titles at the start describe how the film is created “without the help of a scenario (a film without a scenario)” (Vertov, 1929).

In *Man with a Movie Camera*, discrete scenes are combined together in temporal and spatial montages that allow for new narratives to be drawn from the juxtaposition of imagery.

In a similar way, Hypervideo takes the individual videos created in the 24 Frames 24 Hours workshops and combines them together in temporal and spatial montages to allow for narratives and connections to be formed.

2.3.2 Soft Cinema

Soft Cinema is a new media project headed by Lev Manovich. Launched in 2002, Soft Cinema presents software which dynamically generates movies based on algorithmic editing principles.

Writing in *Soft Cinema: Concepts*, Manovich outlines four concepts that are behind the Soft Cinema project: algorithmic editing, database narrative, macro-cinema, and multi-media cinema.

In Soft Cinema, images, audio and video are tagged with keywords which describe the content of the clip as well as their formal properties. Through algorithmic editing, this information is used to determine which media to display next in the sequence, where in the frame it is placed, and for how long the media is played. Because the final sequence of media clips is generated dynamically, no two viewings of Soft Cinema are the same, however, the process of applying these keywords to the clips means that the author retains control over how the sequence unfolds.

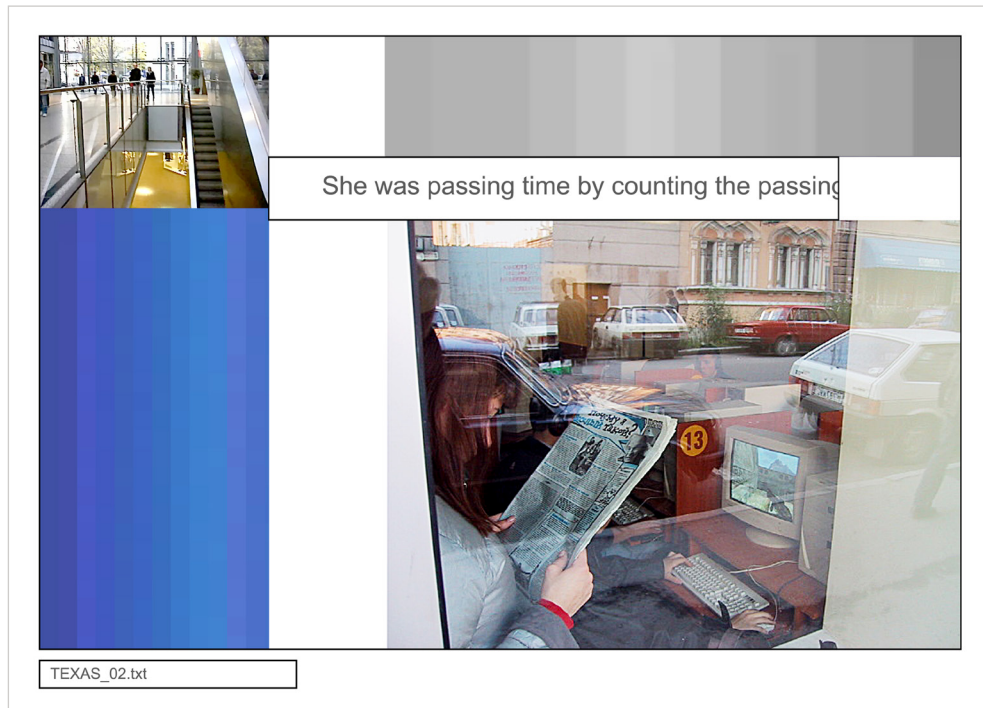


Fig 6. Still image from Texas Soft Cinema.

Even though the Soft Cinema software generates the movie on the fly, the editing rules used in the algorithmic editing system enable an author to retain some control over what ends up in the final output. For example, an author is able to specify that a particular video clip must be included, or that one particular clip must follow another.

The concept of database narrative is that rather than starting with a predefined narrative and generating content to fit, Soft Cinema is able to start with a large database of content and generate narratives through the algorithmic editing process.

The concept of macro-cinema is expressed by the division of the screen into multiple rectangular regions. This division is coordinated by the algorithmic editing. In this way, the system of rules set in place by the author are

responsible for both the temporal and spatial organisation of the work, i.e. both the sequencing of clips and their composition.

“Although the idea of spatial montage emerged rather early in film history, it was realized only in a few avant-garde films. Recently, however, it has become commonly used in the medium of television with many programs presenting several independent streams of information simultaneously on the screen.” (Kratky & Manovich, 2005)

Manovich describes the sources of inspiration for the idea of macro-cinema as graphical user interfaces used in computer systems, contemporary cultural sites (e.g. news and financial broadcasts, computer games), and the evolution of video production and distribution.

The multi-media cinema concept in Soft Cinema refers to the use of other modes of representation: 2D animation, motion graphics, stills, etc., used alongside “normal” video.

Hypervideo draws inspiration from Soft Cinema, and Manovich’s four concepts can be found within the interface. Hypervideo uses algorithmic editing based on metadata to determine which video clips follow one another. The workshop videos are all stored individually within YouTube’s database but combined together within the interface to create database narratives. The videos are displayed within resizing, movable frames which express the concept of macro-cinema. Hypervideo uses metadata to display other information about the video, for example, who the author is and where the video was taken which is a development of the multi-media cinema concept to include social media links between participants.

2.4 Interactive Documentary

In interactive documentaries, the viewer is able to interact with the system to influence their experience of the documentary. In some cases this means the user chooses what content to watch or in which order they watch it, which is the case with Korsakow-films. Other examples of interactive documentaries present other streams of information (such as images or hypertext pages) that the user interacts with while watching a traditional documentary video, which is the experience typically offered by Popcorn powered projects.

“The interactive documentary is partially authored by its collaborators, it is the result of layers of individual inputs that have generated other inputs within a technical environment that has facilitated a creative process.”
(Gaudenzi, 2010)

Interactive documentary is a genre which is rapidly evolving in response to changes in technology and new media culture. Writing in *Representing Reality: Issues and Concepts in Documentary*, documentary theorist Bill Nichols describes documentary as “a representation of the world we occupy” (1991). By making documentaries interactive, viewers become participants and are therefore encouraged to be part of the experience rather than passive consumers. This enables what documentary filmmaker Kent Bye describes as “collaborative sensemaking”, where by becoming part of the process, a collaborator is enabled to engage more closely with it (Bye, 2006).

There are a number of tools used for making interactive documentaries: Korsakow⁶, Klynt⁷, 3WDOC⁸, Popcorn⁹. Each tool allows the interactive documentary maker to specify which media is linked together and how the user interacts with the media.

6 <http://korsakow.org>
7 <http://www.klynt.net>
8 <http://www.3wdoc.com>
9 <http://popcornjs.org>

2.4.1 Korsakow



Fig 7. Korsakow logo

Florian Thalhofer's Korsakow (2000) is a software system used to create non-linear video narratives, known as Korsakow-films or K-films. These K-films are interactive films where a video clip is played to the user in a video frame, while a number of linked videos are displayed as thumbnail images to the side which Korsakow refers to as "clickable previews". The Korsakow System uses the SNU (smallest narrative unit) concept to apply a specific set of metadata to media to determine which clips can link to one another within an interactive film.

In applying this metadata to the clips (Korsakow uses the terminology SNU-ifying), the author of the documentary is able to specify which clickable previews might be shown and therefore influence the sequence of video clips that the end user will watch. However because of the randomness built into the Korsakow algorithms, the author never has full control over the edit the user sees.

In the 24 Frames 24 Hours interactive documentary, participants engage in mobile filmmaking workshops to produce video clips which shape representations about themselves and the communities they live in. These workshop generated videos feed into a larger collaborative documentary which positions participants in a larger context of creative collaboration. The documentary is generated within the Hypervideo interface through the users interactions and an algorithmic process by using metadata associated with the individual video clips.

Hypervideo shares some similarities with interactive documentary creation tools such as Korsakow and Klynt, in that they are both used to create an interactive documentary through a series of fixed and variable parameters.

2.5 Online viewing platforms

2.5.1 YouTube

Google's YouTube is the most popular video sharing platform in the world. Since its beginnings in 2005, YouTube now receives more than 1 billion unique views per month and has over 100 hours of video uploaded every minute (YouTube, n.d.). While YouTube has played a large role in shaping the contemporary online video field, the current interface does not accommodate the aspect ratios present in mobile video well.

YouTube's interface design primarily revolves around the video itself, the video's metadata (i.e. the title, author username, publishing date, video description), comments related to the video, and related videos. The videos themselves are always displayed in a landscape frame and use letterboxing or pillarboxing when the video image does not match the frame. This point and its significance for mobile media interfaces is discussed further in 2.6 Interface Design for Mobile Media: Portrait, Landscape and Square.

Browsing videos is generally performed through keyword based searches, where results relevance is determined by an algorithm which measures audience engagement using a number of metrics. Generally, videos which are more popular (i.e. have more views, are watched all the way through by viewers, have more "thumbs ups", among other metrics) will have more prominence in searches. These popular videos also gain more prominence when YouTube suggests related videos.

While watching a video on YouTube (and again at the end of the video), a set of related videos are presented. Which videos get promoted as related are calculated in a similar fashion to the search algorithm used when browsing. These related buttons can be compared to the clickable previews used by Korsakow in that they present linked media. The difference between them is that where YouTube related videos are algorithmically determined by popularity, Korsakow clickable previews are algorithmically determined using the SNU-concept.

The way that YouTube uses popularity as a primary driver these systems works to promote relevant videos works well, as evidenced by the popularity of YouTube. However, this system can work against cultural projects such as 24 Frames 24 Hours, where individual, personal stories can get lost amongst the popularity of more mainstream channels.

2.5.2 The Original 24 Frames 24 Hours Platform

At the beginning of this project, the 24 Frames 24 Hours videos were being displayed across a number of online services. The videos were displayed on a Wordpress site that displayed the videos in Vimeo video frames and on social media through Facebook.

The established infrastructure and interface afforded through Vimeo, Wordpress and Facebook offer an easy and immediate method of viewing and sharing the videos but do not make accommodations for unique aspects of mobile media such as geolocation or portrait video.

Displaying the workshop videos in the default Vimeo video frame forces the video into a landscape frame which results in pillarboxing when the video dimensions do not match. Furthermore the current interface made no real accommodation for grouping or sorting through the videos, making finding one specific video difficult.

24 Frames 24 Hours workshop participants use social media platforms such as Facebook¹⁰ as well as a Vimeo group¹¹ to facilitate collaboration. The Hypervideo interface presents links back to the author's profile (currently either YouTube or Facebook) as a way of augmenting these existing social networks.

Hypervideo aims to use APIs through interface design in the creation of a new video platform which better accommodates the unique properties of the 24 Frames 24 Hours videos as well as facilitating the collaborative nature of the documentary.

2.6 Interface Design for Mobile Media: Portrait, Landscape and Square

From the very first 4:3 aspect ratio of Thomas Edison's Kinetoscope, filmmaking has generally been performed in a landscape orientation. This in turn influenced cinema screens, televisions and computer monitors to follow this same format. On the other hand, the physical shape of mobile phones lends themselves to being used in portrait orientation which in turn means that many pictures and videos taken with mobile phones are also shot in the portrait orientation.

The video frame refers to the container that the video image is displayed in. In the case of a cinema projection or a traditional television set, the screen itself is the video frame and is fixed in its dimensions and orientation. In mobile phones and tablet computers the screen itself is the video frame and is fixed in its dimensions, but the orientation is changeable by physically rotating the device. In a web environment, the video frame is often a smaller container framed within one or more other frames (for example, browser windows).

¹⁰ The 24 Frames 24 Hours Facebook group can be found at <https://www.facebook.com/groups/222696181104451>

¹¹ <http://vimeo.com/groups/24f24h>

Videos look “correct” when the video image dimensions and video frame dimensions match as closely as possible. In the case of cinema screens, television sets and computer monitors, the dimensions and aspect ratio are fixed in their dimensions and aspect ratios. Because of this, video content has rightly been created with similar dimensions to ensure that the image displayed fits the intended video frame as closely as possible.

On the other hand, mobile phones and tablet computers also have fixed dimensions but have a variable orientation (because the device itself can be rotated), so video content is created without such rigid guidelines on orientation. In mobile media it is common to see video created in portrait, landscape and square orientations.

In the web environment, the video frame is a smaller element contained within a larger frame (the web browser window) which is itself contained within another frame (the monitor or laptop screen)¹². The video frame dimensions are specified in the design of the interface and could theoretically be of any dimension or orientation.

The majority of the complaints rallied against portrait videos can be attributed to the (landscape) frames they are displayed in, rather than the portrait video format itself. In the article *The Suppression of Vertical Videos*, James Watt writes that the “reason for disliking vertical videos was problems with playback.” (Watt, 2013). When video shot in portrait is then shown on a “traditional” landscape frame the options are to either crop the top and bottom of the image or to pillarbox the image.

Cropping the top and bottom edges of the video is undesirable because some of the image is lost. Another drawback to this method is that the resulting

12 When a web video is viewed in full screen, the monitor (or laptop screen, or projection, etc.) becomes the video frame itself.

image will either have much smaller dimensions or must be upscaled, resulting in a lowered image quality due to the enlargement of the image to fill the width.

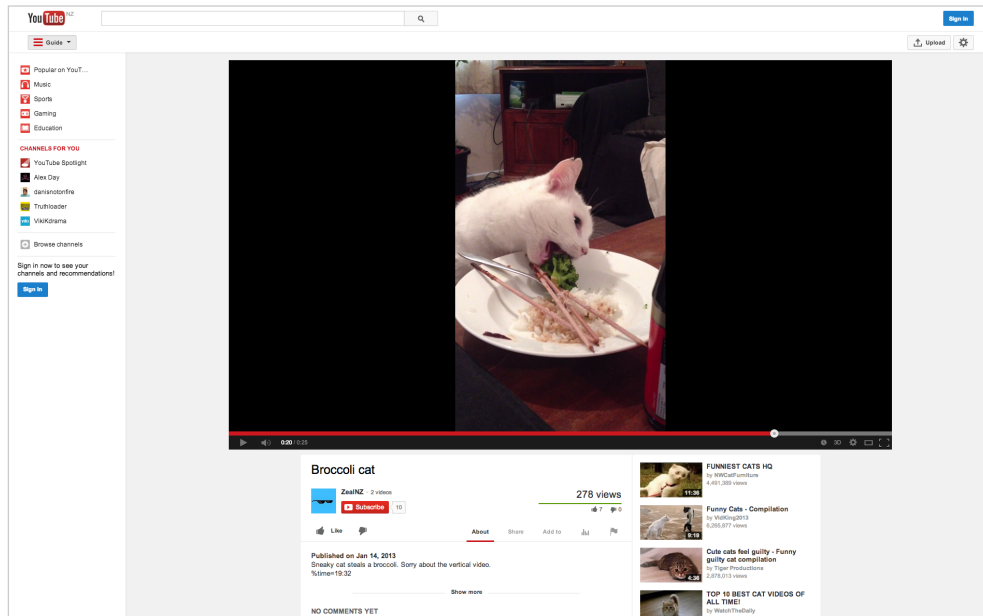


Fig 8. Portrait video viewed on YouTube shows pillarboxing.

The other option is to “pillarbox” the image, which is the method used most commonly on web-platforms such as YouTube and Vimeo. Pillarboxing involves shrinking the image to fit within the vertical bounds of the screen and adding black bars to the left and right of the image. This results in a scaled down image (as the image must be shrunk to allow the vertical height of the image to fit within the vertical height of the screen).

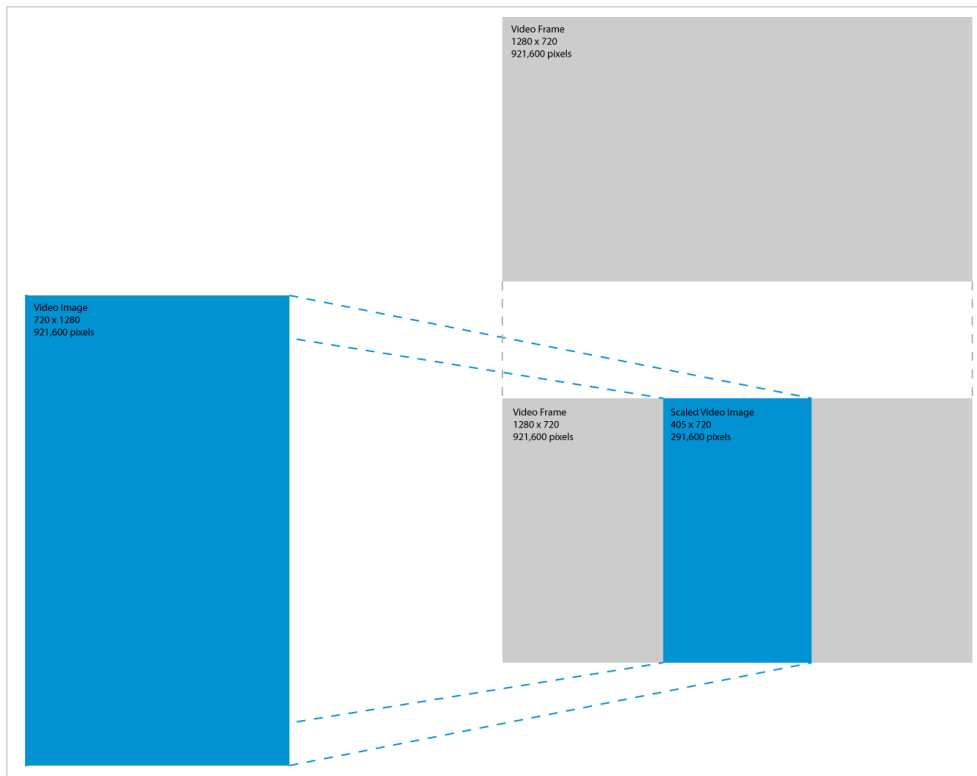


Fig 9. Portrait video resized to fit a landscape frame shows pillarboxing.

As shown above, a video and frame with the same area would require that the video image is scaled down to less than a third of its original size to make it fit the frame.

The pillarboxing that is used on YouTube causes portrait filmmaking to be seen as a faux pas. On almost any given portrait video with enough views it is possible to find comments such as “do you have a vertical display? No, I didn’t think so. Turn the fucking phone sideways!”¹³ and “Such a rare occurrence ruined by vertical shot syndrome”¹⁴.

13 YouTube user Verdoux007 commenting on Street Musician Sings (Onkel Bernis Welt, 2013)

14 YouTube user The Batman commenting on Street Musician Sings (Onkel Bernis Welt, 2013)



Fig 10. Screenshot of Vertical Video Syndrome - A PSA

The term Vertical Video Syndrome (VVS) was coined in a YouTube video titled "Vertical Video Syndrome - A PSA"¹⁵ uploaded by YouTube user gloveandboots on 5th June 2012. Within the first year, the video gained over 3.2 million views. Since then, this video has become a manifesto of sorts against vertical videos. Since then, a number of other videos are being uploaded that support this sentiment, for example, Turn Your Phone! ("No Scrubs" parody with Andrew Huang, DailyGrace, Hannah Hart, Soundlyawake)¹⁶ from YouTube user chescaleigh and Turn Your Phone! (Vertical Video PSA) (Song A Day #1647)¹⁷ from YouTube user Jonathan Mann.

15 <http://www.youtube.com/watch?v=Bt9zSfinwFA>
 16 <http://www.youtube.com/watch?v=eYLL7-rUGPY>
 17 http://www.youtube.com/watch?v=NtQWp_31wul

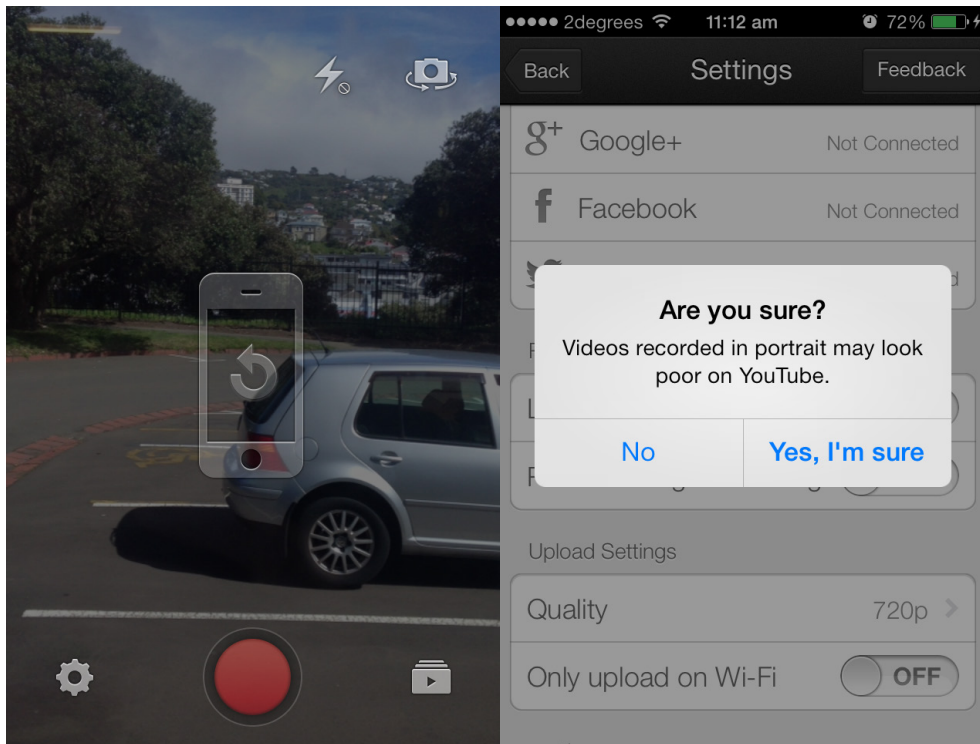


Fig 11. Screenscaps from YouTube Capture app on iOS.

This bias can also be seen in the YouTube Capture app on iOS (as of Oct 2013). The app has a specific option which prevents a video from being recorded while the device is held in portrait. There is an option within the app settings which can be disabled to allow portrait video to be shot, however, the option is enabled by default, and when a user tries to disable it, they are prompted with a warning which asks: “Are you sure? Videos recorded in portrait may look poor on YouTube”.

Shooting in portrait orientation does have some benefits as multiple portrait images can be placed next to each other to create a diptych or triptych to compare multiple videos while still fitting within a “traditional” landscape frame. Three portrait videos (9:16) arranged horizontally approximates a

landscape, 16:9 ratio. This method is explored in the Split iteration of the interface. While this approach can work well with vertical videos, it does not make any accommodations for the landscape or square videos.

Another method of working with media of variable dimensions is to display the videos within a resizable frame, which could theoretically display any aspect ratio. Facebook's interface design allows for a flexible frame which can accommodate different aspect ratios. In the Hypervideo interface, the videos are displayed within frames that have one of three aspect ratios: landscape, portrait and square. This is a technical compromise because it is currently not possible to query the aspect ratio or dimensions of a YouTube video through their API. In Hypervideo, each video is assigned one of the three aspect ratios individually by hand using the interface discussed in section 3.7 Metadata Editor.

There are a number of artists exploring the creative possibilities of portrait filmmaking. Founder of the global music and media foundation MIMA, Christoph A. Geiseler filmed, edited and produced the vertical short film "Curry Power" as a way of exploring the aesthetics of vertical video. Writing in his article *Improvisation: Vertical Videos Offer iPad Users a New Experience*, Geiseler shares his insights about how vertical video and mobile devices such as the Apple iPad "offer new opportunities for filmmakers willing to improvise" (Geiseler, 2012).

The problem with portrait videos does not lie in the format itself, but with the frames that they are displayed in. Videos created on mobile devices pose an interesting challenge for interface design through the variety of aspect ratios they present. The Hypervideo interface embraces the variance in aspect ratios by presenting videos in resizable frames which can then be used to create spatial montages inspired by Manovich's "macro-cinema" concept explored in section 2.3.2 Soft Cinema.

3.0 Hypervideo: Design Experiments and Dynamic Interfaces

“New media art is science fiction. It operates by extrapolating cultural vectors that are technologically inflected”. (Kratky & Manovich, 2005)

Over the course of this project, a number of experimental prototype interfaces have been created to test technologies and concepts in the pursuit of new media forms. I use this section to reflect upon the technological insights; design decisions; and conceptualisation gained and used knowledge in the next iteration of the interface or the next experiment.

The majority of the initial experiments involved technical explorations into contemporary web technologies as a way of examining the existing field of new media interfaces. Later experiments focused on creating an interface targeted at the 24 Frames 24 Hours documentary.

Summaries of the goals and results of the experiments are presented in the following section. More detailed discussions can be found in Appendix A.

3.1 MINA + Snapr App

Technology: iOS, HTML, CSS, LESS, JavaScript, jQuery, jQuery Mobile, GPS

This experiment involved creating an interface for a geolocate app using the Snapr platform. This was an opportunity to explore how mobile media can be geolocated in an interactive interface. Through this experiment, insight was gained into the process of using GPS metadata pulled from image EXIF data to place media within a location, which initiated the investigation into metadata used in other experiments. This experiment also highlighted some of the limitations of web technologies such as lack of support for Flash (used in some APIs) and iOS's inability to play multiple videos at once (which becomes important in later interfaces).

3.2 Popcorn

Technology: HTML, CSS, JavaScript, jQuery, video.js, Popcorn

This experiment explores two JavaScript technologies, Mozilla Popcorn.js and Video.js. These two technologies are recent developments in the area of web video and are explored to see what possibilities they might hold for application in an interactive documentary context.

Popcorn shows potential for enhancing the video content through supplemental media such as images and text. This can be used to provide extra information or context to video. In further experiments, supplemental media is generated from the metadata associated with video media and displayed to the user to provide context for the video (such as the location and time of shooting) and to enable connections between workshop participants (through links to social media).

Video.js can be used to fix inconsistencies in web video between browsers, however, future interfaces use Vimeo and YouTube APIs for video display.

3.3 24 Frames Popcorn

Technology: HTML, CSS, JavaScript, jQuery, video.js, Popcorn, Vimeo API

This experiment applies Popcorn technology to video content created in the 24 Frames 24 Hours filmmaking workshops. Working with a pre-edited video triptych, information related to each video clip used is displayed below while that clip is visible. The information displayed: video title, author (name, social media link and profile picture) and location is used to contextualise the video clips as they appear in the edit. The result is a video interface where the user is able to trace each clip back to its origin.

3.4 24 Frames 24 Hours Vimeo

Technology: HTML, CSS, JavaScript, jQuery, Vimeo API, Google Maps API

This experiment explores the creation of an interface to enable viewing the viewing of the videos shot in the 24 Frames 24 Hours mobile filmmaking workshops. The interface uses the Vimeo API to access the workshop videos that were already uploaded to Vimeo and display them within a custom interface. The interface uses a customised set of metadata, accessed through the Vimeo API to display the videos on an interactive map and timeline.

The use of an interactive map allows the link between the video narrative and the location to be reinforced and contextualises 24 Frames 24 Hours as a global documentary.

The videos are also plotted on a timeline, indicating the time that the video was shot. This allows for another method of interacting with the database of video content, as the user is able to filter their experience of the documentary to only show a specific time, for example.

The videos were displayed in the default, landscape frame offered through the Vimeo API. This causes issues with the 24 Frames 24 Hours footage which was (at the time) mostly filmed in portrait. This issue sparked much of the research discussed in section 2.6 Interface Design for Mobile Media: Portrait, Landscape and Square.

This experiment revealed that the interactive map was the most successful element and that users tended to interact with the map over the timeline. Future interfaces place a higher priority on the location of media over the time of media.

3.5 YouTube API Experiments

Technology: HTML, CSS, JavaScript, jQuery, YouTube API

The next three experiments were initiated to explore the YouTube API, to learn its limitations and suitability for application to the 24 Frames 24 Hours documentary. These experiments all focus on examining how and what metadata can be accessed through the YouTube API rather than the design of interfaces for an end user.

These experiments confirmed that the YouTube API would provide a more suitable platform for the 24 Frames 24 Hours documentary.

The YouTube API provides a wider range of metadata than the Vimeo API that was used in previous experiments. One of the main points of difference is that YouTube supports geolocation for videos, which provides an established infrastructure for storing geolocation metadata.

3.6 24 Frames 24 Hours YouTube

Technology: HTML, CSS, JavaScript, jQuery, YouTube API, Google Maps API

This experiment began by taking the interface from the 24 Frames 24 Hours Vimeo experiment and changing the backend to use the YouTube API instead.

This interface presented the videos in a portrait frame in response to the portrait videos which had been submitted from the 24 Frames 24 Hours workshops. During development of this interface, another round of filmmaking workshops were held in New Jersey, resulting in new videos being added to the database. While the majority of videos were submitted in a portrait orientation, there were also some submitted in landscape and square formats. These caused problems when displayed within the portrait frame used in this interface, prompting a re-think of the video frame aspect ratio problem.

The interface also allows for the creation of playlists of videos. Through this the focus is shifted from the videos acting as individual pieces of content to the videos being part of a larger narrative. This shift initiated the research into database cinema and interactive documentaries discussed in sections 2.3 and 2.4.

3.7 Metadata Editor

Technology: HTML, CSS, JavaScript, jQuery, YouTube API, Google Maps API, PHP, MySQL, Couch DB

In working on the previous experiments, I identified that the metadata that is accessible through the APIs will never be able to accommodate all the metadata that I want to use in the 24 Frames 24 Hours documentary in an efficient, consistent manner. Furthermore, even with APIs that can accommodate the metadata required for the documentary (YouTube's geolocation metadata, for example) there is no guarantee that the author

of the video will actually tag the video with all the required metadata while uploading.

This experiment presents an interface that reads the metadata for a YouTube video and allows a user to edit or add their own metadata, which gets stored in the database and used in the Hypervideo interface.

3.8 24 Frames 24 Hours Split

Technology: HTML, CSS, JavaScript, jQuery, YouTube API, Google Maps API

In *Memory, Schema and Interactive Video*, Legget describes how an “audience, encountering a system in a public place and without prior knowledge, will need to be encouraged and guided within an interactive navigation schema” (2011). This experiment presents a reworking of the 24 Frames 24 Hours YouTube interface which breaks down the process of generating a video playlist into a step by step process designed to make the users options clearer and ultimately make for a more accessible experience.

The resulting video playlist is used to display the videos in a triptych of portrait video frames which resembles the pre-edited triptych discussed in section 3.3 24 Frames Popcorn. This arrangement of videos allows for the juxtaposition of imagery through spatial montage, enabling users to create associative chains between the videos where “meaning and narrative coherence can be created”. (Kratky & Manovich, 2005)

This arrangement of video frames allowed for interesting visual interplay to occur between videos, however, the fixed, portrait frames causes problems with videos of differing aspect ratios.

3.9 Concept

Technology: HTML, CSS, JavaScript, jQuery, YouTube API, Google Maps API

This experiment presents an interface where each video was displayed in its own individual video frame. This was in response to the problems encountered with the fixed video frames in 24 Frames 24 Hours Vimeo, 24 Frames 24 Hours YouTube and 24 Frames 24 Hours Split.

The interface also incorporated algorithmic editing as a way of generating video playlists, rather than requiring the user to generate them. Where Korsakow relies on SNU-ifying content to determine which video clips link to one another, this interface uses the location, time and theme metadata to make links between videos. This allows for links between videos to be generated dynamically, meaning a user only has to choose a single starting video and the interface will automatically select related videos.

3.10 Hypervideo: 24 Frames 24 Hours (previously titled Butter)

Technology: HTML, CSS, JavaScript, jQuery, YouTube API, Google Maps API

This interface is the latest iteration of the project taking into account all of the results of previous experimentation. The goal was to create an interface that enabled viewing of the 24 Frames 24 Hours workshop generated video through an interactive documentary mode which embraces the unique qualities of mobile video.

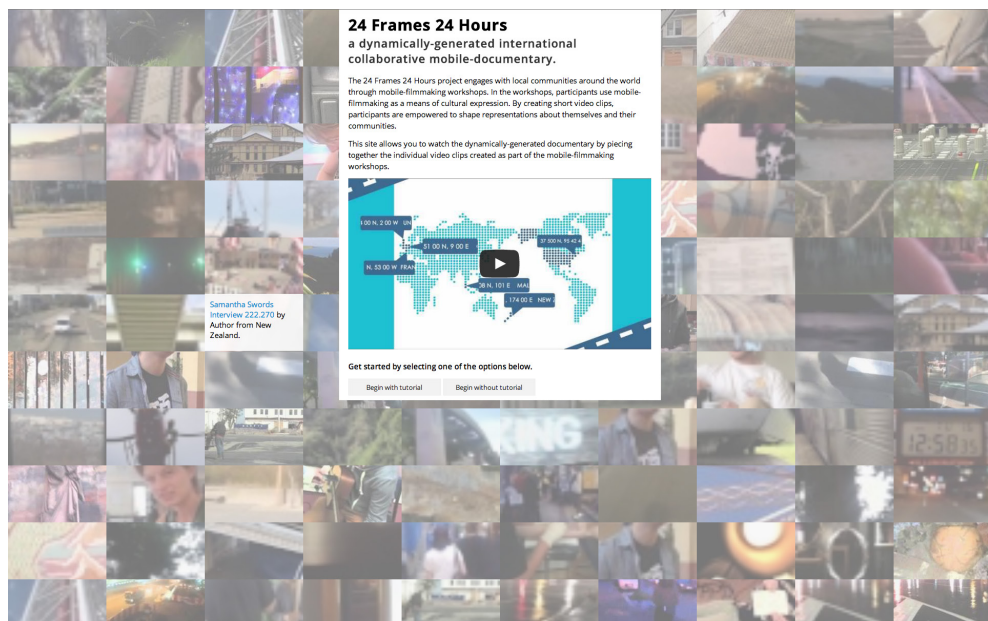


Fig 12. Hypervideo: 24 Frames 24 Hours interface introduction screen

The main browsing and viewing interface is preceded by an introduction and optional tutorial. The introduction outlines the main purpose of the interface and features the 24 Frames 24 Hours trailer video. In the background, image thumbnails from videos in the database are displayed which can be hovered over to reveal some of the metadata associated with that video.

The introduction ends with a pair of buttons where the user chooses whether to begin with the tutorial or to skip it. The button design that is introduced here is carried on throughout the interface in an effort to provide the user with consistent interaction cues.

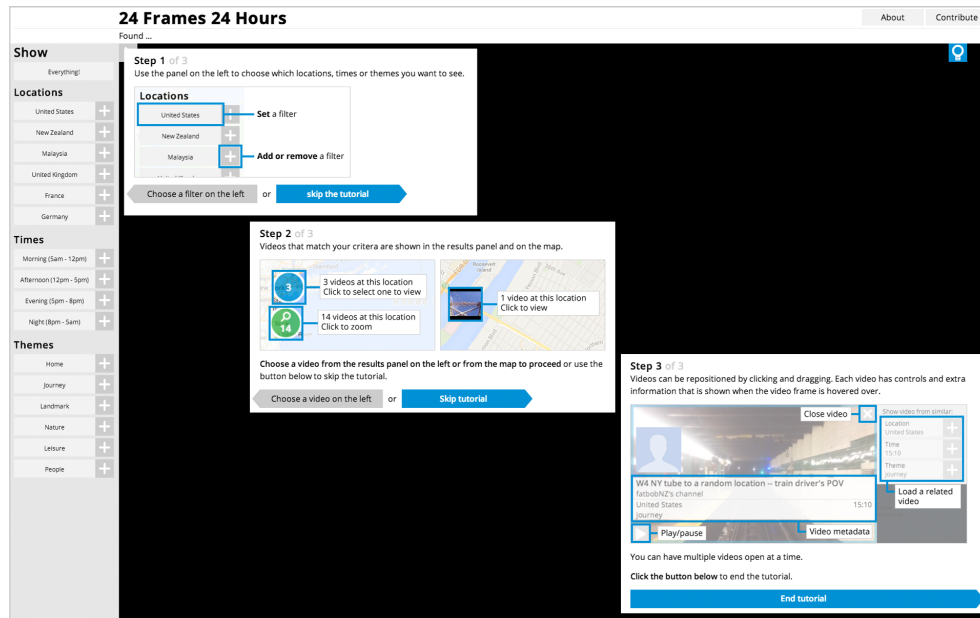


Fig 13. Hypervideo: 24 Frames 24 Hours interface tutorial screen. Tutorial screens displayed simultaneously for illustrative purposes, usually displayed consecutively.

The tutorial breaks down the filter selection process into a step-by-step process in response to positive feedback regarding the similar step-by-step process in the 24 Frames 24 Hours Split interface.

Each step presents the user with a short text instruction and an accompanying image. The user progresses through the tutorial by following the instructions and interacting with the interface itself so that by the time they reach the final step, they have set the interface in motion by selecting some filters and loaded a video.

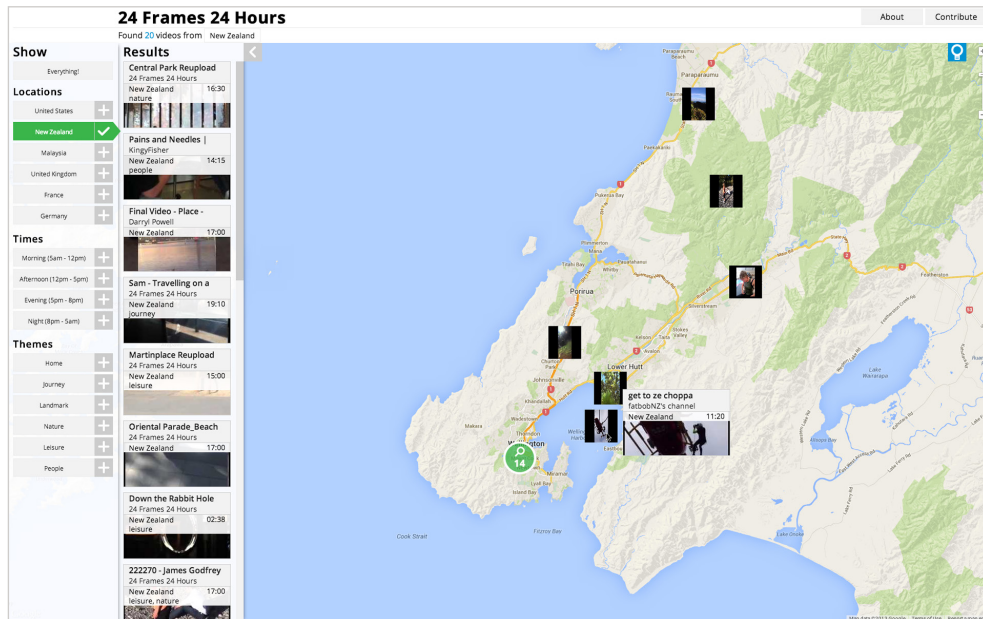


Fig 14. Hypervideo: 24 Frames 24 Hours interface set to show videos from New Zealand. One marker is hovered over to reveal the associated metadata.

The interface is divided into two main sections. The primary element in the interface is the (Google Maps) map which contains markers representing geolocated videos. Each video is represented by a marker on the map that shows a thumbnail image from the video. Where multiple markers occupy the same space, the markers are grouped into a compound marker. These compound markers can be clicked to expand out, showing the markers grouped at that location. Hovering over a marker reveals a larger thumbnail image overlaid with the metadata associated with that video: title, author, time of shooting, location and themes. Clicking a marker opens that video in a new video frame.

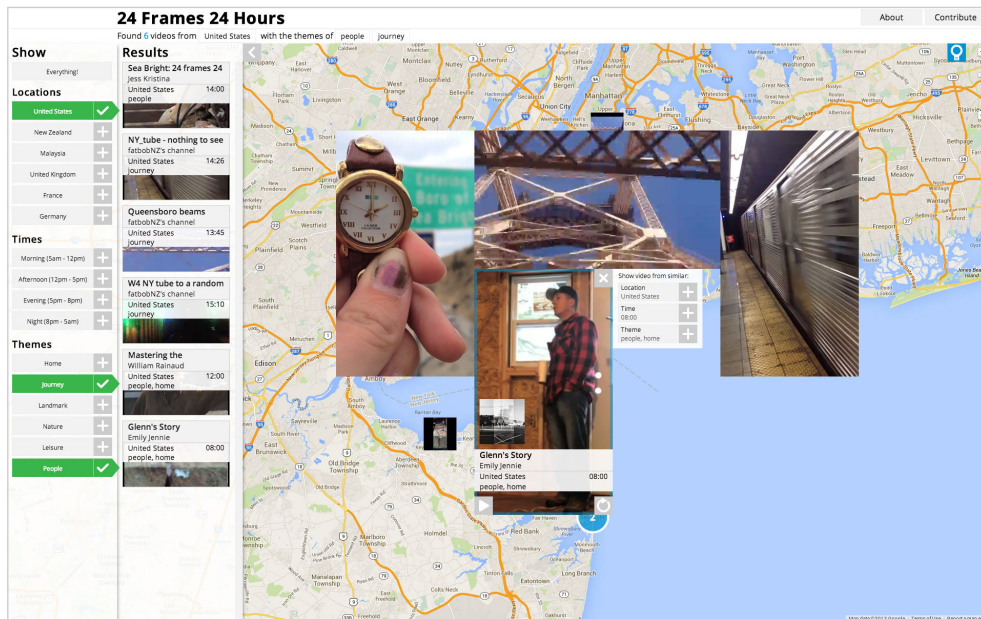


Fig 15. Hypervideo: 24 Frames 24 Hours interface set to show videos from United States with themes of journey and people. 4 video frames open. Active video frame (bottom centre) has panel with related video buttons expanded.

Each video is displayed in its own individual, floating frame. By making the frames independent from one another, the frame's aspect ratio can be adjusted to accommodate any video, portrait, landscape or square. This format also creates opportunities for interesting serendipitous juxtapositions to occur, where two videos might be from opposite ends of the world, yet show similar themes or feature similar movements within the video. This format is inspired by the "macro-cinema" concept used in Soft Cinema, where Manovich uses multiple frames as a way of exploring how cinema might utilise the idea of multiple windows encountered in contemporary graphical user interfaces.

When multiple video frames are presented, the last frame to be interacted with becomes the active video in terms of audio, a mechanism similar to the one used in Split. This is indicated by a semi-transparent blue border.

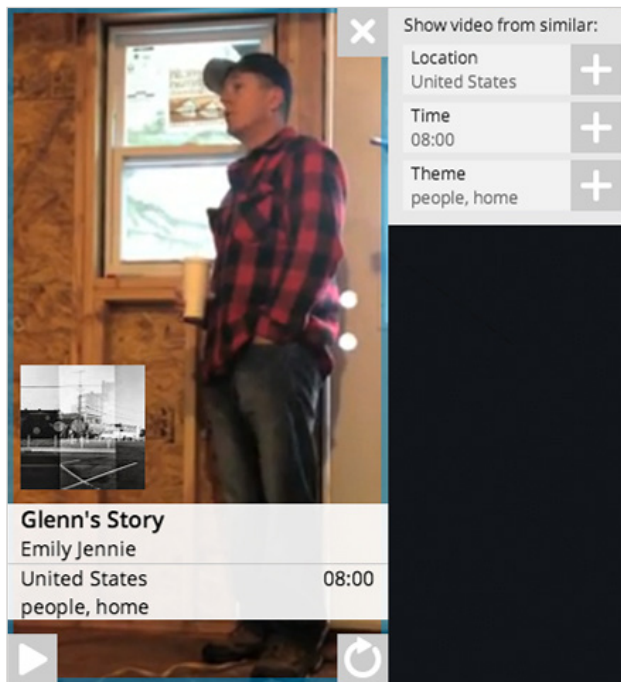
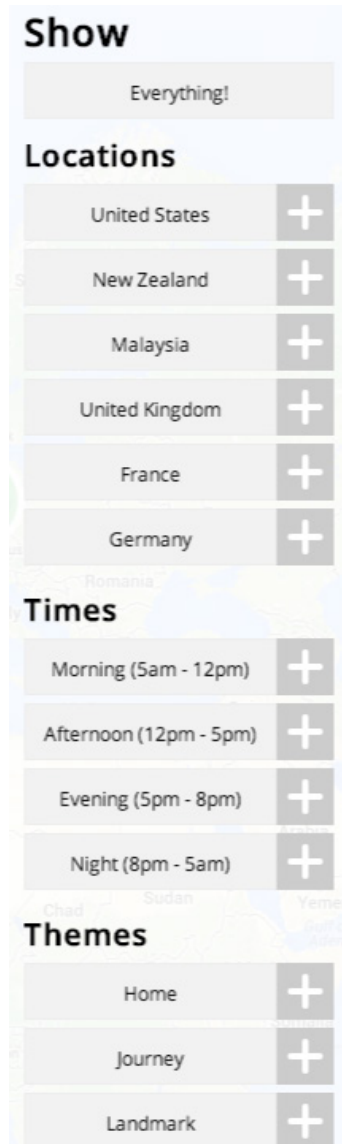


Fig 16. Hypervideo: 24 Frames 24 Hours interface, zoomed in on one video frame to show metadata and controls.

Hovering over a video frame will reveal the metadata for that video: title, author, time of shooting, location and themes. The metadata displayed shows the author's name and their YouTube portrait which are hyperlinked to the author's YouTube channel or Facebook page which enables connections to be formed through the interface, using already established social media infrastructures.

In addition to the metadata, a set of controls for the video are shown: a toggle button for playing and pausing; a toggle button to control whether the video loops when it is finished or whether it moves to a new one; a button to close the video frame; and a set of buttons to load related videos (shown to the top right of the video frame).

These 'related video' buttons serve a similar purpose to the buttons (Korsakow uses the terminology "clickable previews") presented in Korsakow-films. In Korsakow, image previews are presented to the user which act as clickable links to other videos. The buttons in Hypervideo are also used as links to other video, clicking one will load a randomly selected related video from the video database into the active video frame.



The left side of the interface contains two panels that relate to searching or filtering videos. The leftmost panel presents locations, times and themes that can be used to filter the videos that are included in the documentary. The locations relate to the locations that filmmaking workshops have been held: New Zealand, Malaysia, United States of America, Germany, France and United Kingdom. The times relate to the time of day the video was shot, categorised into one of four sections of the day: morning, afternoon, evening or night. The themes respond to the types of content found in the workshop videos: home, journeys, people, nature, landmarks.

Each category for location, time or theme, is comprised of a compound button with two sections allowing a user to either set the filter to that category by clicking the main section of the button, or toggle add/remove that category by clicking the smaller, inset button on the right. This enables the quick selection of a single category, or for any combination of categories for comparison and juxtaposition.

Fig 17. Hypervideo: 24 Frames 24 Hours interface filter panel.

The overall experience offered to users is one where they are able to browse the database of videos either through the map or by using the panel on the left to narrow their search first. When a video is loaded into a video frame, the user can interact with it to explore the metadata attached to it, gaining insight into where it was created and whose story it is telling.

When a video finishes, another video from the same location is automatically loaded into the frame, creating a continuous stream of video content. The user is able to intervene by using the related video buttons attached to each frame to divert the story down a different, yet related, path.

Ideally, the interface would provide the ability to contribute to the project by either uploading a video, through this site by utilising the YouTube API; or by choosing a video that is already uploaded to YouTube and editing its metadata. However, as discussed in section 3.7 Metadata Editor, the underlying database architecture, API and associated security and privacy concerns mean that the interface in its current state does not allow for this functionality.

The interface can be accessed online at www.24frames24hours.org.nz.

4.0 Conclusions

Hypervideo generates innovative interactive video forms through the combination of interface design and the underlying algorithms by using key technologies of today: mobile devices, metadata and web technologies such as HTML5, JavaScript and web APIs. These new technologies are used to explore Soft Cinema concepts of algorithmic editing, database narrative, macro-cinema, and multi-media cinema.

The creation of a custom web video database and associated metadata enables for the exploration of database narrative through algorithmic editing. Hypervideo uses custom metadata generated in collaboration with the 24 Frames 24 Hours documentary as the primary means of driving the database narrative to enable the dynamic generation of global community stories.

The 24 Frames 24 Hours documentary uses the mobile phone as a tool for creative storytelling which “enables ordinary, non-artists and non-performers to become artists” (Baker, 2013). By harnessing the geolocate metadata that is inherent in mobile media, connections between workshop participants can be facilitated through the use of pre-existing social networks. By using geolocate metadata as the primary method of browsing video content, the locative nature of mobile video is emphasised, cementing each video within a place and empowering the author to make a statement about a location.

The unique aesthetic of the mobile video format, characterised by varied video aspect ratios and typically lower image resolutions can be used as a tool for expressing authenticity within a new media form. The different aspect ratios present in mobile media can be used as a counterpoint to the rigid standards found in mainstream broadcast media and used in the creation of spatial montage inspired by macro-cinema which allows for serendipitous juxtapositions and connections to be made between videos.

APIs such as those provided by YouTube and Vimeo offer exciting possibilities for exploring database cinema concepts when combined with metadata and algorithmic editing. Where Korsakow and Soft Cinema use a fixed database of content, Hypervideo enables the collaborative generation of video content through the established infrastructure of video sharing platforms such as YouTube.

The design research explored in Hypervideo extends beyond the interface design that the user interacts with. Its underlying architecture and algorithms demonstrate innovation in interface design for mobile collaborative documentaries. Using APIs alongside JavaScript code enables all of the systems beneath the surface of the Hypervideo interface to operate in a way which would otherwise not be possible. At the point where YouTube comes to its limitations, Hypervideo enables new connections. Through the unique combination of interface and algorithm design used in Hypervideo, the project demonstrates how creative coding can generate new media forms.

4.1 Future Developments and Research

The next step in developing this interface would be to change the underlying database to use a more robust, developed architecture such as Apache MySQL, CouchDB or MongoDB. While this would not result in any major changes to the surface interface, it would allow individual users to submit their own videos and edit their own metadata. This would mean that I would not need to operate as a facilitator to getting content into the Hypervideo interface and provides workshop participants with an unhindered workflow from shooting to editing to viewing. Making this change would require collaboration with a developer experienced with one of the above databases to ensure a robust, secure solution.

Further developments to the interface could provide a more engaging experience for users by making the content more easily accessible. Because the interface presents a unique interactive experience, tutorialisation is an important aspect of enabling users to understand how to use the Hypervideo interface. Further user testing around the latest iteration of the interface would enable development of a more accessible interactive system.

An ongoing collaboration between myself as an interface designer and the 24 Frames 24 Hours workshop participants can continue to evolve the interface as more content is generated. The database architecture set in place and tested with current content allows for the easy expansion of the project to include more videos across more locations and with more themes.

As APIs are updated and the data available through them changes, the functionality available in further interfaces also changes. For example, if the YouTube API were updated to include video image dimensions in the video metadata it would be possible to create an interface which automatically reconfigures itself to best accommodate the video. This would eliminate the need for users to specify the aspect ratio metadata for the video manually and enable an ideal match between video image size and video frame size.

24 Frames 24 Hours and Hypervideo have been presented at a number of conferences and festivals including the 2nd Mobile Creativity and Mobile Innovation Symposium 2012 in Wellington, the 3rd Mobile Creativity and Mobile Innovation Symposium 2013 in Auckland and Mobile HCI 2013 in Munich, Germany (where the project was awarded 2nd place in the design competition). Engagement with these communities has been invaluable for informing the design theory surrounding the project. Through further engagement with creative communities such as i-docs and SXSW, Hypervideo can contribute to the discussions surrounding new media genres.

Finally, the Hypervideo interface could have applications for creative, collaborative projects other than 24 Frames 24 Hours. While the interface design and metadata used in this iteration of Hypervideo is specifically targeted towards the 24 Frames 24 Hours documentary, the underlying architecture is flexible enough that it could be applied to other contexts.

5.0 References

- Aston, J., & Gaudenzi, S. (2012). About i-docs. Retrieved from <http://i-docs.org/>
- Bagley, P. (1968). Extension of Programming Language Concepts. Philadelphia, Pennsylvania. University City Science Center, University City Science Institute.
- Baker, C. (2013). Exploring Mobile Media Performance and Embodying the Network: Visual Communication Re-Mixed. Creative Mobile Media Practices. Brunel University, London.
- Bye, K. (2006). Building a Theory of Collaborative Sensemaking. Echo Chamber Project. Retrieved from <http://www.echochamberproject.com/>
- Cooper, E. (2013). Video of fight causes outrage. 3news Blog. Retrieved from <http://www.3news.co.nz/Video-of-fight-causes-outrage/tabid/423/articleID/327359/Default.aspx>
- Gaudenzi, S. (2010). Interactive Documentary: Exploring new ways of Documenting Reality. Retrieved from http://www.interactivedocumentary.net/wp-content/2009/07/sg_panel-yr3_2009_ch1.pdf
- Geiseler, C. (2012). 10 Vertical Video Insights & Tips. Retrieved from <http://www.verticalvideos.com/>
- Geiseler, C. (2012). Improvisation: Vertical Videos Offer iPad Users a New Experience. Huffington Post. Retrieved from http://www.huffingtonpost.com/christoph-a-geiseler/ipad-vertical-videos_b_1540901.html

- Google. (2005). The world is your JavaScript enabled oyster. Google Official Blog. Retrieved from http://googleblog.blogspot.co.nz/2005/06/world-is-your-javascript-enabled_29.html
- Google. (2013). Google Developers. Retrieved from <https://developers.google.com/>
- International Telecommunication Union. (n.d.). ICT Aggregate Data 2006 - 2013. Retrieved from <http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>
- Joyce, M. (n.d.). Did Ted Nelson first use the word “hypertext” at Vassar College? What is Vassar’s claim?. Retrieved from http://faculty.vassar.edu/mijoyce/Ted_sed.html
- Kozel, S. (2010). Mobile Social Choreographies: choreographic insight as a basis for artistic research into mobile technologies. *International Journal of Performance Arts & Digital Media*, 6, 137-148.
- Kratky, A., & Manovich, L. (2005). *Soft Cinema: Navigating the Database*. Cambridge, Massachusetts. The MIT Press, Massachusetts Institute of Technology.
- Larson, N. & Shindelman, M. (2009) *Geolocation: Tributes to the Data Stream*. Retrieved from <http://larson-shindelman.com/geolocation/geolocation-overview/>
- Leggett, M. (2011). Memory, schema and interactive video. In L. Candy & E. Edmonds (Eds.), *Interacting: Art, research and the creative practitioner*, pp. 282-294

- Manovich, L. (2002) Soft Cinema: Concepts.
Retrieved from <http://www.softcinema.net/>
- MINA. (2013). MINA. Retrieved from www.mina.pro
- Milne, E. J., Mitchell, C., & De Lange, N. (Eds.). (2013). Handbook of Participatory Video. United Kingdom. AltaMira Press.
- Mozilla. (n.d.). Mozilla. Retrieved from <http://www.mozilla.org>
- Mozilla. (2011). Popcorn launch and world premier of one millionth tower. The Mozilla Blog. Retrieved from <http://blog.mozilla.org/blog/2011/11/05/popcorn-1-0-launch-and-world-premier-of-one-millionth-tower/>
- Nichols, B. (1991). Representing Reality: Issues and Concepts in Documentary. Indiana. Indiana University Press.
- Raby, F. & Dunne, A. (1998). Project #26765: FLIRT.
Retrieved from <http://www.dunneandraby.co.uk/>
- Schleser, M. (2011). Mobile-mentary (mobile documentary) 2.0. Expanding Documentary 2011: Conference Proceedings. Vol. 1, No. 2, December 2011. pp. 174
- Thalhofer, F. (2013). Korsakow. Retrieved from <http://korsakow.org>
- Vimeo. (n.d.). Vimeo APIs. Vimeo Developers.
Retrieved from <https://developer.vimeo.com/apis>
- Watt, J. (2013). The Suppression of Vertical Videos. Retrieved from <http://jameswatt.me/2013/02/09/the-suppression-of-vertical-videos>

YouTube. (n.d.). Statistics - YouTube.

Retrieved from <http://www.youtube.com/yt/press/en-GB/statistics.html>

YouTube. (2012). Changes to Related and Recommended Videos. Creators: The YouTube official partners and creators blog.

Retrieved from <http://youtubecreator.blogspot.co.nz>

5.1 Video references

- Bradley, T. [Todd Bradley] (2007). HDTS Promo Video [Video File].
Retrieved from <https://www.youtube.com/watch?v=73xxfseM9IQ>
- chescaleigh. (2013). Turn Your Phone! ("No Scrubs" parody with Andrew Huang, DailyGrace, Hannah Hart, Soundlyawake) [Video File].
Retrieved from <http://www.youtube.com/watch?v=eYLL7-rUGPY>
- Glove and Boots. (2012). Vertical Video Syndrome - A PSA [Video File].
Retrieved from <http://www.youtube.com/watch?v=Bt9zSfinwFA>
- Mann, J. [Jonathan Mann]. (2013). Turn Your Phone! (Vertical Video PSA) (Song A Day #1647) [Video File].
Retrieved from http://www.youtube.com/watch?v=NtQWp_31wul
- Onkel Bernis Welt. (2013). Streetmusician sings - Original singer comes along and joins him [Video File].
Retrieved from http://www.youtube.com/watch?v=Y_DWWE3cjgg
- Vertov, D. (1929). Man with a Movie Camera [Documentary]. Ukraine.
- zidyboby. (2013). CAUGHT ON CAMERA: Fertilizer Plant Explosion Near Waco, Texas [Video File].
Retrieved from <http://www.youtube.com/watch?v=ROrpKx3aljA>

5.2 Image references

Fig 1. Example of metadata used in Hypervideo in JSON format.

Fig 2. Example of a YouTube Video and metadata in JSON format.

Fig 3 and 4. Still images from videos produced in the 24 Frames 24 Hours filmmaking workshops.

Fig 5. Still from New Zealand's 3 News showing user generated content, retrieved from <http://www.3news.co.nz/Video-of-fight-causes-outrage/tabid/423/articleID/327359/Default.aspx>

Fig 6. Still image from Texas Soft Cinema, retrieved from <http://www.softcinema.net>

Fig 7. Korsakow logo, retrieved from <http://korsakow.org/>

Fig 8. Portrait video viewed on YouTube shows pillarboxing.

Fig 9. Portrait video resized to fit a landscape frame shows pillarboxing.

Fig 10. Screenshot of Vertical Video Syndrome - A PSA

Fig 11. Screencaps from YouTube Capture app on iOS.

Fig 12. Hypervideo: 24 Frames 24 Hours interface introduction screen.

Fig 13. Hypervideo: 24 Frames 24 Hours interface tutorial screen. Tutorial screens displayed simultaneously for illustrative purposes, usually displayed consecutively.

Fig 14. Hypervideo: 24 Frames 24 Hours interface set to show videos from New Zealand. One marker is hovered over to reveal the associated metadata.

Fig 15. Hypervideo: 24 Frames 24 Hours interface set to show videos from United States with themes of journey and people. 4 video frames open. Active video frame (bottom centre) has panel with related video buttons expanded.

Fig 16. Hypervideo: 24 Frames 24 Hours interface, zoomed in on one video frame to show metadata and controls.

Fig 17. Hypervideo: 24 Frames 24 Hours interface filter panel.

Appendix A

Hypervideo Design Experiments and Dynamic Interfaces

A more detailed description of the Hypervideo experiments are outlined in the following section. The numbers next to the titles correspond to the sections used in the main exegesis body.

3.1 MINA + Snapr App

Technology: iOS, HTML, CSS, LESS, JavaScript, jQuery, jQuery Mobile, GPS

This experiment involved creating an interface for a geolocate app using the Snapr platform¹. Snapr is a web platform for creating photography apps and is also a photography app in itself. Apps created using the Snapr platform generally focus on allowing a user to take photos using the camera built into their mobile device and share them with other users through the app's built in social network or through other popular social networks (Facebook, Twitter, Tumblr). The Snapr platform's defining characteristics are its emphasis on time- and geo-tagging of photographs. By accessing the EXIF data (a specific type of metadata used in images and some sound files) a Snapr app is able to retrieve the time and location (assuming a GPS enabled mobile device is used) where a photo was taken.

The apps are built as HTML5 web-apps and wrapped into Phonegap² to produce apps that can then be distributed through the App Store (for iOS devices) or Google Play store (for Android devices).

The main purpose of building the MINA app was to gain knowledge around the technology behind developing a mobile app and to learn about the processes behind a locative media system. Apps using the Snapr platform are built using HTML, CSS and jQuery Mobile for the interface and JavaScript, jQuery and Backbone.js for the functionality.

At the time of this experiment being conducted the Snapr platform was still in a developmental phase and not yet open to the public. This meant that I was working from a build still in development and there was not a lot of documentation or technical troubleshooting communities.

1 <http://sna.pr>
2 <http://phonegap.com>



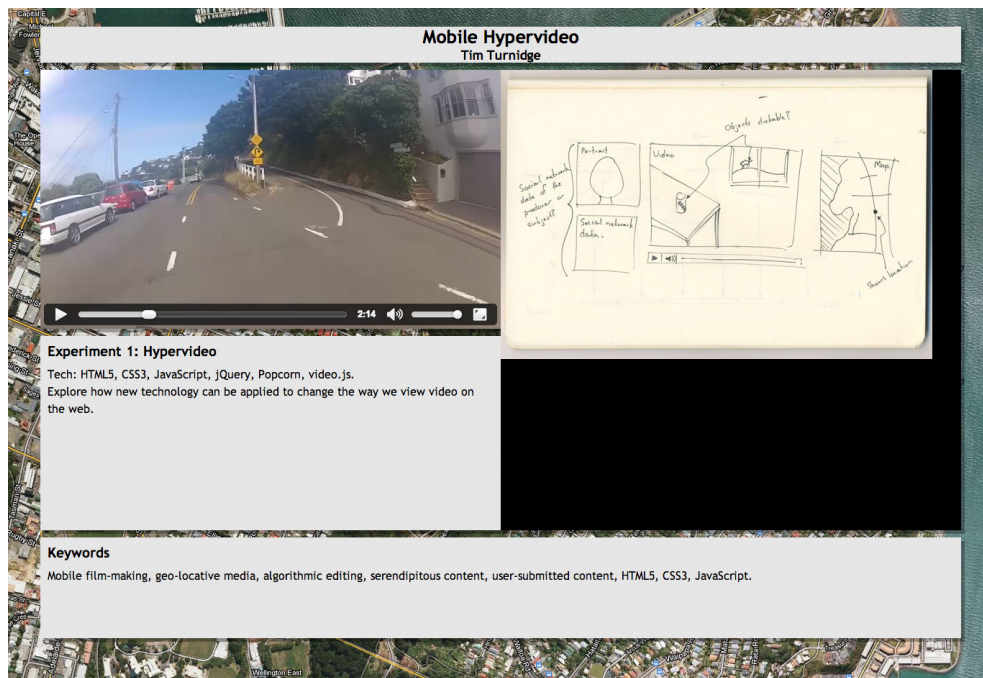
Screenshots from MINA Snapr app showing title screen (left), map interface with geolocated images (centre) and interface for sharing images (right).

Through this experiment I gained insights into current methods of displaying media on an interactive map which involves using the GPS metadata associated with an image to place an image thumbnail on a map. The process of publishing an app also highlights some limitations of designing interfaces for mobile devices such as the lack of Flash support (which some YouTube video player APIs used in later experiments rely on) and iOS's inability to play multiple videos at once (which becomes important in later interfaces).

3.2 Popcorn

Technology: HTML, CSS, JavaScript, jQuery, video.js, Popcorn

The aim of this experiment was primarily technical exploration into two JavaScript technologies: Popcorn.js and Video.js. These two technologies are recent developments in the area of web video and are explored to see what possibilities they might hold for application in an interactive documentary context.



Popcorn interface showing video frame using video.js. Related image and two dynamic text areas are powered by Popcorn.

Mozilla's Popcorn, launched in 2011, is an HTML5 media event framework written in JavaScript (Mozilla, 2011). Mozilla is an organisation whose mission is to "promote openness, innovation & opportunity on the Web" (Mozilla, n.d.). Popcorn expands the functionality of video embedded on a webpage by enabling it to control other elements on the page, for example, playing a video could trigger an image to appear elsewhere on the page or an element could be displayed or hidden depending on how far through a video a viewer is.

The main workflow when working with Popcorn involves writing Javascript to link a certain time-code within a video file to trigger some other action. Popcorn features a number of plugins to automatically display information from a number of contemporary web platforms such as Facebook, Twitter or Wikipedia. In addition to this there is a plugin to execute generic Javascript code which allows for the widest range of creative possibilities.

Video.js is an open source JavaScript library for HTML5 video that is created to extend the functionality of HTML5 video, fix cross-browser inconsistencies and provides an API for working with HTML5 video. In the context of the Hypervideo project, Video.js eases the workflow by fixing cross-browser inconsistencies within web video players and enables an easy way of styling the video player.

My experimentation with these technologies involved the creation of a web video presentation enhanced with Popcorn and Video.js. I set out to produce a vlog (video blog) that would sum up the state of the project so far in a style inspired by motovlogging. Motovlogging is vlogging while on a motorcycle and has gained popularity with the advent of portable action (video) cameras such as the GoPro and Drift Innovation cameras.

The video was embedded as an HTML5 video into a web page designed to accommodate the video, two sets of text and an image. Using Popcorn key points that were described in the audio commentary in the video were displayed within the interface, for example, quotations talked about could be displayed as text within the interface and related images could be shown.

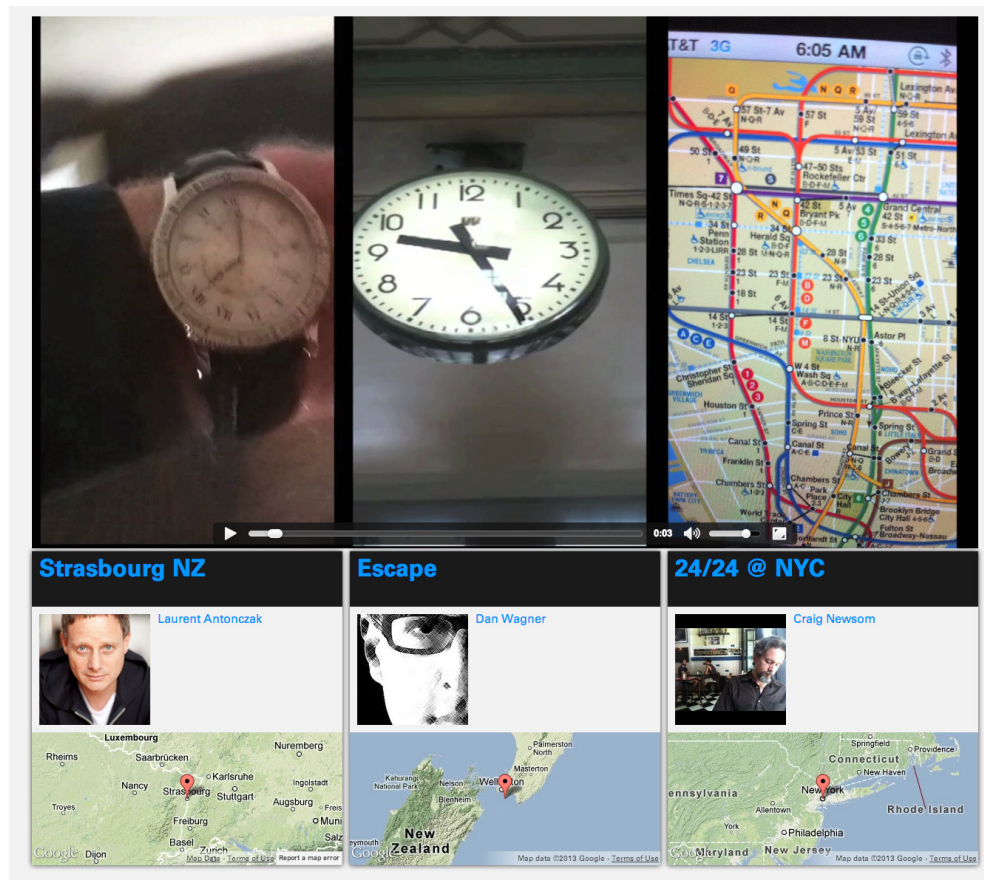
This experiment with Popcorn reveals promising possibilities for enhancing video content by showing additional related material. In the context of user generated or workshop generated content, Popcorn could be used to show metadata relating to that content as a way of providing context for the clip. This idea is explored further in section 3.3 24 Frames Popcorn.

Since this experiment, the online tool Popcorn Maker (accessible at <http://mozillapopcorn.org>) has been released, allowing for Popcorn to be applied to media without having to manually write JavaScript code. This makes Popcorn much more accessible and should encourage more Popcorn powered projects.

3.3 24 Frames Popcorn

Technology: HTML, CSS, JavaScript, jQuery, video.js, Popcorn, Vimeo API

This experiment involved the application of knowledge gained from the previous experiment to the 24 Frames 24 Hours documentary. Through the collaboration with 24 Frames 24 Hours I was supplied with a triptych montage video edited by Rachelle Smith (a 24 Frames 24 Hours collaborator) featuring video clips filmed on mobile devices from multiple authors in multiple locations. The goal in this experiment was to apply the knowledge gained from working with Popcorn in the previous experiment to content generated from the 24 Frames 24 Hours workshops.



24 Frames Popcorn interface.

Using Popcorn, I was able to display the metadata associated with the clips within the triptych for the duration that the clip was displayed. The result of this was that as the video triptych was played, the metadata below each third would update with every cut within the video.

The metadata displayed was the title of the video as it appears on Vimeo, the name of the author and the approximate location where it was shot. The title is hyperlinked to the original video in Vimeo and the author is hyperlinked to the author's Vimeo account page. The location that the video was shot is displayed within an interactive Google Maps map.

Displaying the metadata for the corresponding video clips contextualises them and helps the audience to make sense of the content being displayed. Through this metadata the audience is enabled to make connections between the images being shown and the location or person that it relates to.

3.4 24 Frames 24 Hours Vimeo

Technology: HTML, CSS, JavaScript, jQuery, Vimeo API, Google Maps API

This experiment is the first that targets the 24 Frames 24 Hours documentary directly. In this experiment the aim to to create an interface that allows a user an innovative means of viewing the videos shot in the 24 Frames 24 Hours mobile filmmaking workshops.

The interface uses the Vimeo Simple API for retrieving metadata; the Vimeo Froogaloop player API, to display the videos and metadata; and the Google Maps API to display the map and markers.

The decision to use the Vimeo API within this experiment is based on fact that the existing 24 Frames 24 Hours videos were uploaded to this platform.

The metadata available through the Vimeo “Simple” API contains the following information about a video (an example of the raw data can be found in Appendix B):

- An ID
- Title
- Description
- URL to the video
- Upload date and time
- Thumbnails at 100px, 200px and 640px
- User ID
- User name
- User profile URL
- User portraits at small 30px, medium 75px, large 100px and huge 300px
- The number of likes for the video
- Number of plays
- Number of comments
- Duration (seconds)
- Width and height
- Tags
- Privacy settings

It’s worth noting that the Vimeo API (and Vimeo in general) does not support geolocative metadata, so there is no built-in method of storing the location where a video was shot. When contacting the Vimeo development team online about whether this feature was going to be implemented, the response was that “yes, we are interested in geo data and hope we can implement it in some smart fashion in the future” (Vimeo staff username: Soxiam, 2012), see Appendix C.

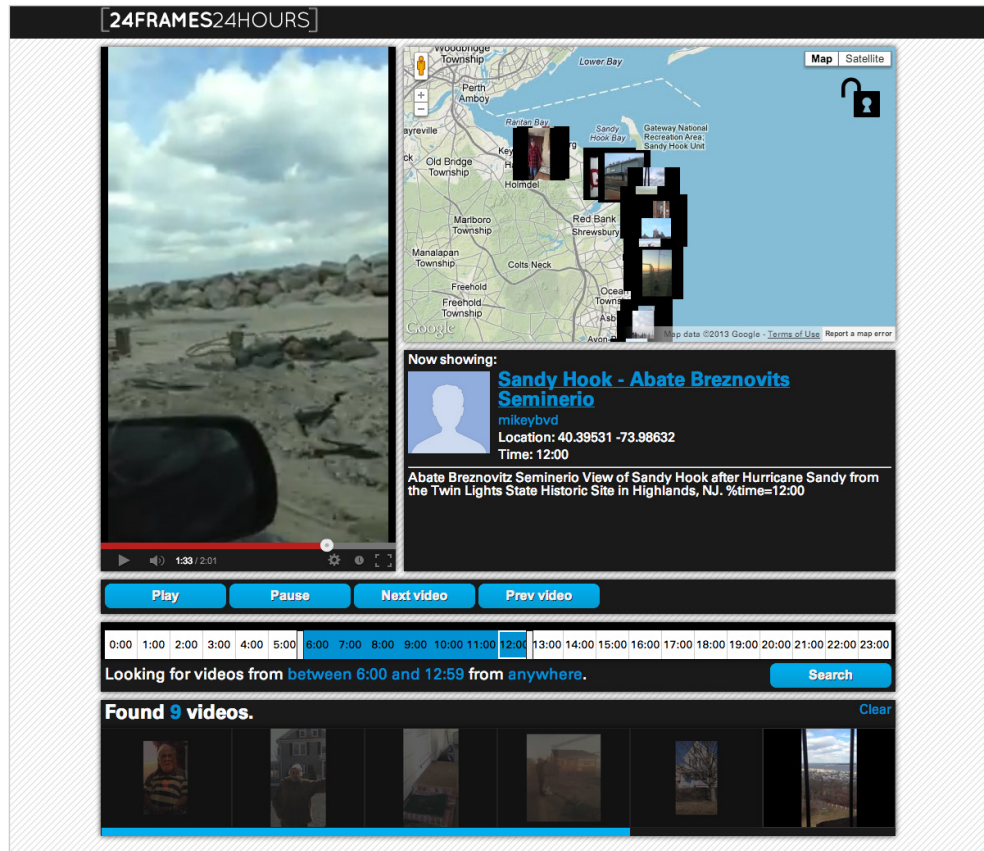
One way of getting around this limitation is to use the “tags” metadata. Usually these tags are used in Vimeo’s search algorithm to show related results, but by making the 24 Frames 24 Hours Vimeo interface respond to a specific tag syntax, I was able to repurpose them for geo-location.

A system was implemented where a user could enter the longitude and latitude as a Vimeo tag by using the prefixes “%lat:” and “%long:” respectively. For example, if the video was shot in Wellington, New Zealand, it could be geo-located by writing the tags: %lat:-41.288126, %long:174.77658. Using a prefix with a “%” character ensured that it was possible to distinguish between a deliberate reference to a coordinate and a regular descriptive tag, for example, without using these special characters, it would be much more difficult to ensure the system did not recognise someone tagging their video with “late for work” as a latitude of “e for work”.

Because the tags must be added by the uploader of the video, the author must log in to Vimeo and update the tags to have the coordinates included with the correct syntax. I made a guide to outline how to get latitude and longitude coordinates from Google Maps and the correct syntax for tagging videos with location coordinates to be distributed to workshop attendants to use while uploading their videos. By the time of the next workshops, I had already progressed to the next iteration of the interface which used the YouTube API instead which involves a different way of inputting geolocate coordinates, so this guide never needed to be used.

Because latitude and longitude coordinates are not common knowledge, I also explored geocoding which is the process of converting a text address into latitude and longitude coordinates. By using the Geocoding Service that is part of the Google Maps API, I implemented a test system where a user could use a custom “%loc” tag when uploading their video to provide a location where the video was shot. This tag would be read by the interface, geocoded to produce a latitude and longitude, and then those coordinates would be used to plot the video on the map. Because this method requires accessing the Google Maps API to geocode for each video that uses the “%loc” tag, it would not be suitable for a large scale project, as each video in the database, multiplied by each user viewing the database would cause an API call. An API call is when your website requests data using the API. Most platforms offering an API will specify limits on how many API calls can be made within a set timeframe.

Vimeo also does not have a metadata entry for the time a video was shot, only when it was uploaded, which is not particularly useful for the 24 Frames 24 Hours documentary. As a temporary workaround, I used the upload time metadata in place of the time of shooting in order to progress the experiment. The interface in this iteration presents the user with three primary methods of selecting a video to view: the map, the timeline and the playlist.



24 Frames Popcorn interface.

The GoogleMaps v3 API is used to display the interactive map as well as the markers which represent videos within the database. By displaying the video as a marker on the map in the location where it was shot the link between the place and the story being told are reinforced.

The timeline is presented as an interactive histogram where each video is represented by a rectangular button which sits under the heading for the two hour period it was filmed in. This system is a reflection of the idea that each video produced in the mobile filmmaking workshops is meant to provide a snapshot of a two hour period, eg. from 2pm - 4pm.

The playlist section at the bottom presents six videos from the database as thumbnail images. Clicking one will load that video into the video player and begin playing immediately.

The videos are displayed in a video frame with a fixed, 16:9 landscape frame which is similar to the default frame shown on Vimeo.

Reflecting on this experiment revealed that the lack of geolocate and timestamp metadata in the Vimeo API poses a challenge. Customised tags such as “%lat” could be one way of overcoming these shortcomings, but manually tracking down latitude and longitude coordinates and entering them as tags in the correct syntax is a hurdle for engagement with the project.

Informal user testing revealed that in this interface the interactive map was the most successful element. Users commented that it was easy to see the relationship between the video and the location that it was shot through the map and that this information was more interesting than the relationship between a video and the time it was shot. Future interfaces place a higher priority on the location of media over the time of media.

3.5 YouTube API Experiments

The next three experiments were initiated to explore the YouTube API, to learn its limitations and suitability for application to the 24 Frames 24 Hours documentary. These experiments all focus on examining how and what metadata can be accessed through the YouTube API rather than the design of interfaces for an end user.

3.5.1 YouTube API Explorer

Technology: HTML, CSS, JavaScript, jQuery, YouTube API

The goal with this experiment was to explore whether YouTube would be a more suitable platform for hosting the videos by exploring what functionality the YouTube “Simple” API provides. The experiment involved creation of a web interface where I could enter the YouTube video ID of a video and the interface would display the video and any relevant metadata available.

The metadata available through the YouTube API contains the following information about a video (an example of the raw metadata can be found in Appendix D):

- Author username
- Author URL
- Video category
- Video URL
- URL to comment data
- Rating (number of likes and dislikes)
- Geolocation coordinates as latitude and longitude
- A unique ID
- URL to video file at different resolutions and formats
- Credits
- Description

- Licence information
- Thumbnails at various sizes and from various times within the video
- Title
- Time and date of last update
- Permissions data for whether a video can be commented on, rated, have video responses or be embedded on a third party site
- Date of recording
- Time and date of upload
- Duration
- Number of views
- Number of favourites

In general, the metadata available through the YouTube API is similar to the metadata available through the Vimeo API with a few notable exceptions. Firstly, the YouTube API does support geolocation, which is an important element in the 24 Frames 24 Hours documentary. When a user uploads a video to YouTube, the video metadata is analysed to see if geolocate data is present, if it is, the location is immediately passed to the YouTube video metadata. If the video does not include geolocate data, then the user has the option of adding this manually either by entering coordinates, an address (which is then geocoded to coordinates) or by using an interactive Google Maps map. The YouTube API does have a metadata entry for the date of recording, but not the time of recording, so a custom solution would have to be implemented.

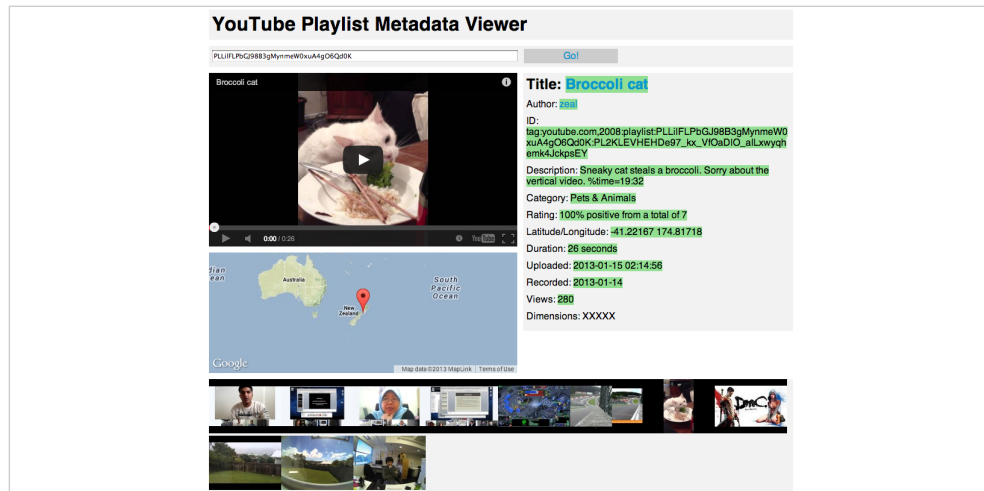
While the YouTube API contains metadata about the author's username and URL, unlike the Vimeo API, the author's profile image is not embedded in the video metadata. This means that an additional API call needs to be made to retrieve the author's profile image.

3.5.2 YouTube Playlist Explorer

Technology: HTML, CSS, JavaScript, jQuery, YouTube API

YouTube has the ability to create playlists of videos within their interface. This experiment took the functionality from the previous experiment, YouTube API Explorer, to read metadata from a YouTube playlist ID rather than a YouTube video ID.

This experiment revealed that making YouTube playlists is an efficient method of selecting videos to include in the 24 Frames 24 Hours documentary. It would be possible to load all the videos produced in the filmmaking workshops into one playlist and use that as a way of loading in the metadata for multiple videos simultaneously, as opposed to making multiple API requests.



Youtube Playlist Explorer interface showing test playlist of 11 videos, metadata for one video shown.

3.5.3 YouTube Comments Explorer

Technology: HTML, CSS, JavaScript, jQuery, YouTube API

This experiment was conducted to explore how comments posted to a YouTube video can be accessed and displayed back to gauge feasibility for use within further experiments.

The results of this experiment are that it is possible to do through the YouTube API with one API request retrieving the comments. Because of the way the API is structured, an additional API request would need to be made for each comment to retrieve the YouTube portrait for the account that made the comment. This has some performance implications as retrieving 50 comments can result in needing to contact the YouTube API 50 times, possibly causing performance issues if not handled correctly.

3.6 24 Frames 24 Hours YouTube

Technology: HTML, CSS, JavaScript, jQuery, YouTube API, Google Maps API

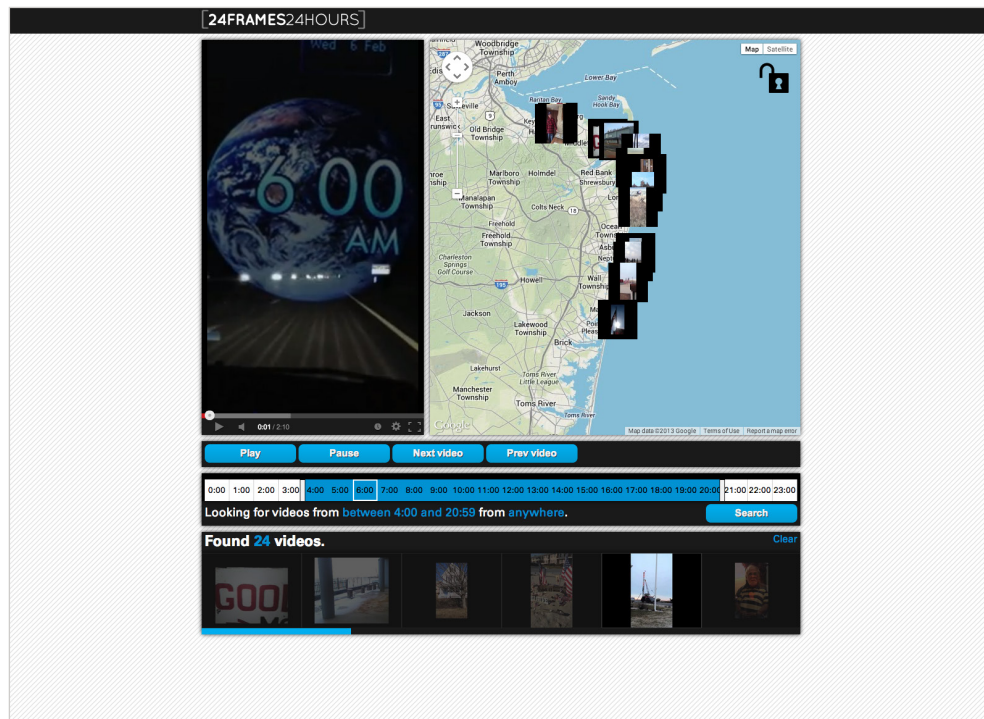
After experimentation with the YouTube and Vimeo APIs, I decided that the YouTube platform would be most suitable for further exploration and application to the 24 Frames 24 Hours documentary. The geolocate metadata built in to the YouTube API is a useful feature because it means that there is an established method of geotagging the videos.

This experiment began by taking the interface from the 24 Frames 24 Hours Vimeo experiment and changing the backend to use the YouTube API instead.

At the time of starting this experiment, participants in the 24 Frames 24 Hours filmmaking workshops were asked to create a video portrait of themselves or a place that they feel a connection with. Through chance or direction, most of the videos that were generated in the workshops around this time were filmed

in portrait orientation. As a response to this, I designed the interface to use a video frame with a portrait orientation as a way of accommodating the portrait videos as best as possible.

In this interface I also experimented with an alternative method of choosing which videos to view. In previous interfaces, clicking a video would load that video into the player and begin playing immediately. In the 24 Frames 24 Hours YouTube interface, a playlist system is implemented where users can queue up which videos they would like to see and then those videos in the playlist are loaded into the video frame sequentially. Playing the videos sequentially creates a temporal montage where the user is able to create their own connections between the videos.



24 Frames 24 Hours YouTube interface showing videos submitted from New Jersey filmmaking workshop.

As discovered in the YouTube API Explorer experiment, the YouTube API does not store metadata for the time that a video was shot. As a way of getting around this limitation, I used a custom tag similar to the latitude (%lat) and longitude (%long) tags used in the 24 Frames 24 Hours Vimeo interface. In this experiment, the author of a video is able to tag the time a video was shot by using the prefix “%time=”, for example, a video shot at 2:30pm would have the tag: %time=14:30. Similar to the Vimeo experiment, a guide was produced and distributed to workshop participants outlining how to tag their video with the correct syntax.

While developing this interface, another round of filmmaking workshops was conducted in New Jersey, USA, resulting in a new batch of content being added to the database. In these videos there were a number of issues which needed to be addressed in further experiments. Firstly, while the majority of videos were in the portrait orientation, there were a number shot in landscape and some in square format. Because the interface was designed with a portrait video frame, videos shot in landscape or square would be letterboxed which does not present an optimal viewing experience. Secondly, while the majority of users used the %time tag correctly, there were some videos which had incorrect syntax (for example, using %time:14:30 instead of %time=14:30) which meant the time metadata could not be applied correctly. Finally, not all videos were tagged with a location in YouTube, so would come through to the interface without any geolocate metadata.

The way that the interface and underlying algorithms are designed require that the metadata is formatted correctly in order for the interface to work. Ensuring that the interface is able to gracefully handle incorrectly formatted metadata is another aspect of this, but the focus at this stage was to see how metadata could be utilised in an ideal situation. In this experiment a guide was supplied to workshop participants outlining how to correctly format the metadata, however, this still did not guarantee a perfect outcome. The following experiment, Metadata Editor, explores creating an interface that ensures correctly formatted metadata.

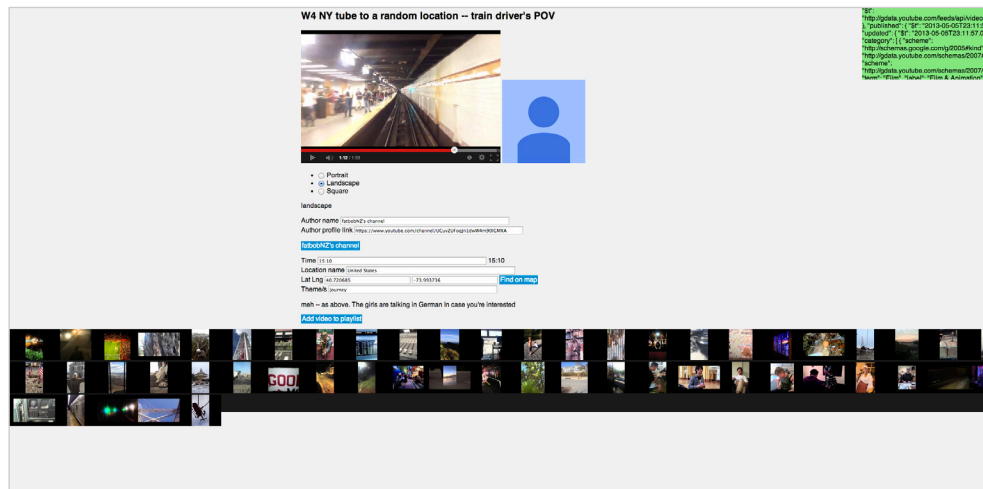
By changing the interface to focus on the creation of playlists rather than on viewing individual videos, the focus is shifted from the videos acting as individual pieces of content to the videos being part of a larger narrative. With the shift in focus from single video clips to multiple video clips shown in sequence, research into algorithms to determine which video clips are shown in sequence was initiated. While in this iteration of the interface the order of videos is determined by the user, future interfaces use metadata and algorithmic editing concepts adapted from Korsakow and Soft Cinema to determine the sequencing of video clips.

3.7 Metadata Editor

Technology: HTML, CSS, JavaScript, jQuery, YouTube API, Google Maps API, PHP, MySQL, Couch DB

In working on the previous experiments, I identified that the metadata that is accessible through the APIs will never be able to accommodate all the metadata that I want to use in the 24 Frames 24 Hours documentary in an efficient, consistent manner. Furthermore, even with APIs that can accommodate the metadata required for the documentary (YouTube's geolocation metadata, for example) there is no guarantee that the author of the video will actually tag the video with all the required metadata while uploading.

My solution to this is to create a custom set of metadata that is applied to videos that are to be included in the project. By doing this, I am able to store the metadata from the YouTube API along with any custom metadata that needs to be included to overcome limitations of the API (for example, the lack of an entry for the time the video was shot).



Metadata editor interface.

The Metadata Editor experiment presents an interface that reads the metadata for a given YouTube ID (see 3.5.1 YouTube API Explorer) and allows a user to edit or add their own metadata. As the user is editing the metadata within the editor, it is only editing the metadata that is used in the 24 Frames 24 Hours documentary, not the metadata stored on YouTube. This means that it is possible to overwrite some metadata if it comes through incorrectly from the YouTube API, for example, if the original upload did not use the correct syntax for tagging the time the video was shot.

The interface presents a system for generating a series of videos and associated metadata to use in the 24 Frames 24 Hours documentary. Entering the video ID for a YouTube video pulls in the metadata for that video and displays it back to the user in a readable, editable format. The user is then able to make any changes required such as changing the title or changing the author name (because YouTube usernames might be an alias whereas they want to use their real name, or vice versa). The user is also able to use an interactive map to specify the location that the video is shot which eliminates the need for a user to go to an external site to work out what the latitude

and longitude coordinates should be. The interface also allows for adding the time a video was shot without having to use a custom tag which reduces opportunities for error. Finally, the interface allows users to specify the aspect ratio for the video, which is currently not possible through the YouTube API.

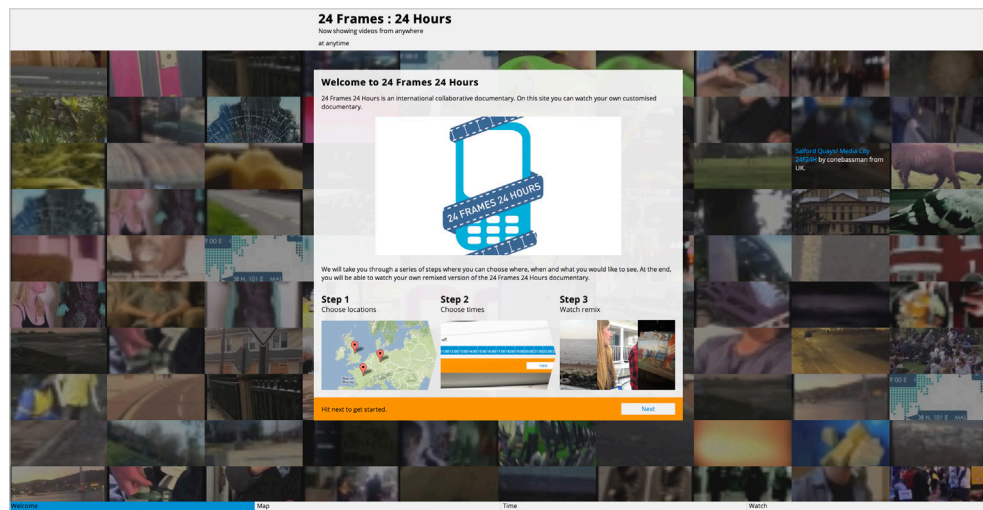
The Metadata Editor allows for multiple videos to be edited and the resulting metadata is output in JSON format and stored as a local JavaScript file. Storing the metadata in this way allows for a rapid development process, but means that I am the only person able to edit the metadata. This limitation is discussed further in 4.1 Future Developments and Research.

3.8 24 Frames 24 Hours Split

Technology: HTML, CSS, JavaScript, jQuery, YouTube API, Google Maps API

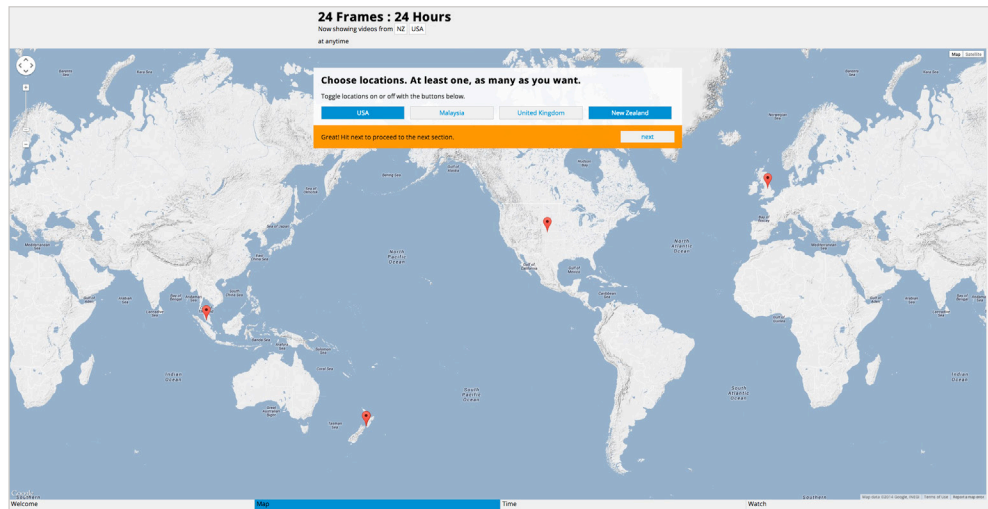
In the previous interface: 24 Frames 24 Hours YouTube, all of the user's options are laid out immediately. The map, timeline, playlist, video and metadata are all constantly visible because the whole interface is contained on a single page. When showing the interface to colleagues, comments indicated that the interface might be presenting too many elements to focus on at once, potentially distracting from the experience or overwhelming users with options.

This experiment presents an interface designed to lead the user through a series of steps where they choose which types of videos they want to see. In Memory, Schema and Interactive Video, Legget describes how an "audience, encountering a system in a public place and without prior knowledge, will need to be encouraged and guided within an interactive navigation schema" (2011). This step-by-step system is designed to make the users options clearer and ultimately make for a more accessible experience.



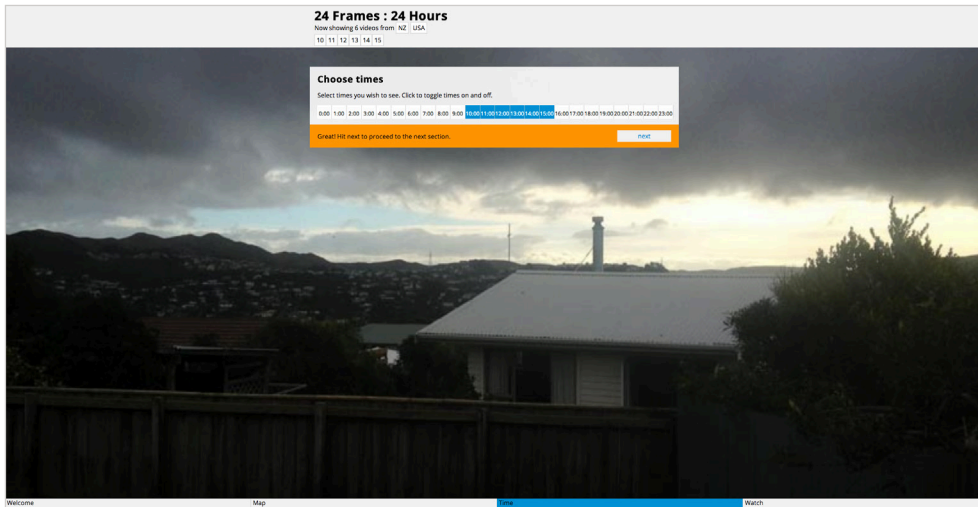
24 Frames 24 Hours Split interface introduction screen.

In the interface, the user is presented with an introductory page which outlines the 3 steps that follow. As the user moves from page to page a horizontal pan effect is used to indicate progression through the steps.



24 Frames 24 Hours Split interface location selection screen.

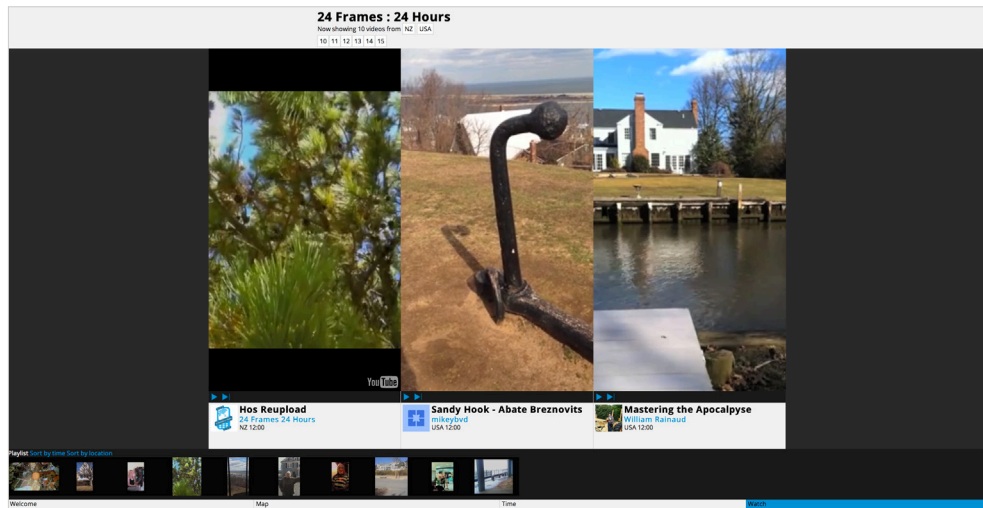
The first step is to select the location or locations the user wants to include in their version of the documentary. Each 24 Frames 24 Hours filmmaking workshop location (so far) is represented as a location that can be selected with a marker on the map and a button in the overlay. As locations are selected, they are added to the “Now showing videos from” phrase at the top of the interface which allows them to keep track of what filters are applied as they progress through the tutorial. When the user has selected at least one location, an option is revealed that allows progression to the next step.



24 Frames 24 Hours Split interface time selection screen.

The second step is to select the range of times the user wants to include in their version of the documentary. Here I experimented with different elements that allow for selection of a range of values. The final element allows a user to click on an individual button to toggle that time on or off, or to click and drag to toggle a range of times.

The final step takes the locations and times selected in previous steps and generates a playlist of videos which match those criteria.



24 Frames 24 Hours Split interface video watching screen.

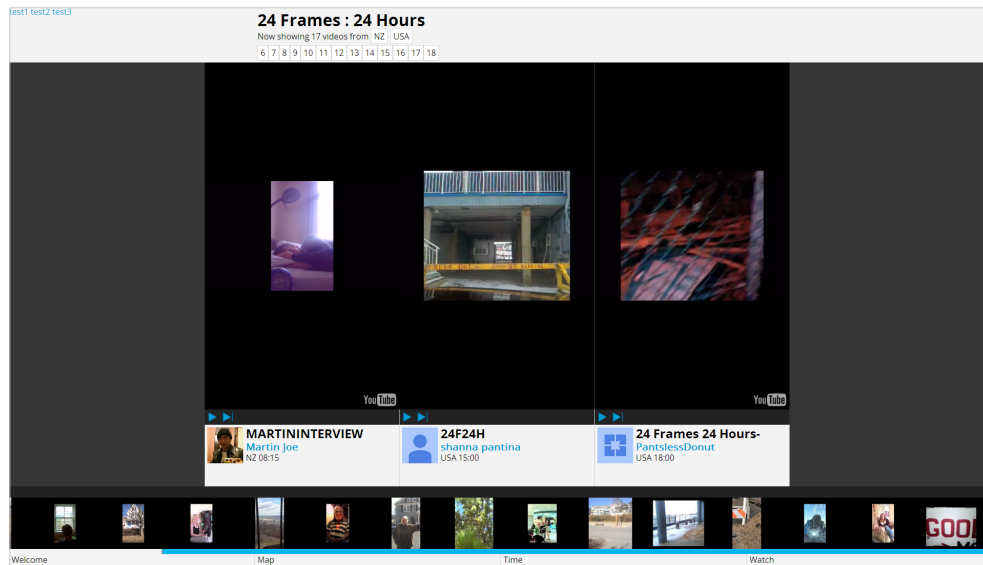
The videos are displayed in a triptych of portrait video frames using the YouTube API in a way which resembles the remix of workshop videos facilitated by Max Schleser that is used in the 24 Frames Popcorn experiment. This arrangement of videos allows for the juxtaposition of imagery through spatial montage, enabling users to create associative chains between the videos where “meaning and narrative coherence can be created”. (Kratky & Manovich, 2005)

The video frames resize in response to the browser window dimensions by using jQuery so that they are as large as possible while retaining the same aspect ratio. Displaying three portrait videos in a triptych conveniently approximates a landscape frame which makes it easily fit within a “traditional” landscape browser window.

Because the interface is now showing three videos simultaneously, the audio needs to be considered as each video has its own audio track. My solution to this was to implement a system where the audio for the video that the user is hovering over is played while the audio for the other two videos are muted.

Feedback on this feature was positive but it was commented on that there should be some visual indicator or tutorialisation to make this interaction explicit. Future interfaces use a highlighted border around the videos to indicate which video is currently active in terms of audio.

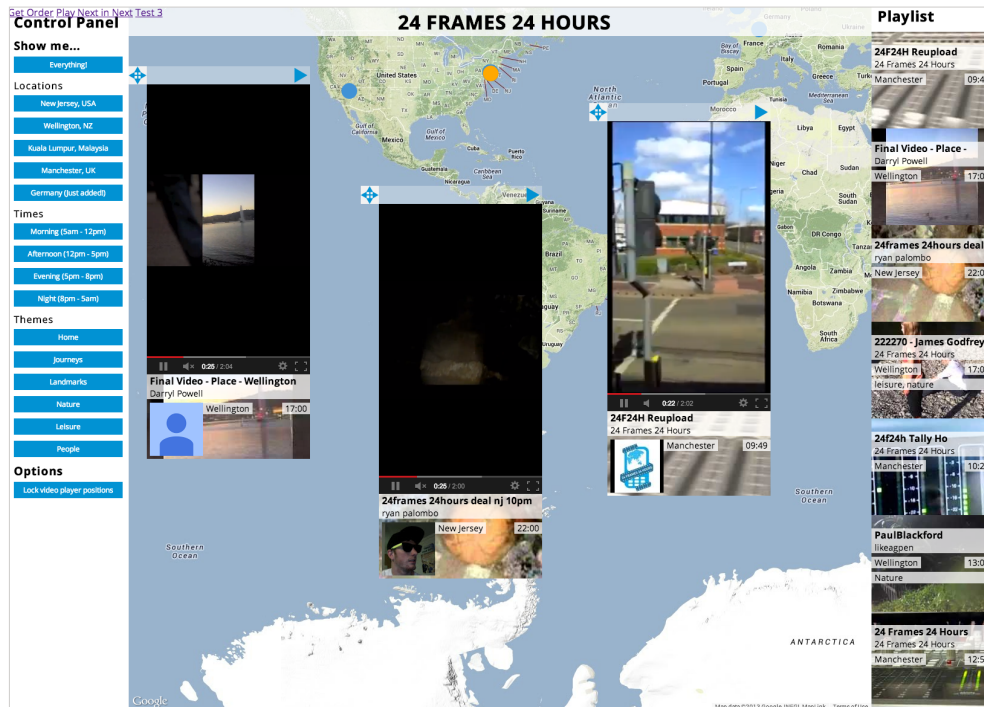
As videos began to populate the database, it became clear that displaying them in fixed, portrait frames would cause problems similar to those faced in previous experiments. While the majority of videos were shot in portrait, there were some that were shot and uploaded in landscape or square orientations meaning that the videos would be letterboxed when shown in the portrait frames. This could be fixed by making the video frames scalable or of a variable aspect ratio but would necessitate a redesign of the interface to accommodate.



24 Frames 24 Hours Split interface video watching screen showing portrait video uploaded as landscape (left) and two square videos (centre and right).

3.9 Concept

Technology: HTML, CSS, JavaScript, jQuery, YouTube API, Google Maps AP



Concept interface showing 3 videos playing.

In this experiment I wanted to try out a different type of interface where each video would be displayed within its own individual video frame. This was a response to the limitations I was encountering with the 24 Frames 24 Hours Split interface portrait triptych and through a desire to allow for a “macro-cinema” (Manovich, 2002) spatial montage to occur.

This interface centres around an interactive map where each video is represented by a marker shown in the location where it was shot.

On the left of the interface is a list of buttons listing the locations, times and themes related to the videos in the database. The locations listed are the places where filmmaking workshops have been held; the times relate to the time of day the video was shot, categorised into one of four sections of the day: morning, afternoon, evening or night; and the themes respond to the types of content found in the workshop videos: home, journeys, landmarks, nature, leisure, people. Clicking any of these buttons causes the interface to search for videos that match this criteria, filling the playlist on the right with matching videos, and loading these videos into the video players.

The interface presents three portrait video frames displayed on top of the map. Each video frame displays the metadata for the currently playing video below it as a way of contextualising the video by displaying who created it, where and when it was shot and what theme/s it contains. This metadata is shown to reinforce the link between the location, the person who filmed it and the story they are telling. The video frames implement the same system as the previous Split interface where the last hovered video frame plays its audio while other videos are muted.

Each frame also has an expandable menu where three related videos are displayed. By reading the metadata associated with the currently playing video and searching through the database, the user is presented with a video from the same location, same time and same theme. These are presented to the user as buttons adjacent to the video in a similar format to that used in Florian Thalhofer's Korsakow system. Clicking one of these related videos loads that video into the frame and begins playing immediately.

Where Korsakow relies on the SNU-ifying of media content to determine which video clips link to one another, this interface uses the location, time and theme metadata to make links between videos.

The position of the video frames on the screen are variable in a way which reflects the locative nature of the mobile video media. When a new video is loaded into the frame, the frame changes position on screen to match the location on the map where it was filmed. This is used as a way of reinforcing the connection between the video and the place it was shot. In addition to this, each video frame can be moved, allowing the user to construct their own spatial montage. This adds an additional layer of interactivity to the project and enables the juxtaposition of imagery within the individual videos through spatial montage.

In user testing, users found the way that the video frames moved whenever they moved the map a bit overwhelming, commenting that while it makes sense that the videos would initially start on the location where the video is from, the video should not always move when interacting with the map.

Appendix B

Raw Vimeo metadata

```
{ "id": 37776572, "title": "Strasbourg  NZ", "description": "24 Frames 24 Hours collaborative mobile film produced in Wellington, Auckland (NZ) and Strasbourg (France).", "url": "http://vimeo.com/37776572", "upload_date": "2012-03-01 20:57:46", "thumbnail_small": "http://b.vimeocdn.com/ts/259/660/259660088_100.jpg", "thumbnail_medium": "http://b.vimeocdn.com/ts/259/660/259660088_200.jpg", "thumbnail_large": "http://b.vimeocdn.com/ts/259/660/259660088_640.jpg", "user_id": 7655126, "user_name": "24Frames24Hours", "user_url": "http://vimeo.com/user7655126", "user_portrait_small": "http://b.vimeocdn.com/ps/221/303/2213033_30.jpg", "user_portrait_medium": "http://b.vimeocdn.com/ps/221/303/2213033_75.jpg", "user_portrait_large": "http://b.vimeocdn.com/ps/221/303/2213033_100.jpg", "user_portrait_huge": "http://b.vimeocdn.com/ps/221/303/2213033_300.jpg", "stats_number_of_likes": 3, "stats_number_of_plays": 63, "stats_number_of_comments": 1, "duration": 103, "width": 1280, "height": 720, "tags": "%lat:-36.88676,%long:174.898224", "embed_privacy": "anywhere" }
```

Vimeo response to geolocation feature request



Appendix D

Raw YouTube video metadata

```
{ "gd$etag": "W/\\"YDwqeyM.\\"", "id": { "$t": "tag:youtube.com,2008:playlist:PLLilFLPbGJ98B3gMynmeW0xuA4gO6Qd0K:PL2K-LEVHEHDe97_kx_VfOaDlO_aILxwyqhemk4JckpsEY" }, "published": { "$t": "2013-01-15T02:12:20.000Z" }, "updated": { "$t": "1970-01-01T00:00:00.000Z" }, "category": [ { "scheme": "http://schemas.google.com/g/2005#kind", "term": "http://gdata.youtube.com/schemas/2007#video" }, { "scheme": "http://gdata.youtube.com/schemas/2007/categories.cat", "term": "Animals", "label": "Pets & Animals" } ], "title": { "$t": "Broccoli cat" }, "content": { "type": "application/x-shockwave-flash", "src": "http://www.youtube.com/v/RHE_msXeaC4?version=3&f=playlists&app=youtube_gdata" }, "link": [ { "rel": "alternate", "type": "text/html", "href": "http://www.youtube.com/watch?v=RHE_msXeaC4&feature=youtube_gdata" }, { "rel": "http://gdata.youtube.com/schemas/2007#video.related", "type": "application/atom+xml", "href": "http://gdata.youtube.com/feeds/api/videos/RHE_msXeaC4/related?v=2" }, { "rel": "http://gdata.youtube.com/schemas/2007#mobile", "type": "text/html", "href": "http://m.youtube.com/details?v=RHE_msXeaC4" }, { "rel": "http://gdata.youtube.com/schemas/2007#uploader", "type": "application/atom+xml", "href": "http://gdata.youtube.com/feeds/api/users/VcH04ohsO6qlxF13xpR5Sw?v=2" }, { "rel": "related", "type": "application/atom+xml", "href": "http://gdata.youtube.com/feeds/api/videos/RHE_msXeaC4?v=2" }, { "rel": "self", "type": "application/atom+xml", "href": "http://gdata.youtube.com/feeds/api/playlists/PLLilFLPbGJ98B3gMynmeW0xuA4gO6Qd0K/PL2KLEVHE-
```

```

HDe97_kx_VfOaDlO_aILxwyqhemk4JckpsEY?v=2"]], "author": [{"name": {"$t": "zeal"}, "uri": {"$t": "http://gdata.youtube.com/feeds/api/users/ZealNZ"}, "yt$userId": {"$t": "VcH04ohsO6qlx-F13xpR5Sw"}]], "yt$accessControl": [{"action": "comment", "permission": "allowed"}, {"action": "commentVote", "permission": "allowed"}, {"action": "videoRespond", "permission": "moderated"}, {"action": "rate", "permission": "allowed"}, {"action": "embed", "permission": "allowed"}, {"action": "list", "permission": "allowed"}, {"action": "autoPlay", "permission": "allowed"}, {"action": "syndicate", "permission": "allowed"}], "gd$comments": {"gd$feedLink": {"rel": "http://gdata.youtube.com/schemas/2007#comments", "href": "http://gdata.youtube.com/feeds/api/videos/RHE_msXeaC4/comments?v=2", "countHint": 0}}, "georss$where": {"gml$Point": {"gml$pos": {"$t": "-41.22167 174.81718"}}, "yt$hd": {}, "media$group": {"media$category": [{"$t": "Animals", "label": "Pets & Animals", "scheme": "http://gdata.youtube.com/schemas/2007/categories.cat"}], "media$content": [{"url": "http://www.youtube.com/v/RHE_msXeaC4?version=3&f=playlists&app=youtube_gdata", "type": "application/x-shockwave-flash", "medium": "video", "isDefault": "true", "expression": "full", "duration": 26, "yt$format": 5}, {"url": "rtsp://r5---sn-a5m7zu7s.c.youtube.com/CiULENy73wIaHAKuaN7Fmj9xRBMYSANFEgGUglwbGF5bGlzdHMM/0/0/0/video.3gp", "type": "video/3gpp", "medium": "video", "expression": "full", "duration": 26, "yt$format": 1}, {"url": "rtsp://r5---sn-a5m7zu7s.c.youtube.com/CiULENy73wIaHAKuaN7Fmj9xRBMYESARFEgGUglwbGF5bGlzdHMM/0/0/0/video.3gp", "type": "video/3gpp", "medium": "video", "expression": "full", "duration": 26, "yt$format": 6}], "media$credit": [{"$t": "zealnz", "role": "uploader", "scheme": "urn:youtube", "yt$display": "zeal"}], "media$description": {"$t": "Sneaky cat steals a broccoli. Sorry about the vertical video."

```

```
%time=19:32","type":"plain"},"media$keywords":{},"media$license":{"$t":"youtube","type":"text/html","href":"http://www.youtube.com/t/terms"},"media$player":{"url":"http://www.youtube.com/watch?v=RHE_msXeaC4&feature=youtube_gdata_player"},"media$thumbnail":[{"url":"http://il.ytimg.com/vi/RHE_msXeaC4/default.jpg","height":90,"width":120,"time":"00:00:13","yt$name":"default"},{"url":"http://il.ytimg.com/vi/RHE_msXeaC4/mqdefault.jpg","height":180,"width":320,"yt$name":"mqdefault"},{"url":"http://il.ytimg.com/vi/RHE_msXeaC4/hqdefault.jpg","height":360,"width":480,"yt$name":"hqdefault"},{"url":"http://il.ytimg.com/vi/RHE_msXeaC4/sddefault.jpg","height":480,"width":640,"yt$name":"sddefault"},{"url":"http://il.ytimg.com/vi/RHE_msXeaC4/1.jpg","height":90,"width":120,"time":"00:00:06.500","yt$name":"start"},{"url":"http://il.ytimg.com/vi/RHE_msXeaC4/2.jpg","height":90,"width":120,"time":"00:00:13","yt$name":"middle"},{"url":"http://il.ytimg.com/vi/RHE_msXeaC4/3.jpg","height":90,"width":120,"time":"00:00:19.500","yt$name":"end"}],"media$title":{"$t":"Broccoli cat","type":"plain"},"yt$duration":{"seconds":"26"},"yt$uploaded":{"$t":"2013-01-15T02:14:56.000Z"},"yt$uploaderId":{"$t":"UCVcH04ohsO6qlxF13x-pR5Sw"},"yt$videoid":{"$t":"RHE_msXeaC4"},"gd$rating":{"average":5,"max":5,"min":1,"numRaters":7,"rel":"http://schemas.google.com/g/2005#overall"},"yt$recorded":{"$t":"2013-01-14"},"yt$statistics":{"favoriteCount":"0","viewCount":"280"},"yt$rating":{"numDislikes":"0","numLikes":"7"},"yt$position":{"$t":7}}
```

Appendix E

Notification of Low Risk Research/Evaluation Involving Human Participants

Document attached on following pages.



Massey University

Te Kunenga ki Pūrehuroa

NOTIFICATION OF LOW RISK RESEARCH/EVALUATION INVOLVING HUMAN PARTICIPANTS

*(All notifications are to be typed)
(Do not modify the content or formatting of this document in any way)*

SECTION A:

1. Project Title	Mobile Hypervideo	
Projected start date for data collection	01/08/2012	Projected end date 01/03/2014

(Low risk notifications will not be processed if recruitment and/or data collection has already begun.)

2. Applicant Details *(Select the appropriate box and complete details)*

ACADEMIC STAFF NOTIFICATION

Full Name of Staff Applicant/s			
School/Department/Institute			
Region <i>(mark one only)</i>	Albany <input type="checkbox"/>	Palmerston North <input type="checkbox"/>	Wellington <input type="checkbox"/>
Telephone	Email Address		

STUDENT NOTIFICATION

Full Name of Student Applicant		Tim Turnidge	
Postal Address		57B Stewart Drive, Newlands, Wellington, 6037	
Telephone 0211637855	Email Address t.turnidge@massey.ac.nz		
Employer (if applicable)		Massey University	
Full Name of Supervisor(s)		Max Schleser	
School/Department/Institute		College of Creative Arts	
Region <i>(mark one only)</i>	Albany <input type="checkbox"/>	Palmerston North <input type="checkbox"/>	Wellington <input checked="" type="checkbox"/>
Telephone ext 62384	Email Address m.schleser@massey.ac.nz		

GENERAL STAFF NOTIFICATION

Full Name of Applicant			
Section			
Region <i>(mark one only)</i>	Albany <input type="checkbox"/>	Palmerston North <input type="checkbox"/>	Wellington <input type="checkbox"/>
Telephone	Email Address		
Full Name of Line Manager			
Section			
Telephone	Email Address		

3 Type of Project (provide detail as appropriate)

Staff Research/Evaluation:

Academic Staff

General Staff

Evaluation

Student Research:

Name of Qualification

Credit Value of Research

(e.g. 30, 60, 90, 120, 240, 360)

If other, please specify:

MDes
120

4. Describe the process that has been used to discuss and analyse the ethical issues present in this project.
(Please refer to the Low Risk Guidelines on the Massey University Human Ethics Committee website)

Attended Massey University Human Ethics seminar presented by Lyn Garrett on 29th February 2012. I discussed the ethical issues with my supervisor Dr. Max Schleser and we could see no ethical problems arising from this project. Completing the screening questionnaire confirmed that a low risk notification would be most appropriate. In addition to this, Dr. Schleser informed me that he has also applied for ethics approval, as there will be some links between my research and his project.

5. Summary of Project

Please outline the following (in no more than 200 words):

1. The purpose of the research, and

The research undertaken is for a Master of Design degree which aims to explore the following research questions:

- What is the role of audience in ubiquitous film-making?
- How can algorithmic mobile video editing be used for storytelling in user-generated content?
- Can mobile video provide a means to creating serendipitous content?

2. The methods you will use.

The project will involve practice-led research centred around mobile filmmaking and mobile photography. Part of the research project will involve the collection and remixing of user-submitted photography and video.

Participants contributing video and pictures will be informed about the purpose of the research and intended outcome before they choose to engage with the project. Information about the project and any ethical issues will be provided on the website where users will engage with the project and at any reasonable opportunity.

(Note: ALL the information provided in the notification is potentially available if a request is made under the Official Information Act. In the event that a request is made, the University, in the first instance, would endeavour to satisfy that request by providing this summary. Please ensure that the language used is comprehensible to all)