Media richness, interactivity and retargeting to mobile devices: a survey

Joel Ross*, Robert Simpson and Bill Tomlinson

Department of Informatics,
University of California, Irvine,
Irvine, CA 92697, USA
E-mail: jwross@uci.edu
E-mail: rmsimpso@uci.edu
E-mail: wmt@uci.edu
*Corresponding author

Abstract: Mobile devices, such as mobile phones, are becoming more ubiquitous and gaining more capabilities, leading to their increasing use as portable media players. However, the great majority of produced visual media is targeted towards devices with larger displays and greater hardware capabilities, making it difficult for these media to be presented on mobile devices. These difficulties will only increase as multimedia becomes more and more interactive, such as with increasingly popular mobile games. In this paper, we survey past methods for retargeting visual media to mobile devices, considering how such methods have been applied to text, still images and video content. We identify a shift from interactive retargeting techniques towards more automated methods as media richness increases. We discuss implications of this trend for the retargeting of 3D games and virtual worlds to mobile devices, particularly the need to focus on the retargeting of experience rather than solely visual content.

Keywords: mobile devices; mobile phones; multimedia; retargeting; games; visual media; experience; survey; text; images; video; mobile games; interaction.

Reference to this paper should be made as follows: Ross, J., Simpson, R. and Tomlinson, B. (2011) 'Media richness, interactivity and retargeting to mobile devices: a survey', *Int. J. Arts and Technology*, Vol. 4, No. 4, pp.442–459.

Biographical notes: Joel Ross is a PhD student in Informatics at the University of California, Irvine. His research interests include ubiquitous computing, human–computer interfaces and pervasive games. He received an MS in Informatics from the University of California, Irvine and a Bachelor's degree in Mathematics and English from Colorado College.

Robert Simpson is an Undergraduate at the University of California, Irvine. Currently, he is pursuing a BS in Computer Science with a minor in Philosophy.

Bill Tomlinson is an Associate Professor of Informatics at the University of California, Irvine and a Researcher in the California Institute for Telecommunications and Information Technology. He studies the fields of environmental information technology (Green IT), human–computer

interaction, multiagent systems and real-time graphics. He holds an MFA in Experimental Animation from CalArts and a PhD in Media Arts and Sciences from the MIT Media Lab.

1 Introduction

The rapid proliferation of mobile phones has made these devices ubiquitous in many parts of the world, accompanied by support for a wide variety of uses beyond simple voice calling. Modern mobile phones are capable media platforms, on which users can surf the internet, watch movies, and play increasingly advanced graphical games. Mobile phones have traditionally been a somewhat limited hardware platform when compared to stationary desktop computers – a lack of processing power and small displays with limited colour gamuts greatly restricted the type of media that could be presented on these devices. Today some of these limitations are disappearing (Capin et al., 2008): e.g. many phones now have full-colour displays with a resolution of 480×320 pixels or larger. Yet even as these capabilities increase, the form factor of mobile phones puts a finite limit on the overall size of displays, restricting screens to only a few inches so that users can carry the devices around. Even as technology advances and miniaturises, the form factor requirements of mobility limit the presentation of visual media and ensure that, for the foreseeable future, there will be profound differences in technology and interaction between mobile and stationary computing devices.

Because of these differences in display capabilities, many types of visual media need to be altered in order for users to view them on mobile platforms. This process of adjusting media for display on different devices is called media retargeting (Bouillon and Vanderdonckt, 2002; Liu and Gleicher, 2006; Rubinstein et al., 2008; Setlur et al., 2007). In its most common form, retargeting is simple image scaling: either reducing the number of pixels used to display an image such as by subsampling (averaging neighbouring pixels), or changing the size of each pixel to maintain the same display resolution across a larger area. Indeed, interacting with media has made this process seem almost natural to us – no one is surprised that the same television programme can show up on different sized TV sets. Nevertheless, in the process of retargeting to a smaller digital display, some information is almost always lost: as the size of a display decreases, the information bandwidth provided by that display also decreases. For example, scaling down an image often causes fine details to be lost, as a smaller number of pixels are used to show the same content. Thus a variety of advanced techniques are employed in retargeting to reduce the amount of information lost in the translation – to retain the 'sense' of the media content as much as possible.

Advanced retargeting methods become more of a concern as media richness (Daft and Lengel, 1986) increases. Media richness indicates the rate and quality of information conveyed by a medium, and that information's subsequent ability 'to change understanding within a time interval' (Daft and Lengel, p.650) – in a way, measuring the complexity of media forms. As mobile phones become more capable platforms, the richness and complexity of the media consumed on these devices also increases. While mobile phones were previously used to display only text and the occasional image, mobile devices are now marketed as portable platforms for videos and gaming.

To understand the effects of this increasing richness in mobile media, we conduct a broad survey of methods for retargeting textual, still image and video content to mobile devices (due to space limitations, we focus on retargeting research rather than commercial products, though commercial systems have similar characteristics). We consider retargeting techniques in terms of their interactivity – how much real-time input and control users have on the process of retargeting – in order to consider the application of these types of methods to richer and more interactive media, such as mobile games. This survey identifies a shift, as media becomes richer, from interactive (user-controlled) retargeting towards more automatic techniques. Informed by this trend, we consider implications for the design and retargeting of 3D games and other interactive media to mobile devices.

2 Retargeting, interactivity and media richness

Media retargeting is the process of adjusting media for display on different devices. In particular, retargeting visual media to small displays (such as those on mobile phones) involves displaying only a limited portion of the media content. This output can be seen as a kind of 'summary', as it shows only a representation of the full content – the reduced display size means that content cannot be presented in its entirety all at once. As such, smaller displays have smaller data bandwidths because of their size and so cannot show as much information, whether measured in words or pixels. Thus retargeting techniques often aim to select the most important information to display, ignoring the less important data. For example, while scaling an image through subsampling removes information (pixels) at regular intervals without regard to the image's content, cropping the edges of an image removes pixels that may not be part of the main subject of the image and thus may be less important for viewing. In this way, a fundamental concern in media retargeting is figuring out how to decide which data are the most important and should be included in the retargeted 'summary' – including the question of who gets to make that decision: the computer displaying the media or the user viewing it.

We consider retargeting techniques in terms of their level of interactivity – the amount of interaction and control the user has in the retargeting process. While some retargeting systems attempt to automatically determine the most important data to show on the small screen, other systems aim to let the user choose which data are important and should be displayed. For example, a system may automatically crop an image (Chen et al., 2003), or allow the user to choose which region of the image to display, such as by selecting and zooming in on a region (Rosenbaum and Schumann, 2005). Thus we arrange retargeting techniques along a continuum from those that are more automatic to those that are more interactive, allowing the user to control the retargeting. This distinction will help us to consider retargeting methods in relation to user behaviour surrounding a particular type of media: certain media forms may lend themselves to different levels of interactivity.

We also consider the distinction between whether the summary output of the retargeting is informative or indicative (Hahn and Mani, 2000). Informative summaries act as a replacement for the media content, providing the same overall information but in a manner suitable for the reduced size of the display. This is often the classic definition of 'summary', such as what may be found in a paper's abstract: a kind of recap of the content. In this way, informative summaries function to provide the desired media

content in a compressed form. Indicative summaries, on the other hand, provide a kind of overview of the media content, letting the user know where to find important or relevant information to view in more detail. For example, an index or a table of contents in a book would be a kind of indicative summary, serving as a map to more detailed views of the media content rather than attempting to provide all the information itself. Indicative summaries are tightly linked to interactive methods of media retargeting: such summaries by definition require some kind of user interaction to choose what content details to view, and thus are generally the product of more interactive retargeting techniques. Indeed, a retargeting method's level of interactivity can somewhat be seen in whether the summaries it creates are indicative (and so more interactive) or informative (and so more automatic).

In this paper, we survey retargeting methods applied to three forms of visual media: text, still images and video. We consider these forms based on the ideas of media richness theory (Daft and Lengel, 1986) – in which richness indicates the rate and quality of information conveyed by a medium, though we focus more on the medium itself rather than its use in communication. Thus for our purposes, we emphasise factors such as 'cues and channels utilised' and 'language variety' over 'rapid feedback' and 'personalisation' in considering whether a medium is rich or not. By these standards, we consider text, images and video to be progressively richer media: images add visual cues and symbols that are not present in plain text, and video adds a temporal channel to a series of still images (frames). This is not to say that the content of one medium is inherently richer than that of another – text can semantically contain as much or even more complex information than images or video, and may be able to transmit information faster and more clearly than other media. Nevertheless, we see an ordering in the progression from text, to images, to video that is reflected in the advancement of mobile device capabilities, and which we will thus use to organise and frame our analysis.

3 Retargeting text

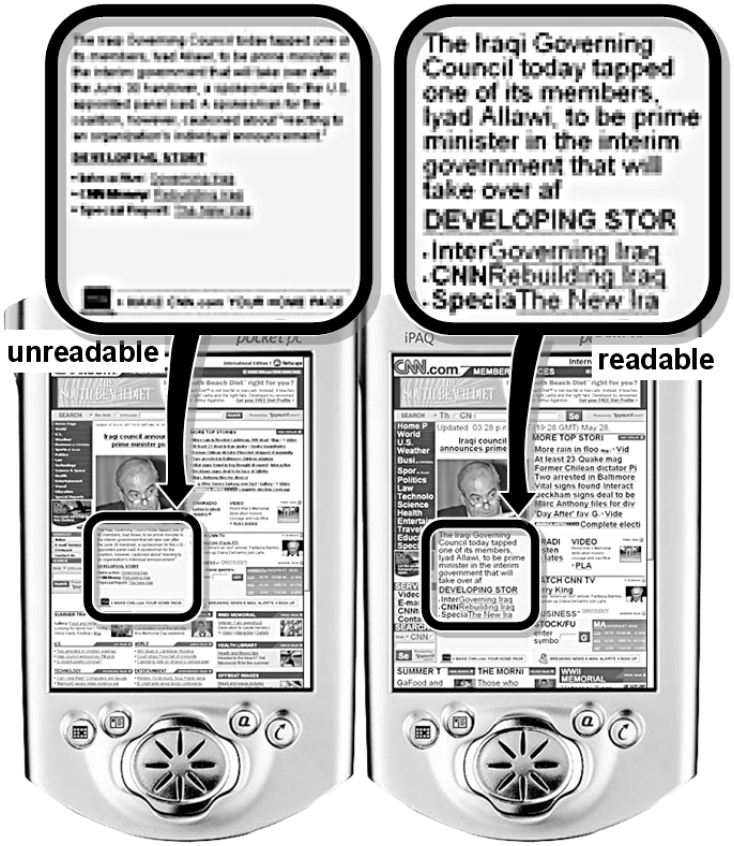
Text is among the most well established forms of visual media, from the earliest forms of writing to the creation of the Gutenberg Bible to more recent research in digital typography (Forlizzi et al., 2003). As such, text is one of the most common forms of media viewable on mobile phones and other mobile devices. Indeed, before mobile phones became ubiquitous, cellular pagers often had to contend with displaying textual messages on displays with very small resolutions. Text was also the medium of choice for early mobile internet browsing, before the advent of more multimedia intensive websites. Mobile phones without the wireless bandwidth for extensive image or video content would view the internet primarily as a textual medium, a paradigm that has persisted even with the increase in modern, multimedia-enabled phones.

In many ways, the internet has been the main focus of techniques for retargeting textual content to mobile devices, with efforts aimed at improving users' abilities to navigate web content. For example, the Power Browser (Buyukkokten et al., 2002) automatically creates an expandable, hierarchical view of summaries of web page text and navigational elements, while Chen et al. (2005) use the semantic structure explicit in a web page's HTML source to separate the page into sections that are viewable on a small display. By taking advantage of the hierarchical nature of HTML, such systems can easily determine how a website is organised and present that organisation as an indicative

summary to the user. However, HTML does not necessarily specify the most important content in the organised text, causing summarising the content of the text to be more difficult than summarising its organisation. Nevertheless, some retargeting methods do provide indicative summaries of the textual content itself, in addition to a summary of the organisation. The Summary Thumbnail system (Lam and Baudisch, 2005) retargets web pages by only providing snippets of text at a readable size, allowing users to choose which sections of the page to view in more detail (see Figure 1).

Other retargeting methods, such as the Minimap system (Roto et al., 2006), create scaled overviews of web pages that users can navigate and then zoom in on to view. Such systems may focus on reducing extraneous space between textual elements, thus easing and speeding up the process of navigating the zoomed target area. These techniques use the visual layout of the web page to help organise navigation through the text, treating the whole page as a kind of text-filled image that users can browse. Indeed, the interaction modes used to view images, described in Section 4, are also used in this manner to view web pages on small displays. Thus textual content is frequently retargeted to mobile devices using a 'pan + zoom' (Cockburn et al., 2008) form of interaction.

Figure 1 Summary Thumbnails create indicative summaries that allow users to interactively retarget web content



Source: Image from Lam and Baudisch (2005), used with authors' permission.

These text retargeting techniques all involve user interaction: although the text is automatically summarised into viewable portions, user interaction is still required for choosing which portion of the content to view in greater detail. These methods create indicative summaries of the text, prompting the user to interact with the system and navigate the retargeted page before viewing the text at its full resolution (i.e. the complete text in a readable size). In this way, such indicative summaries elicit significant user interaction in the viewing experience. However, not all textual retargeting methods involve creating indicative summaries requiring user interaction. For example, adaptive grid-based templates (Jacobs et al., 2003) can be used to retarget the text layout itself, so that the displayed sections of the content remain organised across different display resolutions and aspect ratios. Nevertheless, large amounts of textual content may not be displayable in full on the small screen while remaining readable, and thus may still require user interaction in the form of pagination – itself a kind of retargeting. Thus overall, the detail needed for text to be readable causes such content to be retargeted primarily through methods that are more interactive.

Interestingly, while summarising web pages for navigation on mobile devices involves retargeting previously published content, text is also retargeted when authored on mobile devices. For example, the shortened word forms commonly found in SMS messaging (i.e. 'textese' – see e.g. Grinter and Eldridge, 2003) can be seen as retargeted text. Here unimportant data – in this case at letter resolution – is discarded as the content is being authored. The already short messages are further summarised, perhaps in part because of the effort of writing messages on phones with limited keypads. In this way, media content is retargeted not just because of restricted output displays, but also because of restricted input controls. Indeed, the very act of picking SMS as a communication medium also works to retarget the content of that communication, as the character limit may restrict the length and subject of a conversation (Grinter and Eldridge, 2001). Thus textual content is interactively retargeted to mobile devices not only because of the constraints imposed by the mobile form factor, but also because of how the mobile form factor shapes people's use of and interaction with these devices.

4 Retargeting images

Another medium commonly retargeted to mobile devices is still images. In our taxonomy, images represent a richer medium than text, with a wider variety of cues and channels used – images generally make more extensive use of spatial layout and colour channels than plain text. Some previous research in retargeting still images has focused on such colour channels, altering images so that they may be displayed on colour-limited displays. For example, Rist and Brandmeier (2002) use machine-learning algorithms to determine the best colour settings to use in retargeting colour images to mobile, black and white displays. Indeed, although recent mobile devices have displays with large colour gamuts and so can display full-colour images, the mobility of such displays means that images will be viewed under a variety of different environment lighting conditions. Thus dynamic colour balancing (such as with the cameras commonly found on mobile phones (Lindsay et al., 2008)) remains important for retargeting and displaying images on mobile devices.

Nevertheless, the majority of retargeting methods focus on modifying the spatial layout of images in order to display them on the small screens of mobile devices. Interfaces for navigating visual media (Cockburn et al., 2008) allow users to view and navigate images on mobile devices, providing a form of interactive retargeting. For example, Rosenbaum and Shumann (2005) use a grid to support 'pan + zoom' interaction, in which the user moves a window of interest in order to view different sections of the image – in effect dynamically choosing a subregion of the image to display as a detailed summary. 'Pan + zoom' is one of the most common interaction modes for viewing retargeted images and spatially organised text (which approaches images in terms of richness), and can be found in the web browsers of most popular web-enabled mobile phones, such as the Apple iPhone.

Other systems for mobile web browsing (e.g. Lam and Baudisch, 2005; Roto et al., 2006) may use an 'overview + detail' interaction method to display images. With this method, the user is shown an indicative overview map of the image, along with a zoomed-in 'detail' view of part of the image. Interesting, 'overview + detail' is also often used for playing games and navigating virtual worlds on full-size displays, where the virtual world is still too expansive to be shown all at once – instead, an overview map of the world is simply shown in one corner. Finally, fisheye retargeting (Gutwin and Fedak, 2004) is an example of a 'focus + context' interaction, in which the retargeted summary (the in-focus area) is presented along with information about the location or context of that summary (the distorted region). Thus there are a variety of systems that use a range of methods for interactively retargeting images, with the user choosing how to display the summarised images.

In addition to techniques for enabling interactive, user-driven image viewing on mobile devices, there are also a wide variety of advanced methods for automatically retargeting images to small displays. Such methods commonly use measures of saliency – or what parts of the image are most interesting or important, usually people and objects – in order to determine which regions of the image to include in the retargeted summary (Itti et al., 1998). Saliency may be determined through a combination of edge detection, colour region location and possibly face recognition in order to identify objects that may be important to the image's meaning and so should be displayed in the retargeted images.

After saliency has been determined, the less salient regions may be cropped out of an image (Chen et al., 2003) in order to show the image on a small mobile display. Alternatively, Liu and Gleicher (2005) automatically create fisheye views around the salient region of an image – such a method also demonstrates how more interactive retargeting methods can potentially be automated. To handle salient regions being spread out across an image, Setlur et al. (2007) 'cut' objects from the image and then 'paste' them onto a resized background, resulting in a retargeted image (though no longer photorealistic) that contains all the important content (see Figure 2). These types of retargeting techniques enable significant regions to automatically be detected and displayed to the user on a small display, rather than requiring the user to manually locate regions of interest to view.

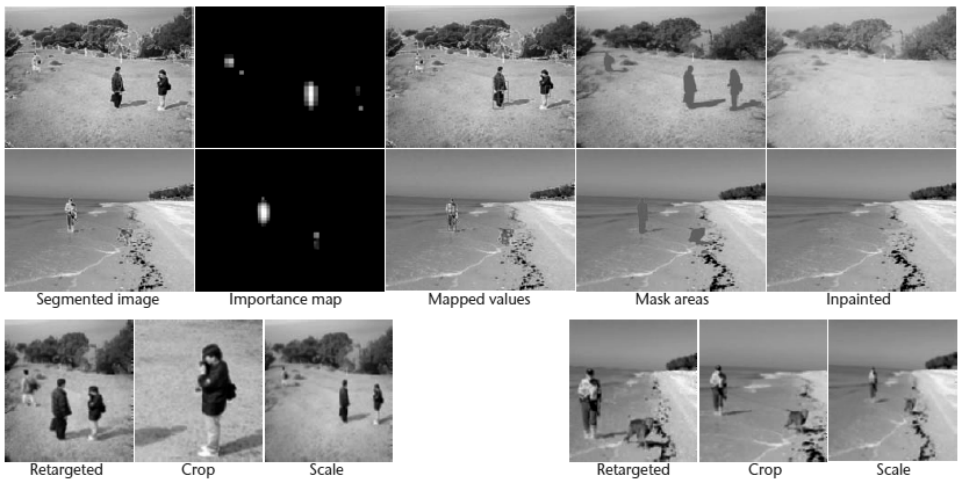
Other automatic retargeting methods focus on retargeting the image in a non-homogenous manner, rather than on selecting a salient region of interest. In these systems, images are altered so that the content remain clear, rather than focusing on finding the 'most important' region to display. For example, seam carving (Avidan and Shamir, 2007) removes 'seams' of low-energy pixels in an image – effectively cropping the image from within by removing a connected arc of pixels. In this way, region saliency

is not considered; instead, the value of a particular pixel is determined in order to decide which pixels to display in the retargeted summary. Although the layout of the image has changed, the overall meaning of the content remains the same.

Similarly, in mesh-based image warping (Wang et al., 2008) tiny quadrilateral regions of the image (containing only a few pixels) are warped and compressed so that the majority of the image distortion takes place where it is least noticeable, thereby allowing images to be resized for small displays. In a different vein, Simakov et al. (2008) automatically retarget images by computing a retargeted summary image of a given size that is most similar to the original image. Here similarity is measured in terms of completeness (all elements, or patches, of one size image are in the other) and coherence (there are no new visual artefacts), and can potentially produce better results (i.e. less artefacts) than seam carving. These kinds of techniques are able to coherently and automatically retarget images to a wide variety of different sized displays, making them highly suitable for the heterogeneous world of mobile devices. However, many automated retargeting systems require significant computational power and may potentially be difficult to run on processor-limited mobile phones.

Thus retargeting images for mobile devices uses a mix of interactive and automatic techniques, as opposed to the primarily interactive methods for retargeting textual content. Indeed, Santella et al. (2006) use interactive methods in order to configure automatic retargeting: users' gaze paths are used to determine salient regions, which are then the subject of automatic cropping based on the target image size. In this system, images are interactively retargeted at authoring time and then automatically retargeted at view time. So while text is primarily retargeted interactively, images are retargeted both interactively and automatically. Avidan and Shamir (2007) even describe how seam carving can also be used to interactively alter and retarget an image, such as through removing particular objects or elements. Although such a user interface may not be suitable for mobile devices, it further demonstrates how methods for retargeting images have mixed levels of automation and interactivity.

Figure 2 Automatic 'cut and paste' retargeting keeps the important parts of the image without user interaction, at the expense of photorealism



Source: Image from Setlur et al. (2007), used with authors' permission.

This increase in automation in retargeting techniques for images may be due to the increased richness of images as a medium. As information is increasingly found in non-semantic channels, users are able to understand the content without needing its full details (viewing an informative summary rather than an indicative summary) – an image with missing pixels may be easier to understand than text with missing words. Thus images can be cropped and scaled and still be comprehendible (instead of needing to be navigated), in a way that is more difficult with the primarily semantic information in text.

5 Retargeting video

Viewing video media on mobile devices and mobile phones is becoming more common as such devices become more advanced. Video is a richer medium than images or text – while images add more spatial information channels to the semantic information channels of text, video adds a significant temporal channel to images, displaying a series of still images over time. As such, video can be summarised (retargeted) in two different domains: the temporal and the spatial.

Temporal video summarisations (e.g. see reviews by Adami et al., 2006; Truong and Venkatesh, 2007) aim to automatically decrease the length of the video, so that the content can be viewed in a smaller amount of time. For example, the most important parts of a sports event like a football game may be selected and joined together to provide an overall summary of the game. Techniques for performing this summarisation commonly involve either selecting key frames and creating a 'storyboard' summary (presenting a broad overview but potentially losing temporal continuity (e.g. Ferman and Tekalp, 2003)), or selecting key shots and creating a kind of 'highlight' summary (presenting contained narrative pieces but potentially glossing over large portions of the content (e.g. Rui et al., 2000)). These key frames or shots are often automatically selected based on saliency or user attention models (Ma et al., 2005), just as important regions of still images are identified (although the audio track usually found in video can also be harnessed for determining important segments of the video (Hua et al., 2003) and indeed is a key component of the medium, due to space limitations we only consider the visual aspect of video in terms of retargeting). In this way, video can be summarised temporally much the same way as images are summarised spatially.

However, temporal video summarisation is not exclusive to the small screens of mobile devices, as such summarisation can be important for video content viewed on any size display. Thus we focus on how video content is spatially summarised and retargeted to small screens and mobile devices. These spatial retargeting techniques are primarily automatic in nature, providing a resized view of the video without requiring user navigation or interaction. For example, Liu and Gleicher (2006) automate the process of 'pan-and-scan', in which a window of interest pans and scans across the video in order to show it on a smaller display or aspect ratio (such as to show a widescreen (16:9) movie on a full-screen (4:3) display without letterboxing). This process is similar to the 'pan + zoom' interface for navigating still images, though it is automated rather than interactive. Indeed, this system also adds virtual pans and cuts to deal with the centre of the action shifting across the frame over the course of a shot, thus adapting the image retargeting technique to deal with the added temporal channel in the richer medium of video.

Other image retargeting techniques have also been adapted to retargeting video media. Wolf et al. (2007) use image warping (similar to that used in images, described above) in order to perform non-homogeneous rescaling on streaming video – a form of the medium common to memory-limited mobile devices. With this technique, each frame is warped as if it was a still image, but consecutive frames are checked to make sure that the video remains smooth. Similarly, seam carving has also been extended to video (Rubinstein et al., 2008), removing planar seams of pixels from the time-space volume of the video and thereby maintaining video continuity (see Figure 3). Thus video retargeting techniques are often equivalent to those for retargeting images, with the further requirement of retaining the consistency of the added temporal channel. Video retargeting can be seen as a direct extension of automatic image retargeting. Furthermore, as with images, automatic retargeting techniques for video often focus on conveying the general idea and information of the media content, rather than maintaining photorealism (a phenomenon also common in temporal summarisation, in which retargeted video may no longer present concurrent images in real time).

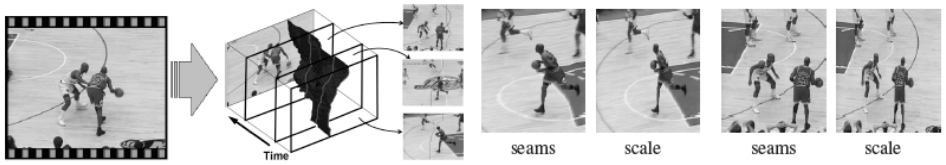
These examples help indicate how video retargeting – whether for temporal or spatial summarisation – is primarily an automatic process, building on automatic image retargeting techniques. Nevertheless, some previous research has considered more interactive techniques for retargeting video content. For example, Fan et al. (2003) describe a system where users are able to pick regions of interest in a video on which to focus while the video is playing (navigating the spatial channel, but not the temporal), similar to an interactive 'pan + zoom' interface for images. However, this system is also able to automatically choose regions of interest to show to users in order to improve ease of use – the complexity of the rich medium can be difficult for people to navigate, and thus automatic retargeting may be desired more than interactive methods when dealing with video.

Indeed, the majority of methods for interacting with video content focus on video browsing, where users navigate temporally through a video in order to locate the time segment they wish to view – a process familiar to anyone who has worked a VCR or DVD player. Previous research (e.g. Sun and Hürst, 2008) has considered a variety of interfaces for browsing video on mobile devices, including the use of sliders, pens and scroll-wheels. These interfaces help to enable browsing on devices with limited inputs, but such interaction still lies primarily along the linear temporal dimension of the video, instead of interacting with the videos spatially. It is likely that these interfaces focus on the temporal dimension rather than the three-dimensional time-space volume of the video content because of the difficulty people have learning to navigate 3D virtual spaces (Fitzmaurice et al., 2008) – a difficulty that is only complicated by the limited inputs available to mobile phones. Furthermore, as described above, this temporal navigation is not exclusive to retargeting video content to mobile devices; it is used on both large and small displays. Thus spatial retargeting of video so that it is viewable on the small displays of mobile devices remains a primarily automatic process rather than an interactive one.

Research by O'Hara et al. (2007) into how users view video on mobile devices may provide further insights into why video retargeting on these devices is primarily automatically in nature. The study describes how users on the go (and so using their mobile phones) require a sufficient amount of time in order to get out their devices and start viewing video content, yet such videos are often used to fill in particularly small gaps of available time. Indeed, this paper remarks on how mobile video is often not

suitable for transitions between spaces (such as for getting on a bus or train), because visual attention is often needed. As such, interactive retargeting techniques – which require even more of a user's awareness – may not be suitable for viewing video on mobile devices because mobile users may not be able to focus on interacting with the content. The limits of human attention may play a role in the need to automate the retargeting of increasingly rich media content. Furthermore, O'Hara et al. note how mobile devices may be used to share media content, with groups of people gathering around the tiny display to watch a video. This style of shared viewing may also make interaction difficult, as it would require extensive coordination for multiple people to interactively retarget the content. Thus video content may be increasingly automatically retargeted in part because of the ways in which people consume mobile video as they go about their lives, as well as the difficulties involved in interactively navigating the temporal and spatial channels of this rich media form.

Figure 3 Seam carving video treats the content as a spatial-temporal volume in dealing with the richer medium



Source: Image from Rubinstein et al. (2008), used with authors' permission.

6 Increased richness leads to increased automation

Our review of retargeting research reveals a trend: as media richness increases, the current techniques for retargeting that media content to mobile devices become more automatic (see Figure 4). Moving from text to still images to video increases the number and variety of available cues, symbols and channels – images add spatial layout to plain text, and video adds a temporal dimension to sequences of images – thereby signifying an increase in what we term richness. And as richness increases, the methods used to retarget media to the small screens of mobile devices become less interactive and more automated. Thus text retargeting tends to involve some level of user interaction (as users choose content to view in detail from indicative summaries), while retargeting images involve a mix of interactive and automatic systems (both navigational interfaces and automatic resizing algorithms), and video retargeting is primarily automatic (with interactive systems focusing primarily on video browsing rather than retargeting to small displays). In this way, an increase in media richness appears to correspond with a decrease in interactive retargeting systems.

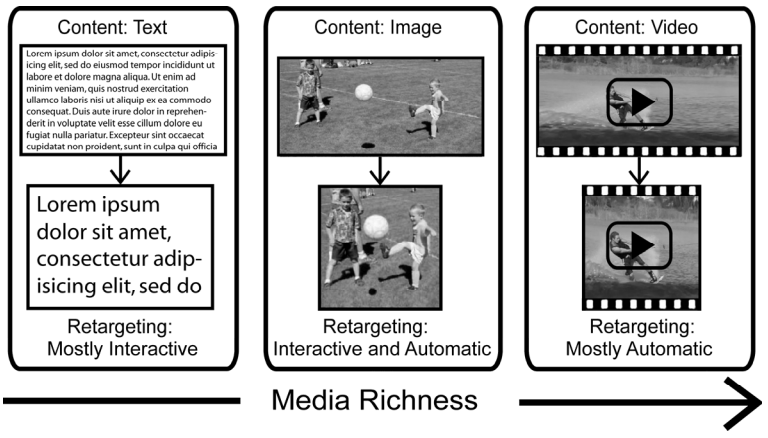
Note that while the discussion here has focused on retargeting research, commercial products also appear to currently follow this trend, providing interactive retargeting capabilities for textual content and automatic retargeting for video content. For example, textual and image-based media such as the internet are increasingly navigated with interactive 'pan + zoom' style full web-browsers, such as on the Apple iPhone, rather than more automatic (from the user's perspective) summarisations that could be found in

WML-based browsers (in which the appearance of the content on mobile platforms was predefined). As the hardware capabilities of mobile devices improve, people continue to want increasing amount of interactivity in how they view the web. At the same time, video content shown on mobile devices continues to be automatically subsampled and resized, with all the loss of detail that implies. Thus less rich media are interactively retargeted, while more rich media are automatically retargeted.

Nevertheless, these automatic retargeting techniques often fail to properly account for the subjective nature of the semantic content they are retargeting. Processing semantic data is difficult, so automatic retargeting techniques instead often focus on spatial data, with less consideration for semantic information – images and video are cropped or subsampled, with semantic information only approximated through saliency measures. Yet these limited, low-level automatic techniques continue to be used (particularly in commercial retargeting systems), and may in fact be necessary because of the difficulty of retargeting media interactively with the limited inputs of mobile devices.

Because of the small size and compact design required by a mobile form factor, mobile phones usually have only a small set of inputs – often just a few buttons, though increasingly mobile phones have touch screens or full (if miniature) keyboards. Yet even with this new hardware, the relatively limited forms of input may restrict the amount of interaction that a user can have with media and in particular users' ability to control retargeting interactively. As media content becomes richer, retargeting needs to occur along a wider variety of channels – interacting with this greater number of channels may require a greater amount of user input. For example, while images may be navigable in two dimensions (possible with a small keypad), video adds a temporal channel – effectively requiring controls for complex 3D navigation – as well as potentially needing inputs to control retargeting of an audio track. Mobile devices may not have the inputs required to support interactive retargeting methods, even if those methods are more capable of dealing with semantic content. For this reason, richer media may in fact require rich media content to be automatically retargeted – as demonstrated by the trend towards more automatic methods for richer media – because the limited affordances of a mobile form factor do not effectively support interactive methods.

Figure 4 As the richness of a medium increases, retargeting methods shift from mostly interactive to mostly automated



Source: Examples taken from Setlur et al. (2007) and Rubinstein et al. (2008).

7 Retargeting 3D games and virtual worlds

The link between media richness and the interactivity of retargeting methods offers potential insights into the future retargeting of other media forms, such as graphical games and virtual worlds in which the user navigates an avatar through a 3D virtual space shown on the display. For example, consider playing Second Life or World of Warcraft on a mobile phone (Figure 5). 3D gaming content is becoming increasingly popular on mobile devices, particularly as display and processing capabilities have improved. Although not all mobile games (or even the majority) are adapted versions of games for other, non-mobile platforms, with the rise in popularity of massively multiplayer games and online virtual worlds, users may desire such systems to be accessible on their increasingly powerful mobile phones. For example, players may want to access their Second Life avatars while on the move, perhaps to mirror their own real-life travels in the game's world. Similarly, the popularity of non-mobile games or particular franchises may make mobile versions of games commercially lucrative. Thus we believe that it is valuable to consider the process of retargeting 3D games and virtual worlds to mobile devices.

Mobile 3D games have many of the same design requirements as games developed for other platforms, but with added constraints based on potential restrictions of the mobile form factor (Korhonen and Koivisto, 2006). While such games continue to need to be fun and engaging, the mobile nature of the platform also means that the characteristics of the environment (lighting, ambient noise, etc.) may change frequently as the user moves around. As with video, users may only be able to play a game in small blocks of time because of their mobile behaviour, which can have significant effects on how games are played. Furthermore, the limited processing power, restricted number of inputs and small displays of mobile phones may also affect how games are presented on these devices, just as with other media forms. Thus – like text, images and video – 3D games may require a level of retargeting in order to be presented and played on mobile devices.

Figure 5 Users may wish to play 3D games or interact in virtual worlds through mobile devices, as demonstrated in this mock-up

Building on our earlier taxonomy, we see 3D games and virtual worlds as representing a richer form of visual media than text, images or video. While images add a spatial channel to text and video adds a temporal channel to images, 3D games add a new set of channels to video in the form of user interactions – playing a game involves user input and control in addition to the spatial and temporal presentation of video – as well as providing increased feedback and personalisation to users. Indeed, this added channel of interactivity has significant influence on the spatial and temporal presentation of the content as the user controls characters or elements in the game (just as, with video media, the temporal movements of actors in a movie shapes the spatial presentation of the cinematography). Mobile games may also add an additional information channel in the form of real world localisation and actions (e.g. Barkhuus et al., 2005), further increasing the potential rate of communication and hence the richness of the medium. In this way, 3D games and virtual worlds can be seen as a next step in the progression of increasingly rich media available on mobile devices.

Based on the trend identified in this paper, 3D games' position as a richer medium would suggest the use of more automatic retargeting techniques. The limited input capabilities of mobile devices combined with the increased complexity of the medium's form may cause 3D games to rely on automatic methods rather than interactive ones. Indeed, media such as 3D games and virtual worlds may be well suited for automatic retargeting methods, as they often already include natively digital and programmatic models of their content. Unlike with other media forms, little effort may be needed to understand and interpret the content of the media (such as parsing the natural language of text or identifying objects in images or video) because of the highly authored and artificial nature of virtual worlds. It is possible for retargeting characteristics – such as what objects to display or what details to lose when viewing a virtual world on a particular screen size – to be defined at the time of authorship. Designers and artists could specify hints for how best to automatically retarget a particular 3D game to the heterogeneous requirements of mobile platforms, much as semantic metadata may be used to assist automatic web page retargeting.

However, the interactive nature of 3D games and virtual worlds may problematise the use of automatic retargeting methods. Often a significant portion of 3D game presentation is user controlled: users implicitly choose what content is shown on the screen as they navigate through a virtual world, or even explicitly choose though interactive camera controls (e.g. Gleicher and Witkin, 1992) – a form of interactive retargeting. 3D games and virtual worlds may contain distinct content in the form of the interactions they enable, in addition to their visual content – the process of interactively retargeting a game's content through playing the game can be a significant part of the fun of the game. So although the increased complexity of this richer medium may suggest a need for more automatic retargeting methods in order to lessen the burden placed on the user, a decrease in the interactivity of the retargeting may detract from the content and enjoyment of the game.

A potential answer to this conflict may be found in how many advanced automated retargeting techniques for richer media do not aim to maintain photorealism or the original composition of the content. Methods such as seam carving and image warping alter images and video to remove unimportant regions, and even automated pan-and-scan adds new camera movements and cuts (themselves part of the content), all in order to preserve the overall information found in the content. The trend towards increased automation as media richness increases thus accompanies a shift away from the need to

preserve the exact details of media content – the overall idea is important, not the exact form or composition. When applied to mobile 3D games, this idea suggests that in order to avoid potentially reducing the enjoyment of a game by simply automating visual retargeting, retargeting processes should focus on maintaining the experience of playing a 3D game or virtual world rather than its visual content.

The goal in retargeting media to mobile devices is to reduce the amount of information lost because of limited output capabilities, and thus to retain the 'sense' of the media content as much as possible. In text, images, and video, this essence can usually be found in the visual details of the content and the information it presents. However, with interactive media such as 3D games or virtual worlds, it is usually the experience of that interaction that defines the content, not its visual presentation. The interaction and the way a game is played separate that game from a video, and it is the experience – the feeling – of playing a game that defines it (Federoff, 2002). It is this experience that should be preserved in the retargeted content. For example, in retargeting an online game like World of Warcraft, one should maintain the experience of exploring a fantastical world even at the expense of that world's graphical splendour, or of communicating with other avatars even if those avatars are not graphically detailed. Such experiences can be found even in text-based virtual worlds (Turkle, 1997), a less rich medium than 3D graphical games, suggesting that the graphical nature of the game is not necessarily a requirement for the overall idea of playing that game. The overall idea of many games may be found in how users interact with them, not how they are presented visually – thus the visual presentation may be automatically retargeted to mobile devices, without compromising the experience of playing the game.

8 Conclusion

In this paper, we have surveyed methods for retargeting visual media (specifically text, images and video) to mobile devices. This review has identified a trend towards more automated retargeting methods rather than interactive techniques as media becomes richer. We have discussed potential implications of this trend for retargeting interactive media in the form of mobile 3D games and virtual worlds, particularly the need to transfer and maintain the experience of the game rather than its visual content. Such retargeting has the potential to increase users' enjoyment of games and other forms of multimedia when presented on their mobile devices. Furthermore, such retargeting techniques may be applicable to a wide variety of other areas, from search applications to computer-supported learning or computer-mediated communication involving a large number of people.

Indeed, as mentioned above, the portable nature of these mobile platforms means that users often use them to share media content with others – a sharing that may be a key component of multimedia for both work and entertainment. This sense of sharing is particularly present in online multiplayer games and virtual worlds, for which the feeling of playing with others is in fact a core component of the experience (Ducheneaut et al., 2006). Being able to retarget multimedia to a wider variety of platforms may be able to increase the diversity of people who can participate in and experience these systems, as well as providing a basis for a range of new interactive multimedia experiences for multiple users on multiple platforms.

Because of the multiplicity of current (and likely future) devices, retargeting is an important technique in the creation of high quality media content. Whether providing universal access, enabling cross-media franchises or developing new kinds of heterogeneous multidevice systems, retargeting is likely to play a growing role in the creation and propagation of multimedia content and experiences. Enabling technology designers and content creators to have an understanding of retargeting – as well as improved tools to perform retargeting effectively – can enhance the ways in which media are propagated across ubiquitous mobile platforms.

Acknowledgements

This research has been supported by NSF grant IIS-0644415, the California Institute for Telecommunications and Information Technology (Calit2) SURF-IT program and the Donald Bren School of Information and Computer Sciences. The authors thank Garnet Hertz and the members of the Social Code Group.

References

Adami, N., Benini, S. and Leonardi, R. (2006) 'An overview of video shot clustering and summarization techniques for mobile applications', *Proceedings of the 2nd International Conference on Mobile Multimedia Communications*, Alghero, Italy: ACM, pp.1–6.

Avidan, S. and Shamir, A. (2007) 'Seam carving for content-aware image resizing', *ACM SIGGRAPH 2007 Papers*, San Diego, CA: ACM, p.10.

Barkhuus, L., Chalmers, M., Tennent, P., Hall, M., Bell, M., Sherwood, S. and Brown, B. (2005) 'Picking pockets on the lawn: the development of tactics and strategies in a mobile game', *UbiComp 2005: Ubiquitous Computing*, pp.358–374.

Bouillon, L. and Vanderdonckt, J. (2002) 'Retargeting of web pages to other computing platforms with VAQUITA', *Proceedings of the Ninth Working Conference on Reverse Engineering (WCRE'02)*, IEEE Computer Society, p.339.

Buyukkokten, O., Kaljuvee, O., Garcia-Molina, H., Paepcke, A. and Winograd, T. (2002) 'Efficient web browsing on handheld devices using page and form summarization', *ACM Transactions on Information Systems*, Vol. 20, No. 1, pp.82–115.

Capin, T., Pulli, K. and Akenine-Moller, T. (2008) 'The state of the art in mobile graphics research', *Computer Graphics and Applications, IEEE*, Vol. 28, No. 4, pp.74–84.

Chen, L., Xie, X., Fan, X., Ma, W., Zhang, H. and Zhou, H. (2003) 'A visual attention model for adapting images on small displays', *Multimedia Systems*, Vol. 9, No. 4, pp.353–364.

Chen, Y., Xie, X., Ma, W. and Zhang, H. (2005) 'Adapting web pages for small-screen devices', *Internet Computing, IEEE*, Vol. 9, No. 1, pp.50–56.

Cockburn, A., Karlson, A. and Bederson, B.B. (2008) 'A review of overview+detail, zooming, and focus+context interfaces', *ACM Computing Surveys*, Vol. 41, No. 1, pp.1–31.

Daft, R.L. and Lengel, R.H. (1986) 'Organizational information requirements, media richness and structural design', *Management Science*, Vol. 32, No. 5, pp.554–571.

Ducheneaut, N., Yee, N., Nickell, E. and Moore, R.J. (2006) 'Alone together? exploring the social dynamics of massively multiplayer online games', *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Montréal, Québec, Canada: ACM, pp.407–416.

Fan, X., Xie, X., Zhou, H. and Ma, W. (2003) 'Looking into video frames on small displays', *Proceedings of the Eleventh ACM International Conference on Multimedia*, Berkeley, CA, USA: ACM, pp.247–250.

Federoff, M.A. (2002) 'Heuristics and usability guidelines for the creation and evaluation of fun in video games', Master's Thesis, Indiana University.

Ferman, A. and Tekalp, A. (2003) 'Two-stage hierarchical video summary extraction to match low-level user browsing preferences', *Multimedia, IEEE Transactions*, Vol. 5, No. 2, pp.244–256.

Fitzmaurice, G., Matejka, J., Mordatch, I., Khan, A. and Kurtenbach, G. (2008) 'Safe 3D navigation', *Proceedings of the 2008 Symposium on Interactive 3D Graphics and Games*, Redwood City, CA: ACM, pp.7–15.

Forlizzi, J., Lee, J. and Hudson, S. (2003) 'The kinedit system: affective messages using dynamic texts', *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Ft. Lauderdale, FL, USA: ACM, pp.377–384.

Gleicher, M. and Witkin, A. (1992) 'Through-the-lens camera control', *SIGGRAPH Computer Graphics*, Vol. 26, No. 2, pp.331–340.

Grinter, R. and Eldridge, M. (2001) 'y do tngrs luv 2 txt msg?' *Proceedings of the Seventh Conference on European Conference on Computer Supported Cooperative Work*, Bonn, Germany: Kluwer Academic Publishers, pp.219–238.

Grinter, R. and Eldridge, M. (2003) 'Wan2tlk? everyday text messaging', *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Ft. Lauderdale, FL, USA: ACM, pp.441–448.

Gutwin, C. and Fedak, C. (2004) 'Interacting with big interfaces on small screens: a comparison of fisheye, zoom, and panning techniques', *Proceedings of Graphics Interface 2004*, London, ON, Canada: Canadian Human–Computer Communications Society, pp.145–152.

Hahn, U. and Mani, I. (2000) 'The challenges of automatic summarization', *Computer*, Vol. 33, No. 11, pp.29–36.

Hua, X., Lu, L. and Zhang, H. (2003) 'AVE: automated home video editing', *Proceedings of the Eleventh ACM International Conference on Multimedia*, Berkeley, CA, USA: ACM, pp.490–497.

Itti, L., Koch, C. and Niebur, E. (1998) 'A model of saliency-based visual attention for rapid scene analysis', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 11, pp.1254–1259.

Jacobs, C., Li, W., Schrier, E., Bargeron, D. and Salesin, D. (2003) 'Adaptive grid-based document layout', *ACM Transactions on Graphics*, Vol. 22, No. 3, pp.838–847.

Korhonen, H. and Koivisto, E. (2006) 'Playability heuristics for mobile games', *Proceedings of the 8th Conference on Human–Computer Interaction with Mobile Devices and Services*, Helsinki, Finland: ACM, pp.9–16.

Lam, H. and Baudisch, P. (2005) 'Summary thumbnails: readable overviews for small screen web browsers', *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Portland, OR, USA: ACM, pp.681–690.

Lindsay, C., Agu, E. and Wu, F. (2008) 'Dynamic correction of color appearance on mobile displays', *Proceedings of Graphics Interface 2008*, Windsor, ON, Canada: Canadian Information Processing Society, pp.25–31.

Liu, F. and Gleicher, M. (2005) 'Automatic image retargeting with fisheye-view warping', *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology*, Seattle, WA, USA: ACM, pp.153–162.

Liu, F. and Gleicher, M. (2006) 'Video retargeting: automating pan and scan', *Proceedings of the 14th Annual ACM International Conference on Multimedia*, Santa Barbara, CA, USA: ACM, pp.241–250.

Ma, Y., Hua, X., Lu, L. and Zhang, H. (2005) 'A generic framework of user attention model and its application in video summarization', *Multimedia, IEEE Transactions*, Vol. 7, No. 5, pp.907–919.

O'Hara, K., Mitchell, A.S. and Vorbau, A. (2007) 'Consuming video on mobile devices', *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, San Jose, CA, USA: ACM, pp.857–866.

Rist, T. and Brandmeier, P. (2002) 'Customizing graphics for tiny displays of mobile devices', *Personal and Ubiquitous Computing*, Vol. 6, No. 4, pp.260–268.

Rosenbaum, R.U. and Schumann, H. (2005) 'Grid-based interaction for effective image browsing on mobile devices', *Multimedia on Mobile Devices*, San Jose, CA, USA: SPIE, pp.170–180.

Roto, V., Popescu, A., Koivisto, A. and Vartiainen, E. (2006) 'Minimap: a web page visualization method for mobile phones', *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Montréal, Québec, Canada: ACM, pp.35–44.

Rubinstein, M., Shamir, A. and Avidan, S. (2008) 'Improved seam carving for video retargeting', *ACM Transactions on Graphics*, Vol. 27, No. 3, pp.1–9.

Rui, Y., Gupta, A. and Acero, A. (2000) 'Automatically extracting highlights for TV Baseball programs', *Proceedings of the Eighth ACM International Conference on Multimedia*, Marina del Rey, CA, USA: ACM, pp.105–115.

Santella, A., Agrawala, M., DeCarlo, D., Salesin, D. and Cohen, M. (2006) 'Gaze-based interaction for semi-automatic photo cropping', *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Montréal, Québec, Canada: ACM, pp.771–780.

Setlur, V., Lechner, T., Nienhaus, M. and Gooch, B. (2007) 'Retargeting images and video for preserving information saliency', *Computer Graphics and Applications, IEEE*, Vol. 27, No. 5, pp.80–88.

Simakov, D., Caspi, Y., Shechtman, E. and Irani, M. (2008) 'Summarizing visual data using bidirectional similarity', *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference*, pp.1–8.

Sun, Q. and Hürst, W. (2008) 'Video browsing on handheld devices – interface designs for the next generation of mobile video players', *IEEE MultiMedia*, Vol. 15, No. 3, pp.76–83.

Truong, B.T. and Venkatesh, S. (2007) 'Video abstraction: a systematic review and classification', *ACM Transactions on Multimedia Computer Communication Applications*, Vol. 3, No. 1, p.3.

Turkle, S. (1997) *Life on the Screen: Identity in the Age of the Internet*. New York: Simon & Schuster.

Wang, Y., Tai, C., Sorkine, O. and Lee, T. (2008) 'Optimized scale-and-stretch for image resizing', *ACM Transactions on Graphics*, Vol. 27, No. 5, pp.1–8.

Wolf, L., Guttmann, M. and Cohen-Or, D. (2007) 'Non-homogeneous content-driven video-retargeting', *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference*, pp.1–6.