Glyph: Lightweight and Evocative Looping Images in the News

by Savannah Niles
B.F.A., B.A. Southern Methodist University (2013)

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for a degree of

Master of Science in Media Arts and Sciences

at the

Massachusetts Institute of Technology

June 2015

Signature redacted

Author.....

Savannah Niles
Program in Media Arts and Sciences
May 15th, 2015

Signature redacted

Certified by.....

Andrew B Lippman
Senior Research Scientist, Program in Media Arts and Sciences
Thesis Supervisor

Signature redacted

Accepted by.....

Patricia Maes
Associated Academic Head, Program in Media Arts and Sciences

Glyph: Lightweight and Evocative Looping Images in the News

by Savannah Niles

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, on May 15, 2015, in partial fulfillment of the requirements for a degree of Master of Science in Media Arts and Sciences

Abstract

There is an emotional dimension to the informative function of the news. When we read the news we participate in a collective emotional experience– whether that is grief, celebration, worry, or wonder. News video is a crucial vector for these shared emotional experiences, which can propel civic action. But video comes at a high cost in time and attention, and is thus unsuitable for high-volume news and social feeds, and mobile and wearable devices. On these interfaces, there is value in presenting video in a way that's immediately evocative, preserving the richness of video while atomizing it to an excerpt as "glanceable" as a still image. This thesis proposes Glyph, a tool for creating expressive, seamlessly-looping GIFs from video. The tool integrates open-source software for video/image manipulation and loop detection into a simple, web-based authoring interface. Glyph allows a user to automatically detect perfect loops that occur in a video, create the appearance of seamless motion in a non-looping clip, still some regions of movement in a clip to highlight others, or imbue a still frame with subtle dynamism. The part-automated, part-manual editing tool thus allows users to quickly build non-literal excerpts from video that can immediately crystalize an affective quality or crucial moment, suspending and intensifying its semantic or emotional content through continuous, seamless looping. This thesis additionally explores applications for this class of image, called glyphs.

Thesis supervisor: Andrew B Lippman, Senior Research Scientist, Program in Media Arts and Sciences

Glyph: Lightweight and Evocative Looping Images in the News

by

Savannah Niles

Signature redacted

Thesis Advisor .....

Andrew B Lippman

Senior Research Scientist, Program in Media Arts and Sciences

Signature redacted

Thesis Reader .....

V. Michael Bove, Jr.

Principal Research Scientist, MIT Media Lab

Signature redacted

Thesis Reader ...

Ethan Zuckerman

Principal Research Scientist, MIT Media Lab

Acknowledgements

With immense gratitude, I thank my advisor, Andy Lippman, for the opportunity to join the Media Lab, and for the conversations, provocations, and support that he's given me in the last two years.

I also thank my thesis readers: Ethan Zuckerman, for encouraging artistic and expressive dimensions to this research, and Mike Bove, for ultimately sparking the ideas that became this thesis work and for ample encouragement through challenges.

I thank my teammates and friends in Viral Communications– Eric Dahlseng, Vivian Diep, Rob Hemsley, Amir Lazarovich, Travis Rich, Jeremy Rubin, Dan Sawada, Jonathan Speiser, Thariq Shihipar, and Tomer Weller. I especially thank Vivian, for her sweetness and support, and Travis Rich, for practical guidance, mentorship, and cheerleading.

I thank my best friends for sometimes knowingly, sometimes unknowingly keeping me going: Julie Legault, Alexis Hope, Deepak Jagdish, and the rest of Funfetti; my Chaseface bbs John, Joey, Max, Taylor, Chase, Lindsay, Jacob, Peggy, and Brad; and my best bb Katrina.

I thank my whole family for their love and support, especially my parents, Ileana and Chris, and my siblings, Liv, Ava, and Jack.

I thank Keira and Linda for all of their help in the last two years.

I thank 1369 Coffeehouse, where this entire thesis was written, for their generous hospitality.

And I thank Ryan McGinnis, for everylilthing.

Contents

1. Introduction

1.1. Motivation

The job of the news is both to inform and engage. The news informs by reporting the facts of a story or the multiple dimensions of an issue. But for an informed society to become invested in the outcome of events and feel empathy for those affected by them, the news must emotionally engage audiences with stories as they unfold. We don't want to be a part of a society that's blindly hysterical, nor do we want to live in a callously knowledgeable world. The news thus works along two axes, one intellectual and one affective, to communicate events across the world.

Figure 1-1: News websites and apps integrate many forms of media– including text, images, and video – to offer content of varying depth– from long articles to short headlines. From left to right: *The New York Times*, *Buzzfeed*, and Circa News[1,2].

News is accessed through wide-ranging platforms on the web, and the emotional impact of these broadcasts varies. Hard news is reported in terse and efficient language, and headlines and breaking stories are written to immediately inform. Twitter has become a legitimate and important platform for news broadcasting, but it atom-

---

1 Summers, Nick. "Circa Reworks News App Around Daily Brief Catchups." *The Next Web*, 24 Sep '14. http://thenextweb.com/apps/2014/09/24/circa-refocuses-bite-sized-news-app-around-daily-brief-catchups-wire-story-updates/.

2 Circa News. https://circanews.com/.

izes news to 140-character units. Headlines are high-velocity: "hard news" (breaking stories and immediately-reported events[3]) disseminates quickly through social networks. But in-depth and evocative analyses of the news still rely on longer-form content. Long-form journalism and feature stories have always had the word count to engage their audiences more deeply with their content, but most readers read only a fraction of an article[4], and scan the rest of the page for visual information[5]. So longer-form content has adapted to capture readers' fleeting attentions and minimize bounce rates, and, consequently, video journalism and photo-essays now proliferate in online news media. Images and moving images are thus crucial access-points for users to engage with news stories. Video is an especially rich medium to elicit emotional engagement, grounded as it is in real places and centering attention on faces, voices, and first-person experience. Video journalism can engage through expressive, narrative, and multi-sensory dimensions that headlines and hard news lack.

---

3 Reinemann, Carsten, James Stanyer, Sebastian Scherr, and Guido Legnante. "Hard and Soft News: A Review of Concepts, Operationalizations and Key Findings." Journalism, November 11, 2011, 1464884911427803. doi:10.1177/1464884911427803.

4 Manjoo, Farhad. "You Won't Finish This Article." Slate, June 6, 2013. http://www.slate.com/articles/technology/technology/2013/06/how_people_read_onli ne_why_you_won_t_finish_this_article.html.

5 Kenneth Holmqvist, Jana Holsanova. "Reading or Scanning? A Study of Newspaper and Net Paper Reading," in R. Radach, J. Hyona, and H. Deubel (Eds.), The mind's eye: Cognitive and applied aspects of eye movement research (pp. 657 - 670). Amsterdam: Elsevier.

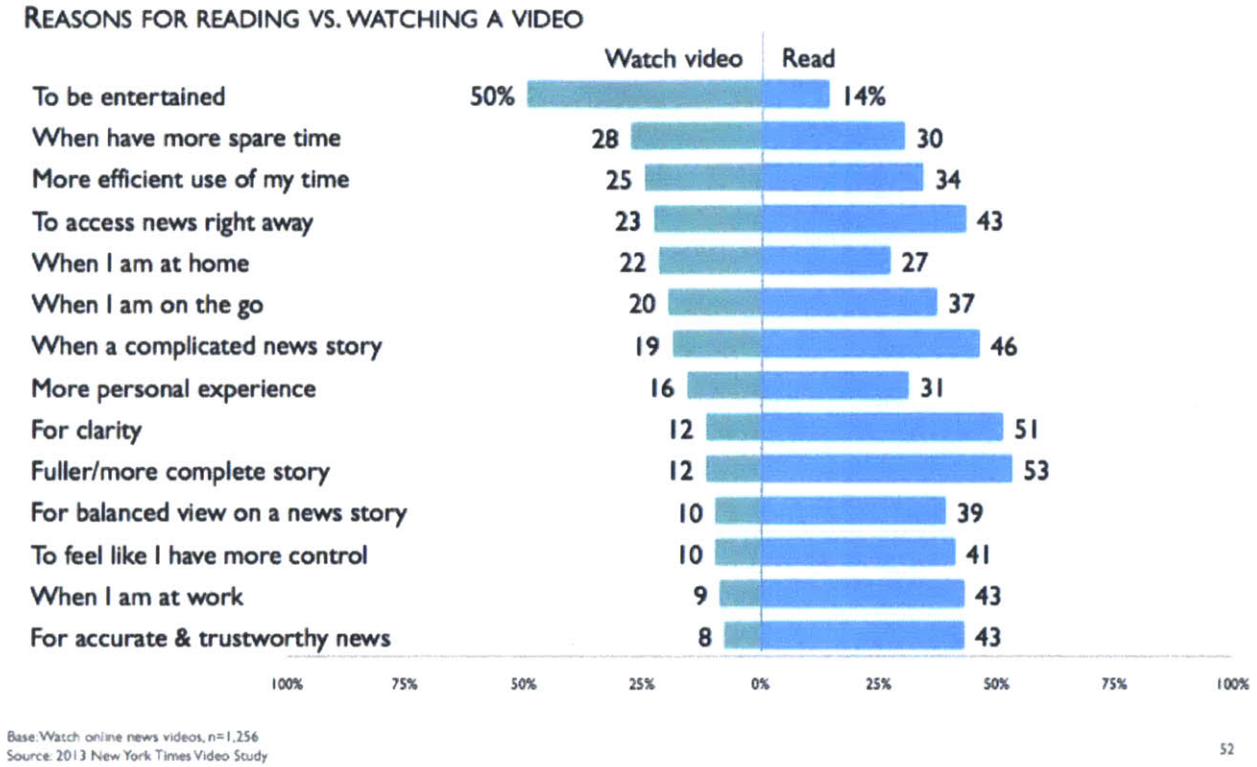


Figure 1-2: In a 2013 survey on preferred methods of information access, the New York Times found that in all cases except entertainment, people preferred to read an article over watching a video, and this preference carried over to news access[6].

But despite the increase in video accompanying online news, much of it goes unwatched (see figure 1-1). Video requires commitment in time, attention, and bandwidth. Mobile devices, social networks, and browser tabs allow us to interface with news intermittently and in the context of multitasking. In our environments, our engagement with news is also challenged by our limited attention bandwidths. Through-

6 The New York Times Customer Insight Group. "Results of The New York Times Customer Insight Group's 2013: The Year of Video Survey." *The New York Times Company Blog*, October 17, 2013. http://www.nytco.com/results-to-the-new-york-times-customer-insight-groups-2013-the-year-of-video-survey/.

out the day, rich visual material around events in the news propagates through the increasingly ubiquitous display surfaces in our lives and on our bodies. Cumulatively, among phones, wearables, and ambient screens, it is not unusual to find hundreds of square inches of pixels in our immediate environments. These devices are often designed to monitor or be responsive to developing headlines; we are surrounded by real-time news on high frame rate surfaces.

On our browsers, on our devices, and in our environments, news increasingly persists in our periphery as a second-order information stream. But the problem is that the moving images of the news are time-, attention-, and bandwidth-intensive, and unsuitable for mobile and peripheral access. Headlines are quickly informative, and news images are quickly engaging. But the richer emotional engagement of news video costs more in time and attention. How can we optimize the evocative effect of the moving images of news for interfaces accessed under conditions of divided attention? In these contexts, there is value in presenting the visual media of news somewhere between the high emotional impact, high cognitive load of video, and the "glanceability" of still images. A solution could be graphically manipulating news video into un-intrusive, visual encapsulations that function more like headlines: just as headlines are immediately informative, "video headlines" are immediately engaging.

One way to begin to develop a model for video headlines is to look to an existing, successful paradigm for representing high-bandwidth visual material– animated GIFs. GIFs are designed for immediacy: they're short, looped, silent, and utilize the power of suggestion. GIFs are usually created from a number of sequential video frames to create a looped excerpt from a video independent of the whole file. But GIFs don't always relate to their original source through start and end times. As GIFs become increasingly important as a viral form of media, creators derive GIFs from video frames with greater creative license. Authors may manipulate each frame of a video clip by hand in image-editing software—reducing or diminishing movement in some parts of each frame, and highlighting or altering movement in others—resulting in a moving image with no direct analogue in its source. Such manipulations are often used to produce seamless loops, cinematic compositions, and strategic reductions in movement. The resulting images are non-literal excerpts: their visual effects efficiently extract the semantics of a video, the possibilities of an image, or the affective qualities of visual material.

Figure 1-3: GIFs are incredibly versatile staples in news media, with widely varying impact on the stories they enhance. On the left, *Buzzfeed* uses GIFs from the *Jurassic Park* movie to orchestrate readers emotional reactions through an explanation of Egypt's 2011 revolution[7]. Center, a *Medium* story on an earthquake is illustrated with an animation of a shaking house designed for the feature[8]. Right, *the New York Times* homepage features a looping GIF of a wind turbine[9].

The expressive capacity and transmissibility of animated GIFs increasingly situates them in news stories and alongside headlines. Lightly dynamic images are staples in today's news diet, but these types of images are not generalizable. Creating GIFs, especially the kind of highly editorialized, looping GIFs that are integrated with increasing prominence in online feature stories, requires time-intensive, frame-by-frame image manipulation. If the tools to reduce video down to single, evocative moments were faster, easier, and more accessible, journalists, bloggers, and others actively engaging with visual material in the news could quickly extract from news video a looped moving image with the emotional cogency of a video headline.

---

7 Staff, Benny Johnson BuzzFeed. "The Story Of Egypt's Revolution In 'Jurassic Park' Gifs." *BuzzFeed*. Accessed May 10, 2015. http://www.buzzfeed.com/bennyjohnson/the-story-of-egypts-revolution-in-jurassic-park-gifs.

8 Mock, Rebecca. Animation for The Aftershock. "Matter", *Medium*, August 24, 2014. https://medium.com/matter/the-aftershocks-7966d0cdec66.

9 Barbaro, Michael. "With Farms Fading and Urban Might Rising, Power Shifts in Iowa." *The New York Times*, October 20, 2014. http://www.nytimes.com/2014/10/21/us/politics/iowa-senate-election.html.

1.2. Purpose

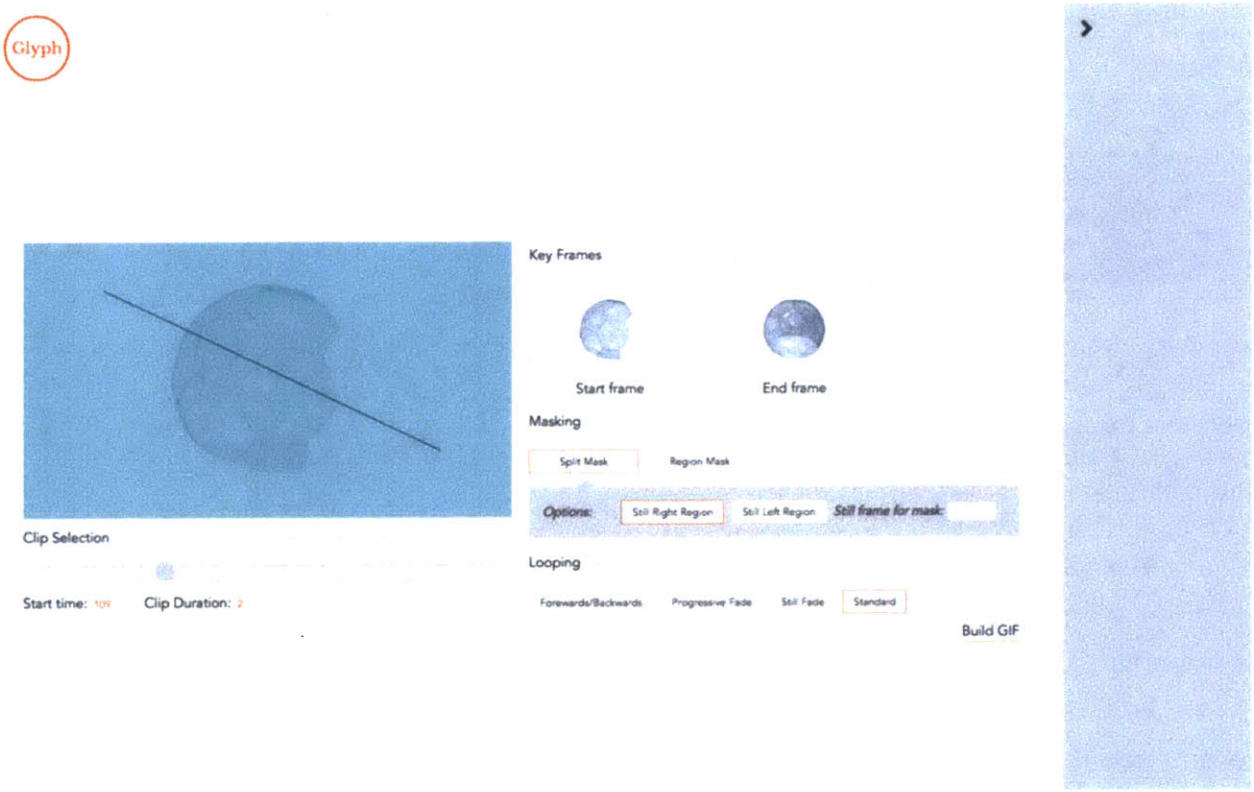


Figure 1-4: The Glyph authoring tool.

This thesis proposes Glyph, a tool for creating lightweight and evocative looping GIFs from video. The tool integrates open-source software for video/image manipulation and loop detection into a simple, web-based authoring interface. Glyph allows a user to automatically detect perfect loops that occur in a video, create the appearance of seamless motion in a non-looping clip, still some regions of movement in a clip to highlight others, or imbue a still frame with subtle dynamism. The part-automated, part-manual, cross-platform editing tool thus allows users to quickly build non-literal excerpts from video that can immediately crystalize an affective quality or crucial moment, suspending and intensifying its semantic or emotional content through continuous, seamless looping. The purpose of Glyph is to provide a solution for creating a generalized class of images with these qualities, referred to in this thesis as *glyphs*.

This thesis focuses on the creation of glyphs from news video. But a solution for creating evocative, looping excerpts from video can be applied to content beyond the

news. The value of Glyph as a tool scales to the inventiveness of its users and encourages self-expression with many forms of existing media.

1.3. Contribution

The first contribution of this thesis work will be integrating existing methods of video and image manipulation into a Python-based system and connected web interface to invent a tool for the creation of highly editorialized, seamlessly-looping GIFs. A second contribution will be investigating uses and applications for this class of moving images, particularly in the context of news media.

1.4. Overview of Thesis

The next two sections situate this work in relation to historical and cultural contexts and related research. Sections 4 through 6 outline the design and technical implementation of the Glyph application. Section 7 describes a series of explorations into applications for glyphs. Section 8 compares a viewer's attention and affective response to still images and non-looping GIFs to their reactions on those same axes to glyphs to evaluate the tool's output. Finally, section 9 concludes the thesis with a discussion of future work. Among the appendices included in this thesis is a "best practices" guide to using GIFs in visual journalism, co-written with Audrey Cerdan.

2. Background

This section aims to outline the cultural and historical context for seamlessly-looping images, with particular emphasis given to the use of looped images in news media.

2.1. Seamlessly-Looping Images

Figure 2-1: A zoetrope[1]

Many of the first moving images were short, looping animations. Early proto-cinema technologies of the late 19th century such as the phenakistoscope[2], zoetrope[3], and the praxinoscope[4] were devices that rotated a sequence of images through a

---

1 Wikipedia contributors, "Zoetrope," Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/w/index.php?title=Zoetrope&oldid=661705257 (accessed May 11, 2015).

2 Wikipedia contributors, "Phenakistoscope," Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/w/index.php?title=Phenakistoscope&oldid=653291460 (accessed April 26, 2015).

3 Wikipedia contributors, "Zoetrope," Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/w/index.php?title=Zoetrope&oldid=655715709 (accessed April 26, 2015).

4 Wikipedia contributors, "Praxinoscope," Wikipedia, The Free Encyclopedia, http://en.wikipedi-

viewer to create the illusion of motion. Over the next 150 years of development in the technologies of cinema, digital video, and high-bandwidth media, these short, looped images thrive in visual culture as animated GIFs.

The continuously-looping image sequences of GIFs are part of the visual vernacular of the internet. In recent years, the GIF has undergone a renaissance[5,6] of creative and technological development. Specifically relevant to the work in this thesis, new typologies of GIFs have emerged in recent years embracing the endless looping that distinguishes GIFs from other video compression schema. Communities on the web have formed around "perfect loops"[7], a kind of GIF that loops seamlessly, with a smooth, continuous animation throughout the sequence, including the transition from the last frame back to the first frame. An increasingly visible subset of seamlessly looping GIFs, called *cinemagraphs*[8,9] or *moment imagery*[10], emphasizes the GIF's place somewhere between an image and a video. These images juxtapose static and dynamic elements in a continuously-looping image, resembling a subtly-dynamic picture, like a digital lenticular postcard, more than a video clip.

---

a.org/w/index.php?title=Praxinoscope&oldid=594293785 (accessed April 26, 2015).

5 Friedman, Ann. "What Journalists Need to Know about Animated GIFs – Really," August 8, 2012. http://www.poynter.org/how-tos/writing/183802/what-journalists-need-to-know-about-animated-gifs-really/.

6 Weiner, Jonah. "Christina Hendricks on an Endless Loop." Slate, October 14, 2010. http://www.slate.com/articles/arts/culturebox/2010/10/christina_hendricks_on_an_endless_loop.html.

7 "r/PerfectLoops". Reddit. http://www.reddit.com/r/perfectloops/.

8 r/Cinemagraphs. Reddit. http://www.reddit.com/r/cinemagraphs/.

9 Flock, Elizabeth. "Cinemagraphs: What It Looks like When a Photo Moves." The Washington Post - Blogs, July 13, 2011. http://www.washingtonpost.com/blogs/blogpost/post/cinemagraphs-what-it-looks-like-when-a-photo-moves/2011/07/08/glQAONez3H_blog.html.

10 Tompkin, J., F. Pece, Kartic Subr, and J. Kautz. "Towards Moment Imagery: Automatic Cinemagraphs." In Visual Media Production (CVMP), 2011 Conference for, 87–93, 2011. doi:10.1109/CVMP.2011.16.

Figure 2-2: Examples of *cinemagraphs* by the fashion and lifestyle photographers Jamie Beck and Kevin Burg who coined the neologism and popularized the use of this class of images[11]. On the right, milk is poured endlessly in an otherwise still scene. On the left, the image of the woman is still except for a lock of hair blowing in the wind; images like this have led some to refer to cinemagraphs as "hair gifs"[12].

Looping GIFs and Cinemagraphs embody particular affective qualities of GIFs that have received more scholarly attention in recent years. The artist and critic Tom Moody famously argued that "Animated GIFs have evolved into a kind of ubiquitous 'mini-cinema'"[13]. The artist and critic Sally McCay echoes Moody by characterizing GIFs as images of self-enclosed affect, intensified by the intimate, tight viewing distance between eyes and the monitor, and argues that GIFs' mesmerizing loops sets them apart from our lived temporality[14].

2.2. Looping Images in Journalism

GIFs as a medium have outgrown the retro aesthetic that characterized them in the days when bandwidth was a serious obstacle to transmitting moving images on the in-

---

11 Burg, Kevin and Beck, Jamie. Cinemagraphs. http://cinemagraphs.com/.

12 Moody, Tom. "Hair GIFs and the Male Gaze at Tom Moody." Blog of New York Artist and Musician Tom Moody, March 17, 2012. https://www.tommoody.us/archives/2012/03/17/hair-gifs-and-the-male-gaze/.

13 Brian Massumi, "The Autonomy of Affect," in Parables for the Virtual: Movement, Affect, Sensation (Durham & London: Duke University Press, 2002) p. 29-30

14 McKay, Sally. "The Affect of Animated GIFs (Tom Moody, Petra Cortright, Lorna Mills)." Art & Education, 2009. http://www.artandeducation.net/paper/the-affect-of-animated-gifs-tom-moody-petra-cortright-lorna-mills/.

ternet: the increasingly broad and visually sophisticated use of GIFs in journalism effectuates this coming-of-age.

Pushing the medium beyond the recycling of the memes and reactions GIFs volleyed on image boards, publishing bodies born on the web like *Medium* and *Buzzfeed* have embraced a use of looping GIFs and cinemagraphs to evocatively illustrate footage of an ongoing event[15] or intensify the emotional power of a feature[16,17]. A story published in *Medium*'s "Matter" column on the Ebola outbreak in Liberia was organized around a series of seamlessly looping GIFs of individuals impacted by the outbreak. These evocative, arresting portraits offer more than a still image could through their mesmerizing dynamism, simulating a shared gaze between the subject and the viewer, or adding 3-dimensionality to the pictured environment. But these effects are achieved more immediately and compressively than they could be in an embedded video.

Figure 2-3: Left, water flows continuously from the faucet in this looped animation; right, as the camera continuously circles a child she meets our gaze

15 Phelps, Andrew. "The 2012 Summer Olympics Are Turning into a Giant Coming-out Party for the Animated GIF." NiemanLab, August 3, 2012. http://www.niemanlab.org/2012/08/the-2012-summer-olympics-are-turning-into-a-giant-coming-out-party-for-the-animated-gif/.

16 Z., Svetlana. "Sex Is Sex. But Money Is Money. – Matter." Medium. Accessed November 9, 2014. https://medium.com/matter/sex-is-sex-but-money-is-money- e7c10091713f.

17 Yan, Wundan and Francois Beaurain. "World War E: How Ebola Turned Liberia into a Zombie Movie – Matter." Medium. Accessed November 9, 2014. https://medium.com/matter/world-war-e-99dc164a257d.

Increasingly, more established publications like *The New York Times* and *The Boston Globe* experiment with this kind of lightly dynamic visual media. In 2012, *The New York Times* published the Pulitzer-prize winning feature "Snow Fall", which opens with a slowly looping image of windblown snow[18]. Critics praised *The Times'* innovative use of looping moving images as injecting "a subtle, atmospheric quality"[19] into the feature, achieving evocative and cinematic depth[20]. The same year, *The Times* did a series called Still Life[21], in which writers described their favorite summer settings around New York. A series of subtle, contemplative moving scenes accompanied and sentimentalized each story. In the last two years, *The Times* has continued to use gifs, and especially looping gifs, in feature stories and alongside headlines, and a growing number of online publications have followed.

Figure 2-4: Seamlessly-looping imagery increasingly appears in *The New York Times* alongside stories and features.

---

18 Branch, John. "Snow Fall: The Avalanche at Tunnel Creek." The New York Times, December 2012. http://www.nytimes.com/projects/2012/snow-fall/.

19 Duenes, Steve. Erin Kissane, Andrew Kueneman, Jacky Myint, Graham Roberts, Catherine Spangler. "How We Made SnowFall," Open Source, January 30, 2013. https://source.opennews.org/en-US/articles/how-we-made-snow-fall/.

20 Sonderman, Jeff. "How The New York Times' 'Snow Fall' Project Unifies Text, Multimedia," Poynter. December 20, 2012. http://www.poynter.org/mediawire/top- stories/198970/how-the-new-york-times-snow-fall-project-unifies-text-multimedia/.

21 Jennings, Dana. "Still Life: The Meadowlandscapes." The New York Times, July 10, 2012, sec. Arts / Art & Design. http://www.nytimes.com/2012/07/11/arts/design/still-life-gallery-of-the-meadowlands.html.

2.3. GIFs and Authorship

As derivations from video, GIFs often create issues in authorship and attribution.

Newsrooms with visual journalists on staff such as *The New York Times* and *Medium* have the option of creating GIFs and looping video from their own content, or even shooting content specifically for GIFs in the field, avoiding the risk of copyright violation or plagiarism. But citizen journalists, bloggers, and activists usually do not have access to original content, and instead excerpt and re-contextualize content from another source.

GIFs embody the culture of exchange, remixing, and collective authorship that characterizes visual media on the web. Conventional practice online treats embedded web video as fair game for creating and sharing GIFs. GIF creators have historically embraced the appropriation of other content, and remain free to do so, as GIFs are highly transmissible images that erase their sources and are difficult to take down (unlike a YouTube video, for example, which can be removed with a DMCA notice[22]). Because GIFs are often short excerpts or parodies of their source, GIFs often do not violate intellectual property law in many countries. And because their creators are hard to track and often not profiting from the images, few copyright violations involving GIFs have been brought to courts. This could change in the future, as GIFs become an increasingly powerful medium in the attention economy. In the 2014 FIFA World Cup, for example, a blogger creating GIFs of goals in near-realtime received a cease and desist[23].

---

22 Phelps, Andrew. "The 2012 Summer Olympics Are Turning into a Giant Coming-out Party for the Animated GIF." NiemanLab, August 3, 2012. http://www.niemanlab.org/2012/08/the-2012-summer-olympics-are-turning-into-a-giant-coming-out-party-for-the-animated-gif/.

23 Brouze, Emilie. "Copyright: Could You Be Attacked for an Animated GIF?" *Rue89*, April 2, 2015.

3. Related Work

The section covers related work in the research areas of automatic video summarization, authoring tools for making GIFs from video, and the automated creation of looping GIFs and cinemagraphs.

3.1. Automatic Video Summarization

Previous work in video summarization has explored methods for taking a longer video and generating shorter clips or key frames to summarize its content[1]. Many video summarization algorithms extract representative footage by detecting motion and scene changes across the video[2,3]. Recent work expands the field by developing models of user attention to extract the essential clips from a video [4]. These excerpts are literal; like a direct quotation, they correspond to a time interval in their source content. Research has also been done in summarizing video through non-literal excerpts, with no direct analogue in their source. Video synopsis algorithms segment video into objects and activities, then transpose moving objects into a temporally condensed scene[5]. Past work at the Media Lab on "Salient Stills" looked at representing the temporal progression of video within a single-frame[6].

---

1 Truong, Ba Tu, and Svetha Venkatesh. "Video Abstraction: A Systematic Review and Classification." *ACM Transactions on Multimedia Computing, Communications, and Applications*. 3, no. 1 (February 2007). doi:10.1145/1198302.1198305.

2 Gianluigi, Ciocca, and Schettini Raimondo. "An Innovative Algorithm for Key Frame Extraction in Video Summarization." Journal of Real-Time Image Processing 1, no. 1 (March 1, 2006): 69–88. doi:10.1007/s11554-006-0001-1.

3 Liu, Tianming, Hong-Jiang Zhang, and Feihu Qi. "A Novel Video Key-Frame-Extraction Algorithm Based on Perceived Motion Energy Model." IEEE Transactions on Circuits and Systems for Video Technology 13, no. 10 (October 2003): 1006–13. doi:10.1109/TCSVT.2003.816521.

4 Ma, Yu-Fei, Lie Lu, Hong-Jiang Zhang, and Mingjing Li. "A User Attention Model for Video Summarization." In Proceedings of the Tenth ACM International Conference on Multimedia, 533–42. MULTIMEDIA '02. New York, NY, USA: ACM, 2002. doi:10.1145/641007.641116.

5 Wikipedia contributors, "Video synopsis," *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/w/index.php?title=Video_synopsis&oldid=660532840(accessed May 4, 2015).

6 Teodosio, Laura, and Walter Bender. "Salient Video Stills: Content and Context Preserved." In Proceedings of the First ACM International Conference on Multimedia, 39–46. MULTIMEDIA '93. New York, NY, USA: ACM, 1993. doi:10.1145/166266.166270.

3.3. GIF Authoring Tools

Creating video excerpts with GIFs is an increasingly common practice for manually summarizing videos on the web. Many online tools allow a user to enter a YouTube URL, choose a start and an end time, and export an animated GIF[7]. Some of these tools allow users to add text captions to the GIF or adjust its frame rate. Some of these tools add a watermark to the final image they export. Recently, image and video hosting sites Imgur[8], Giphy[9], and YouTube[10] have rolled out their own online GIF making tools (see figure 3-1).

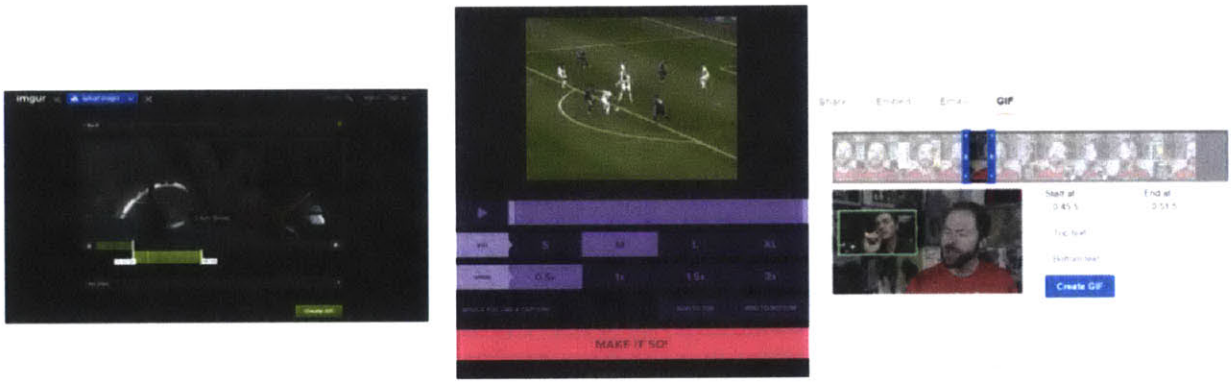


Figure 3-1: From left to right, online GIF-making interfaces by Imgur, Giphy, and YouTube.

Desktop software exists for creating GIFs with more advanced manipulations. Giffing[11] is a tool for Windows operating systems for creating GIFs with advanced effects, like cropping, compositing clip sequences, and reversing playback. Adobe Photoshop, Premier, and AfterEffects also support frame-by-frame manipulations to GIFs

---

7 Moynihan, Tim. "How Imgur's New GIF-Maker Stacks Up Against Other Tools Out There." *Wired*, January 30, 2015. http://www.wired.com/2015/01/gif-tools/.

8 Schaaf, Sarah. "Introducing Video to GIF." *Imgur Blog*, January 29, 2015. http://imgur.com/blog/2015/01/29/introducing-video-to-gif/.

9 Giffffr by Giphy. http://blog.giffff.fr/.

10 Kumparak, Greg. "YouTube Gets A Built-In GIF Creator." TechCrunch. Accessed May 4, 2015. http://social.techcrunch.com/2014/12/11/youtube-gif-maker/.

11 Giffing. http://www.giffingtool.com/.

and scripted editing of image sequences[12] (see figure 3-2).

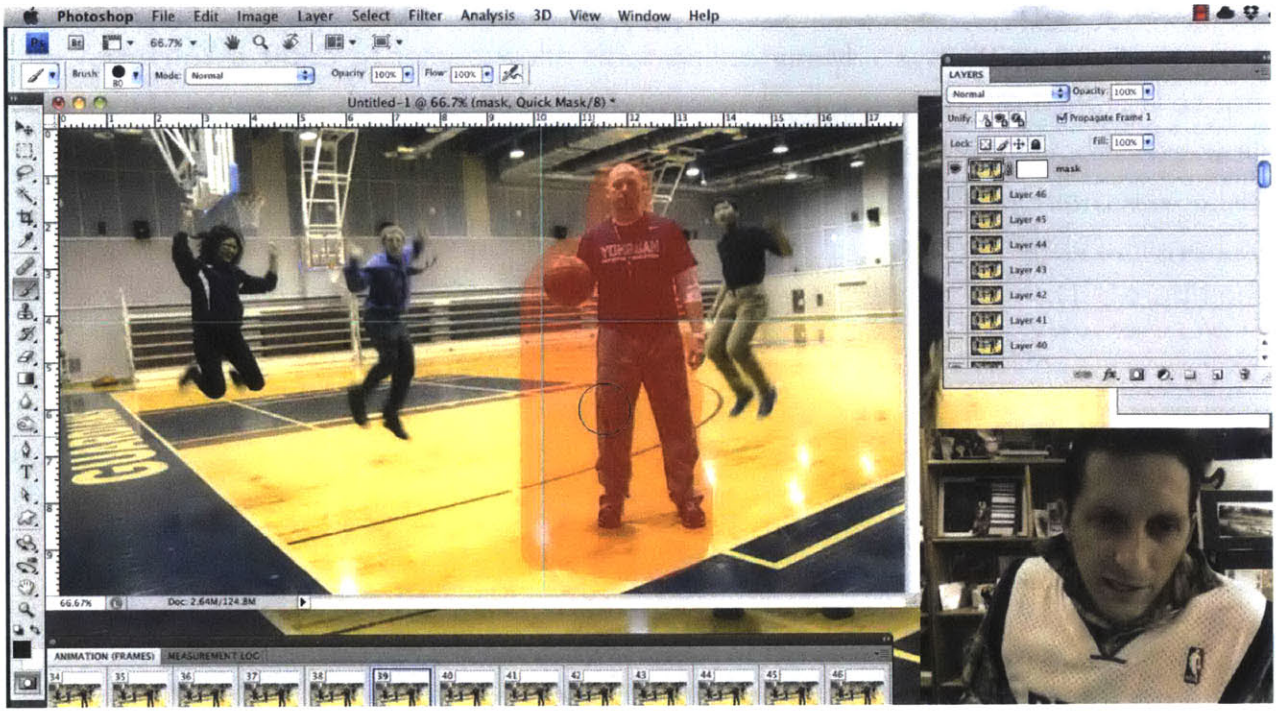


Figure 3-2: Many tutorials online demonstrate the complicated process of creating looping GIFs and cinemagraphs in desktop image-manipulation software like Photoshop[13].

12 Adobe. http://www.adobe.com/.

13 Uploaded by mouseflip. "Simple GIF Cinemagraph Video Tutorial in Adobe Photoshop". YouTube, Jan 24 2013. https://www.youtube.com/watch?v=UKY3scPIMd8.

3.3. Making Looping GIFs and Cinemagraphs Computationally

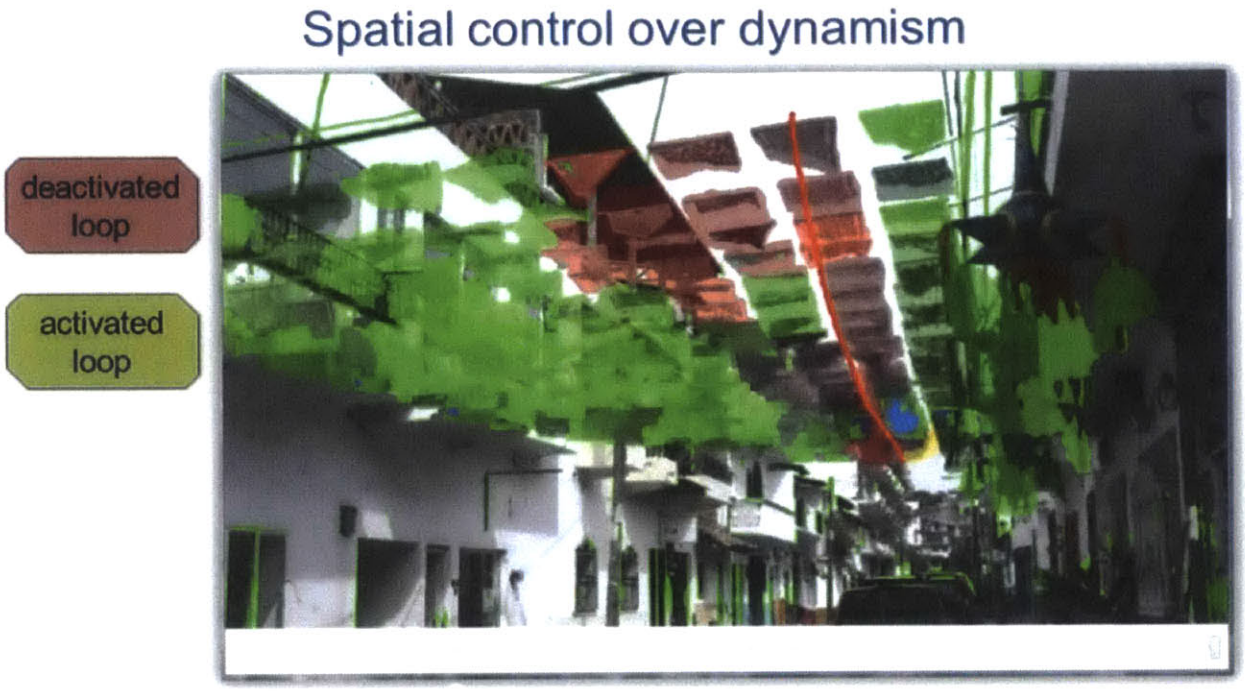


Figure 3-3: Interface developed at Microsoft Research to composite static and dynamic video segments into a looping scene.

Developing methods for automatically extracting looped clips from video is an increasingly active research area. Open source software, including MoviePy[14] and LoopFindr[15], implement basic loop detection algorithms. More advanced research at Microsoft Research involves creating perfectly looping clips by identifying the independent looping periods of dynamic segments in the clip, and compositing these asynchronous segments at user-specified levels of dynamism[16] (see figure 3-3). Re-

14 Zulko. "An Algorithm to Extract Looping GIFs From Videos". __del__(self) / Eaten by Python. 2015 Feb 1. http://zulko.github.io/blog/2015/02/01/extracting-perfectly-looping-gifs-from-videos-with-python-and-moviepy/.

15 Burger, Colin. LoopFindr. http://loopfindr.tumblr.com/.

16 iao, Zicheng, Neel Joshi, and Hugues Hoppe. "Automated Video Looping with Progressive Dynamism." ACM Trans. Graph. 32, no. 4 (July 2013): 77:1–77:10. doi:10.1145/2461912.2461950.

search has also been done in compositing static and and dynamic imagery to create cinemagraph-like "moment imagery"[17,18].

[17] Tompkin, J., F. Pece, Kartic Subr, and J. Kautz. "Towards Moment Imagery: Automatic Cinemagraphs." In Visual Media Production (CVMP), 2011 Conference for, 87–93, 2011. doi:10.1109/CVMP.2011.16.

[18] Joshi, Neel, Sisil Mehta, Steven Drucker, Eric Stollnitz, Hugues Hoppe, Matt Uyttendaele, and Michael Cohen. "Cliplets: Juxtaposing Still and Dynamic Imagery." In Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology, 251–60. UIST '12. New York, NY, USA: ACM, 2012. doi:10.1145/2380116.2380149.

4. Approach

This section describes the approach toward building Glyph, a web-based tool for creating expressive, looping GIFs from videos on the web. The tool's development divides into two phases: 1. the design of its web-based user interface, and 2. a set of video-processing functions with API endpoints.

4.1. From GIFs to Glyphs

Glyph is designed to create a specific type of moving image, referred to here as a glyph. Like regular animated GIFs, a glyph is a looping clip derived from a video. But glyphs are designed to loop seamlessly with continuous dynamism throughout the image sequence. Many of the transformations Glyph enables a user to make on a video clip are designed to create this continuous transition from the last frame to the first frame. These transformations can result in a clip made from a sequence of frames that never appear in their source video. In this case, glyphs can relate to their source video as *non-literal* or *paraphrased* video excerpts.

Figure 4-1: Glyphs loop seamlessly, which means there is a seamless transition between their first and last frames.

Figure 4-2: Examples of glyphs made by applying various transformations to non-looping video excerpts.

4.2. User Interface

Glyph provides a visual interface for creating glyphs. While section 5 details the implementation of this interface, this section provides an overview of the UI and describes the interaction concepts behind the system's design.

4.2.1. Overview of UI

Figure 4-3 provides an overview of Glyph's authoring interface[1]. The blue labels superimposed on the figure describe the four basic steps for creating a glyph in the interface:

1. *Select a clip:* A user selects a clip from the video to work with. This can be done manually or semi-automatically using Glyph's automatic loop detection.
2. *Preview the selected clip:* The embedded video player loops the selected clip, and the first and last frames of the clip are previewed as thumbnails.

---

[1] The video used in these figure is of the Silk Pavilion by the Mediated Matter group at the MIT Media Lab. It's citation is included in the listing of video samples used in evaluations in Appendix B.

3. *Add additional effects:* A user can optionally add looping effects and masks to the image.
4. *Build the image and iterate:* A user builds the image and views it in the output area.

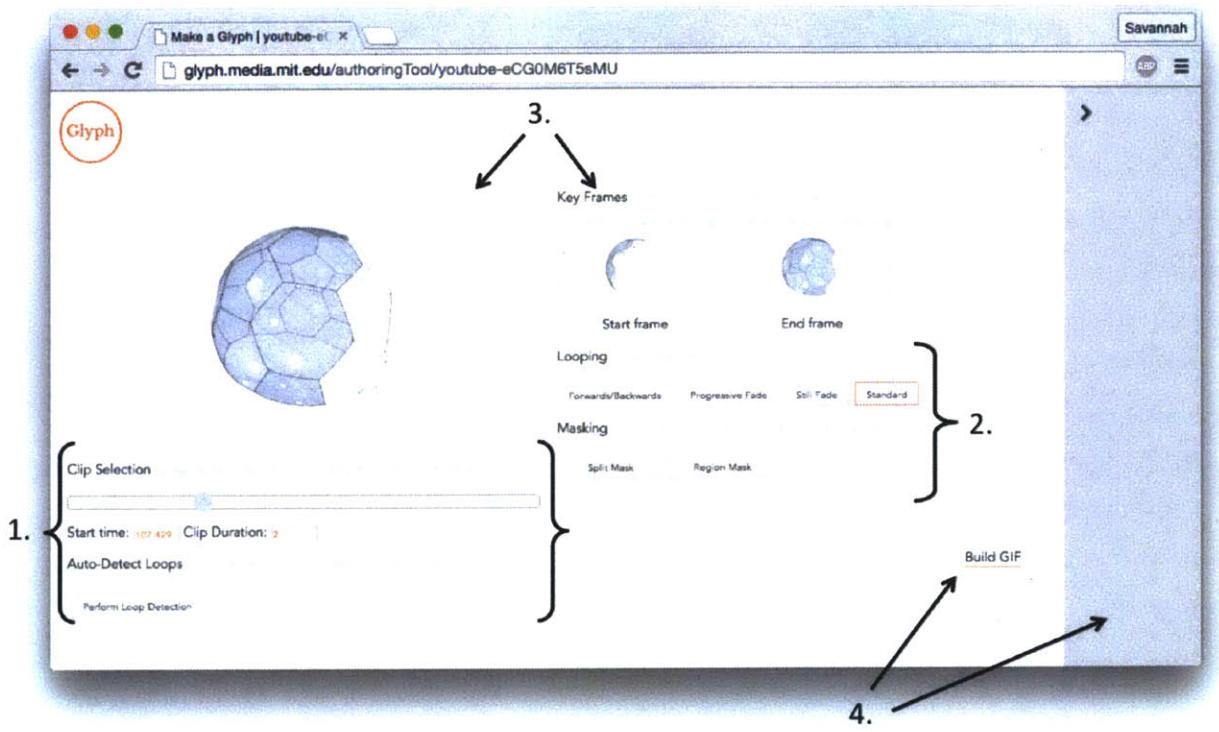


Figure 4-3: Overview of the Glyph UI

4.2.2. Concept of Interaction

Glyph is designed to afford *simple*, *expressive*, and *iterative* interactions.

Like the image and video editing software suites described in Section 3.2, such as Photoshop and AfterEffects, Glyph affords a similarly robust set of possible transformations on video clips. But Glyph is designed to be a simpler alternative, enabling a user to perform these transformations in just a few clips through a cross-platform web tool. Rather than requiring users to create or download content, and then upload it into the system, Glyph allows users to bring editable video content into the system

simply through a video URL. Though Glyph could theoretically support many sources of web video, Glyph currently only supports video URLs from YouTube. This limitation enhances the simplicity of the system: the playback of an embedded YouTube player can be controlled on the client with JavaScript, allowing users to scrub through a video for clips and preview looping excerpts (for further discussion, see sections 5.2 and 5.3).

Glyph is designed to enable a user to expressively excerpt videos. Glyph simplifies the process for creating seamlessly-looping or non-literal excerpts from video, but it does not automate the selection of the clip or the effects applied to it. The interface defers to the editorial intent of the user by allowing a user to manually adjust, select and combine clips and their transformations. Glyph thus gives users latitude to create many kinds of output, including images that are not seamlessly-looping or "glyph-like". To address this, the visual design of the tool was developed to encourage and inspire users to create beautiful, thoughtful, and evocative glyphs. An earlier visual design for Glyph evoked the playful design of many existing GIF creation and discovery websites, such as GIFFFR[2] and Giphy[3] (see figure 4-4). But user feedback on this earlier design revealed that the playfulness of the interface felt dissonant with the serious artistic and journalistic intentions that users were bringing to the tool. The interface was redesigned visually to reflect and encourage the a use of the tool for serious work, rather than memes.

---

2 GIFFFFR. http://giffffr.giphy.com/.

3 Giphy. https://giphy.com/.

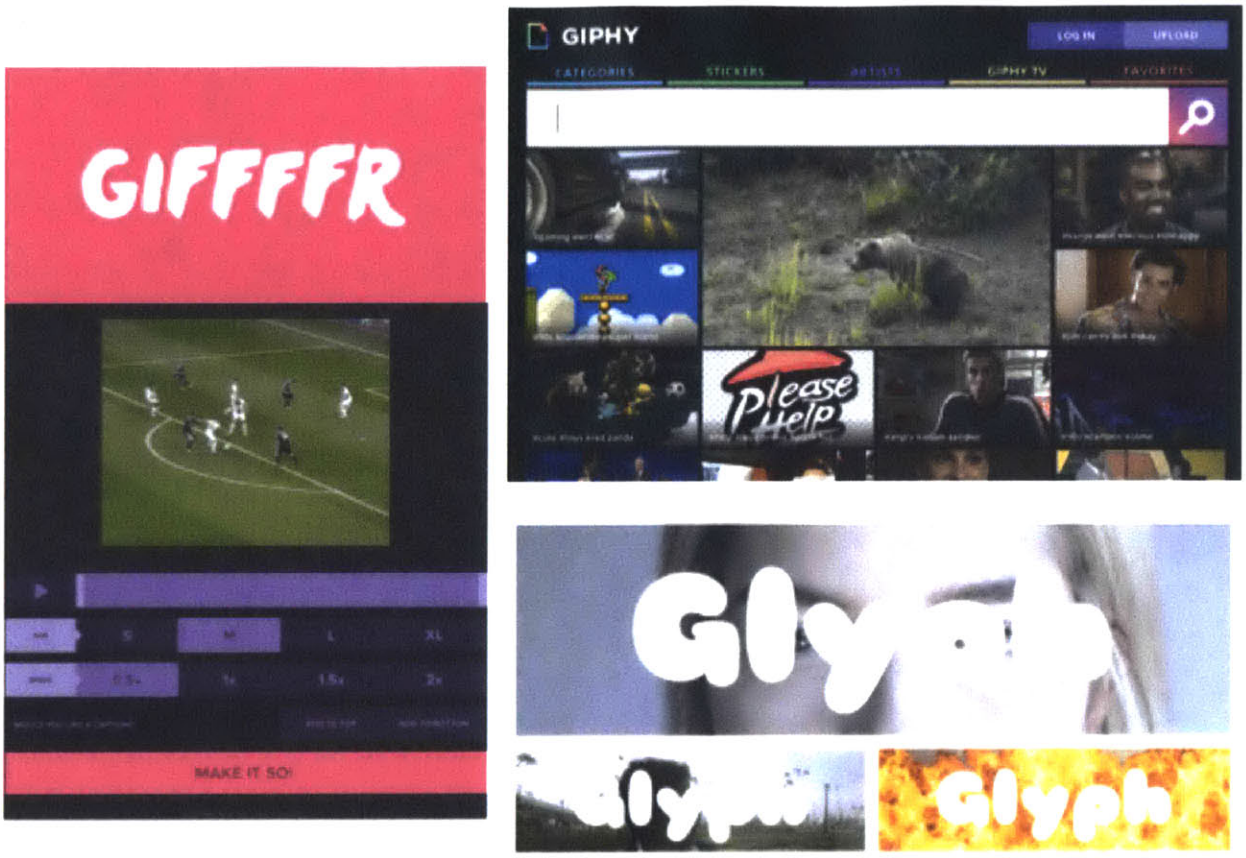


Figure 4-4: The original Glyph design (bottom right) was meant to evoke the playful designs of existing GIF hosting and creation sites, such as Giphy (top right) and GIFFFR (left).

Glyph is thus designed to support an iterative workflow. Applying frame-by-frame transformations to a video clip and processing it into a GIF can take a few seconds. Therefore, Glyph is not a WYSIWYG[4] interface: users must output the GIF to see masks and looping effects applied. But building highly designed and highly evocative images in Glyph may require several rounds of micro-adjustments to the clip's start time or duration, or to the dimensions of a mask. A user might also apply several types or combinations of transformations to a clip before choosing the best image. Because all of these adjustments would require a user to output a new glyph, the output request is handled asynchronously and on the same page as the authoring tool, so that a user

4 What You See Is What You Get. http://en.wikipedia.org/wiki/WYSIWYG.

can begin making new adjustments before the glyph is ready. A user can slide back and forth between the authoring tool and the output area during workflow. When the *Build GIF* button is pressed, all settings in the authoring tool retain their state and are not reset, to support further adjustments to the image in a subsequent iteration. The most time-intensive processes in Glyph– downloading a video input by a user and performing loop detection on the video– are cached in Glyph's data store, so if the user navigates away from the interface and then returns, she can immediately begin working.

4.3. Video Processing System and API

Glyph's video processing functions and API power its visual interface. Section 6 details the implementation of this process engine. This section provides an overview of the system's architecture and discusses its dependencies.

4.3.1. System overview

Glyph performs all of its video and image processing server-side in response to asynchronous requests from the client application. Figure 4-5 provides an overview of these functionalities and how they relate to user interaction with the tool.

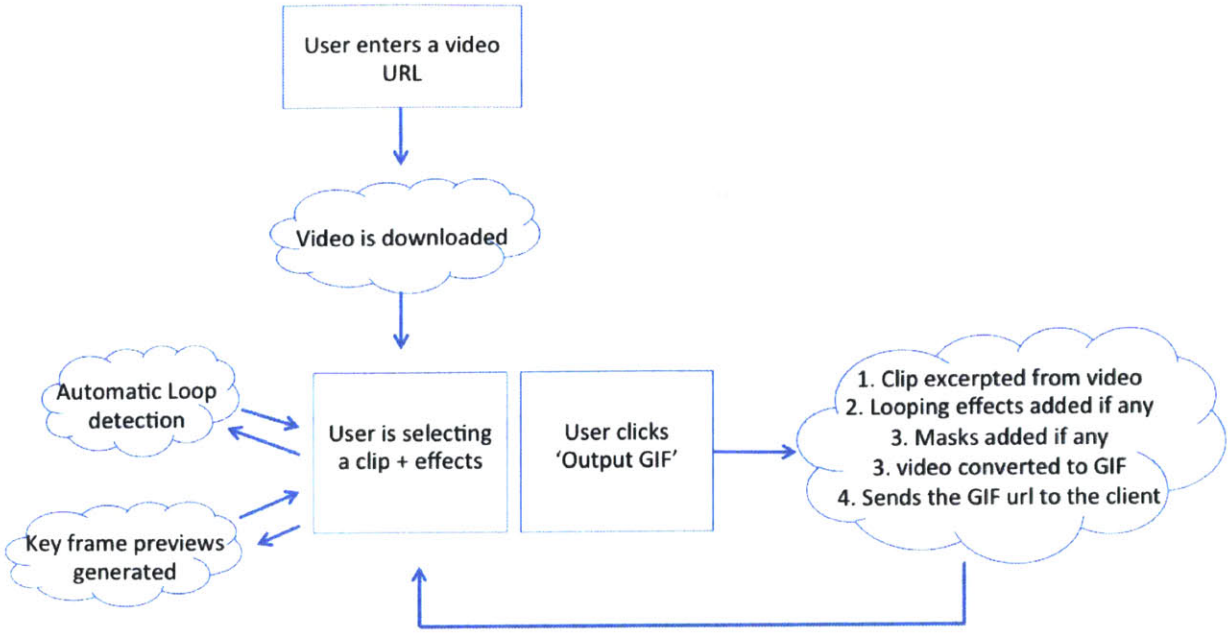


Figure 4-5: A flow-chart of server-side video processing functions available in Glyph

4.3.2. Dependencies

Glyph is written in Python[5]. Its client application and API are served through a Flask[6] application. Glyph's video processing functions make use of the video and image processing capabilities of three open-source frameworks: FFmpeg, ImageMagick, and MoviePy.

FFmpeg[7] is a cross-platform framework for converting and manipulating video and audio. FFmpeg can quickly excerpt a clip or extract image frames from video.

---

5 Python. https://www.python.org/.
6 Flask. http://flask.pocoo.org/.
7 FFmpeg. https://www.ffmpeg.org/.

ImageMagick[8] is a software suite for image manipulation. ImageMagick enables cropping and compositing images in a variety of formats, and performing color adjustments.

MoviePy[9] is a Python framework for video editing and compositing. It provides a Python wrapper to a suite of video and image manipulations enabled by FFmpeg and ImageMagick, simplifying complicated image and video transformations, such as masks, crossfades, and the composition of video clips and still images. An active community supports the MoviePy project. A subreddit[10] dedicated to the library pointed to many technical references used in the development of Glyph. In particular, a blog post by the developer of MoviePy describing a number of procedures for manipulating frames in GIFs with the MoviePy library served as a key starting point for the development of Glyph's suite of looping and masking effects[11].

4.4. Assumptions and Limitations

Glyph demonstrates what a highly-accessible tool for quickly creating evocative, seamlessly-looped moving images might look like. And as discussed in section 7, Glyph also supports explorations in deploying seamlessly-looping video excerpts to a variety of visual interfaces by enabling the rapid creation of these images. But Glyph necessarily accepts several assumptions and limitations related to the technical and legal constraints of video processing on the web.

4.4.1. Video Stabilization

With occasional exceptions, Glyph's capabilities for auto-detecting and creating seamlessly-looping clips from videos depend on stabilized footage. Glyph assumes that videos are stabilized, and does not offer automatic stabilization. The unavailability of a cross-platform solution for stabilizing video fast enough to support a simple and iterative web-based authoring tool made implementing automatic stabilization unfeasible

---

8 ImageMagick. http://www.imagemagick.org/.

9 MoviePy. http://zulko.github.io/moviepy/.

10 r/MoviePy: Pythonic Video Editing. http://www.reddit.com/r/moviepy/.

11 Zulko. "Making GIFs From Video Files With Python". __del__(self) / Eaten by Python. 2014 Jan 23. http://zulko.github.io/blog/2014/01/23/making-animated-gifs-from-video-files-with-python/.

and beyond the scope of this thesis. Furthermore, this thesis emphasizes the creation of seamlessly-looping video excerpts from news video, which is often shot with tripods or stabilized post-production.

4.4.2. Image Quality

Glyph assumes that a user always wants to create a looping image in GIF format. Because GIFs can be embedded as images in browsers, they are highly transmissible. But as a video compression schema, the GIF format entails a loss of image quality, since each frame of the animation references a limited 256-color table. If Glyph created output in .mp4 format, it could create higher resolution, more colorful looping sequences. But these clips would have to be embedded onto webpages as looping, autoplaying HTML5 video elements, and they could not as easily be displayed on social feeds and image boards. This iteration of Glyph assumes that users favor transmissibility of the output over its quality, and therefore converts all clips to GIFs. However, all of Glyph's functions are performed on .mp4 files, and the edited clip is only converted to a GIF as a final step. Glyph could easily support both output formats in a future iteration.

4.4.3. Use of YouTube Videos

Finally, Glyph assumes that addressing the copyright issues around video remixing on the web is beyond the scope of this thesis. Glyph makes possible– and, by design, encourages– the creation of images that fall within fair use parameters: Glyph limits video excerpts to short clips, and the visual effects it facilitates have transformative effects on the created images. The tool could easily preserve the provenance of the content by adding a small watermark noting the source video's YouTube ID to the GIFs it created.

However, Glyph may still risk copyright violation. Glyph downloads the YouTube videos supplied by users to cut excerpts and manipulate the frames of those excerpts in generating looping GIFs, in spite of the fact that downloading YouTube videos violates YouTube's Terms of Service[12]. This violation may or may not pose a copyright vio-

[12] YouTube Terms of Service. https://www.youtube.com/t/terms.

lation: Glyph is one of many emerging technologies operating within the gray legal environment of accessing and remixing media. Practices as mundane as searching for and downloading images from Google Image Search could technically pose a copyright violation[13], and the legality of many popular websites and services online is still to-be-determined. In 2012, YouTube sued YouTube-MP3[14], a site that allowed users to save audio from YouTube videos as MP3s, which operated by downloading YouTube videos to convert their audio[15]. YouTube-MP3 was found in violation of copyright law not because the site downloaded YouTube videos, but because it cached the video files to increase the speed of the service for many users making similar requests. The site operates today, and it continues to download YouTube videos in the conversion process, but these videos are no longer cached. The case suggests that if Glyph deleted all video content after processing, it could likewise be exempt from copyright violation.

Sites like YouTube-MP3 and online GIF-making tools like Glyph are necessary challenges to a dusty legal infrastructure that lacks robust enforcement mechanisms and fails to respond to emerging practices and technologies.

---

13 LuAug, Denise. "14 Illegal Things You're Doing on the Internet." Mashable, August 22, 2013. http://mashable.com/2013/08/22/illegal-internet-activity/.

14 YouTube-MP3. http://www.youtube-mp3.org/.

15 "Andy." "YouTube MP3 Converter Loses Court Battle But The Music Plays On." TorrentFreak, October 23, 2013. https://torrentfreak.com/youtube-mp3-converter-loses-court-battle-but-the-music-plays-on-131023/.

5. User Interface

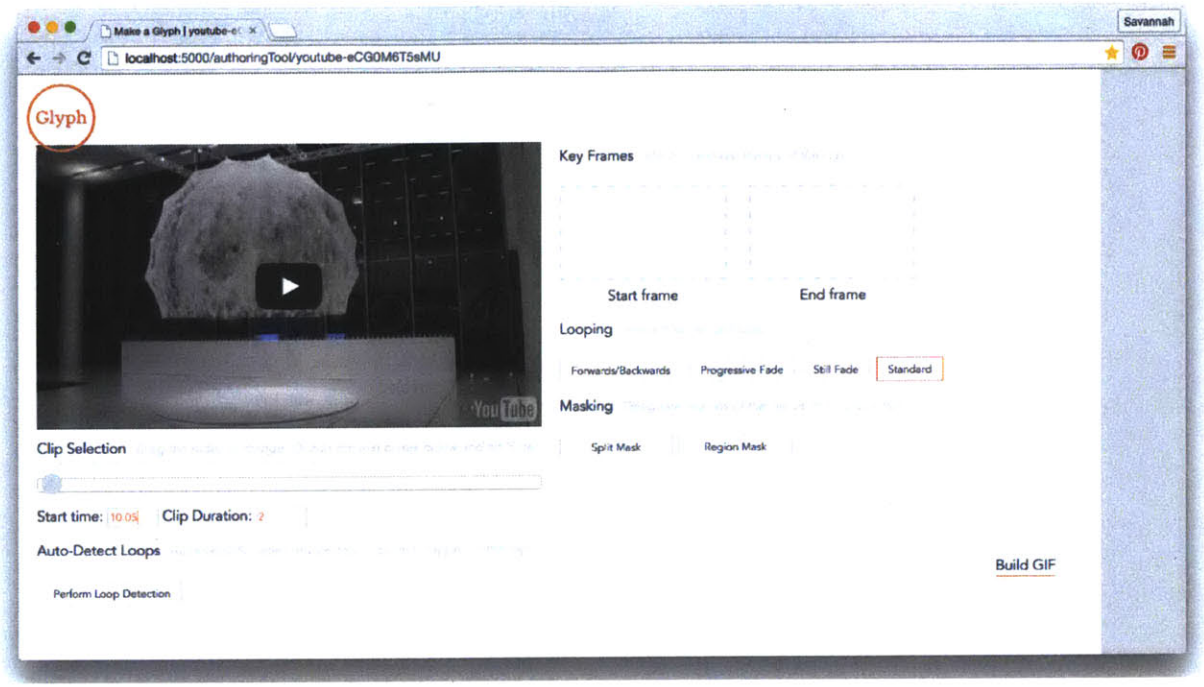


Figure 5-1: The Glyph User Interface.

Glyph provides a simple web interface for processing YouTube videos into looping GIFs. The interface allows a user to input a YouTube URL, excerpt a clip, and then optionally add a number of stylistic effects to control the looping and level of dynamism in the final image. This section discusses the design and technical implementation of this editing interface.

5.1. Overview of Implementation

Glyph is implemented as a web application for JavaScript-enabled browsers. Glyph makes use of the YouTube IFrame Player API[1] for embedding a YouTube player in the authoring interface and controlling its playback with JavaScript (further detail in sec-

1 YouTube IFrame Player API. https://developers.google.com/youtube/iframe_api_reference.

tion 5.3.1). Glyph also uses JQuery[2], a JavaScript library to simplify client-side scripting, and JQuery UI[3] for building UI elements in the interface, including the clip selection slider and radio buttons. Glyph uses the CSS toolkit Font Awesome[4] for icons.

5.2. Starting with a YouTube URL

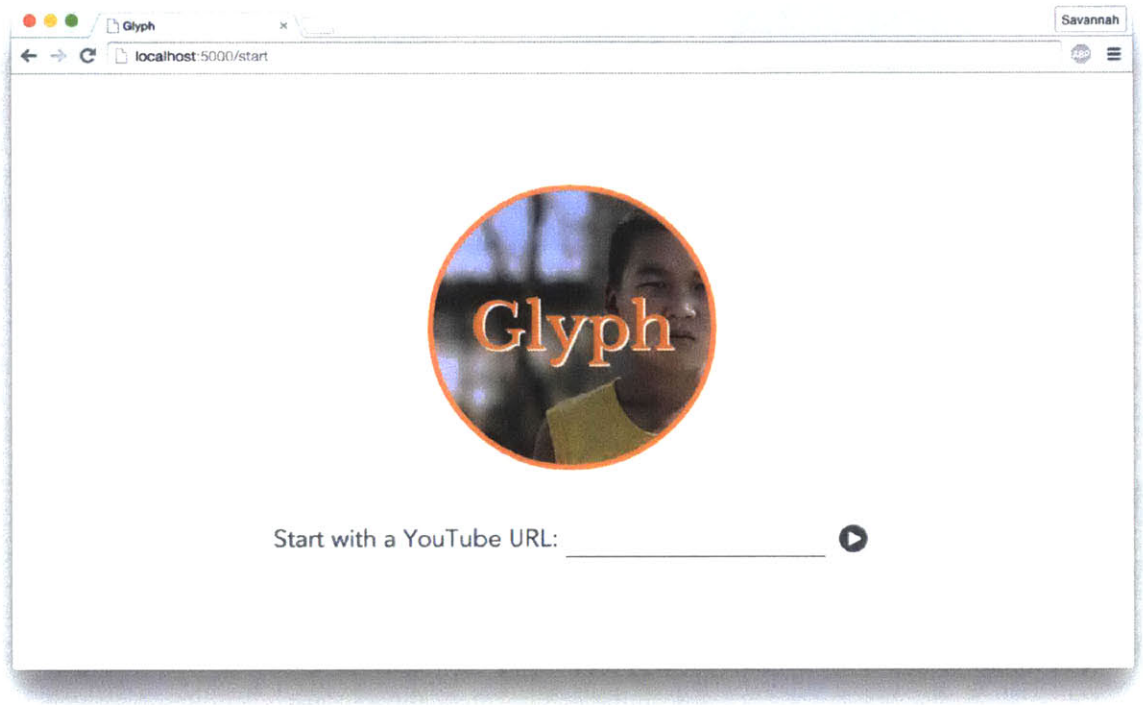


Figure 5-2: Glyph start page

Users begin creating glyphs by entering the URL of a web video. As discussed in section 4.2.1, the Glyph user interface is optimized for YouTube video URLs, in order to take advantage of YouTube's embedded player and IFrame player API.

The background image of the Glyph header on the start page is randomized to show a glyph from a curated collection of output created with the tool. This is a cue to

2 JQuery. https://jquery.com/.
3 JQuery UI. https://jqueryui.com/.
4 Font Awesome. http://fortawesome.github.io/Font-Awesome/.

the user that demonstrates the kind of looping image that the Glyph tool is designed to create.

Figure 5-3: Glyph header

Entering a YouTube URL redirects the user to the main authoring interface. Here, a user can make a number of editorial decisions in the process of building a glyph.

5.3. Selecting a Clip

To start, a user selects a clip from their selected video. A clip is an excerpt from the video, and is defined by a start time and duration in seconds. A user can use a slider element in the *Clip Selection* section of the interface to set the clip start time by scrubbing through the video. A user can also set the start time by manually entering a value in seconds in the accompanying text box. In the same way, the user can set the clip duration to any value between 0 and 15 seconds. In order to support fast processing times, Glyph does not allow processing clips longer than 15 seconds, and will display an error message on screen if a user exceeds this duration.

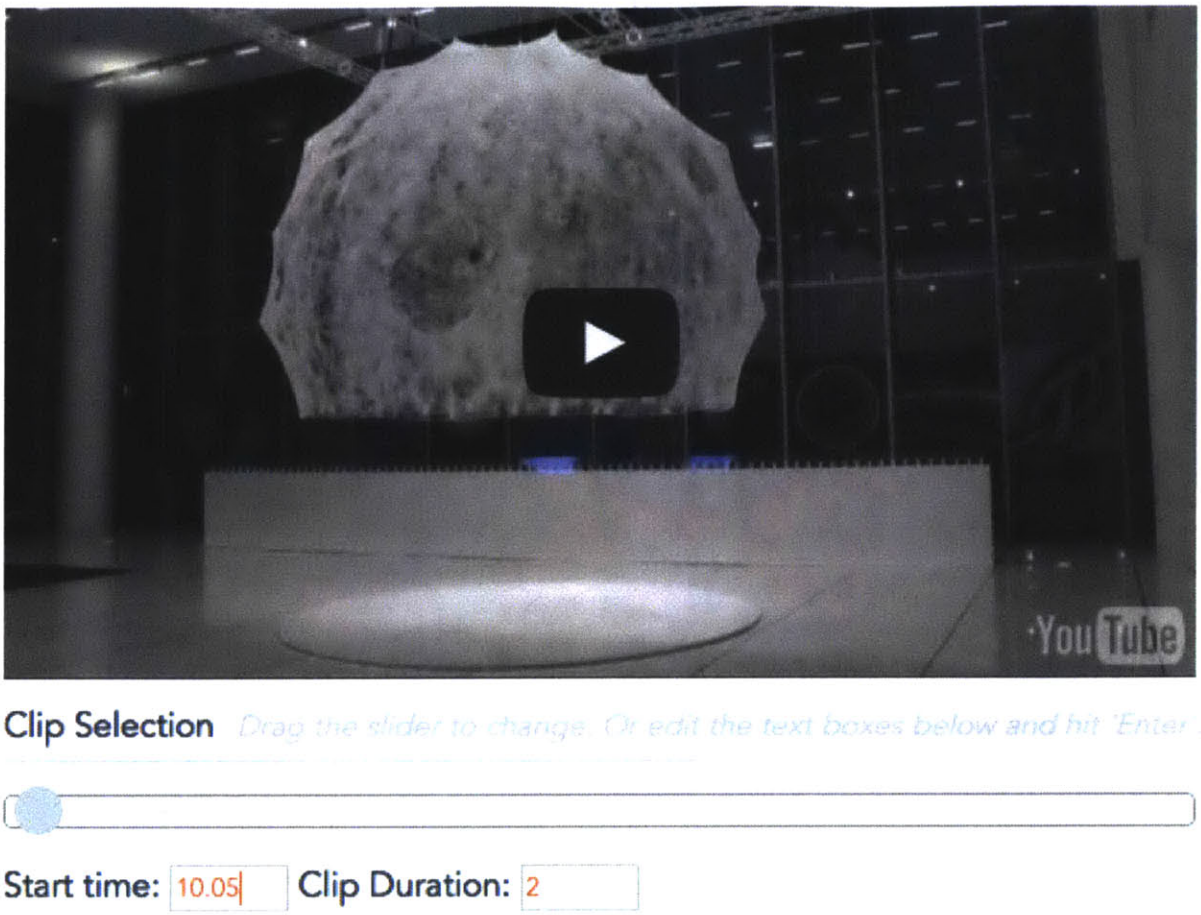


Figure 5-4: Clip Selection

Two features in the Glyph interface provide responsive feedback as the user is selecting a clip: the embedded video looping preview and the key frame preview.

5.3.1. Looping Preview

As a user is defining a clip by adjusting start time and duration, the embedded YouTube player plays the updated clip on loop, thus showing a preview of what the clip will look like as a looping GIF. The YouTube IFrame Player API supports this level of dynamic control over the video's playback.

5.3.2. Key Frame Preview

Image previews of the first and last frame of the currently selected clip give the user an additional level of feedback whiles adjusting the clip's start time and duration. These thumbnails are updated automatically when the start time or duration is updated. In an earlier iteration of the user interface that did not have key frame previews available, user testers had difficulty making fine-tune adjustments quickly to the clip time, such as beginning or ending a clip right on a scene change, or choosing the best pair of start and end frames to create the appearance of seamless looping. The key frame preview is designed to support frame-level clip time adjustments that are difficult to see at the video's normal frame rate.

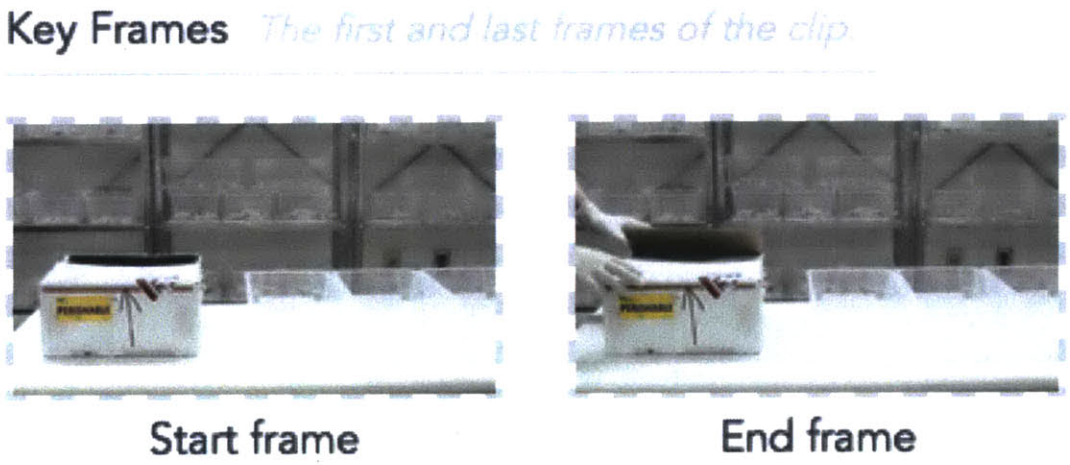


Figure 5-5: Key frame preview

5.4. Auto-detecting Loops

In addition to manually selecting a clip from the video, users can use Glyph to automatically detect seamless or near-seamless looping segments that occur in the video. To do so, users can click the *Perform Loop Detection* button in the *Auto-Detect Loops* section of the interface. This initializes a server-side call to perform loop detection on the video (see Section 6.4). Because this analysis can take a few minutes, this call is asynchronous, so the user can continue editing clips from the video while waiting for results to populate the section. When loop detection is complete, pairs of start times

and durations that correspond to loops in the video populate the section in a horizontal list. A user can scroll through the list and click on a pair to select that clip.

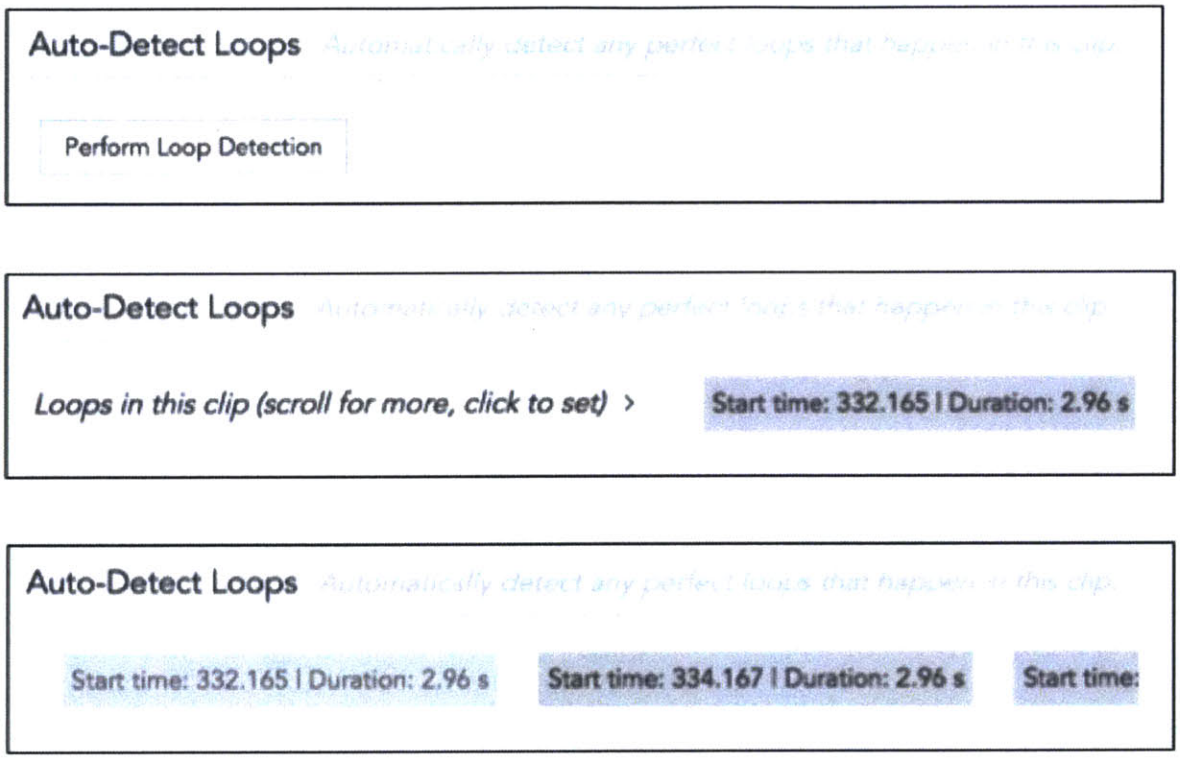


Figure 5-6: Loop detection interface

5.5. Adding Looping Effects

Some videos have many seamless loops that a user can uncover with Glyph's automatic loop detection. But often, a video or a key scene does not have any naturally occurring loops. In these cases, Glyph allows users to create the appearance of seamless looping. In the *Looping* section of the interface, users can control how their glyph loops from four patterns: standard, forwards/backwards, progressive fade, and still fade.

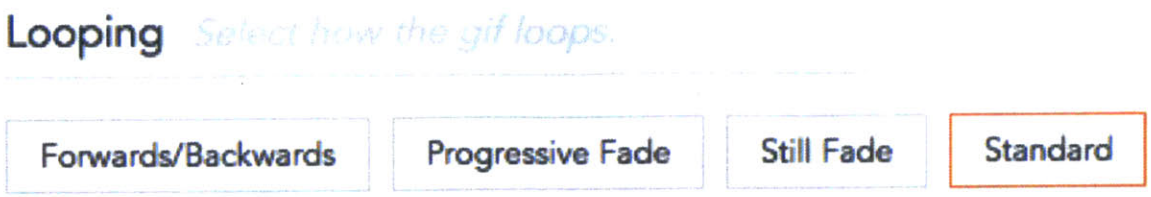


Figure 5-7: Selecting looping effects

5.5.1. Standard

Standard is the default looping pattern for glyphs, in which no additional transformations are performed on the clip's frames. In a glyph created with standard looping, the frames of the clip play sequentially from beginning to end, and then repeat from the beginning of the sequence, like a typical animated GIF. Standard looping is often ideal for clips whose first and last frame are similar enough that the clip appears to loop seamlessly, with no additional editing needed. For example, a user might select a seamlessly looping clip from the results of automatic loop detection, and keep the default standard looping option when outputting the final glyph.

5.5.2. Forwards/Backwards

Forwards/backwards looping is a looping pattern that plays the original sequence of frames, and then plays them in reverse, repeating this alternating pattern. This process is called time symmetrization, and keeps the dynamism across the timeline of a clip smooth and continuous by avoiding a potentially jarring transition as the clip's playback cuts from the last frame back to the first frame.

Time symmetrization always yields a clip that loops seamlessly. But the extent to which a time symmetrized clip's motion appears natural and convincing varies. The dynamism in the clip needs to make semantic sense when its direction reverses for this looping pattern to yield a convincing image. For example, a clip with falling snow or a crowd of people walking would not make sense when time symmetrized, since snow doesn't fall up and crowds of people usually don't walk backwards. Likewise, a clip can appear unnatural when several independently moving regions all reverse their direction of movement simultaneously. Clips that already have dynamism that is roughly symmetric across some axis often yield a natural and convincing looped image with

this effect applied. For example, time symmetrization is often successful when applied to clips with human gestures, such as the gentle motion of a person's body as they hold relatively still, or a series of upper-body gesticulation as a person speaks. These movements happen along the axis of a body, and therefore make semantic sense when played in the reversed direction.

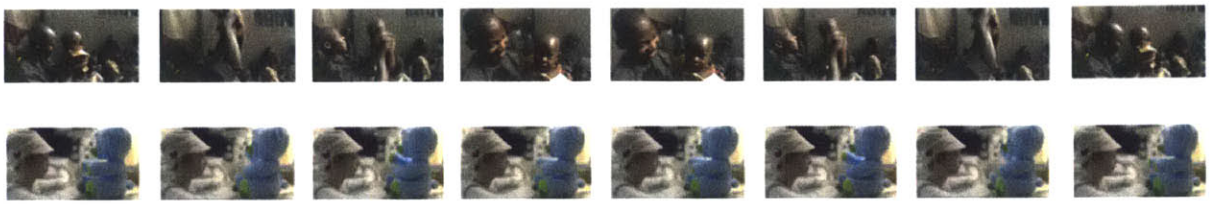


Figure 5-8: Time symmetrization often successfully loops human gesture. Above, a man and a baby play with seamless continuity. Bottom, a girl and a robotic bear take turns interacting.

5.5.3. Progressive Fade

A progressive fading is a looping pattern that gradually fades in the beginning of the clip as the end of the clip plays. The effect creates a montage of the clip overlapping itself in a time-offset sequence that fades out the clip's jump cut from its end back to its beginning, appearing to extend the action continuously.

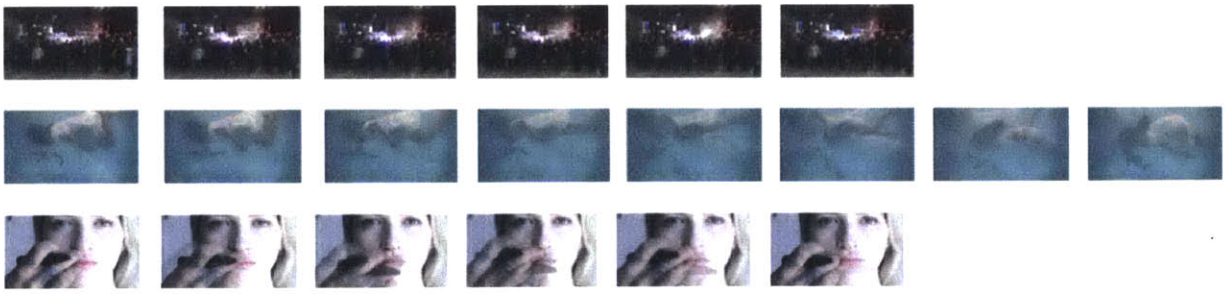


Figure 5-9: Progressive fading. Top, a protest in suspended animation; middle, layer effects loop lateral motion; bottom, a a woman holds our gaze as a makeup artist's hands blur with activity.

The progressive fade is designed to create the appearance of a seamless loop in clips with large fields of small, independently moving regions–like a landscape with

moving trees and water, rainfall, or a crowd of people walking. When applied to a clip with small regions of movement, the progressive fade creates the appearance naturally looping dynamism, and the ghosting artifacts that accompany the progressive crossfades are less noticeable. When applied to clips with large regions of movement, progressive fading yields a more stylized loop, as its ghosting effect across the image is obvious.

5.5.4. Still Fade

Whereas the progressive fade transforms every frame in the clip, the still fade effect plays the clip from beginning to end, and then adds a still crossfade from the last frame to the first frame. This effect is designed to disguise a slight pixel difference between the last and first frames that would otherwise register as discontinuous when looped. In a clip with a small amount of motion, the still fade can appear to erase the transition between the clip's end and its beginning. But because the still fade stills the image when crossfading between the last and first frame, this transition offers less utility when attempting to create a loop with a very dynamic clip.

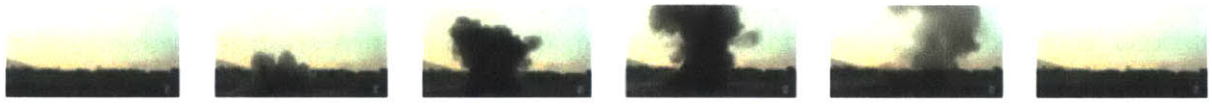


Figure 5-10: Above, the still fade effect is apparent in the second to last frame, when the effect fades out the explosion to transition seamlessly back to the first frame.

5.6. Masking a Clip

Figure 5-11: Split (left) and Regional (right) Masking interfaces

Sometimes, a region of movement in a clip disrupts an otherwise seamlessly looping image. Or, a region of erroneous movement in the clip distracts from the area of intended focus. In these cases, Glyph enables users to designate which regions of the clip are static and which regions are dynamic through masking. When a region of a clip is masked, the pixels of a designated frame in the clip do not change in this region in the final looping glyph.

Glyph offers two methods for defining a mask: split masks and region masks. When a user clicks the split mask button in the Masking section, a transparent, HTML5 canvas overlays the embedded video player. On this canvas, users can draw a line anywhere in any direction to split the clip into two regions. In the options menu that appears in the Masking section, the user selects whether to still the region to the left or to the right of the drawn line. Similarly, the region mask button allows the user to draw a rectangular region, and select whether to still the inside of this region or to still the entire clip except this region. The frame used for the stilled region of the mask defaults to the first frame in the clip. But in the options menus for both masking types, users can designate a time in seconds from 0 to the clip's duration to select which still frame to use in the mask.

Figure 5-12: Top, a split mask stills erroneous motion in the windows, while a progressive fade extends the protest and communicates its ongoingness . Bottom, a region mask freezes a woman fixing her car to reduce a source of distracting motion in the looping roadside scene.

5.7. Outputting Glyphs

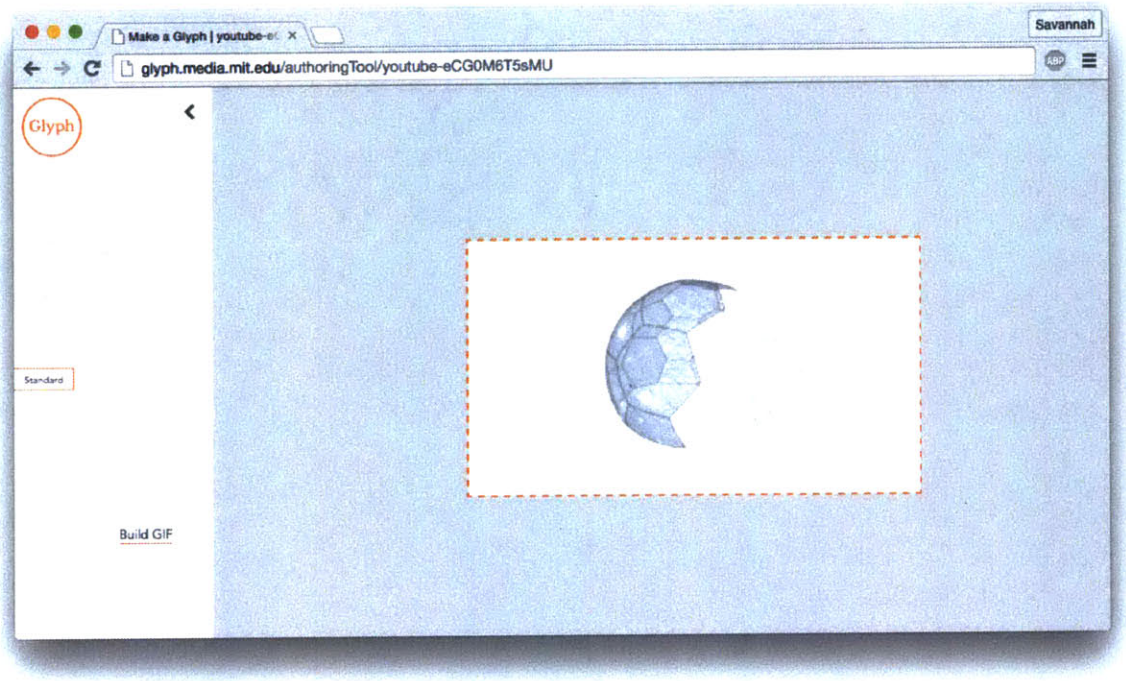


Figure 5-13: Output area

Once the user is satisfied with her choice of clip and has selected any additional effects, she can click the *Output GIF* button, which slides the authoring interface to the left to reveal the output area. After a few seconds of server-side processing, the finished glyph appears in this area of the interface. This sliding interface gives space to a single page interface for authoring and viewing output, thus supporting an iterative workflow.

6. Video Processing System and API

The Glyph API connects the client application to a backend system that supports the video processing and image creation that the interface affords. While section 4.3 provides an overview of the structure and dependencies of this system, this section discusses its technical design and implementation in detail.

6.1. Video Download and Metadata Acquisition

As discussed in section 5.2, the process of creating glyphs begins for a user with the selection of a YouTube video. When a user submits a YouTube URL, Glyph's client application sends a POST request to the Glyph API to validate the URL, and downloads the video and its metadata.

For each video processed in Glyph, Glyph downloads and stores metadata from YouTube, such as the video's author, title, and description. This metadata is preserved so that authorship and content information for a video processed by Glyph is retrievable if the video is later deleted.

Appendex A is an example of a metadata file stored for a processed video. The video's metadata is acquired through an open-source command-line program called YouTube-dl[1]. YouTube-dl provides a command-line flag to scrape metadata from a YouTube video page and save this information as a .json object. YouTube-dl is then used to download the selected YouTube video as an .mp4 video file.

6.1.1. Directory Structure

When a video's metadata is scraped, Glyph uses the video's YouTube id from this metadata to generate its unique identifier in Glyph's data store. The naming conventions for videos in Glyph follow the pattern `[extractor]-[id]`, where the extractor is the source of the video (i.e. YouTube) and the id is the video id at its source.

After downloading the video's metadata and creating its identifier, a directory tree is created to store this metadata and the downloaded video (see figure 6-1). Because

---

1 YouTube-dl. http://rg3.github.io/youtube-dl/.

this content is stored, if the same URL is input into Glyph, the user will be immediately directed to the authoring tool, as Glyph does not need to re-download the content. The video's directory also contains a subdirectory called `/gifs` for collecting all gifs generated from the video.

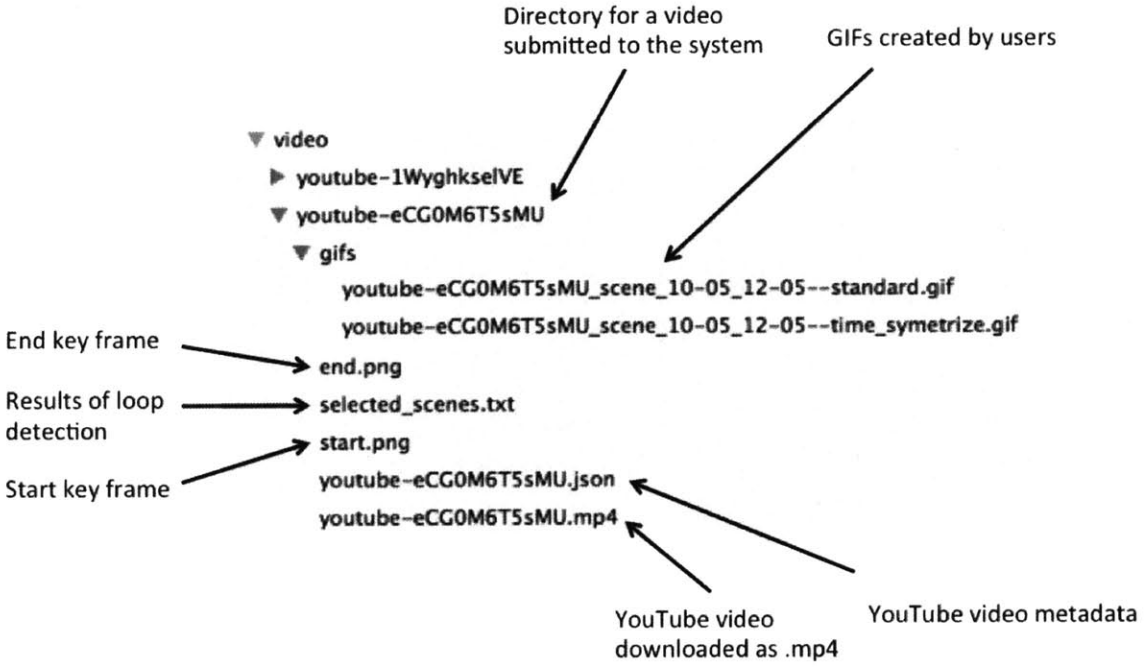


Figure 6-1: Directory structure for Glyph video, image, and metadata store

6.2. Key Frame Previews

As discussed in Section 5.3.2, the Glyph user interface supports a user making fine-tune adjustments to a clip's start time and duration by continuously updating key frame previews for first and last frame of the currently selected clip. The Glyph API supports the asynchronous call to generate these frames. The function `createThumbnail` quickly builds a preview frame from the video for each time value using MoviePy and FFmpeg.

6.3. Loop Detection

The Glyph API includes an endpoint to perform loop detection on a video, an automatic process generating a list of clips that can be looped seamlessly with no additional manipulation. The implementation of loop detection in Glyph is based off of a blog post written by the developer of MoviePy[2], which describes an algorithm for quickly detecting all matching pairs of frames in a video and filtering the results.

A seamlessly-looping GIF can be created from a short clip whose first and last frame are almost identical, i.e. having a pixel difference below a noticeable threshold. The MoviePy library includes a toolkit for making automated cuts in a video clip, including the `FramesMatches` class for automating the detection of matching pairs of frames[3]. Glyph implements the `FramesMatches` algorithm with the default maximum percent pixel difference of .04% and sets the duration for matched clips to between 3 and 5 seconds. The algorithm compares every framec to all the frames 3 to 5 seconds before it, checking for matches. It reduces the computation time of this matching process by three methods: (1) reducing the dimensions of a clip before it's processed, (2) generalizing the pixel difference between two frames as a geometric *distance* between two frames, then using triangular inequalities to detect matches among a range of frames without computing the distance between every pair, and (3) improving the efficiency of the formula for calculating frame distance.

The `FramesMatch` algorithm finds many more matching pairs of frames in a video than there are compelling loops. This is because the algorithm registers still clips as matches, and may detect many nearly identical, slightly-time-offset segments from an extended clip with repetitive motion. These results have to be filtered to be meaningful to the user. The MoviePy toolkit for automated cuts provides methods for filtering a list of clips by a variety of parameters. Using these filtering methods, Glyph returns clips within the following parameters: clips with a duration of at least 1.5 seconds; clips that are not static, with at least one frame in the sequence of frames having at least a .5% pixel distance from the the first and last frames; clips that occur greater

---

2 Zulko. "An Algorithm to Extract Looping GIFs From Videos". __del__(self) / Eaten by Python. 2015 Feb 1. http://zulko.github.io/blog/2015/02/01/extracting-perfectly-looping-gifs-from-videos-with-python-and-moviepy/.

3 MoviePy.video.tools.cuts.py. https://github.com/Zulko/moviepy/blob/1ddd608ccc28cf604a-da17c85e2c436dc8ab61b2/moviepy/video/tools/cuts.py.

than 1 second before the start and end of other matching clips. These results are calculated and sent to the client as an array of clip start times and durations.

6.4. Building a Basic GIF

A user can build a glyph after defining a clip and optionally selecting any looping or masking effects. The Glyph API handles all requests from the client to process a GIF from a video with one route that minimally requires a video identifier, a start time, and an end time. The `/makeGif` endpoint also accepts optional parameters that define looping effects and masks (see table 6-1).

Parameter	Required?	Description	Examples
videoId	yes	Video's unique identifier	'youtube-eCG0M6T5sMU'
start	yes	Clip start time in seconds	'7'
end	yes	Clip end time in seconds	'11'
loop	no	Type of looping effect if any	'still_fade', 'progressive_fade', or 'time_symmetrize'
maskType	no	Type of mask if any and region to still	'maskRight', 'maskLeft', 'maskInner', or 'maskOuter'
mask	no	Coordinates of mask in pixels in the form 'x1,y1,x2,y2'	'100,150,300,375'

Table 6-1: Required and optional parameters for creating a GIF from a video

All glyphs are processed with the same basic procedure, using MoviePy, FFmpeg, and ImageMagick. The process begins by creating a unique file name for the to-be-processed glyph. Glyphs are stored with a specific naming convention that describes all parameters used in the glyph's generation (see figure 6-2).

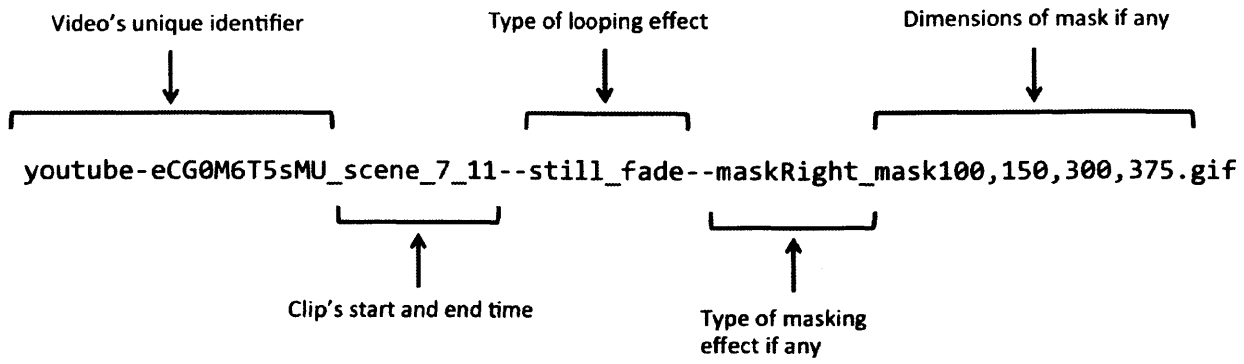


Figure 6-2: Glyph naming convention

Next, the clip is generated from the video using the `VideoFileClip` class[4] included in the MoviePy library. This class provides a Python wrapper for a set of FFmpeg video writing functions, including resizing and excerpting. The video is excerpted into a clip at the start and end time provided. The clip is then resized to 600 pixels in width to reduce load time and reduce the size of the final glyph.

After creating a clip, its frames can be further manipulated if looping and masking parameters are provided in the request. Implementation details on handling these cases are discussed in the following sections. If no additional processing is called for–i.e., the user selected the default "standard" looping method and did not create a mask–then the clip is immediately processed into a GIF. MoviePy's `VideoFileClip` class includes a `write_gif` function that is a Python wrapper for writing a GIF from a video file with either ImageMagick or FFmpeg.

Both ImageMagick and FFmpeg were tested for speed, compression, and image quality. Because Glyph is designed to be an iterative tool (see section 4.3.2), enabling users to process a video into many GIFs very quickly, the speed of these tools, or the time it takes for each tool to process a GIF from a video clip, is critical. But the effectiveness of Glyph as a tool for generating evocative, transmissible images depends on the compression and image quality of its output. FFmpeg builds GIFs slightly faster than ImageMagick, but at the cost of poor color optimization, which results in worse

[4] VideoFileClip documentation. https://zulko.github.io/moviepy/_modules/moviepy/video/VideoClip.html

image quality and larger output files. ImageMagick, though slower than FFmpeg, provides several features for optimizing the compression of the file and quality of the image[5], and is therefore used to build GIFs in Glyph.

6.5. Looping Effects

As discussed in section 5.5, Glyph enables users to select from three looping effects to create the appearance of seamless looping in a non-looping clip. These effects are transformations and compositions performed on the frames of the video clip before it is output as a GIF.

6.5.1. Time-Symmetrization

Time-symmetrization refers to the process of compositing a clip's sequence of frames with a reversed sequence of its frames, resulting in a clip twice the duration of the original, and composed of two symmetrical frame sequences (for further discussion, see section 5.5.2). To accomplish this, Glyph makes a copy of the clip, and applies an effect provided in the MoviePy video effects library[6], `time_mirror`, to reverse the ordering of the frames. This clip is then concatenated with the original clip.

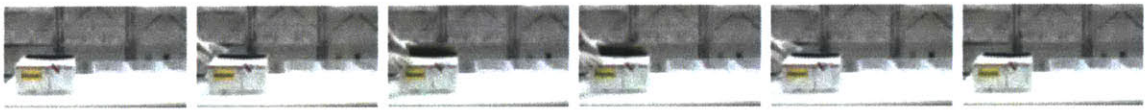


Figure 6-3: Time symmetrization

Because time-symmetrization creates a clip that's twice the duration of the original clip, this transformation is the most time intensive of all of Glyph's clip transformations.

---

5 "ImageMagick V6 Examples: Image to GIF, Optimization Summary". http://www.imagemagick.org/Usage/video/#gif.

6 MoviePy.video.fx (vfx). http://zulko.github.io/moviepy/ref/videofx.html.

6.5.2. Progressive Fade

Progressive fading creates a clip that progressively fades in the beginning of the clip as it ends, resulting in a montage of continuously looping, time-offset clips that obscure the transition from the clip's end to its beginning (for further discussion, see section 5.5.3). To build a clip with progressive fading, a transformation provided by the MoviePy video effects library, `crossfadein`, is applied to the first half of the clip's duration. This transformation creates a clip that begins transparent, progressively fading to full opacity by the middle of the clip's duration.

Three copies of this transformed clip are then composited at a time offset of half the clip's duration. So, for a clip with a duration of 2 seconds, a copy of the clip would start at t=0, then a second copy of the clip would start at t=1, then the last copy of the clip would start at t=2. This composition is then cut at t=d/2 and t=3d/2, where d is equal to the original clip's duration, resulting in a progressively-faded clip equal in duration to the original clip.

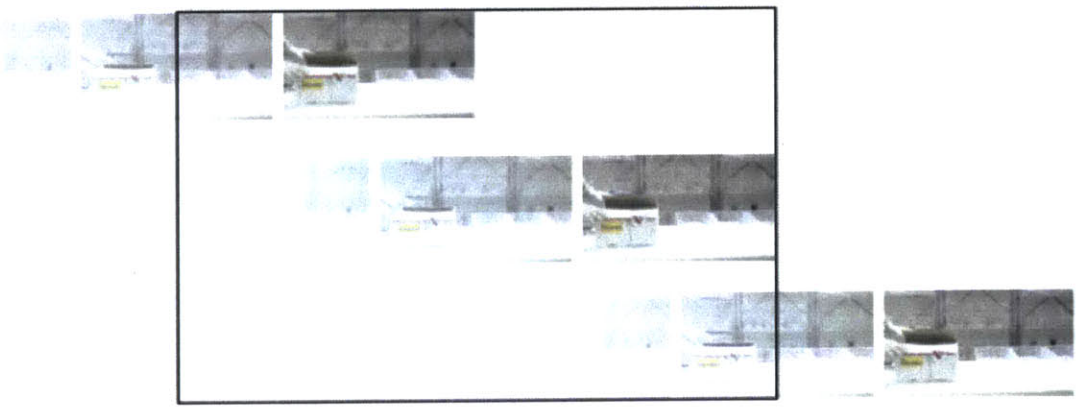


Figure 6-4: Progressive fading

6.5.3. Still Fade

A still fade is a limited crossfade between the last frame to the first frame, designed to "blend" a small pixel difference in the transition between these two frames (for further

discussion, see section 5.5.4). To achieve this, Glyph uses the `ImageClip` class[7] included in the MoviePy library to create an image that can be composited with the rest of the clip based on a given start time, duration, and position. The image is given a duration of 1/6 of the original clip's duration, and a `crossfadein` effect is applied to this still clip. The transformed still clip is then composited with the original clip, beginning at the start time 5d/6, where d is the duration of the original clip.

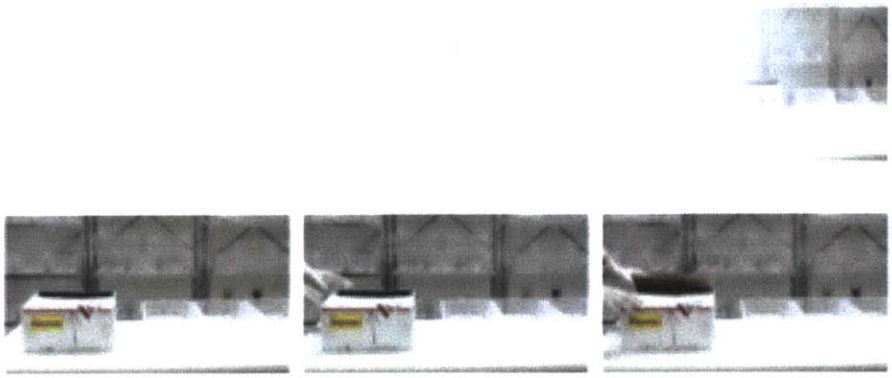


Figure 6-5: Still fading

6.6. Masking

As discussed in section 5.6, masking is a powerful feature for building semantic or affective emphasis in seamlessly looping gifs. Masking is performed in Glyph by using the MoviePy `ImageClip` object to build cropped, still images that are composited as overlays onto the clip. Masks can be combined with other looping effects. Linear and

7 MoviePy.Video.ImageClip. http://zulko.github.io/moviepy/_modules/moviepy/video/ImageClip.html.

regional masking are the two currently supported mechanisms for defining a mask in the Glyph API.

6.6.1. Linear Masking

The MoviePy library supports building a linear mask from the `VideoFileClip` class. A mask is a type of video clip. But, where each pixel in a video clip has three values (RGB) from 0 to 255, each pixel in a mask has just one value per pixel, 0 or 1, designating the pixel as visible or transparent.

Building a linear mask in the Glyph user interface defines a line that splits the area of the video clip (see section 5.6). A request to process a glyph with a linear mask thus includes two coordinates that define this user-drawn line. The request also includes selections from the user indicating which region (either the region to the right of this line or to the left of it) and which frame in the clip will be used to build the still region of the mask. This information is used to build the linear mask. A mask is created from the clip, and split by a line defined by the two coordinates included in the request. To create a natural division between the static and dynamic regions of the final clip, this boundary is given a 5-pixel gradient. All the pixels on one side of the boundary are given a value of fully visible or fully transparent depending on which region is specified as still in the request. This mask is used to build an `ImageClip` using the frame specified in the request, or if not specified, using the first frame as a default. The pixels in the `ImageClip` are assigned the visibility values of the mask. The image– now half visible, half transparent– is given a duration equal to the clip's duration. The image is composited with the original video clip to create a partially static, partially dynamic composition.

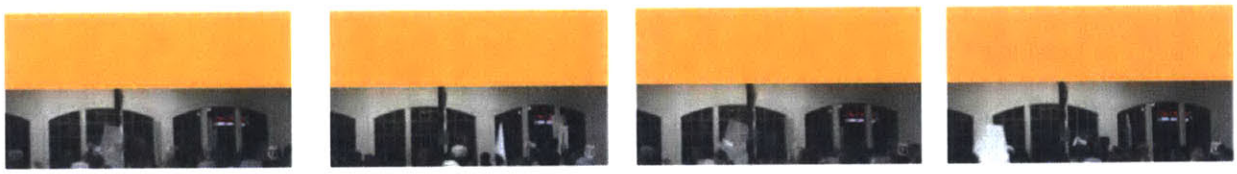


Figure 6-7: Linear Masking

6.6.2 Regional Masking

Regional masks allow users to designate a rectangular area anywhere on the video clip as static or dynamic (see section 5.6). Although regional masking enables adjustments to the dynamism of regions in a clip much like linear masking, regional masking is implemented without the use of masks. Instead, Glyph builds regional masks through use of basic cropping and composition operations available in MoviePy's `VideoFileClip` class.

After a user defines a region in the clip, the user selects whether to still the area inside the region or the rest of the clip outside the region. A request to process a glyph with a regional mask thus includes two coordinates that define this user-drawn rectangle, and which region to still.

To still the inner region, an `ImageClip` of the video at the first frame in the clip (or the frame specified in the request) is generated and cropped to the dimensions selected by the user. The duration of the cropped image is set to the duration of the clip, and its position is set to the top left coordinate specified in the request. This static region is then composited with the clip.

The procedure for stilling the outer region of the clip differs slightly. In this case, after an `ImageClip` of the video is created, the pixels within the specified region are subtracted from the image. The result of this operation creates the image that will be extended and composited with the clip.

Figure 6-8: Regional Masking

7. Applications and Explorations

Glyph makes it possible to produce seamlessly looping images quickly and in large numbers. Glyph thus supports explorations in integrating these kinds of images into interfaces on the web and off. This section describes several applications and interfaces that glyphs improve or make possible.

7.1. Glyphs in News on the Web

Glyph can create compelling looped images on all kinds of stabilized video, but glyphs are particularly suited for integration alongside news stories and headlines. As GIFs, glyphs make embedding video onto webpages and in social feeds lightweight and as easy as embedding and sharing images. The broad editorial control Glyph allows users when creating images supports crafting polished, precise images that reflect the seriousness of headlines or long-form features more appropriately than the jarring jump-cuts and retro aesthetics of most GIFs. The tool allows journalists to create dynamic images that crystallize a particular affective quality of a video to visually set the emotional tone of a story. The light, continuous dynamism of glyphs enlivens scenes and may enhance the sense of ongoingness or simultaneity of the events in the news, or evoke a reader's empathy for the subject of a story.

An article titled "What You Need to Know about the Cuban Thaw"[1] was written alongside the development of this thesis work to explore a close integration of glyphs with text. The article was written on FOLD, an authoring platform developed at the Center for Civic Media at the MIT Media Lab[2,3]. FOLD allows an author to build modular stories supported by multimedia context cards. The article "What You Need to Know about the Cuban Thaw" uses these context cards to distribute glyphs throughout an explainer on US-Cuban foreign policy. Some of the glyphs depict Havanan scenes with more life than a still image, grounding the text in an actualized setting

---

1 Niles, Savannah. "What You Need to Know about the Cuban Thaw." *FOLD*, April 21, 2015. https://readfold.com/read/sannabh/what-you-need-to-know-about-the-cuban-thaw-CST8k5cg.

2 Hope, Alexis. *FOLD*. https://readfold.com/.

3 Andrew, Liam. "FOLD Wants to Keep You from Tumbling down Link Rabbit Holes." *Nieman Lab*, July 2, 2014. http://www.niemanlab.org/2014/07/fold-wants-to-keep-you-from-tumbling-down-link-rabbit-holes/.

without interrupting the reading experience like an embedded video would require[4]. Integrated throughout the article are glyphs of Fidel Castro, Raúl Castro, President Obama, and Pope Francis II that provide the text with richer characterizations of its principal actors. A glyph captures President Obama's characteristic gestures as he addresses congress, and another extends a handshake between President Obama and Raúl Castro. Providing a salient backdrop to a section on the Cuban revolution, one glyph captures and prolongs a young Fidel Castro's aggressive gestures and the wisps of smoke from his cigar as he addresses fellow revolutionaries (see figure 7-1). The image offers the story an evocative portrait of Castro's radicalism and charisma that would be lost if depicted in an image and ignored if buried in a video.

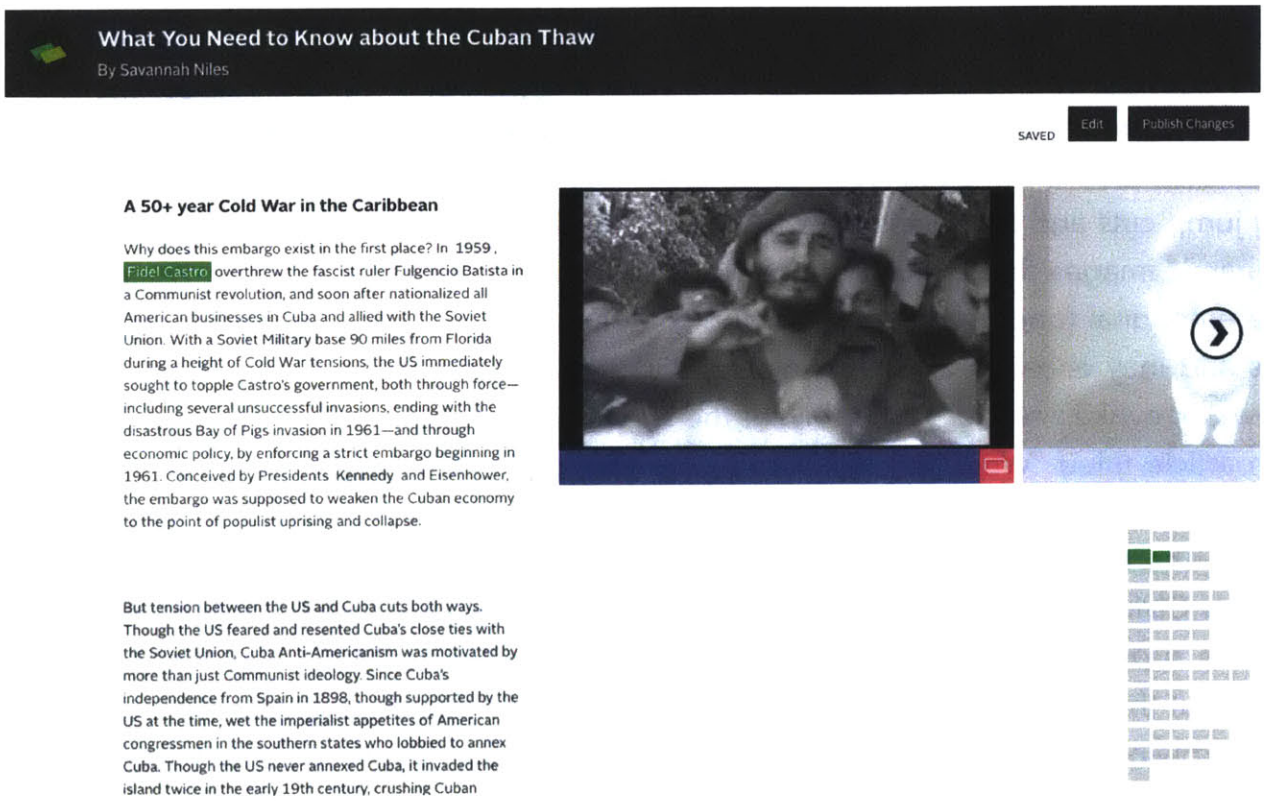


Figure 2-1: "What you need to know about the Cuban Thaw", FOLD

4 Zuckerman, Ethan. "Introducing FOLD, a New Tool (and a New Model?) for Storytelling." *...My Heart's in Accra*, April 22, 2015. http://www.ethanzuckerman.com/blog/2015/04/22/introducing-fold-a-new-tool-and-a-new-model-for-storytelling/.

Section 2.2 discusses other examples of integrating seamlessly-looping images within news stories. GIFs– especially the stylized, looping GIFs that Glyph facilitates creating– are emerging as a genre in visual journalism. But there exists little literature outlining (1) the best practices for editing GIFs for particular journalistic purposes, (2) shooting for GIFs in the field, and (3) exporting GIFs or short embedded video files to optimize image quality and transmissibility.

Included in Appendix C in this thesis is a best practices guide for incorporating GIFs into visual journalism, co-written with Audrey Cerdan[5], a visual journalist based in France. The guide aims to describe the precedent for using GIFs in journalism and differentiate among typologies of GIFs based on emotional and informational qualities. The guide also provides overviews of the technical considerations for shooting video in the field for creating GIFs and the available technologies for processing them.

7.2. Glyphs Beyond the Browser

Glyphs have applications beyond the browser. As pixels increasingly occupy our peripheries through mobile devices, wearables, and ambient displays, video's high costs in attention, time and bandwidth tether it to single-task, non-mobile interfaces, such as televisions and computers. But Glyph makes it feasible to produce dynamic, cogent images suitable for the ubiquitous and diverse display surfaces in our lives. As moving images that are compressive in information and bandwidth, glyphs suggest a paradigm for visually manipulating video for small displays in our environments. They avoid distraction through their limited dynamism and repetition. They're "responsive images" that retain their meaningfulness both at high-resolutions and as bits in the second-order information streams that occupy the peripheries of our attention.

The periphery of our awareness and attention is an increasingly significant frontier in human computer interaction. In 1995, Mark Weisser and John Seely Brown at Xerox PARC described the need to design interfaces that both "encalm and inform" as potentially "the most important design problem of the twenty-first century"[6]. They outlined a vision for *calm technology* that moves seamlessly between the foreground and

---

5 Audrey Cerdan. *Rue89*. http://riverains.rue89.com/audrey-cerdan

6 Weiser, Mark and Brown, John Seely. "Designing Calm Technology". Xerox PARC, December 21, 1995. http://www.ubiq.com/hypertext/weiser/calmtech/calmtech.htm.

the background of our experiences, "informing without overburdening" us. Glyph offers a paradigm for visually manipulating video for such interfaces by supporting the creation of video excerpts that maintain informational and emotional density without overburdening. Put another way, Glyph is a platform for creating content for displays on our bodies and in our environments that are responsive to vast information and news streams, but scaled to the limited scope of our attention.

The following sections speculate how glyphs support the design of attention-conscious media interfaces at two scales: at the small scale of objects and wearables, and at the larger scale of environments and architecture.

7.2.1. NanoVideo

Small visual media interfaces designed with a conscientiousness for 'calmness' "[seem] a necessity when imagining an everyday world in which technology pervades our bodies and environments and is embedded in even the most ordinary experiences"[7]. With small screens now common to our environments, there are opportunities to explore visual paradigms for integrating high-bandwidth, informationally dense content into the world, such as news video.

NanoVideo is a simple exploration of the use of glyphs to integrate video subtly, meaningfully, and socially into the background of our working and living environments. The project involves two common objects– a watch and a plastic block– both equipped with small screens. Both the NanoVideo block and watch are built with Moto 360s, an Android-based smart watch[8]. They utilize the Android applications aGif[9] for GIF support and Swipify[10] for overriding the screens' automatic shut-off.

The NanoVideo watch displays a glyph on loop on the OLED watch face at all times, demonstrating how glyphs can add dynamic, visual headlines to wearable displays. Wearables are self-centric devices, and this manifests in their primary content– inboxes, calendars, and health data (see figure 2-3). But wearables are also objects of

---

7 Veel, Kristin. "Calm Imaging: The Conquest of Overload and the Conditions of Attention". Ed. Fuller, Matthew. Throughout: Art and Culture Emerging with Ubiquitous Computing. Edited by Ulrik Ekman. Cambridge, MA: The MIT Press, 2012.

8 Moto 360. Motorola. http://moto360.motorola.com/.

9 aGif. https://play.google.com/store/apps/details?id=src.schimi.gifwatchface&hl=en.

10 Swipify. https://play.google.com/store/apps/details?id=com.wingoku.swipify.phone&hl=en.

fashion, positioned to signal quietly to others the things that we're concerned about. In the case of news content, wearables can display images from the news with dramatically increased relevance to the wearer's lived experience and social interactions. Evocative, lightweight, un-distracting moving images like glyphs are particularly suited for such displays.

Figure 2-3: Currently, most wearables are exclusively responsive to a user's personal information, such as the yellow activity tracker shown above. Glyph supports the creation of content for wearables that reference events far away and experiences outside our own, such as protests in Ferguson, MO.

The NanoVideo block explores how small displays embedded in objects in our environment can become interfaces to visual media. The NanoVideo block is a plastic block with an embedded round display that plays a looping glyph. The block materializes glyphs into a physical token that can be placed in a personal space, like a nightstand, or a social space, like a desk at work, to spatially organize visual media.

Whereas video would be distracting on such an interface, glyphs visually represent an event, a place, or a feeling with subtlety and cogency. Whereas video content is ephemeral (its fleeting scenes locked within a browser window that is just a click away from closure), the NanoVideo block gives a moment in a video the persistence of a physical object. The block thus supports new patterns of awareness and reinforced concern for issues in the news. In a browser, "keeping tabs" on a developing event or issue in the news literally requires keeping browser tabs open to search results and news feeds; the NanoVideo blocks demonstrates that the news media collapsed in tabs could be materialized onto small, persistent interfaces arranged spatially in our peripheries.

The NanoVideo block begins to demonstrate how Glyph creates opportunities for designing new objects that are integrated with and responsive to high-bandwidth visual media. With small, low-cost, disposable screens and heads-up displays on the horizon, pixels will map to many more objects in our lives beyond our phones and laptops. Video is too attention-intensive for these distributed display surfaces, but glyphs allow moving digital images into our environments less invasively. Glyphs combined with low-cost display technology could make media-responsive objects and packaging possible: a glyph on a medicine's packaging could communicate universally how it should be taken; glyphs on produce stickers could provide a small, dynamic picture of the place in which they were grown; a glyph inside a makeup compact could enhance the experience of using the product through subtle, narrative imagery; a glyph-equipped disposable coffee cup could quietly broadcast a visual morning headline. Glyphs thus make possible further explorations into the design of objects with new relationships to high-bandwidth media and to our attentions.

7.2.2. Environments and Ambient Video

Figure 2-5: Multimedia displays in urban environments documented by the Urban Screens Initiative[11]

Glyphs could also imbue larger surfaces with media-responsiveness while avoiding information pollution. Large video screens already occupy many of our public and personal environments. In our homes and work places, televisions and computer monitors offer pixel real estate even when we're not seated in front of them. In urban spaces, building facades and billboards broadcast video continuously: the commons are increasingly digital. But all of these large media surfaces threaten to pollute our environments with an overabundance of digital imagery. Malcolm McCullough in *Am-*

11 The Urban Screens Initiative. http://www.urbanscreens.org/about.html.

bient Commons: Attention in the Age of Embodied Information addresses the threat of "visual pollution" and "attention theft" in our increasingly pixelated environments[12]. He argues that "information deserves its own environmentalism", but advocates against the legal measures some cities such as Amsterdam and São Paulo have taken to ban outdoor advertising and large information displays. He instead embraces the possibilities of technologically-augmented environments and calls for a "technologically mediated but humane urbanism"[13] in designing media displays for these ambient contexts.

Designing ambient digital media requires precise trade-offs between information density and attention cost. Glyph makes it possible to create images with this necessary precision by facilitating precise control over fluid, undisrupted dynamism in the clip. McCullough posits that our attentions are adapting to our changing information-environments, embodying a state, which he calls the Ambient, of "an increasing tendency to perceive information superabundance whole"; if this is the case, glyphs are videos scaled to our ambient cognition, with the capability of referencing rich media stories with cogency and immediacy.

Glyphs could be integrated into "mixed-reality" architecture, a growing field that designs architectural interiors and facades with rich visual media[14,15]. Many examples of ambient video integrated into the periphery of our environment are either ignored (such as a 24 hour news program playing muted in a waiting room) or enthralling (in the case of a cinematic projection mapping show in a public event). This use of video can be disruptive and intrusive, as anyone who has tried to carry a conversation while a nearby TV plays has experienced. Glyphs can reduce video for un-intrusive dynamic displays. Their fluid, layered dynamism reflect "specific trends within a broader direction" of cinematography developing in response to "new screen technologies [that]

---

12 McCullough, Malcolm. Ambient Commons: Attention in the Age of Embodied Information. Cambridge, Massachusetts: The MIT Press, 2013.

13 Morozov, Evgeny. "Only Disconnect." *The New Yorker*, October 28, 2013. http://www.newyorker.com/magazine/2013/10/28/only-disconnect-2.

14 Schnadelbach, H., A. Penn, S. Benford, and B. Koleva. Mixed Reality Architecture: Concept, Construction, Use. Report. Nottingham: University of Nottingham, 2003. http://discovery.ucl.ac.uk/1021/.

15 Schieck, Ava Fatah gen. "Towards an Integrated Architectural Media Space." First Monday 0, no. 0 (February 6, 2006). http://firstmonday.org/ojs/index.php/fm/article/view/1550.

support and mandate a strong shift to the pictorial"[16]. Glyphs thus add to a new genre of ambient, moving imagery developing at the intersection of film studies, architecture, design, and video processing.

16 Bizzocchi, Jim. "Ambient Video: The Transformation of the Domestic Cinematic Experience." *In Digital Tools in Cultural Contexts: Assessing the Implications of New Media.* Eds. Byron Hawk, David M. Rieder, and Ollie Oviedo. University of Minnesota Press, 2007.

8. Evaluations

This section describes the evaluation of Glyph's output, called glyphs. Through a series of studies, glyphs are evaluated across a variety of dimensions to assess their distractingness, their engagingness, and their informational and emotional density.

8.1 Overview of Study Designs

Three studies were designed to evaluate glyphs: a study on the Quantify platform to assess viewers' subjective responses to glyphs (section 8.2), an engagement evaluation (section 8.3), and a selection-task evaluation (section 8.4). In each study, glyphs are compared to non-seamlessly-looping GIFs and images across various qualitative dimensions.

8.1.1 Evaluated Image Set and Controls

Figure 8-1: Example of a glyph selected for the evaluated image set and two generated controls.

A common image set was selected and used in the three evaluations discussed in this section. This image set began with a selection of 61 seamlessly-looping glyphs[1]. For each of these glyphs, two controls were generated from the same video: a still image was created with the same start time, and a GIF was created with the same start time and duration, without the transformations of the original glyph, such as masking or looping effects[2]. The resulting image set thus included 183 images– 61 glyphs, 61

---

1 The listing of videos used to source these gifs is listed in Appendix B.

2 As shown in figure 6-2, Glyph's naming convention preserves all information used in creating the

still images, and 61 GIFs. The studies discussed in 8.3 and 8.4 also utilize the source video for each glyph.

8.1.2 Overview of Limitations

The inherent difficulty in quantifying qualitative differences within a dataset creates limitations common to each of the studies.

Qualities such as distractingness, engagingness, and informational/emotional density are highly dependent on the viewer of the glyph and the glyph's formal and semantic properties. Glyph makes the creation of moving images with a set of common formal qualities like seamless looping or stylized dynamism. These studies investigate whether these generalizable formal qualities correspond to generalizable responses in viewers. However, the responses are also contingent on the glyph's semantics, i.e. the settings and subjects it represents. The semantic characteristics of an image could override, counteract, or enhance a viewer's response to its formal qualities, and vice versa, making an evaluation of glyphs' generalizable formal qualities difficult.

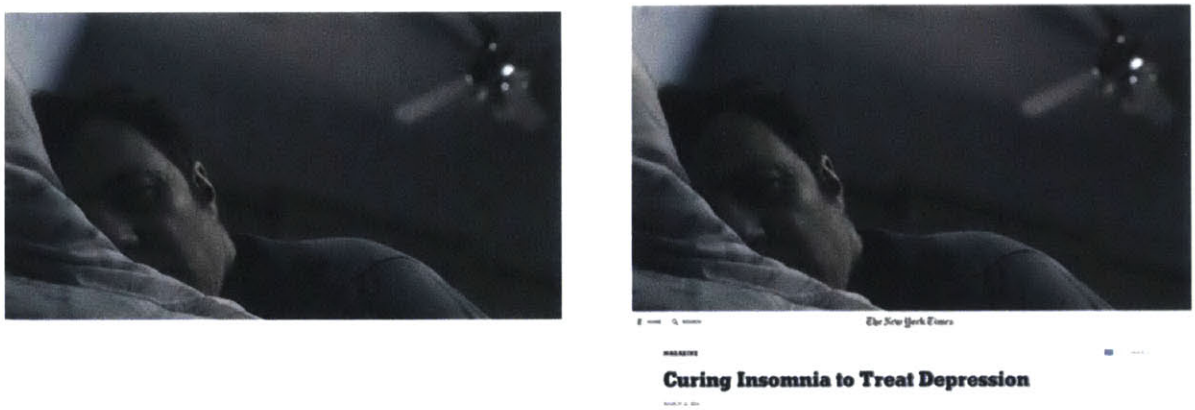


Figure 8-2: The GIF of this man is almost static, except for an occasional blink. The image becomes more evocative when contextualized within a story about insomnia and depression[3]: the formal quality

---

image, including the YouTube ID, the clip's start and end time, and the effects added to the glyph. This allowed controls to be generated precisely and automatically.

3 The Editorial Board. "Curing Insomnia to Treat Depression." The New York Times, November 23, 2013. http://www.nytimes.com/2013/11/24/opinion/sunday/curing-insomnia-to-treat-depression.html.

of repetitive looping reflects our semantic understanding of the ongoingness of the man's health conditions.

Viewer responses evoked by glyphs are also highly dependent on the context in which they're viewed. A glyph embodies a set of editorial decisions: though Glyph is an interface for performing a variety of computationally-driven transformations on a video excerpt, these transformations are user-driven. A user might choose particular effects when creating a glyph for a variety of reasons– to make a more subtle clip, to emotionally elevate an excerpt, to call attention to a particular gesture, or to emphasize a particular affective quality– and for a variety of contexts– alongside a headline, embedded in an article, as a mobile notification, or pushed to a wearable. Evaluating glyphs as more or less successful than the same content excerpted as non-seamless-looping GIFs or stills is difficult when the images are decontextualized from real-world contexts and user behaviors.

8.2. Quantify Evaluation

The Quantify evaluation aimed to assess how viewers responded to glyphs across a variety of subjective dimensions.

8.2.1. Method

The study was conducted on the Quantify platform[4]. Quantify is a platform developed at the MIT Media Lab by Travis Rich in the Viral Communications group that offers a generalized framework for the comparative measurement and multi-dimensional mapping of the subjective qualities of a corpus of content. The framework allows for uploading content (such as still images and GIFs) and a list of subjective metrics (opinion-based modifiers, like "evocative" and "beautiful"). Through a web interface, the framework then crowd-sources comparisons between pairs of content based on these modifiers. For example, Quantify might present a user with two images and the question "Which is more evocative?". The user can select either of the images or 'neither', and this vote is counted as a data point in the system. As an increasing number of votes are collected, each imagea is scored on each subjective characteristic. The result

[4] Rich, Travis. Quantify. MIT Media Lab. http://quantify.media.mit.edu/.

is a corpus of content in which each item is comparatively scored across these subjective metrics.

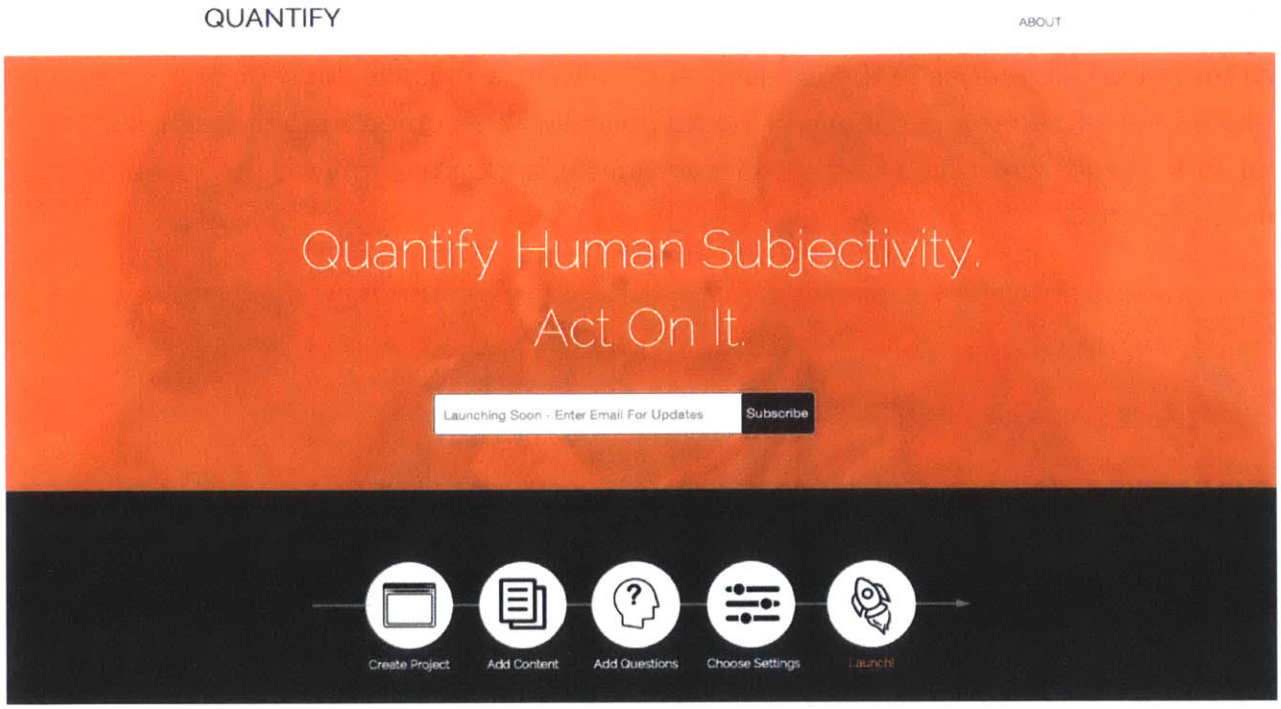


Figure 8-3: Quantify

This study used Quantify to compare the subjective characteristics of glyphs to corresponding GIFs and still images. A Quantify web application was created with this image set and twelve subjective qualities: *evocative*, *interesting*, *informative*, *powerful*, *beautiful*, *boring*, *confusing*, *exciting*, *distracting*, *subtle*, *surprising*, and *engaging*. These terms were chosen to reflect the characterizations glyphs often received compared to their non-looping or still counterparts based on earlier observations and interviews. The terms were intentionally varied in their emotional valence.

The study crowdsourced votes from hired Mechanical Turk workers. Mechanical Turk[5] is a crowdsourcing Internet marketplace from Amazon Web Services that facilitates hiring and coordinating large numbers of people to perform short tasks that

computers are unable to do, such as identifying or comparing photographs, or writing product descriptions.

In the study, Mechanical Turk workers were given a link to the study and asked to vote on which image from a random pair best embodied a given subjective quality. The system paired all images in the dataset randomly, regardless of content or image class: all images had an equal chance of being paired with all other images. Each Mechanical Turk worker was given one subjective quality, and asked to vote on 25 pairs to complete the task.

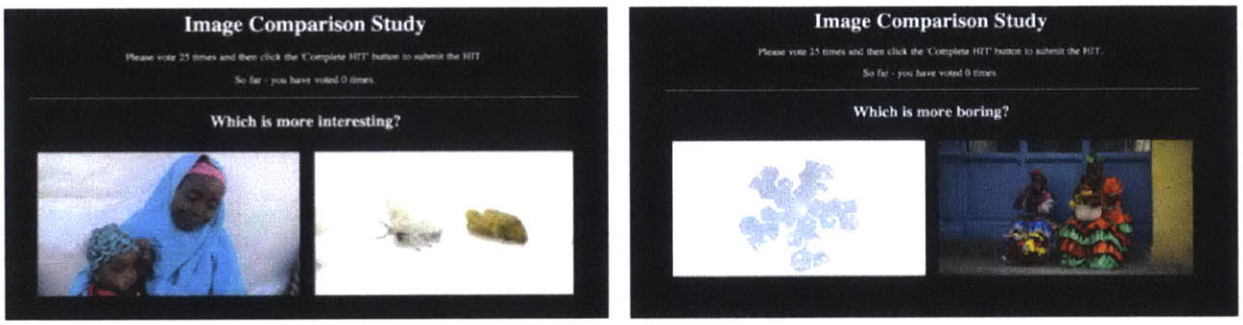


Figure 8-4: Examples of the Quantify-based evaluation application's voting interface.

894 participants voted in the study, and 14,912 votes were received. With an dataset of 183 images, each evaluated on 12 subjective qualities, each image received 7 votes per quality[6]. Each subjective quality thus received about 1243[7] votes total, by about 74 participants[8].

8.2.2. Results

Within each of the 12 subjective dimensions, a distribution of scores was observed for each of the 183 evaluated images, where μ is the mean of the image's ranking within that dimension relative to the other images on a calculated range of 0 to 50, and Σ is the variance of the distribution. Within each dimension, a distribution of the means for each image class was calculated. The distribution of scores for glyphs were compared

---

6 14,912 / 183 / 12 = ~7
7 14,912 / 12 = ~1243
8 894 / 12 = ~74

to the distributions of scores for GIFs across each subjective dimension. Table 8-1 illustrates the difference of means for glyphs and GIFs, with positive scores indicating that glyphs scored higher than GIFs in that dimension. In 11 of the 12 evaluated dimensions, scores differed only slightly between GIFs and glyphs. The greatest difference observed was subtle, where glyphs ranked significantly higher than GIFs.

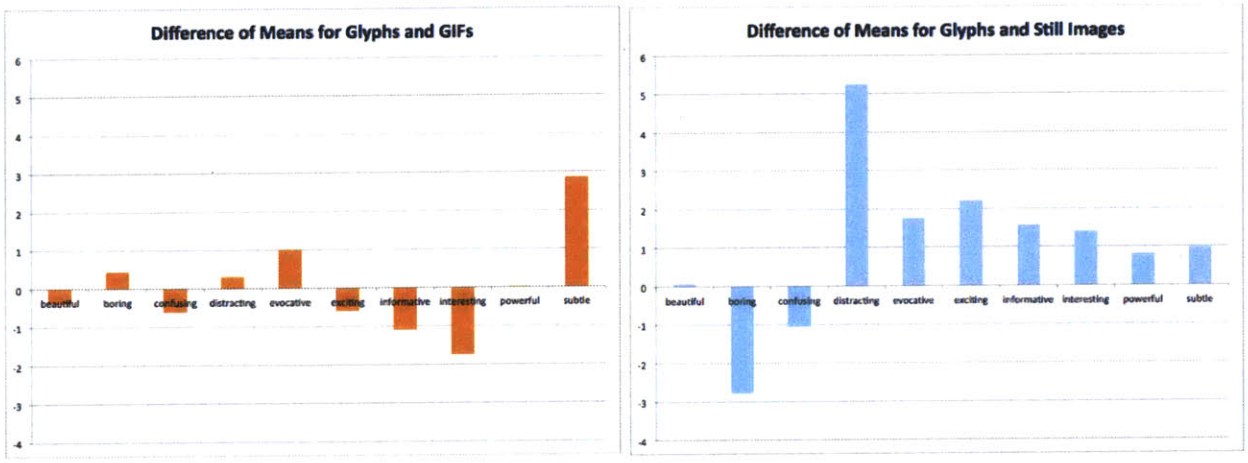


Table 8-1: Difference of means

The distribution of scores for glyphs were also compared to the distributions of scores for still images across these dimensions, also illustrated in Table 8-1. The greater difference in scores observed here likely reflects the more obvious formal differences between images, which are static, and glyphs, which are dynamic. Glyphs ranked as slightly more evocative, exciting, informative, interesting, powerful, and subtle than still images, and they ranked as significantly more distracting. They ranked as less boring and significantly less boring than still images and slightly less confusing.

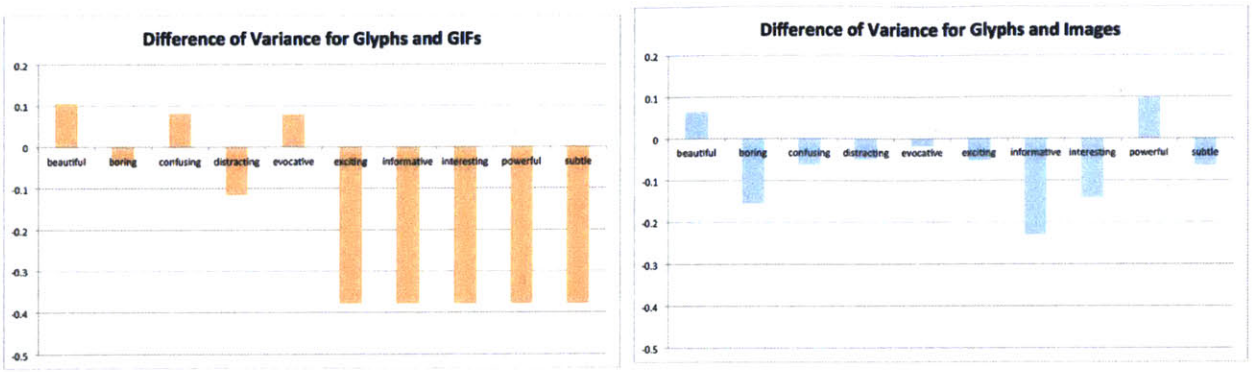


Table 8-2: Difference of variances

Table 8-2 illustrates the difference of variance for glyphs and GIFs, and for glyphs and images, indicating that the confidence observed across all categories is uniform.

8.2.3. Considerations

The following considerations limit the usefulness of this study's results. First, given that the distribution of the ranking of each image was calculated from just 7 votes, this study would need significantly more votes for observations to be made with confidence. Second, it's unclear whether mechanical turk is an appropriate source of participants for this study, or if this service introduces unknown biases that impact scores. For example, since Mechanical Turk workers are paid to finish a given task (in this case, vote on 25 pairs of images), workers might be incentivized to complete tasks as quickly as possible. A participant trying to finish this study as quickly as possible might arbitrarily click the right or left image at every vote, or might click an image before both images have fully loaded. The variances across the categories are low enough to suggest that arbitrary and effectively random voting was not rampant[9], and the Quantify application requires that participants wait at least 2 seconds before a vote can be cast, in order to deter against random voting at maximum speed. But many of the GIFs and glyphs included in this image set were high resolution, and could have taken more than 2 seconds to a load on a slow internet connection: this could have disadvantaged glyphs and GIFs in voting. The limitations discussed at the beginning of this

9 The Quantify platform assigns each image a starting variance of 8.83. If votes were assigned randomly to images, this variance would remain 8.83.

section– including the conflation of formal and semantic qualities throughout the image set, and the de-contextualization of the images from text– cast additional doubt on the conclusiveness of these results. Further evaluation is necessary to assess how viewers respond to glyphs across a variety of subjective dimensions.

8.3. Engagement Evaluation

A second study was conducted to evaluate if (1) users on average looked at glyphs longer than they looked at still images and GIFs, and (2) if users were more likely to watch a video previewed by a glyph than by a still image or a GIF. The study was designed to compare the engagingness of glyphs relative to GIFs and still images.

8.3.1. Method

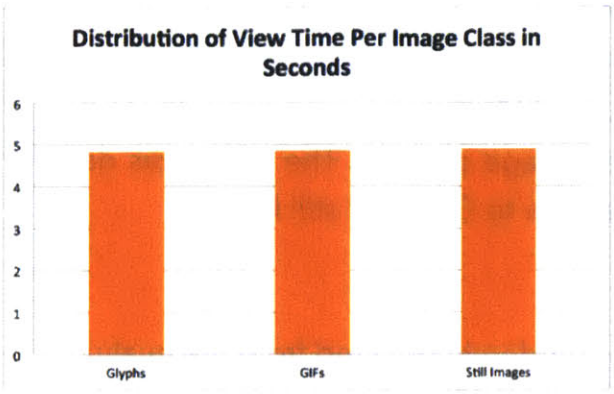
In this study, users were presented with 30 images randomly selected from the evaluation image set, each image shown one at a time in a slideshow format. Users were instructed to advance through the slideshow to view all 30 images, and click on any image if interested in watching the video that the image previewed. The web application recorded the duration of seconds each user spent on each image, and whether the user clicked on the image to watch the video.

Figure 8-5: The engagement evaluation interface. On the left, an image is shown to the user. The user can click on the image to watch the source video, which opens as an orange modal screen shown on the right.

32 participants took the study. Users viewed approximately 320 images in each category of images (still images, GIFs, and glyphs).

8.3.2. Results

As shown in Table 8-3, users on average viewed images of all three categories for approximately 5 seconds. Users clicked on an image to watch its source video at a similar rate of approximately 20% for images in all three categories.



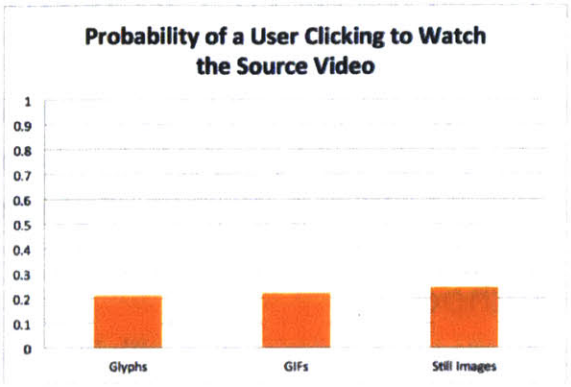


Table 8-3: Results of the engagement evaluation show little observable difference in user behavior among the three categories of images.

8.3.3. Considerations

The design of the testing application could have contributed to the observed consistency of average viewing times across the image categories: the slideshow-like interface could have encouraged a rhythmic clicking pattern as participants progressed through the study. Participants were not encouraged to consider images thoughtfully: all images were shown in a decontextualized manner without associated text, and although context for the images could be gained by clicking the image and watching the video, there was little incentive to do so and therefore prolong the duration of the study. This lack of context could have impacted user behavior during the study or minimized a user's differentiation among images in the evaluation set.

8.4. Multimedia Selection Test

The following study was designed to evaluate whether glyphs facilitate scanning and selecting from a collection of media as succinct previews.

8.4.1. Method

In this study, participants were presented were presented with six matrices of multimedia, shown one at a time. Each matrix contained 12 images, 12 glyphs, or 12 autoplaying videos. For each matrix, a participant was given a question randomly from a set of six questions, chosen to include both identification and opinion questions. The participant answered each question by selecting a media item from the given matrix. The time between the matrix (with all multimedia fully loaded) first appearing to the participant and the participant making a selection was recorded.

Figure 8-6: Three examples of the multimedia matrices shown to users in the study. From top to bottom: a grid of glyphs, a grid of embedded, autoplaying YouTube videos, and a grid of still images.

30 participants took the study, making a total of six selections, resulting in 60 recorded selections per class of media.

8.4.2. Results

As shown in Table 8-4, on average, participants made a selection from a grid of glyphs in 6.9 seconds, from a grid of images in 7.8 seconds, and from a grid of videos in 8.6 seconds.

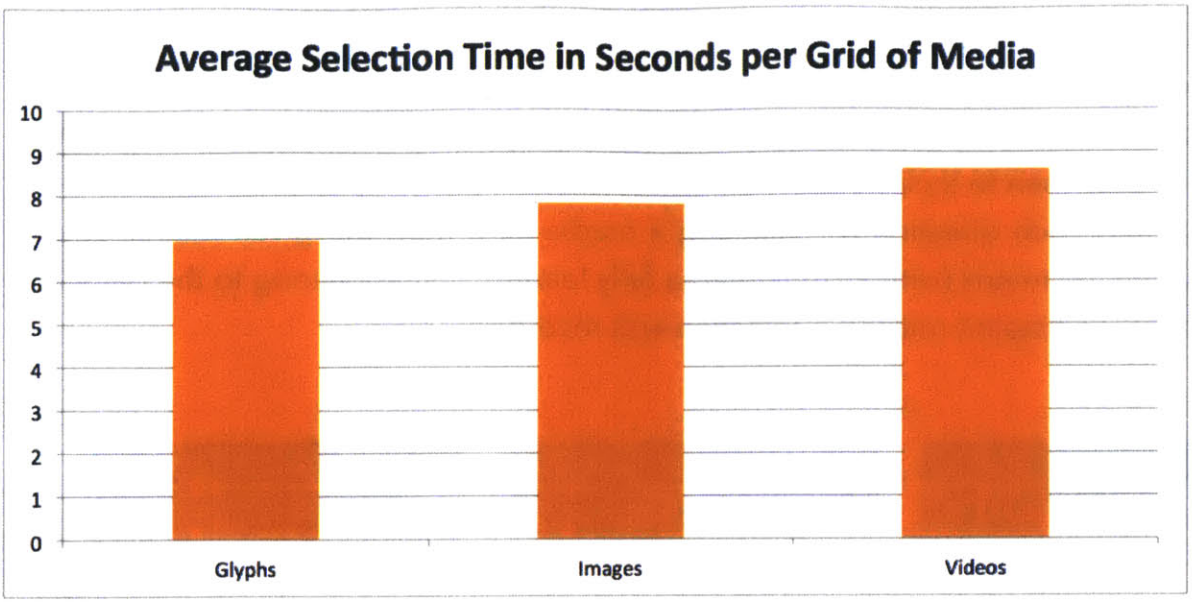


Table 8-4: Multimedia selection test results

These results suggests that glyphs support a user in efficiently scanning and selecting from a group of media more than images and autoplaying video.

8.4.2. Considerations

The study is designed to simulate media interfaces on the web today. Many media site homepages, including *Buzzfeed*, *the New York Times*, YouTube, and Netflix, display grids of video thumbnails on their homepages. If these results maintain across additional studies, it could suggest that glyphs are well-suited for efficiently previewing video content to users. Additional studies should involve more participants, and a wider range of content, different quantities of content. This study does not compare user selection time between grids of glyphs and grids of non-seamlessly-looping GIFs, so additional studies could additionally investigate this comparison.

9. Conclusion

9.1. Overall Achievements

This thesis describes the development of Glyph, a tool that facilitates the creation of lightweight, evocative looping GIFs from video. This work included the development of a set of video-processing functions with API endpoints, as well as the design of a web-based user interface for quickly and iteratively building images based on a user's editorial intentions. Glyph allows users to detect automatically looping clips in a video or create seamlessly-looping sequences from non-looping excerpts. It supports the addition of layering, transition, and masking effects comparable to those offered by sophisticated image-editing software such as Photoshop or AfterEffects, but with the speed, simplicity, and cross-platform compatibility of online GIF-making tools. Glyph allows users to excerpt video with precise modulations of its emotional and informational density– stilling some regions of movement in the clip and emphasizing others to draw out a particular affective or semantic quality.

Through its output, Glyph generalizes a richly-visual, compressive class of moving images, referred to as glyphs. Glyphs were evaluated across a variety of dimensions to assess their ability to distract, engage, inform, or evoke emotional responses from viewers. These evaluations suggested that glyphs do not perform significantly worse than still images or non-seamlessly-looping GIFs in these capacities, and in fact may facilitate more efficient scanning and selecting tasks on media interfaces. But further evaluations are needed to assess the informational, emotional, and cognitive effects of glyphs on viewers and in various contexts.

This thesis explores a variety of applications for glyphs: in web interfaces, on small embedded displays, on wearables, and on ambient displays. Because Glyph allows users to create cogent video excerpts with reduced dynamism, the tool offers a content-creation platform for emerging media interfaces that are responsive to high-volume media streams but conscientious of the limited scope of our attentions.

This thesis gives particular emphasis to the potential impact of glyphs in visual journalism and news media. The emotional dimension of the news is just as important as its informative function. When we read the news we participate in a collective emotional experience– whether that is grief, celebration, worry, or wonder. Visual journal-

ism is a crucial vector for these shared emotional experiences, which can propel civic action. Thus the immediacy, nuance, and evocativeness of the images in the news matter. But footage shot in the field should not limit a visual journalist's ability to derive powerfully evocative imagery. Glyph makes video more malleable, with applications both as a versatile post-production tool in news rooms and as a democratic remixing tool for bloggers and activists. Glyph can support the creation of news imagery with greater emotional and informational reach– lightly dynamic images that look less like renderings and more like windows; images that, when we look at them, appear to look back at us; and images created for the distributed, synthesizing nature of our attentions online, that earn our concern.

9.2. Future Work

The following sections describe opportunities for improving the current prototype of Glyph, and further explorations of its output.

9.2.1. Scaling Glyph for the Web

Glyph is built using web technology, and its video processing functions operate on a server. Future work in making Glyph into a viable, scalable web-service will involve designing a new system architecture than can support numerous simultaneous users requesting sequences of computationally-intensive processes through the client interface. Two potential solutions to this challenge are (1) hosting Glyph server-side computations on a scalable cloud-computing platform like Amazon Web Services[1], and/or (2) removing or throttling the availability of the most computationally-intensive processes to users, such as automatic loop detection. Finally, further research could find that many of Glyph's transformations on video could be done client-side using graphics technologies like WebGL, thereby reducing the computational load that the system would need to support.

---

1 Amazon Web Services. http://aws.amazon.com/.

9.2.2. Glyph as a Desktop Application

Building a version of Glyph as a desktop application could solve many of the challenges faced by the web application. As a desktop application, users could perform very computationally expensive transformations or detections on content without burdening a remote server. Users could also process their own footage on the tool without uploading it online. Thus, whereas the cross-platform ease of a web application of Glyph might have wider availability to citizen journalists and bloggers, a desktop version of Glyph could be more useful to news rooms and professional visual journalists.

9.2.3. Exploring the Integration of Glyphs on a Variety of Interfaces

The usefulness of Glyph is scaled to the inventiveness of its users[2]. As discussed in Section 7, future investigations can explore how glyphs can exist on the web and in the world.

---

[2] In this way, Glyph echoes the research mission of the Viral Communications group at the MIT Media Lab, "to create scalable technologies that evolve with user inventiveness."

Appendix A: Example of a metadata file stored for a processed video


```
{
upload_date: null,
extractor: "youtube",
height: 720,
thumbnail: "https://i.ytimg.com/vi/eCG0M6T5sMU/
maxresdefault.jpg",
like_count: null,
duration: 379,
fulltitle: "Silk Pavillion",
id: "eCG0M6T5sMU",
view_count: 2917,
playlist: null,
title: "Silk Pavillion",
_filename: "Silk Pavillion-eCG0M6T5sMU.mp4",
playlist_index: null,
dislike_count: null,
width: 1280,
age_limit: 0,
annotations: null,
webpage_url_basename: "watch",
display_id: "eCG0M6T5sMU",
description: "The Silk Pavilion explores the relationship
between digital and biological fabrication on product and
architectural scales.",
format: "22 - 1280x720",
uploader: "Swedbrand",
format_id: "22",
uploader_id: "Swedbrand",
subtitles: null,
```

```
stitle: "Silk Pavillion",
url: "https://r8---sn-ab5l6nel.googlevideo.com/
videoplayback?
fexp=900720%2C907263%2C933112%2C934954%2C9406307%2C9406848%
2C9406922%2C9408101%2C9408305%2C9408616%2C947243%2C947255%2
C948124%2C948703%2C948822%2C951703%2C952612%2C957201%2C9614
04%2C961406%2C964740%2C966201&mt=1428254912&ratebypass=yes&
sver=3&expire=1428276633&itag=22&pl=17&ipbits=0&dur=378.136
&ip=65.96.163.201&mm=31&ms=au&mv=m&upn=zg_eXP8xUQ4&source=y
outube&sparams=dur%2Cid%2Cinitcwndbps%2Cip%2Cipbits%2Citag
%2Cmm%2Cms%2Cmv%2Cpl%2Cratebypass%2Crequiressl%2Csource
%2Cupn%2Cexpire&id=o-
AKEzb5TUHCdq5O63Fe6nHhmFYxs1q9ww7ClQX_qczNGK&key=yt5&requir
essl=yes&signature=8F5C6808567C3EFCED9766AF9FEA50103D877544
.
28D7A08FE10E307E11D37E7FDF553DE01C679CC6&initcwndbps=155000
0",
extractor_key: "Youtube",
player_url: null,
ext: "mp4",
webpage_url: "https://www.youtube.com/watch?v=eCG0M6T5sMU",
formats: [...]
}
```

Appendix B: List of Videos Used in Evaluations

Hidalgo, Caesar and Niles, Savannah. "Video", April 14, 2015. https://www.youtube.-com/watch?v=UKY3scPIMd8&feature=youtube_gdata_player.

Unifrance. "Stray Dogs / Les Chiens errants (2014) - Trailer English Subs", March 24, 2014. https://www.youtube.com/watch?v=2Yd11yViVU0.

Playful Systems Group, MIT Media Lab. "20 Day Stranger", May 13, 2014. https://www.youtube.com/watch?v=3a6c7hHbphA.

Jacobson, Joseph M. and Comiskey, Barrett. "E-book display", May 29, 2013. https://www.youtube.com/watch?v=5Q7lX4YLYWA.

Roy, Deb. "Big data for social impact", Mar 24, 2015. https://www.youtube.com/watch?v=9_XRCakyHt8.

TED. "Hugh Herr: The new bionics that let us run, climb and dance", Mar 28, 2014. https://www.youtube.com/watch?v=CDsNZJTWw0w.

Lejeune, Lorrie. "MIT Center for Civic Media - 10/10/2013", Oct 10, 2013. https://www.youtube.com/watch?v=ChtRx3KdoPA.

Mediated Matter Group, MIT Media Lab. "Silk Pavillion", Jul 5, 2013. https://www.y-outube.com/watch?v=eCG0M6T5sMU.

TEDx Talks. "Being there -- sensor data to sensory superpowers: Gershon Dublon at TEDxWarwick 2014", Apr 18, 2014. https://www.youtube.com/watch?v=f2j5cLayLgM.

Into the Gloss. "How To: Highlight For Day & Night", Feb 4, 2015. https://www.y-outube.com/watch?v=gdPpJ11kg04.

ABC News. "US-Cuba Relations: Stage Set for Historic Meeting Between Obama and Castro", Mar 31, 2015. https://www.youtube.com/watch?v=GhhUTVOx33c.

Yang, Hyesoo. "Back to the Desktop 2014 Spring", Apr 21, 2014. https://www.youtube.com/watch?v=GZGZSDzuDxM.

TED. "Ed Boyden: Vitamins for brain function", Jun 25, 2012. https://www.youtube.com/watch?v=hdjauvoc5KQ.

Team Coco. "'Conan In Cuba' Open - CONAN on TBS", Mar 4, 2015. https://www.youtube.com/watch?v=hjfogiltO80.

EnGadget. "MIT Media Lab's Smarter Objects eyes-on", April 29, 2013. https://www.youtube.com/watch?v=hkqXpRxDpfU.

MIT Media Lab. "Mood Meter", Nov 6, 2013. https://www.youtube.com/watch?v=i7PwXjJujt8.

BBC News. "Nigeria elections: Displaced face challenges to cast vote", March 28 2015. https://www.youtube.com/watch?v=iaWCcgz5sNs.

Ira Winder. "City Planning Technology Invented by MIT - CityScope Mark II", February 11, 2014. https://www.youtube.com/watch?v=JEH01UvjzD0.

Object Based Media Group, MIT Media Lab. Infinity-By-Nine", Jun 4, 2012. https://www.youtube.com/watch?v=koLOyFbqFDU.

The New York Times. "Police Shooting Tests New Wisconsin Law", Mar 19, 2015. https://www.youtube.com/watch?v=m7cO-lyaFDw.

Lifelong Kindergarten Group, MIT Media Lab. "Scratch Jr", Mar 18, 2014. https://www.youtube.com/watch?v=mXbOMQ-0WWU.

Tangible Media Group, MIT Media Lab. "Transform: Nature and Machine", Jul 7, 2014. https://www.youtube.com/watch?v=NYW763HaLQE.

BBC News. "US & Cuba: A turbulent relationship", Dec 17, 2104. https://www.youtube.com/watch?v=POUrs_Aoe4s.

TEDx Talks. "Updates from our future city | Kent Larson, Ryan Chin, Caleb Harper and Ira Winder", Nov 5, 2104. https://www.youtube.com/watch?v=pUum3OQul24.

Design Fiction Group, MIT Media Lab. "I Wanna Deliver a Dolphin", Feb 20, 2015. https://www.youtube.com/watch?v=PV7sYqocxGw.

Macro Connections Group, MIT Media Lab. "Immersion", July 3, 2013. https://www.youtube.com/watch?v=QNXh8zQf_B8.

Lifelong Kindergarten Group, MIT Media Lab. "Makey-Makey", Nov 6, 2013. https://www.youtube.com/watch?v=TY8Ky7a7ijw.

The New York Times. "In Yemen, Where Are the Women? Bonus Scene | Times Dispatched", Mar 23, 2015. https://www.youtube.com/watch?v=U4E0pri_t0s.

Wired. "Huggable Robot Befriends Girl in Hospital", Mar 30, 2015. https://www.youtube.com/watch?v=UaRCCA2rRR0.

TED. "Ramesh Raskar: Imaging at a trillion frames per second", https://www.youtube.com/watch?v=Y_9vd4HWlVA.

The New York Times. "TimesCast - March 28, 2012", Mar 28, 2012. https://www.youtube.com/watch?v=Z0RIJyk2KRU.

The New York Times. "Obama State of the Union 2015 Address: President Discusses Next Steps With Cuba", Jan 20, 2015. https://www.youtube.com/watch?v=ZS3GQhmLTis.

Opera of the Future. MIT Media Lab. "Death and the Powers", Sep 10, 2010. https://www.youtube.com/watch?v=zzLaAsyPEc8.

Appendix C: Animated GIFs in Visual Journalism: State of the Art + Best Practices

By Savannah Niles and Audrey Cerdan

News is increasingly visual, and increasingly visually-responsive to a wide-range of platforms on the web and in our connected environments. Headlines are atomized to high-velocity, highly transmissible units on social feeds, but in-depth and evocative analyses of the news and features lose their power by becoming "bite-sized". Images and moving images are thus crucial access-points for users to engage with richer stories. Video is especially suited to render the nuance and humanity of a story, grounded as it is in real places and centering attention on faces, voices, and first-person experience. But video carries a cost in time, attention, and bandwidth. These resources are limited on social networks, browser tabs, and mobile and wearable devices, where we interface with news intermittently and in the context of multitasking. Visual journalists are increasingly required to produce material suitable for the ubiquitous and diverse display surfaces in our lives. They're asked to invent "responsive images" that retain their meaningfulness both at high-resolutions and as bits in the second-order information streams that occupy the peripheries of our attention.

In response, visual journalists are pushing their work beyond and between video and still images, using GIFs and GIF-like looped images as a medium that preserves the engagingness of video with the cogency and transmissibility of still images.

GIFs are designed for immediacy: they're short, looped, silent, and utilize the power of suggestion. GIFs are usually created from a number of sequential video frames to create a looped excerpt independent of the source. But as GIFs become increasingly important as a viral form of media, creators derive GIFs from video frames with greater creative license, manipulating the frames of a clip with precision to extract the semantics of a video, the possibilities of an image, or the affective qualities of visual material.

Much of this experimentation is happening in newsrooms and on front pages. The expressive capacity and transmissibility of animated GIFs increasingly situates them in features and alongside headlines. Visual journalists are beginning to shoot for GIFs in

the field. The GIF as a medium is outgrowing its retro, low-res adolescence and is coming of age in a journalistic and civic context.

This is a short guide on using GIFs in visual journalism.

Types of GIFs

A GIF is a precise moving image that extracts and prolongs one key gesture or a particular affective quality from a video. Historically GIFs were a pragmatic schema for volleying lightweight video clips on image boards and around the dial-up web. But in an age of copious bandwidth and streaming, high-definition video, the continuously-looping image sequences of GIFs thrive as the visual vernacular of the internet. GIFs of course remain a crucial vector for memes and viral media. But in recent years, the GIF has undergone a renaissance[1,2] of creative and technological development. Below, we describe a few classes of GIFs that we've also observed utilized in journalism.

"A moving story compressed to its most essential form"

Publishing bodies born on the web like *Buzzfeed* and (another example here) have embraced GIFs as a means of splintering web video into its funniest or most evocative excerpts to craft rich coverage of an ongoing event. Andrew Phelps writing for Neiman Lab describes much of the coverage of the 2012 Winter Olympics as typifying this use of GIFs and marking a point for the medium in its coming-of-age as a strategic journalistic tool[3]. He examines this coverage to illustrate how GIFs are uniquely capable of excerpting a a video to add explanatory, emotional, or entertainment value to a story. He concludes that ultimately GIFs so successful at this task because they're concise and self-contained: "A GIF is a moving story compressed to its most essential form."

---

1 Friedman, Ann. "What Journalists Need to Know about Animated GIFs – Really," August 8, 2012. http://www.poynter.org/how-tos/writing/183802/what-journalists-need-to-know-about-animated-gifs-really/.

2 Weiner, Jonah. "Christina Hendricks on an Endless Loop." Slate, October 14, 2010. http://www.slate.com/articles/arts/culturebox/2010/10/christina_hendricks_on_an_endless_loop.html.

3 Phelps, Andrew. "The 2012 Summer Olympics Are Turning into a Giant Coming-out Party for the Animated GIF." NiemanLab, August 3, 2012. http://www.niemanlab.org/2012/08/the-2012-summer-olympics-are-turning-into-a-giant-coming-out-party-for-the-animated-gif/.

Perfect Loops

New typologies of GIFs have emerged in recent years embracing the endless looping that distinguishes GIFs from other video compression schema. Communities on the web have formed around "perfect loops"[4], a kind of GIF that loops seamlessly, with a smooth, continuous animation throughout the sequence. Erasing the often jarring jump-cut of the GIF's "seam" (the transition from the last frame back to the first frames as the GIF replays itself) yields a self-referential, editorialized, polished image that denies the glitch-aesthetic of the GIF's history. The effect of perfect loops can sometimes mimic natural movement in the real world like an artificial live stream (see the next section). But often their effect is stylized and memorizing– a perpetual motion machine that defies our internalized understanding of dynamics in the physical world.

This class of GIFs is often deployed in online features and editorial to add emotional depth or evoke empathy from readers. For example, a story published in *Medium*'s "Matter" column on the Ebola outbreak in Liberia was organized around a series of seamlessly looping GIFs of individuals impacted by the outbreak. These evocative, arresting portraits offer more than a still image could through their mesmerizing dynamism. Many of the images simulate a shared gaze between the subject and the viewer, or add 3-dimensionality to the pictured environment. These effects are achieved more immediately and compressively than they could be in an embedded video.

Atmospheric Imagery

"Perfect loops" have a toned-down subspecies in the form of the sometimes long, sometimes seamless clips that illustrate the settings of a story of set its tone. These images often accompany stories as dynamic headers. Increasingly, established publications like *The New York Times* and *The Boston Globe* experiment with this kind of looped image. In 2012, *The New York Times* famously published the Pulitzer-prize winning feature "Snow Fall", which opens with a slowly looping image of windblown snow[5]. Critics praised *The Times*' innovative use of looping moving images as inject-

---

4 "r/PerfectLoops". Reddit. http://www.reddit.com/r/perfectloops/.

5 Branch, John. "Snow Fall: The Avalanche at Tunnel Creek." The New York Times, December 2012. http://www.nytimes.com/projects/2012/snow-fall/.

ing "a subtle, atmospheric quality"[6] into the feature, achieving evocative and cinematic depth[7]. The same year, *The Times* did a series called Still Life, in which writers described their favorite summer settings around New York. A series of subtle, contemplative moving scenes accompanied and sentimentalized each story. In the last two years, *The Times* has continued to use gifs, and especially looping gifs, in feature stories and alongside headlines, and a growing number of online publications have followed.

"Moment Imagery"

An increasingly visible subset of seamlessly looping GIFs, called *moment imagery*[8] or *cinemagraphs*[9,10,11], emphasizes the GIF's place somewhere between an image and a video. These images juxtapose static and dynamic elements in a continuously-looping image, resembling a subtly-dynamic picture, like a digital lenticular postcard, more than a video clip. The images' characteristic atemporality seems particularly valuable for enhancing the preciousness of a moment– they're often used in fashion photograph and created from the scenes of beloved movies. This image is growing in popularity, and will likely begin to appear in new contexts and alongside news stories.

---

6 Duenes, Steve. Erin Kissane, Andrew Kueneman, Jacky Myint, Graham Roberts, Catherine Spangler. "How We Made SnowFall," Open Source, January 30, 2013. https://source.opennews.org/en-US/articles/how-we-made-snow-fall/.

7 Sonderman, Jeff. "How The New York Times' 'Snow Fall' Project Unifies Text, Multimedia," Poynter. December 20, 2012. http://www.poynter.org/mediawire/top- stories/198970/how-the-new-york-times-snow-fall-project-unifies-text-multimedia/.

8 Tompkin, J., F. Pece, Kartic Subr, and J. Kautz. "Towards Moment Imagery: Automatic Cinemagraphs." In Visual Media Production (CVMP), 2011 Conference for, 87–93, 2011. doi:10.1109/CVMP.2011.16.

9 Flock, Elizabeth. "Cinemagraphs: What It Looks like When a Photo Moves." The Washington Post - Blogs, July 13, 2011. http://www.washingtonpost.com/blogs/blogpost/post/cinemagraphs-what-it-looks-like-when-a-photo-moves/2011/07/08/glQAONez3H_blog.html.

10 r/Cinemagraphs. Reddit. http://www.reddit.com/r/cinemagraphs/.

11 Cinemagraphs on Tumblr. https://www.tumblr.com/search/cinemagraphs.

Considerations in designing gifs:

Time

More so than video, GIFs vary in their relationship to time. A GIF might be created to reduce the time it takes to watch a key moment from a video, representing the moment immediately and repeatedly. Some GIFs may prolong a moment in time indefinitely, or through their looping sequence have a mesmerizing power over viewers. The artist and critic Sally McCay characterized GIFs as images of self-enclosed affect, intensified by the intimate, tight viewing distance between eyes and the monitor, and argues that GIFs' mesmerizing loops sets them apart from our lived temporality[12]. But some GIFs, through their continuous dynamism and low resolution, borrow the visual cues of livestreams, and heighten the sense of ongoingness in what they represent.

Emotion and Empathy

Especially with the visual media of the news, looping images can be powerfully evocative. If some GIFs do have a mesmerizing power over us, and hold our attention for longer than a still image, they might allow us more time to invest emotion and empathy in the subject of the image. A GIF can enliven a scene with dynamism and enhance the sense of simultaneity between a viewers experience and what's depicted. GIFs can be very evocative, but they can also over-sentimentalize, lingering on a moment or stylizing dynamism to the point of kitschiness[13].

Attention

GIFs are moving images scaled to the fleeting attention that characterizes online reading behavior. GIFs are more immediate and compressive than embedded videos. In this way, GIFs and looped autoplaying video risk distracting readers from accompanying text. Eye-catching and bite-size, GIFs support visually skimming an article, which

---

12 McKay, Sally. "The Affect of Animated GIFs (Tom Moody, Petra Cortright, Lorna Mills)." Art & Education, 2009. http://www.artandeducation.net/paper/the-affect-of-animated-gifs-tom-moody-petra-cortright-lorna-mills/.

13 Moody, Tom. "Hair GIFs and the Male Gaze at Tom Moody." *Blog of New York Artist and Musician Tom Moody*, March 17, 2012. https://www.tommoody.us/archives/2012/03/17/hair-gifs-and-the-male-gaze/.

may or may not be ideal. But GIFs can also meter distraction through their limited and repetitive dynamism. In this way, GIFs offer a promising paradigm for broadcasting moving images on ambient and wearable displays, where attention is even scarcer than on news and social feeds.

Authorship

As derivations from video, GIFs often create issues in authorship and attribution.

Newsrooms with visual journalists on staff such as *The New York Times* and *Medium* have the option of creating GIFs and looping video from their own content, or even shooting content specifically for GIFs in the field, avoiding the risk of copyright violation or plagiarism. But citizen journalists, bloggers, and activists usually do not have access to original content, and instead excerpt and re-contextualize content from another source.

GIFs embody the culture of exchange, remixing, and collective authorship that characterizes visual media on the web. Conventional practice online treats embedded web video as fair game for creating and sharing GIFs. GIF creators have historically embraced the appropriation of other content, and remain free to do so, as GIFs are highly transmissible images that erase their sources and are difficult to take down (unlike a YouTube video, for example, which can be removed with a DMCA notice[14]). Because GIFs are often short excerpts or parodies of their source, GIFs often do not violate intellectual property law in many countries. And because their creators are hard to track and often not profiting from the images, few copyright violations involving GIFs have been brought to courts. This could change in the future, as GIFs become an increasingly powerful medium in the attention economy. In the 2014 FIFA World Cup, for example, a blogger creating GIFs of goals in near-realtime received a cease and desist[15].

---

14 Phelps, Andrew. "The 2012 Summer Olympics Are Turning into a Giant Coming-out Party for the Animated GIF." NiemanLab, August 3, 2012. http://www.niemanlab.org/2012/08/the-2012-summer-olympics-are-turning-into-a-giant-coming-out-party-for-the-animated-gif/.

15 Brouze, Emilie. "Copyright: Could You Be Attacked for an Animated GIF?" *Rue89*, April 2, 2015.

Trust

GIFs carry associations that might undermine a publication's brand and values, or threaten a reader's sense of its journalistic integrity. GIFs proliferate in online media as click-bait, hot takes, entertainment, appropriation, and the meme-culture of Reddit, 4chan, and other image boards: introducing GIFs in serious reporting and editorial implicitly requires addressing the heritage of the medium.

GIFs, DIY

Shooting for GIFs in the field

Shooting specifically for GIFs in the field will likely involve stabilizing footage either with a tripod or in post-production, and a thoughtful consideration of the dynamism of the footage. A clip with too little movement might make a better photograph, and a clip with too much movement may as well be a video. Shooting for GIFs often means shooting a still scene with just one or a few elements of movement in the frame.

Processing GIFs from Video

Many online tools allow the creation of GIFs from web video. The advantages of using an online tool to create GIFs are the portable, cross-platform functionality and speed/ease of use. Many online tools allow a user to enter a YouTube URL, choose a start and an end time, and export an animated GIF[16]. Sites like GIFsoup and the Chrome plug-in Jiffy allow for building GIFs quickly by simply entering a YouTube URL. MakeAGif support building GIFs from YouTube videos as well as uploading video or images, or creating a GIF from a webcam. The speed and convenience of online tools often involve a tradeoff in quality of output and flexibility of editing, but an increasing number of services are beginning to offer higher-quality GIFs and advanced effects. MemeCenter supports cropping, size-adjustments, and adding text to images, and includes a timeline editor to build a GIF from non-consecutive segments in a video. GIFpal, a swiss-army-knife of GIF-makers, offers many additional advanced editing options,

---

16 Moynihan, Tim. "How Imgur's New GIF-Maker Stacks Up Against Other Tools Out There." *Wired*, January 30, 2015. http://www.wired.com/2015/01/gif-tools/.

such as color and playback adjustments. Recently, sites that host GIFs, like Imgur [17], Giphy[18], and Gfycat, as well as YouTube[19], have rolled out their own online GIF-making tools.

For many visual journalists, online GIF-making tools don't offer the bandwidth or the editing tools necessary to make high-quality GIFs; or, editing lots of high-quality footage is not feasible with an online service. Desktop software is more appropriate for creating GIFs in these cases. Some software are designed specifically for building GIFs. Giffing[20] is a tool for Windows operating systems for creating GIFs with advanced effects, like cropping, compositing clip sequences, and reversing playback. Glyph[21] is a Mac-based tool developed at the Media Lab at MIT for creating expressive, seamlessly-looping GIFs from video. Glyph allows a user to automatically detect perfect loops that occur in a video, create the appearance of seamless motion in a non-looping clip, still some regions of movement in a clip to highlight others, or imbue a still frame with subtle dynamism. Open-source image manipulation software like Gimp and subscription-based suites like Adobe Photoshop, Premier, and AfterEffects also support frame-by-frame manipulations to GIFs and scripted editing of image sequences[22]. These applications support the more flexibility– allowing an editor to build complicated masks to still some regions of motion in the clip, or invent new frame transitions through detailed custom editing.

Compressing GIFs for the Web

Though creating GIFs with these tools and performing sophisticated frame adjustments has a steep learning curve, for visual journalists already skilled in these industry-standard applications, they may offer the best toolkit. But because of the GIF's specific compression schema, creating and editing GIFs optimized for quality and load time requires more strategy than creating and editing video sequences and photos.

---

17 Schaaf, Sarah. "Introducing Video to GIF." *Imgur Blog*, January 29, 2015. http://imgur.com/blog/2015/01/29/introducing-video-to-gif/.

18 Giffffr by Giphy. http://blog.giffff.fr/.

19 Kumparak, Greg. "YouTube Gets A Built-In GIF Creator." TechCrunch. Accessed May 4, 2015. http://social.techcrunch.com/2014/12/11/youtube-gif-maker/.

20 Giffing. http://www.giffingtool.com/.

21 Niles, Savannah. Glyph. glyph.media.mit.edu.

22 Adobe. http://www.adobe.com/.

First of all, GIFs aren't always the answer to getting looping images onto a webpage. Sometimes, other file formats might be appropriate. The same sequence of image frames could be saved as a .mp4 file and as a .gif file, but the two clips could look very different in their quality, colorfulness, or playback speed. Though the content of these clips would be identical, each file would store data with a different *compression schema*. Compression refers to a file's tradeoff between image quality and file size. Creating GIFs from video results in a noticeable loss of image quality, because the GIF video compression schema specifies that each frame of the animation reference only 256 colors. Because of the inherent limitation of quality that GIFs entail, GIFs may not always be the best means of including looping videos in a news story. What GIFs have going for them is they're highly transmissible: they can be embedded as images in browsers and thus shared easily across many platforms:

```
<img src="beautiful-loop.gif">
```

For large and/or high-quality content, an autoplaying, looped embedded video player with hidden controls might be more appropriate. Embedded .mp4s can offer beautiful, more colorful, high-resolution looping videos, but at the cost of a less cross-platform, less viral embedded video player:

```
<video width="320" height="240" controls="false" autoplay
loop>
  <source src="beautiful-loop.mp4" type="video/mp4" />
  <source src="movie.ogg" type="video/ogg" />
  Your browser does not support the video tag.
</video>
```

In the case where a GIF is the most appropriate file format, strategies can be employed in image-editing software to optimize the image's tradeoff between file size and quality. There is no definitive combination of export settings to optimize any GIF—optimization is often a process of trial and error. But reducing the dimensions of the GIF will always reduce file size without reducing image quality, and is therefore often a good first step. Reducing the number of frames in the GIF will also reduce its size without reducing the image quality or, by adjusting the frame delay, the playback. If these

strategies aren't enough, the next step is to think about reducing color across the images frames. In the image export settings, choosing a *selective* reduction of the image's color table is usually a good starting point: this builds a color table by from colors to which the human eye is most sensitive, favoring broad areas of color and web colors[23]. Though not always possible, try to limit the number of colors in the table to 64 colors. Avoid dithering, especially if your animation has a lot of solid color areas and no gradients. If your clip requires dithering to preserve the integrity of its colors, countering with lossiness might reduce the file size without a noticeable reduction of image quality. Blurring background elements or creating an image with very selective regions of dynamism also leads to lighter-weight images. Finally, export your GIF with interlacing if you'd like it to be loaded by the browser with multiple passes.

The Future of GIF-making?

As GIFs continue to earn increasing centrality to visual culture, research is emerging in the automatic creation of beautiful and/or looping GIFs. Open source software, including MoviePy[24] and LoopFindr[25], implement algorithms for automatically detecting perfect loops in video clips. More advanced research at Microsoft Research involves creating perfectly looping clips by identifying the independent looping periods of dynamic segments in the clip, and compositing these asynchronous segments at user-specified levels of dynamism[26] (see figure 3-3). Research has also been done in compositing static and and dynamic imagery to create cinemagraph-like "moment imagery"[27,28]. These research initiatives and others suggest that more sophisticated GIF-

---

23 "10 Ways to Optimize an Animated GIF File - Tuts+ Design & Illustration Tutorial." Design & Illustration Tuts+. Accessed May 13, 2015. http://design.tutsplus.com/tutorials/10-ways-to-optimize-an-animated-gif-file--psd-34649.

24 Zulko. "An Algorithm to Extract Looping GIFs From Videos". __del__(self) / Eaten by Python. 2015 Feb 1. http://zulko.github.io/blog/2015/02/01/extracting-perfectly-looping-gifs-from-videos-with-python-and-moviepy/.

25 Burger, Colin. LoopFindr. http://loopfindr.tumblr.com/.

26 iao, Zicheng, Neel Joshi, and Hugues Hoppe. "Automated Video Looping with Progressive Dynamism." ACM Trans. Graph. 32, no. 4 (July 2013): 77:1–77:10. doi:10.1145/2461912.2461950.

27 Tompkin, J., F. Pece, Kartic Subr, and J. Kautz. "Towards Moment Imagery: Automatic Cinemagraphs." In Visual Media Production (CVMP), 2011 Conference for, 87–93, 2011. doi:10.1109/CVMP.2011.16.

28 Joshi, Neel, Sisil Mehta, Steven Drucker, Eric Stollnitz, Hugues Hoppe, Matt Uyttendaele, and Michael Cohen. "Cliplets: Juxtaposing Still and Dynamic Imagery." In Proceedings of the 25th An-

making tools are on the horizon, and may soon be available as commercial or open-source software for visual journalists and designers.

Summary: Best Practices

The following tips summarize some of the best practices for using GIFs in visual journalism:

- GIFs can be specific, summarizing a story or a key moment in a video, or generalized, performing tone-setting or affective functions.
- GIFs can be powerfully evocative images that elicit emotional, empathetic responses from viewers.
- Consider how your GIF relates to the time and attention of your audience.
- Consider how the provenance of your footage relates to your GIF's authorship.
- When shooting for GIFs in the field, shoot stabilized, lightly dynamic footage.
- You may opt for a looped embedded video over a GIF.
- Play with your GIF's export settings. Start with the following options: Selective, No Dither, 256 or fewer colors, No Transparency Dither, Bicubic.

---

nual ACM Symposium on User Interface Software and Technology, 251–60. UIST '12. New York, NY, USA: ACM, 2012. doi:10.1145/2380116.2380149.