Shadow Figures: An Interactive Shadow Animation Platform for Performance

B. Tyler Parker
Brown University Computer Science Department
Providence, RI 02912
May 2013

Abstract

While front projection is often a viable solution for the creation of large displays, it is limiting for an interactive context. A user standing in front of a projected display results in a shadow and occluded imagery. This paper demonstrates an effective approach for eliminating user shadows and occlusion in a front projection setup. It further elaborates on this approach by presenting new software tools to replace and augment a performer's shadow with pre-recorded and generated imagery for interactive and performative purposes. Despite issues arising from projector latency and calibration imprecision, an effective proof-of-concept system for interactive shadow performance was created.

(a)

Figure 1: Even the acclaimed Metropolitan Opera production of Siegfried [1] encountered occlusion issues with their projection mappings. Spotlights are used not only to draw focus but to wash out the projection cast on a performer, eliminating the imagery projected on them but also creating a stark shadow.

1 Introduction

1.1 Problem Statement

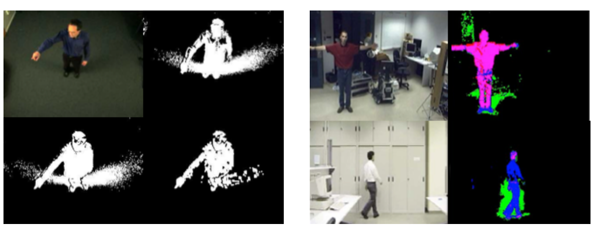
Typically, projectors are used to display imagery on a flat surface for a passive audience. However, problems arise if the projections need to be displayed in an interactive context. The primary issue is occlusion: if a user stands in the way of the projection, it creates not only a shadow behind the user, but a distorted image projected on the user as well. Multiple projectors can be used to alleviate the shadow problem by creating multiple but less noticeable shadows. Performers may wear clothing that either blends with or does not reflect the light projected on them, but this also only goes so far.

A possible solution would be to have a system that detects an occluder and eliminates both the unwanted light projected on the occluder and its shadow.

Detecting, tracking, and removing a shadow in real-time introduces opportunities for innovative interaction and performance. The replaced user shadow (either pre-recorded or generated) can become an independent character. To achieve this effect, custom tools and techniques would need to be developed, with the constraint that they must perform in real-time. Computationally restoring a shadow also poses a unique problem for the shadow removal algorithm, as it needs to recognize a real shadow from a generated one and be unaffected by any other projected imagery.

1.2 Background and Previous Work

The shadow removal problem has been approached in several different ways. One trend is to track an object in a video and use color comparison near that object to determine what pixels belong to the darkened background of its shadow. Both [2] and [3] employ this technique and remove shadows in real-time. Unfortunately, they are too noisy in their results. For these methods to work a near-perfect shadow must be detected and recorded. In addition, having a dynamic background as well as a generated shadow would greatly increase the difficulty of this task. For the purposes of enabling the algorithm to run in real-time, and to have a clean shadow for the sake of the illusion, working with traditional vision techniques in the visual light spectrum appears to be an overly problematic and ultimately ineffective approach.



(a) (b)

Figure 2: Both [2] and [3] have strong artifacts when detecting and segmenting a shadow from the background.

An infrared-based approach circumvents some of the previously discussed problems with visible light. In this case, a projected dynamic background is no longer a problem because it would not affect an infrared reading, and an artificially generated shadow also would not pose a problem for the same reason. In [4], a setup that would potentially be very effective in shadow elimination is demonstrated. The paper purely focuses on the front projection problem (where the projected light "blinds" the presenter) without extensively discussing shadow removal. Useful solutions to certain sub-problems of the project (such as multi-projector image alignment and calibration) are covered by [5], yet involve an overly complex tracking model to eliminate shadows. The approach used by [6] for the most part addresses the needs of the shadow removal component of this project: it is infrared-based which mitigates the issues with a dynamic background and generated shadow, it operates in real-time, and results in a clean shadow image.

2 Description

2.1 Approach

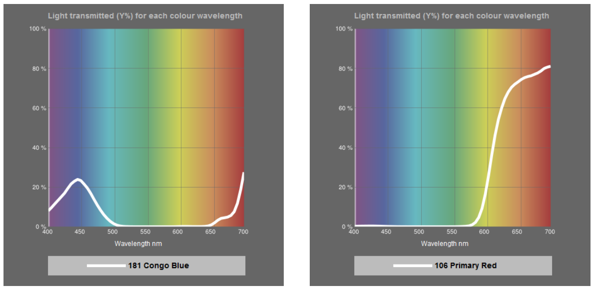
Based on the previous work, an infrared-based approach to the shadow removal/detection problem was used. By occluding an infrared lit backdrop, a user's silhouette is recorded via an infrared camera. This image, effectively a user's shadow, is then used to generate occlusion masks to create two segmented feeds. Using two projectors whose projections are mapped to the same surface, the two segmented feeds are overlaid to create one unified projection that eliminates the shadow.

For the interactive and playback portion, prerecorded shadows and imagery are integrated with the interactive live performance using computer vision tracking and transition techniques.

2.2 Infrared Hardware

An infrared floodlight is created by taking a normal light and using two colored light filters: Congo Blue and Primary Red. Congo Blue blocks green, yellow, and most red light, while Primary Red blocks blue light.

Overlaying the filters effectively multiplies their waveforms, blocking the visible spectrum but allowing infrared light to pass through. For this setup, 3 Congo Blue filters and 1 Primary Red filter for a single light source were used [8]. Additional Congo Blue gels were required as they reduce, but do not completely eliminate, red light (see Figure 3). One floodlight with a modest light throw was assembled for the purposes of the proof of concept, more in-

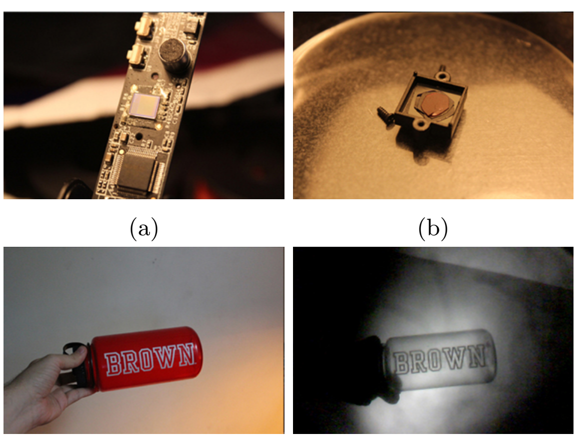


(a) (b)

Figure 3: Light transmission graphs for the two light gels [9]. Cobalt Blue (a) blocks out green and most red, Primary Red (b) blocks out blue and green. Combined, they block out most of the visible light spectrum, allowing infrared light to pass through.

frared lights covering the scene can scale this setup to an arbitrarily large performance space.

Creating an infrared camera utilizes the same principles of creating an infrared light. All digital cameras contain a filter that blocks infrared light, otherwise this light would interfere with the image sensor. By replacing the infrared filter of a webcam with the same combination of light gels only infrared light reaches the image sensor, resulting in a single-channel infrared digital image.



(c) (d)

Figure 4: The disassembled camera with its image sensor (a) and the filters used to block out visible light (b). The same scene viewed with a normal visible light camera (c) and an infrared camera (d).

2.3 Occluder Mask Generation

Each frame recorded from the IR camera is processed to generate two images: an occluder and a non-occluder mask. These two masks are used to split the desired background image into two complementary projections. Overlaying these two projections on one background surface results in a single, unified background projection, despite occluding agents (such as a performer).

The first IR frame recorded is saved for differencing subsequent frames in order to obtain the performer's silhouette. At this point, the software can be put in recording mode and save out the differenced frames for later use as a shadow performance.

The current "shadow frame" is thresholded and then dilated, resulting in a binary image with a buffer around the occluder. This buffer provides an important margin for error when tracking and replacing a moving shadow, mitigating the effect of the various latencies in the system (computational time per frame, IR camera FPS, projector display lag, etc.). The now-dilated image is then blurred in order to blend visible seams when overlaying projector segments. The resulting image and its inverse serve as the occluder and non-occluder masks.

2.4 Projector Roles

Once these masks are generated, they are image multiplied with the desired background image in order to create two piecewise projections.

The first, or source, projector is treated as the primary light source. It projects the segment of the desired background that is generated by image multiplying the background image with the non-occluded mask. The IR camera has been placed as near to the source projector as possible; this allows for a performers detected IR silhouette to effectively act as a detected shadow.

The second, or fill, projector fills in the missing segment left behind by the source projector. Its projection is created by multiplying the background image with the occluded mask (the non-occluded mask inverse).

Warping the homographies of the two projectors


(a)

(b)

(c)

(d)

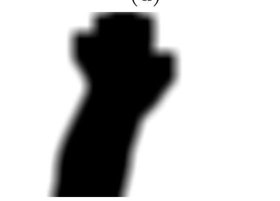
(e)

(f)

Figure 5: IR camera frame (a), differenced shadow (b), thresholded (c), dilated (d), blurred/non-occluding mask (e), inverted/occluding mask (f).


(a)

(b)

Figure 6: The resulting projected piecewise background image.

in order to map their output to the same surface results in a unified piecewise projection. In this manner, the shadow caused by the occluder is effectively eliminated.

At this point the project closely matches the Virtual Rear Projection setup described in [6]; this setup can be used to create a large display, where having an actual screen or rear projection source would be impractical due either to space or cost constraints. Yet this project does not use this technology as an end-all, but rather as a springboard for artistic performance and interactive applications.

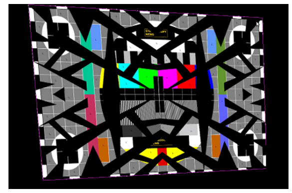
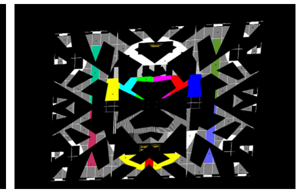
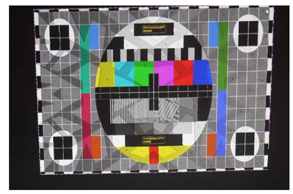

(a)

(b)

(c)

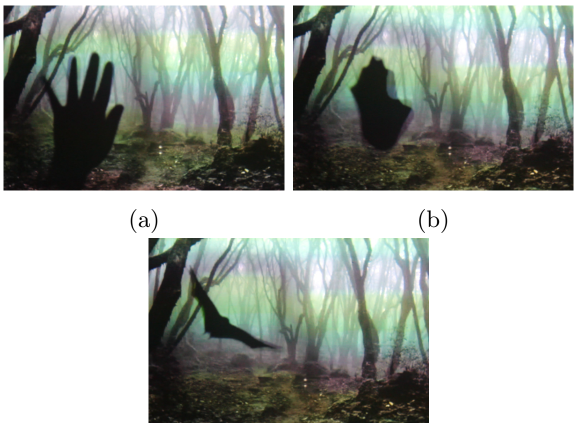
Figure 7: Projections of the calibration images (a) and (b) are aligned by using the software to distort their homographies, creating a seamless overlapped projection.

2.5 Shadow Replacement/Tracking

With the performer's shadow removed, the projected background is a blank slate which can host any arbitrary imagery. The removed shadow can either be computationally added back in, swapped with a previously recorded shadow, or replaced with any other generated content.

By tracking both the removed and recorded shadow's positions, the recorded shadow can have its translatory motion match the current performance. This allows a recording to follow a live performer, creating an interactive dynamic.

Additionally, shadow sequences can be swapped on the fly. In order to blend the transition between live and recorded frames in real-time, the contours of both shadows are computed and the vertices of the live shadow contour are mapped to the vertices of the recorded contour. The resulting "tween" contour morphs between the shapes over a given length of transition frames. During this transition, the tween contour is filled and slightly blurred to give it the appearance of an intermediary shadow. This creates a smoother transition than a jarring jump from one potentially disparate shadow to another.



(a) (b)

(c)

Figure 8: Tweening from the performer's shadow to the replacement imagery (a bat).

Tweening and tracking are just two applications for the shadow detection functionality. Shape detection could be used as triggers for animated sequences, or as drivers for interactive characters such as virtual puppets [11].

2.6 Performer/Shadow Lighting

One technique used in other shadow performance installations is to have a shadow contain imagery and animation, such as in Matreyek's Glorious Visions performance [7]. While Matreyek relied on choreographed motion and poses to achieve the illusion, one can use shadow tracking to achieve this effect with any spontaneous live performance. Using a non-thresholded frame of a computed live shadow or a previously recorded shadow as a mask, as well as optionally using position tracking information, a shadow performance can appear to contain any arbitrary imagery.

Similarly, any imagery or animation can be tracked to and projected on the performer. Normally an occluder would have no light from either projector cast on it, as this would cast a shadow (defeating the purpose of the entire setup). By eroding and blurring the thresholded image of the performer's shadow, a mask that is slightly smaller than the performer's silhouette is obtained. That, combined with the desired imagery, is projected directly onto the performer. Yet, unlike the occlusion masks, this mask has very little margin for error and requires a near fit to be effective; the mask has to be closely matched to the performer at all times. A significantly smaller mask would not cover the performer effectively yet a larger mask may overlap the silhouette and spill light onto the background, ruining the illusion.

3 Results

A user interface for adjusting shadow removal/replacement properties was implemented, as well as the ability to save/load a settings file. Balancing the settings was key for an effective mask.

(a)

Figure 9: The user control panel, able to tweak such settings as dilation, blurring, and playback states.

Dilation and blurring of the masks proved to be the user-controlled properties that had the largest effect on the framerate of the application; too many blur and dilation iterations would lower the number of frames per second considerably. Increasing the mask would provide a larger margin for error, but too large a mask would increase latency, causing additional error, as well as overlap the occluding and non-occluding projections.

Blurring the masks was crucial for the blending of the mask seams in this setup. Minor imperfections in projector calibration caused a slight darkened border to appear around the mask segments, but this was acceptable for the scope of this project. Further methods for the mitigation or elimination of this artifact are discussed below.

With a well balanced set of properties the software algorithm could run upwards of 60 frames per second (FPS), typically between 40-50 FPS with additional tracking and shadow replacement functionality activated. With full performer and shadow animations it ran around 30 FPS. All these framerates are within acceptable bounds for the illusion of continuous motion.

While the software was relatively fast, the projectors and camera used for the project had a distinct latency, and as such, if the performer was moving rapidly, the performer could exceed the projectors' capacity to compensate and remove the shadow. This effect was even more pronounced with front projection on the performer, where the margin for error is narrower. In addition, a stuttering effect was visible with fast motion. For this specific hardware setup, these limitations must be kept in mind; an emphasis on restrained user motion but highly dynamic shadow animation could be an effective way to work with the issue.

For a full demonstration of the shadow tool's functionality, please view the accompanying video.

4 Discussion and Future Work

Building on previous strategies for shadow removal using infrared background illumination and occlusion masking, a new system for interactive performance was developed and a proof-of-concept successfully implemented, providing: 1. a means to calibrate two projector planar homographies, 2. processing of an infrared camera feed to isolate a performer silhouette, 3. recording of a computed shadow, 4. contour/position tracking of a shadow, 5. elimination/replacement of a shadow, 6. shadow transition tweening, and 7. animation overlay/tracking/substitution for either the live performer or recorded/generated shadow. All this functionality serves as a foundation for interactive installations and performance, and could be extended in a myriad of ways.

One way to extend this project would be to refine projector mapping. The project scope assumed that user distorted projector-surface homographies would be an effective means of calibration, but a computational solution would be even more accurate. Techniques also exist to accommodate for the subtle lens radial distortions of a projector, as well as for differences in brightness (such as using Luminance Attenuation Maps (LAMs) [6]). Planar distortions also do not take into account a non-uniform planar surface; an interesting extension would be to pre-calibrate the projector output to warp to any arbitrary 3D surface geometry.

The primary bottleneck in the setup was not the algorithm but the optical hardware. With the application running upwards of 60 frames per second, the IR camera could only reach half that (with a lower resolution setting). The projectors appeared to have even less effective FPS, with chromatic aberration becoming a problem when compensating for any sufficiently speedy motion of the performer. Much of this could simply be fixed with better hardware. There are faster cameras and projectors specifically designed for outputting the required frames per second necessary to virtualize reality. However, this level of technology would have been outside the budget of a student's DIY proof-of-concept, which was sufficiently demonstrated even with the hardware constraints.

A possible extension that could mitigate the "frame-stuttering" of fast motion would be to dynamically generate a motion blur that visually matches the projector frame-rate (if the camera contribution of motion blur is not enough). Additionally, motion prediction could be employed to generate the occlusion masks and computed shadows. Synchronizing the projectors might also alleviate the stuttering.

As for future applications of the interactive and artistic potential of the technology, they are practically unlimited. Detected shadows could be used as a means to play and engage [11] or as merely an interface for interaction with large displays [12].

Performers may use the technology for visual effect and exploration of utilizing a shadow as a complementary character.

Ultimately, the aim of this project was to implement and showcase a technological setup and toolset that could be used for a large variety of applications, be it for artistic expression or practical interactive solutions.

Acknowledgments

I would like to especially thank Barbara Meier for her mentorship and support. I also thank Todd Winkler for his helpful reference and discussion regarding the hardware component of the project.

References

[1] Lepage, Robert. Der Ring des Nibelungen - Robert Lepage. Retrieved 4/25/2013, from http://realisations.net/Robert-Lepage-MET-Opera-NYC-2011

[2] D. Grest, J.-M. Frahm, and R. Koch, A color similarity measure for robust shadow removal in real time, in Proc. of Vision, Modeling and Visualization (VMV), Munich, Germany, Nov. 2003, pp. 253260.

[3] Mohamed Dahmane,Jean Meunier. Real-time moving object detection and shadow removing in video surveillance[J]. 3rd International Conference: IEEE of Sciences of Electronic, Technologies of Information and Telecommunications March 27-31,2005.

[4] D.S. Tan, and R. Pausch. Pre-emptive Shadows: Eliminating the Blinding Light from Projectors. In Interactive poster at CHI 2002 Conference on Human Factors in Computing Systems, Minneapolis, MN, 2002.

[5] S. Audet and J. R. Cooperstock. Shadow Removal in Front Projection Environments Using Object Tracking. In Proceedings of the 2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 07) - Workshops (Procams 2007). IEEE Computer Society, 2007.

[6] Matthew Flagg, Jay Summet, and James M. Rehg, Improving the speed of virtual rear projection: A GPU-centric architecture, in Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR05) - Workshops, 2005, vol. 3, p. 105.

[7] Matreyek, M. (2010, July). Miwa Matreyeks Glorious Visions [Video file]. Retrieved from http://www.ted.com/talks/miwa_matreyek_s_glorious_visions.html

[8] Beaty, Bill. INFRARED GOGGLES FOR UNDER $10: A Human IR Vision Experiment. Retrieved 5/14/2013 http://amasci.com/amateur/irgoggl.html

[9] Lee Filters' color details for Primary Red #106 and Congo Blue #181, from http://www.leefilters.com/lighting/colour-details.html

[10] Sibbe, Scott S. (n.d.). Scott Sibbe: Interactive Art. Retrieved 5/13/2012, from http://www.snibbe.com/

[11] Design I/O LLC, PUPPET PARADE. Retrieved 4/20/2013, from http://design-io.com/projects/PuppetParadeCinekid/

[12] Shoemaker, G., Tang, A., and Booth, K.S. (2007). Shadow reaching: a new perspective on interaction for large displays. Proc. UIST 2007, 5356.