Chronographer: Remixing Segmented Video in Interactive Spaces

Seth Hunter, Eric Rosenbaum
MIT Media Lab
hunters@media.mit.edu

Charlie DeTar, Jeff Lieberman
MIT Media Lab
ericr@media.mit.edu

ABSTRACT
In this paper we present an approach to manipulating time-based content in interactive installations. The *Chronographer* initiative utilizes computer vision techniques to segment foreground objects into independent video streams, which can be combined with other video streams or used to augment live video content with imagery from the past. To illustrate the generative possibilities enabled by this approach we present the design and implementation of three installations developed at different times in the same building by the authors. Our contribution lies in outlining the software techniques used in all three projects and illustrating possibilities enabled by each of the design processes and installation configurations. We conclude with an outline of features and challenges to make these techniques more accessible to interactive designers.

Author Keywords
Segmentation, chronographer, time remixing, video editing, interactive installation, design, temporal media

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
This document outlines and evaluates an approach to designing time based media installations that segment foreground elements into individual time streams that can be manipulated independently. We outline and compare three installations (Figure 1) that use implementations of these software techniques from an initiative we have nicknamed "Chronographer".

The motivation for the initiative started in a discussion with the artist Jeff Lieberman [11], who explained, "In my work in photo and video, especially the high speed video images for the Discover Channel Show *Time Warp* [26], the need for such a set of tools became immediately evident. I keep wishing I could separate subjects from backgrounds and each individual subject from time - in essence, taking an input and fundamentally separating the foreground and background entities to enable time warping of separate subjects relative their movement across the background."

Figure 1. **The Chronographer installations (A,B,C) presented.**

Photo and video software currently provide linear transformation techniques but do not allow the editor to specify their own mapping parameters across multiple frames. By supporting time remapping, an entirely new set of possibilities for creative composition would be available to the general public, widening the playing field of what is possible.

The *Chronographer* initiative fundamentally changed how we approach video as a creative medium. We shifted from thinking of individual frames within a video to individual segmented elements over time relative to each other. Our initial experiments in time remixing began with static images and progressed to layered images, segmented video clips, and reconstructed video streams. Our hope is that by documenting the installations, sharing techniques, and providing a public code repository, others will be inspired to take a similar approach and contribute to the development of the tools and techniques presented.

RELATED WORK
The *Chronographer* project is a conceptual extension of techniques used in film and photography. These are applied to interactive installations that are enabled by computer vision algorithms. Techniques and projects from these three fields inform our approach and provide a foundation for the design of the installations.

Film and Photography
Traditional photography includes many evocative time warping techniques, which influenced our motivation to develop *Chronographer*. Slit scanning involves taking a single row or column of video and stacking them horizontally to create a still image. This historically arose from film cameras' use of an actual slit scanning in front of the film during an exposure. Time-lapse photography

DIS 2012, June 11-15, 2012, Newcastle, UK.

accelerates the perception of time by recording frames at regular but less frequent intervals.

Figure 2. **Slit Scan (Robert Doisneauvia), Time Lapse (Andrew Stawarz), and Photo Averaging (Jason Salvaon)**

More recently artists such as Jason Salavon [21] have used mean averaging of multiple images to create atmospheric representations of photographs in similar contexts over decades of similar images.

Artistic Installations

Computer vision has opened new artistic possibilities for interaction designers to experiment in the context of site-specific responsive installations.

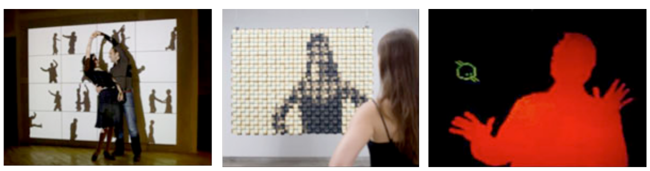

Figure 3. **Scott Snibbe's Deep Walls, Daniel Rozin's mirrors, and Myron Kruger's Artifical Reality Installations**

Artists such as David Rokeby [19], Daniel Rozin [20], Myron Kruger [10], Camille Utterback [27], Lincoln Shatz [25] and Zack Lieberman [12] have presented work in which a representation of the user is projected back into a "magic mirror" where some aspect of the user is transformed or re-contextualized. Some of these designers [24, 25] have used recordings of people to evoke moments from the past, but did not juxtapose the images with live video footage in order to construct fictional histories or create the illusion of historic and live footage occurring simultaneously.

Computer Science

Separating foreground objects into separate streams has many applications in computer science. Surveillance applications [3] and time based media analysis [11] combine multiple camera feeds and condense time-based media. Approaches to overcoming difficulties in tracking due to abrupt motion, changing appearance patterns in the object and scene, and object - object occlusions have been evaluated extensively in computer vision and pattern recognition [29].

The Reeding People Tracker [23], Reflexion [30] by Stephen Agamanolis, and MERL Pedestrian Detetion [9], are all examples of computer vision applications that extract video of people for post processing and analysis with surveillance systems.

Figure 4. **Human Speechome Project Time Stream**

The Human Speechome Project [5] extracted data in order to study the relationship between space and speech acquisition in childhood development. Figure 4 shows extracted contours of the occupant movements over a period of minutes in the fish eye view on the left. Many of the algorithms used in these projects are documented but implementations are not available for general use.

DESIGN PROCESS AND INSTALLATIONS

Our process has been a fairly non-linear one in which occasionally we would revisit old ideas about time remixing in a new context without immediately seeing the connection to previous work. The software has evolved through an iterative cycle of inspiration, experimentation, reflection, and revisal that is rooted in a playful approach to time manipulation. We encourage readers to view the videos [1] of the projects for context.

For example, *Stillness Clock, Motion Clock*, (Figure 5) explores the perception of time, in motion and in stillness. We used a slit-scanning technique to paint live video around a circle, mapping time onto space. The clock on the left responds to the motion of the viewer, leaving a jagged trace of activity over time. The clock on the right is active only when the scene is still, encouraging the viewer to watch in contemplative silence as their image slowly emerges.

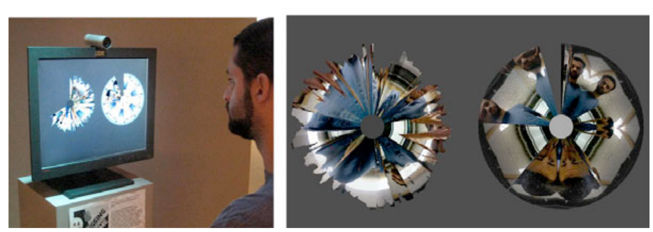

Figure 5. **Stillness clock, Motion clock, 2009**

Representations of time can also be quickly experimented with by pointing a camera at a display with a view of the camera feed. The result is an analog loop of time-delayed images that are evocative but difficult to manipulate. We added software controls in two different settings that demonstrate what happens when you take this effect and manipulate the orientation and delay parameters.

Figure 6. **Feedback simulations during prototyping, 2011**

The explorations (Figure 6) introduce a delay that is inherently interesting. On a conceptual level the viewers can virtually interact with themselves. By extension, viewers can respond to the virtual representations of previous viewers.

Additional explorations that informed the *Chronographer* initiative were extended from time remixing work with commercial tools like Photoshop and After Effects.

Figure 7. **Order from chaos, and events that never occurred.**

The compositions in Figure 7 illustrate time warping concepts such as selectively ordering foreground figures and constructing hypothetical events. The marathon on the left was constructed from hundreds of images of runners and ordered by the color of their jerseys. The image on the right of people at the party never happened, individuals posed one by one over the course of the evening. The results were evocative and promising but the process was time intensive and did not scale well for large datasets or interactive installations.

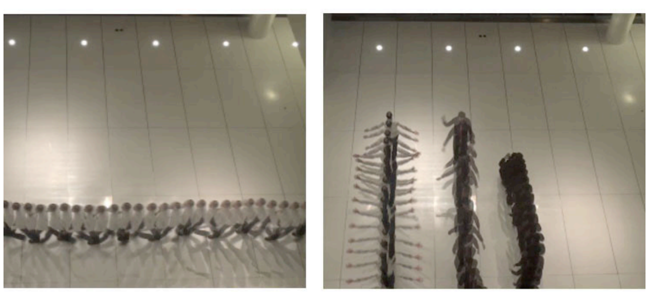

Figure 8. **History trails filmed from multiple frames**

Proof of concept for the first installation began with a series of videos of the authors improvising in the building lobby as if their bodies could leave historic trails through the space. Figure 8 are frames from a composite video that demonstrates what this might look like if visualized in real time from a window with a view from above. Video documentation [1] all installations are available online.

Installation A) Balcony Touch Screen, *Slow Glass*

The first interactive application in the *Chrongrapher* series was positioned in the fifth floor of the building overlooking the third floor. The screen was mounted on the lip of the balcony, providing a view of the space below in the same line of sight as the monitor.

Figure 9. **The pedestal and a screenshot of the atrium in Installation A, slow glass.**

The effect can be described as slow glass, an ongoing view of accumulated history of the space from above, an "eagles perspective" of the atrium below. RGBA png files were mapped to a database using a grid roughly the size of foreground objects to prevent overlap of the foreground objects. These png files were drawn on top of the live video feed. The edges of foreground objects were blurred sufficiently to appear to exist in the same space.

Due to light changes over the course of the day, the images mapped to each grid square corresponding to when they were recorded so that they were within an hour of the current time. The piece was continually adding new foreground objects. Touching the screen triggered a grid object to refresh, but the resolution of the grid was not high enough to convey the sense of drawing the past. Most viewers would glance at the screen and then at the space below and compare views. Throughout the day viewers would glance in passing at the evolving view on the screen as an ambient display.

Installation B: LCD Wall Mirror, *(now (now (now)))*

The second application was built a few months later and mounted on the wall in the public café on the fifth floor. The piece is a "magic mirror" that echoes video recordings of viewers in an exponentially delayed feedback system. The size and resolution of the video clips are scaled proportionately as they recede in time.

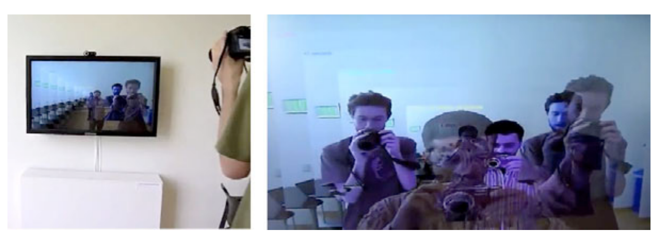

Figure 10. **Installation and screenshot of Installation B, (now(now(now))).**

The authors explain: "*(now(now(now)))* layered foreground video segments from the past over the present background. Video segments from approximately the past 5 seconds, 1 minute, 5 minutes, 1 hour, 1 day, 1 week, and 4 weeks played simultaneously with the present, giving an impression of ghostly images of the past joining with the present. The effect was most striking when the installation was placed at the end of a corridor, so that viewers would see images of people walking down the corridor from several days ago coming up from behind them."

Viewers who paused in front of the monitor would see themselves begin to repeat after a five second duration. Many performed in front of the piece in order to discover the behavior of the system. Occasionally videos from other viewers would play from the distant past while a viewer was performing. The piece did not treat time entirely linearly. It triggered the nearest clips to that layer in time during interstitial moments to increase the perception of this occurrence.

The videos in *(now(now(now)))* were segmented and encoded as five second video clips with magenta encoded in place of the alpha channel to reduce file size and because most video codec with compression do not support an alpha channel. These video clips were linked and replayed using OpenGL shaders in which all magenta pixels are transparent.

Installation C: Full Body Floor Projection, *If the Space Could Speak*

The third application was designed for a floor projection system installed in the atrium of the building on the third floor. The hardware system consists of a four projector software blended image which is 16'x 24' at a resolution of 1024 x 1536. Cameras and IR lights on the ceiling are used to track the position of the viewers and trigger content related to their proximity.

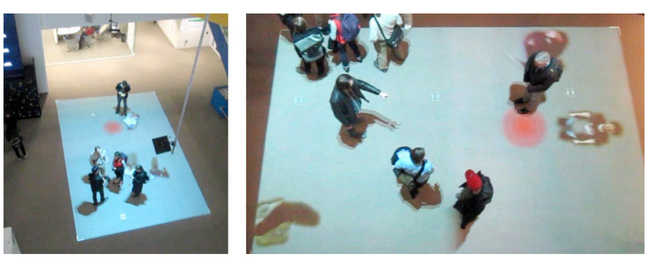

Figure 11. **Overhead views of full body projection with speakers in Installation C, If the Space Could Speak.**

The installation content contains 48 audio interviews that are synced with thirty-second videos of the interviewees walking up to two positions in the space, looking up at the camera, and after 10 seconds exiting the projection area. Audio from the interviews is triggered through ultrasonic speakers with a narrow range hanging 12 feet above the floor above the positions where the videos were recorded.

Each interviewee was asked two open ended questions: "What aspects of this place do you think people miss out on when they visit for one day?" and "How has your relationship with other people been affected by technologically mediated communication?". The recorded answers reflect the attitudes and perspectives of the people who work the building about their work and relationships with others and technology.

The intension of the designers was to convey deeper reflections about the culture that might be apparent at a glance through the video and audio clips from the past. The first instantiation did not explicitly show viewers where to stand to hear audio and many viewers were unfamiliar with the

To compensate we added footsteps, which lead from your position to the red circles where audio can be heard. The circles vary in size based on the FFT of the audio. We also added text animation of the words in the interviews, which spill out of the red circles onto the floor. These steps helped to increase understanding and engagement, and increased the duration of engagement, however the piece seemed be the most successful when viewers were alone because it requires listening and most people do not listen to the audio preferring instead to discuss the piece with others.

SOFTWARE

Although all three installations were developed separately, they use a similar set of functions from OpenCV 2.0 [16]. All applications were written in C++ on Mac OS 10.6 using OpenFrameworks libraries [17] and Quicktime [18]. The third installation was ported to Windows 7 in its final version. Example code is posted online here [2].

The general software processes are outlined in Figure 12. All of the installations model the background, and segment figure from ground. The formats, classification strategies, and playback models differed between installations. The pros and cons of these are outlined in the discussion. This section outlines the approaches used for a general implementation of video stream remixing.

Capturing and Saving Foreground Frames

The instillations obtained images from a Logitech 9000 UVC camera. This camera has hardware settings that allow for control of exposure, white balance, and focus which are essential to maintaining consistent light conditions when combining videos later. To save video we used QuickTime with the Animation codec at a low setting. This reduced the file size sufficiently that the videos could either be preloaded or loaded dynamically without delaying the playback system.

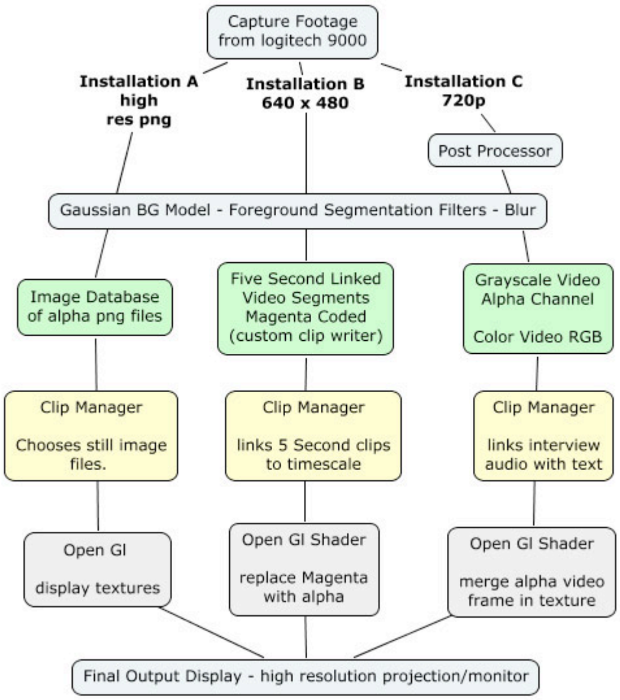

Figure 12. **General software processes from capture to playback**

Background Modeling

Background subtraction is a popular method, due to its simplicity, for motion segmentation, especially in a case like ours where the backgrounds are relatively static. The approaches that were evaluated for our framework were temporal averaging and statistic color modeling.

Temporally averaged image models such as those proposed by Yang and Levine, 2002 [28] construct a background model by taking the media value of the pixel color over a series of images. The median value and a threshold value are determined by using a histogram procedure based on least the median squares method which are used to create a difference image. This algorithm is good for handing gradual inconsistencies due to lighting changes, and is commonly used because an implementation exists in OpenCV.

A more advanced approach to also account for shadows and localized noise in the camera feed in low light conditions is to adopt a statistical method. The statistical approach uses the characteristics of groups of pixels to classify them into foreground or background by comparing current statics to the background model. For example, an adaptive background mixture model [6] models each pixel as a mixture of Gaussians. The gaussian distributions of the adaptive mixture model are evaluated to determine if it is from a background process. This is a more CPU intensive model but still relatively robust. It is better suited for situations with clutter and localized movement.

With most background models letting the application run for a few minutes and setting the model to adapt slowly helps to reduce the chance of slow moving objects being blended into the background model. For our purposes the Gaussian Mixture model produced the best segmentation results.

Figure-Ground Segmentation

In order to segment the foreground elements from the background we used standard methods outlined in Figure 10. These functions are documented in the OpenCV spec [16]. The parameters varied for each function between projects but the order of the filters were very similar.

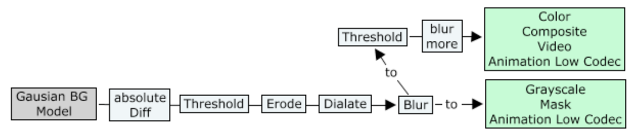

Figure 13. **Post processor filters and output.**

Our algorithm requires that the background image be relatively similar between image capture and the final playback. If you are not choosing images by hand or have a variety of light conditions, consider increasing the blur in the grayscale mask.

Tracking

If multiple foreground objects exist in the same recording a tracking algorithm can be used to differentiate between them. For example in the third installation after finding the contours of objects with an area greater than a minimum threshold, the tracker records the position, velocity, direction, and area of the objects and looks for similarities across frames to keep track of each individual element.

Extensive research has been done in computer vision on more advanced techniques that can be used in cases when blobs overlap or are ambiguous such as Kalman filters [15], Optical Flow [7], Sub-region tracking, and feature tracking such as SIFT and SURF [8]. Although examples of these techniques are available they were not needed for the controlled conditions of our installation. Because they are computationally expensive it might be more beneficial to include these options when post processing outdoor conditions or more complex scenes with multiple people.

Replaying with Transparency

One of the unforeseen problems we encountered when working with video frames was the lack of support for an alpha channel in video compression codecs. Apple's Animation codec supports an alpha channel in its uncompressed setting but the file size for frames can be in excess of a gigabyte per minute of data. In order to load the video without delaying the playback program, file sizes of 10 megabytes or less per minute of video are required.

One implementation strategy to solve which was used in the second installation is to replace the alpha channel with an uncommon color, which can be subtracted when the frame is converted into a texture. The benefits of this approach are a reduced file size, decreased loading times, and perfect playback syncing. However, compression codecs introduce artifacts, especially in shadows, which appear if the filter does not account for a gradient of magenta pixels. An exaggerated version of this problem is evident in Figure 14.

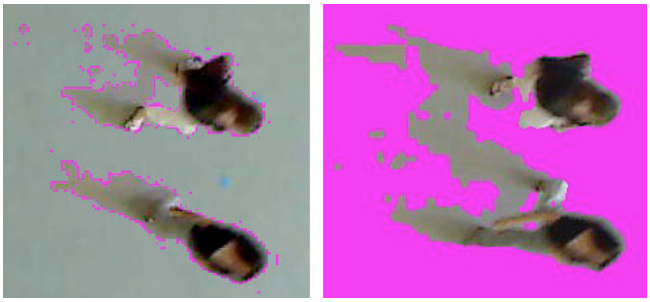

Figure 14. **Artifacts introduced by video compression when Magenta is encoded in place of an alpha channel.**

Writing magenta into the RGB image also prevents a blur from being encoded into the foreground mask. Figure 15 illustrates the second strategy we adopted in order to solve this problem. Saving identically compressed versions of the grayscale mask with a blur, and a simplified version of the RGB image improved the overall aesthetics during playback. Even when images were recorded with very different light conditions and exposure settings, the introduction of a mask helped create the illusion of a blended reality.

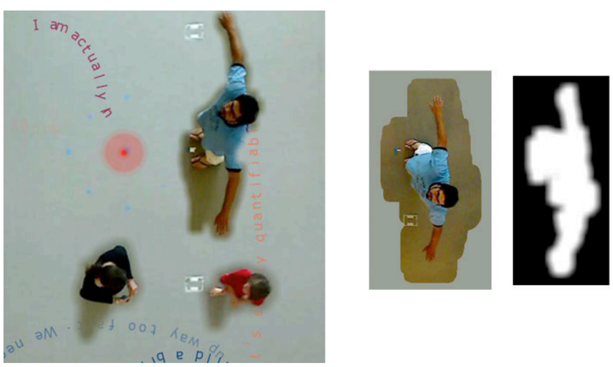

Figure 15. **Example frames from the final composition, a color composite, and a grayscale mask in the third installation.**

A second benefit to using shaders is that you can introduce a color tint or increase the transparency of the foreground video streams to differentiate and blend them together. This technique was used extensively in the installation B. Installation C was projected onto a carpet with low contrast, so we opted to keep the videos at their original saturation.

DISCUSSION

Similar motivations inspired the design processes of the three installations discussed in this paper but the physical layout and user experience varied between the installations. By reflecting on the differences between the installations, user feedback, and software development challenges, we hope to share useful insights with other interaction designers.

Physical Layouts and Viewer Behavior

For our installations the main differences we observed were in user behavior and the perspective point of view. These correlate most directly to two factors: the orientation of the camera to the viewer, and the form factor of the display.

The position, size, and angle of screens in tabletop research [22] elicit different behaviors and contexts in previous studies. The three setups were wall based, and a 45-degree orientation, and floor projection.

The vertical setup was the most successful at eliciting performances from the viewers, many describing it as a time mirror that gives direct feedback. However it's position in the café was not ideal. It was on the right or left of occupants as they passed through the space. The ideal location is opposite a doorway or at the end of a corridor where viewers approach directly.

People responded to the 45-degree balcony setup in the greatest variety of ways, observing at a distance, touching the screen, and peering beyond the screen to the space below. This installation was also the most easily overlooked. Viewers said they had seen it before or didn't notice it as they were walking by because it was integrated into the architecture.

The full body projection is so big that people would either step around the projection, or step in to see how it would react. The expectation for interaction was highest in the mixed reality floor display, as well as confusion regarding how to respond to the videos from the past. Users reported feeling like they were immersed in the installation but that the content was in a parallel space, describing them as "pancake people". Scott Snibbe presents guidelines for designers in large interactive displays [24] that include responsiveness, socially scalability, and using socially familiar content. These suggestions and observation contributed to the design improvements described earlier in the paper.

User Feedback and Key Insights

We conducted informal interviews with occupants of the building as well as videotaping visitors to the building during the duration of the installation. The initial confusion regarding what the installation does was not uncommon, but this confusion is usually followed by an experimental gesture. If this gesture is incorporated into the piece or the piece responded to touch input and movement people would engage for longer durations.

The time remixing ideas behind the project were more evident in installations A and B, which also had title cards with explinations. Installation C did not have a title card but was explained during tours. After this explanation people would listen for an average of 20 seconds. Most users reported that the concepts were very interesting but that they were more interested in the responsiveness of the installation than the content of the interviews.

One exception was a viewer who interpreted the piece by saying: "There is a way of connecting to the institutional memory of a place that is actualized in the physical space. There is nothing quite like the feeling of discovering a space right beneath your feet that you never knew existed that has also been visited by other people that had the curiosity or the insight to look deeper."

When animated text from the interviews in installation C was added on the floor, people could read it from both the fourth and fifth floor balconies. A few users reported enjoying the piece more from a distance because the juxtaposition between the live viewers and the projected content was seamless and they felt like they had a God's eye view of people interacting with piece. One user suggested a hybrid between Installation A and C would allow viewers on the touch screen to puppet content on the floor in response to people in the atrium below.

A few viewers also suggested using actors to construct video streams from the past that would appear to respond directly to actions of viewers. Interactive designers at Disney Research [14] have use animated characters rigs with people behind the scenes triggering actions to convey the appearance of artificial intelligence. This approach could be modified with time remixing to create a short interactive scene where the viewer becomes an actor in a scene that adapts to their presence.

Software Limitations and Challenges

Reconstructing histories using foreground segmentation techniques introduce many of the same challenges that exist in computer vision. In a key overview of tracking methods in 2006 the authors say: "Object tracking, in general, is a challenging problem. Difficulties in tracking objects can arise due to abrupt object motion, changing appearance patterns of both the object and the scene, non-rigid object structures, object-to-object and object-to-scene occlusions, and camera motion. Typically, assumptions are made to constrain the tracking problem in the context of a particular application." [29]

The middle ground for artists and designers is to identify established techniques that work well now, and design situations that account for many of the variables introduced by changing light conditions, occlusion, and camera motion. Established time remixing methods such as green screening and motion tracking with infrared markers require more rigid parameters than the methods described in this paper. The main things to consider are: find diffuse light conditions, make sure the camera is stationary, and if possible minimize situations in which objects occlude each other.

A Framework for Designers

The *Chronographer* explorations are instances of a tool that can create the illusion of events that happened at very different times being replayed as if they occurred simultaneously. Our implementations mixed videos from the past with live footage in the context of installations. It is easy to imagine possibilities for postproduction that would have previously only been possible with stop motion animation. Imagine falling leaves which become text for an instant, or drawing a shape over an empty image of Times Square and then seeing a video of hundreds of people walking up to form that shape.

Plugins for Adobe After Effects such as Mocha [13] support tracking between frames and have powerful rotoscope tools. These still require manual key framing and don't utilize advanced background modeling as part of the image capture process. The *Chronographer* software takes an approach that can be used for image capture, foreground background segmentation, and dynamic playback of time remixing content.

Currently the software is available as a series of C++ examples, presented in this paper. As we develop the tools for future projects our goal is have a generalized capture tool, a segmentation library, and alpha video playback libraries that handle loading and unloading videos. This would make these techniques more available to artists who are not experienced with computer vision techniques.

CONCLUSION

The primary takeaway from this research for us has been an awareness of new design possibilities enabled by tools that allow for individual elements within a video to be parsed into separate histories. Most video editing tools focus on transforming frames relative to each other rather than transforming elements within the video relative to other elements. They also do not provide tools that can be used during the capture process to help identify the foreground objects.

As segmentation techniques improve and capture technologies such as depth sensing and cameras with onboard computers emerge that can do background modeling, more designers and photographers will be able incorporate time remixing techniques on the fly into their work. The *Chronographer* initiative provides an example of three implementations that explore this approach in the controlled conditions of public installations.

Although there are many technical limitations to the software approach we chose, the creative possibilities that arose from the work exceeded our expectations. In the future we hope to explore crowd sourced time remixing, on the fly construction of histories, and narrative histories for interactive installations. We present our current design process, technical challenges and solutions, and software architecture in order to open these tools to others who might be inspired by similar notions of time remixing.

ACKNOWLEDGMENTS
Removed for blind review.

REFERENCES

1. *Chronographer* Installation videos: http://www.perspectum.com/chronographer/index.htm
2. *Chronographer* Code Repository: https://github.com/sethismyfriend
3. Cucchiara, Rita. 2005. Multimedia surveillance systems. In Proceedings of the third ACM international workshop on Video surveillance & sensor networks (VSSN '05). ACM, New York, NY, USA, 3-10
4. Cooper, Matthew. 2004. Video segmentation combining similarity analysis and classification. In Proceedings of the 12th annual ACM international conference on Multimedia (MULTIMEDIA '04). ACM, New York, NY, USA, 252-255
5. DeCamp, Philip and Deb Roy. 2009. A human-machine collaborative approach to tracking human movement in multi-camera video. In Proceedings of the ACM International Conference on Image and Video Retrieval (CIVR '09). ACM, New York, NY, USA, Article 32 Reflexicon, Stephen Agmanolis
6. Grimson, W.E.L. Chris Stauffer, Raquel Romano, and Lily Lee. "Using adaptive tracking to classify and monitor activities in a site," In Computer Vision andPattern Recognition 1998(CVPR98), Santa Barbara, CA. June 1998.
7. Horn, Berthold K.P., and Brian G. Schunck. Determining optical flow. Artificial Intelligence 17, no. 1-3 (August 1981): 185-203.
8. Isard, Michael, and Andrew Blake. "CONDENSATION — Conditional Density Propagation for Visual Tracking." *International Journal of Computer Vision* 29, no. 1 (1998): 5-28.
9. Ivanov, Y.; Sorokin, A.; Wren, C.; Kaur, I., "Tracking People in Mixed Modality Systems", SPIE Conference on Visual Communications and Image Processing (VCIP), Vol. 6508, January 2007
10. Krueger, Myron W., Thomas Gionfriddo, and Katrin Hinrichsen. 1985. VIDEOPLACE\—an artificial reality. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '85). ACM, New York, NY, USA, 35-40.
11. Lieberman, Jeff: http://bea.st/.
12. Lieberman, Zach: http://thesystemis.com/projects/
13. Mocha: http://www.imagineersystems.com/products/mochav2
14. Mine, Mark 2003. Towards virtual reality for the masses: 10 years of research at Disney's VR studio. In *Proceedings of the workshop on Virtual environments 2003* (EGVE '03). ACM, New York, NY, USA, 11-17
15. Oliver, Rosario, Pentland, A Bayesian Computer Vision Sytem for Modeling Human Interactions, IEEE Transactions on Pattern Analysis, Vol 22, No 8, 2000
16. OpenCV 2.2: http://opencv.willowgarage.com/wiki/
17. OpenFrameworks: http://www.openframeworks.cc/
18. Quicktime: http://developer.apple.com/quicktime/
19. Rokeby, David: http://homepage.mac.com/davidrokeby/mirrors.html
20. Rozin, Daniel: http://www.smoothware.com/danny/
21. Salavon, Jason: http://salavon.com/
22. Scott, Stacey D. Karen D. Grant, and Regan L. Mandryk. 2003. System guidelines for co-located, collaborative work on a tabletop display. In *Proceedings of the eighth conference on European Conference on Computer Supported Cooperative Work* (ECSCW'03), 159-178
23. Siebel, Nils T Steve Maybank: "Fusion of Multiple Tracking Algorithms for Robust People Tracking". In Proceedings of the 7th European Conference on Computer Vision (ECCV 2002), København, Denmark, May 2002, volume IV, pages 373—387
24. Snibbe, Scott S. and Hayes S. Raffle. 2009. Social immersive media: pursuing best practices for multi-user interactive camera/projector exhibits. In Proceedings of the 27th international conference on Human factors in computing systems (CHI '09). ACM, New York, NY, USA, 1447-1456.
25. Shatz, Lincoln: http://lincolnschatz.com/
26. Time Warp: http://dsc.discovery.com/tv/time-warp/time-warp.html
27. Utterback, C. 2005. Camille Utterback. In *ACM SIGGRAPH 2005 Electronic Art and Animation Catalog* (Los Angeles, California, August 01 - 04, 2005). SIGGRAPH '05. ACM, New York, NY, 22-23
28. Yang Y. H. and M. D. Levine, "The background primal sketch: An approach for tracking moving objects", Machine Vision and Applications, vol. 5, pp. 17–34, 1992
29. Yilmaz, Alper, Omar Javed, and Mubarak Shah. *Object tracking.* ACM Computing Surveys 38, no. 4 (December 2006): 13-es.