Pixel Motion:
A Surveillance Camera-Enabled Public Digital Game

Holly Robbins[1]
New York University
Game Innovation Lab
h.robbins@tudelft.nl

Katherine Isbister
New York University
Game Innovation Lab
Katherine.isbister@nyu.edu

[1] Author now at Delft University of Technology (work conducted at NYU).

ABSTRACT

In this paper, we present a public digital game with a novel interaction strategy: the use of a surveillance camera and motion flow detection software that analyzes movement patterns. The game was designed to support and inspire creative collaborative public play in a public flow-through area of a science museum. In this paper we report about the design of the game to achieve these end-experience goals, with findings from field observations of the game over a 2-week period on-site at the museum. We also discuss implications for further work on the use of surveillance cameras and motion flow as a public game interaction paradigm.

Keywords
Surveillance cameras; collaborative/cooperative game; public display.

INTRODUCTION

Surveillance cameras are a ubiquitous technology—for example, a 2011 article claimed there was one CCTV camera for every 32 UK residents [12]. Cameras installed by government, private companies, and individuals are meant to protect property and personal safety. Yet there is much resentment about and resistance to their use [10]. The standard mode of operation for these cameras is to capture and store a video stream that an authorized person or persons monitor for danger, either in real time, or more commonly, in retrospect (by looking over video archives) after an issue has been raised by other means. Cameras may be semi-hidden or in plain view, but the camera stream itself is typically not accessible to those being monitored.

Until recently, these cameras were usually part of a closed circuit system that could not be readily reconfigured for public access. However, in the last few years, more and more of the installed security cameras are IP-based, which means their video streams could be relatively readily adjusted for public access.

In our ongoing research project, we ask the question: what if surveillance camera streams were available to encourage spontaneous and collaborative encounters among strangers in public spaces? What if those surveilled were offered ready access to these streams, and provided with tools for repurposing the streams for their own use? Could surveillance cameras follow the path of formerly highly restricted-use defense-oriented technologies, and become a widely and diversely used resource?

In the work presented in this paper, we begin to explore game play strategies that can be used in such a context (without yet tackling head-on some of the political and social implications and challenges). We constructed a game for a semi-public setting that uses the interaction components we would like to use in fully public settings in future. We worked with a local science museum to develop and deploy the game on-site for an extended period so we could see if the game mechanics and interaction strategies we had in mind could work in the ways we hoped.

RELATED WORK

There is broad evidence of a desire to reimagine the role of surveillance cameras in our lives. Surveillance cameras and their streams have been appropriated in unconventional ways by the layperson, corporations, artists, and activists. For example, drivers throughout Russia have taken to equipping their cars with dashboard cameras as a measure against lawlessness on the road [7]. Yet, the footage from these surveillance cameras is being utilized beyond their ability to protect car owners. Remarkable footage from these cameras capturing the good, the bad, and the bizarre

are widely circulated throughout the Internet as entertainment.

Marketers have also made use of surveillance cameras. Advertising campaigns such as Converse's 'Pro Streets' set up cameras in public spaces that individuals could perform in front of, with the promise of the footage being featured on the Converse website [2]. Another ad campaign by Coca Cola shows snippets of surveillance footage of people caught in spontaneous acts of kindness, heroism, or just in quirky circumstances, in a viral marketing exercise [14].

Artists have been repurposing public cameras for some time now—a well-known example is the Surveillance Camera Players [15], a group formed in 1996 in New York City, which staged plays using surveillance cameras in subways and other locations, and which is still active in monitoring surveillance camera abuses and issues.

However to our knowledge, there has not yet been work that uses surveillance cameras as a method for creating public engagement in the manner we are exploring with this line of research.

DESIGN CONTEXT

Our research group was invited by a local children's science museum to create an interactive experience for an exhibit concerned with surveillance technologies. We began our design process with on-site observations of existing visitor interactions and behaviors. The museum has a number of large display-based interactions, for which users predominantly followed a turn-taking protocol that attracted more users to gather and spectate.

When engaging with interactive displays at this museum, visitors typically engaged in a "single actor" mode. The first visitor to arrive, the "initiator," took the lead in participating in the interaction. The presence of the initiator at a site then put into motion the "honey pot effect" [12] attracting other visitors to the site as audience members. The other visitors that gathered by the initiator's station took on more of a spectator role. One of these spectators has the potential to be become the actor, but only when the current actor has left. In interviews, museum docents confirmed that this single actor and honey pot effect occurred at both digital and non-digital exhibitions throughout the museum.

DESIGN INTERVENTION

We began with a set of framing objectives in our design process: *access*, *collaboration*, and *playfulness*. The core premise of the project was the desire to transform the camera stream from a hidden proprietary resource to one that was readily accessible. We also wanted to use the stream to promote a stronger sense of collaboration among those being monitored. Finally, we decided to begin with a game in order to counteract the rather heavy, isolating feeling of being observed.

To achieve these ends, we built a collaborative game around the use of a surveillance camera stream, by adding the following technological augmentations:

1. A means for viewing the footage together in real-time: a large public display. (*access)*
2. A way to engage with the footage in real-time through collaborative action: a simple, motion-based game with no barriers to participation beyond occupying the game space. (*access, collaboration, and playfulness)*
3. A simple and familiar means for appropriating the camera's stream: providing players with the opportunity to memorialize the moment with a snapshot that can be shared via the system itself. (*access, and playfulness)*

We chose to enable collective viewing of the footage through the use of a large shared projection of the stream. In making design decisions about how best to design for the display, we made use of the large body of research on public display interaction design and evaluation [1, 3, 4, 11, 12].

The team had access to a motion-flow analysis tool that could be used to analyze the surveillance camera stream in real-time [13]. We used this tool to track the broad physical movement of those being monitored, which then became their mode of interaction with the video stream. This method has the advantage of not requiring any additional special equipment (smartphones, touch screens) for interacting with the camera stream, supporting *access*.

Finally, we created a web-based repository of images from public interaction with the camera, which those monitored could access and repurpose using two common methods for online sharing of images: Twitter and email.

Design Process

The team began with the first two technological augmentations to the camera's stream: the public display and the motion flow analysis. We built a series of prototype interactions in which we made use of the movement of passersby (as detected using the surveillance camera) in front of a large screen, encouraging them to engage. We conducted extensive internal tests, and also deployed an early version of the interaction at a public festival, making substantial revisions after observations there.

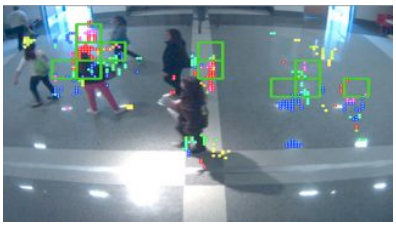

Figure 1. Motion flow overlay on video stream.

The motion flow software functions quite differently than many sensor-based tracking systems. It has no notion of locking onto an individual—it simply indicates regions of the video stream that currently have a significant amount of motion in them (Figure 1). This means that the interaction cannot make use of any notion of individual tracking, and rather must make use of the collective flow of movement. Through extensive (and sometimes painful) experimentation, we realized this could actually become a strength rather than a liability. In the final design, we presented the visualization of what the software 'saw' as a part of the HUD (heads-up display), and created a game goal that was quite simple in nature: passersbys would work to erase pixels that covered the video stream (Figure 2).

Figure 2. An example of players moving in the game space to uncover the video stream.

More motion distributed over the field of view of the video camera would result in more rapid success, and crossing the path of another person did not in any way disrupt this experience. Instead, greater participation would enhance the experience, and encourage *collaboration*. In the final version, we use this core mechanic in a short-cycle game in which those monitored attempt to 'wipe' pixels away from the video stream as quickly as possible with their movement through the camera's field of view. When a threshold percentage is reached, there is a chance to pose for an augmented digital 'postcard' that can then be accessed through the web-based photo stream. High scores are also added to the game's leader board.

Final Design

The final design is titled 'Pixel Motion' (see Figure 4 for game progression). The game cycles through 30-second rounds, in which players must clear about 60% of the screen's pixels to 'win'. There is no required minimum number of players, but it is hard to complete a round working alone, so the game encourages observers to jump in to help, as they watch the progress bar on the round. When players have reached the win state, a collection of props appear on the screen mimicking a vintage postcard scene. Players pose among the props for a snapshot that will go into the photo stream from the game. These digital postcards can be accessed from a nearby kiosk, and are shareable via email or twitter. There are two additional kiosks on that wall, one of which contains educational material about the exhibit itself; the other contains data from the system that the users can manipulate, study, and sift through.

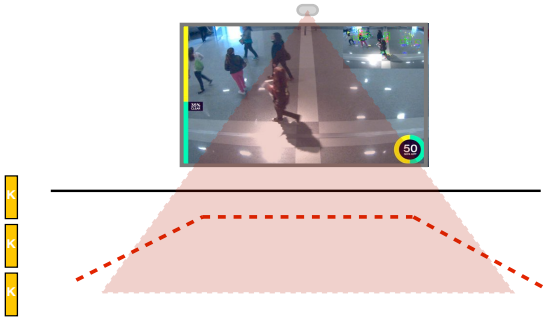

Figure 3. Wireframe of final design. Display is projected onto the wall, above which is surveillance camera. In red: field of camera vision and area within which it detects motion. Three associated kiosks line the left wall.

A projector is mounted in the hallway, throwing a 10-foot tall image onto the wall. The surveillance camera is mounted 11 feet off the ground directly above this projected image, facing outward (Figure 3). The area is also outfitted with speakers for a surround sound effect. Visitors' motion, as detected by the motion flow software, is the primary mode of interaction.

In-between game rounds, the projection shows a leader board, with four postcard snapshots ranked according to which games had been completed the fastest. The most recent snapshot taken is also featured prominently on the leader board for players to admire. The system cycles between the game state and this leaderboard continuously, unless three consecutive rounds have passed without the system detecting any movement in the play space. In this case, the game enters an idle state and the display remains on the leaderboard until motion is detected in the room, at which point the game immediately shifts into a game round (see Figure 4).

Figure 4. Pixel Motion game progression. First: the heads-up display counting down to a round. Second: in-round gameplay, players move through the play space to clear pixels from the screen to reveal the live camera feed. Third: when a round is won, props appear on-screen for players to pose among while a snapshot is captured. Fourth: an example of the postcard snapshot image. Fifth: the leaderboard.

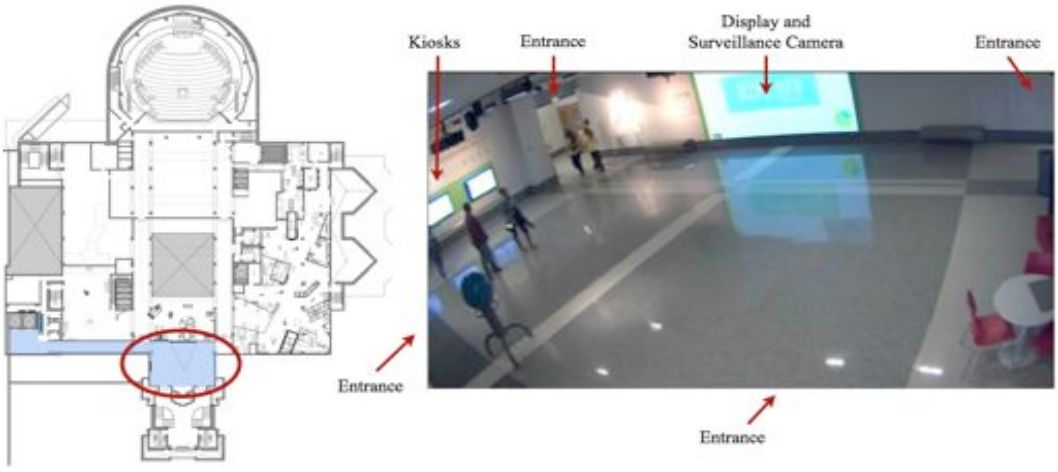

Figure 5. Left: floor plan of museum, exhibit space in blue. View from the red circle on the right; Right: the exhibit space functions much like a public plaza as a transitory space connecting several different entrances.

Installation Environment

As mentioned earlier, Pixel Motion was a custom design and installation at a museum. This museum attracts a wide range of visitors; predominantly children aged 3-15 years old visiting with their families or on school trips. Attendance on weekends ranges from 2,500-4,000 a day, and on weekdays, from 2,000-3,000.

The space that our game was deployed in is a transitory, in-between space (Figure 5). The room where we staged the game functions as an anti-room or hallway before the internal exhibit hall, and also is where the stairs and elevators lead. Beyond this anti-room area, sealed off by heavy doors that remain closed, are the exhibition halls where the majority of the exhibits, displays, and activity are. People move through this space, lingering to talk or sometimes to have a snack, while in transit from one location to another. In this way it functions like a public plaza setting, despite being indoors.

EVALUATION

We conducted preliminary evaluation of the game before installation and during a trial period at the museum after which we made modifications to the interface, timings, an difficulty levels. At the end of this process, observations and semi-structured interviews with visitors of various ages, (adults and children) confirmed that the different HUD components and the game mechanic were understood.

Figure 6. In early designs, lower interface elements were misinterpreted as the game interaction.

One noteworthy change we made to the game's interface during the last stages of testing was the readjustment of where the motion flow screen appeared, from the bottom right (Figure 6) to the top right (Figures 2 and 4). We found that players would crowd close to the screen and physically interact directly with the lowest interface elements, not realizing that they were also playing a game on the larger projected image. Changing the positioning of this element helped re-position players the proper distance from the screen—the optimal distance was what anthropologist Edward Hall characterizes as the 'social' distance zone for interaction [6], about 4-12 feet.

We collected a two-week sample of interactions with the game in order to examine whether and how the installation achieved our intended design objectives, which were to:

- Make passersby aware of the video stream and encourage them to engage it (*access*).
- Encourage and reward collaborative and playful interaction with the video stream (*collaboration* and *playfulness)*.
- Encourage appropriation of the web images taken from the video stream (*access*).

The game ran continuously during the museum's operating hours on weekdays and weekends. Over the course of the 14 days we selected to observe, 35,185 visitors patronized the museum; 102.5 hours of game play occurred over the course of 11 weekdays and 3 weekends. 7270 rounds were played, and 4264 rounds were won. Of the photos taken in the win state, 475 unique photos, and 717 photos total, were sent; 57 unique photos, and 60 photos total were tweeted.

We collected the following data streams: log files from the game itself; logs of images from the game sent via email and Twitter; video stream from the camera facing the players

used to power the game; video stream from a camera situated behind the game, from a perspective that allowed us to see the game screen as well as the players.

Given the vast amount of data, we analyzed a subsample that was representative of the range of visitor density and the range of movement detected by the game within any given round. We were also interested in rounds that resulted in high volume sharing of images via Twitter and email.

To represent the range of visitor density, we used museum records of attendance and peak hours to make selections from our dataset. To capture the full range of movement scenarios within game rounds, we worked with a colleague to create a simple visualization of movement per game round using the game logs (Figure 7), to select appropriate rounds for analysis. We used Twitter and email logs to select sample rounds that resulted in a high volume of image sharing. The final samples selected for analysis were as follows:

- (1) Museum's peak visitation weekday timeframe.
- (1) Museum's peak visitation weekend timeframe.
- (1) Museum's low visitation weekday timeframe.
- (1) Museum's low visitation weekend timeframe.
- (1) High density email sharing of winning photos.
- (1) High density twitter sharing of winning photos.
- (2) Weekday periods of consistently high activity detected by game logs.
- (1) Weekend period of consistently high activity detected by game log.
- (1) Weekday period of consistently medium activity detected by game log.
- (1) Weekend period of consistently medium activity detected by game log.
- (2) Weekday periods of consistently low activity detected by game logs.
- (1) Weekend period of consistently low activity detected by game log.
- (1) Random sample.

The total sample examined includes 2:49 hours of game play in which 240 rounds took place, 180 of which were won. Of the photos taken in the win state, 30 unique photos, and 36 photos total, were sent; while a total of 2 unique photos were tweeted

Analysis and Findings

First, the entire video sample set was viewed and transcribed (written notes on what transpired as observed—there was not discernable audio recording of conversation). Video transcription notes were then combined with an examination of photos from the win states that occurred during the sample period, and information on whether and how those images were shared. Working from this record set, we then created a coding scheme that derived both from our design aims and from observed behaviors, using a grounded theory approach [5]. We then annotated and tallied examples of behaviors that mapped to the coding scheme, the results of which are reported here.

Awareness and engagement

It was immediately apparent in the video when attention was caught by the system. The person would look at the screen while still in transit (albeit at a slower, distracted pace), stop completely to observe the screen, or immediately run up to the space that was in front of the display and start engaging with the system[2] (Figure 9). Awareness of the game was also apparent in how visitors would stop and gather around the display, even if they were not participating in the game play (Figure 8). Overall, in the video segments observed, there was almost universal awareness of the video stream, and a relatively high conversion rate from awareness to engagement for almost all passersby (making a movement, whether or not they persisted through an entire game round).

Figure 7. Visualization of game log files. Each frame represents a 30-second game round. In red: the summary of how motion was distributed and concentrated in each round with darker red showing areas of higher concentrations of movement, and lighter red, low concentrations of movement.

Figure 8. Visitors' attention was readily drawn to the exhibit, often collecting spectators (circled in red).

[2] Note that the audio from the game surely heightened the sense that there was something active in the space, and that the sounds associated with game actions helped to create a link between the actions of visitors in the space and the imagery on the screen.

Figure 9. Visitors' attention was easily drawn to the display. This group is immediately drawn to participate in the game round as soon they enter the exhibit space.

Figure 10. Players often mimicked one another's gestures--these boys mirrored one another's dramatic flailing.

Figure 11. Photo bombing: visitors who had not played in a round sometimes entered the camera's field of vision to have their picture taken in the digital post card at the game's win state.

Figure 12. Visitors often observed from a location near the screen and *within* the camera's gaze, even if not engaging in game play.

Figure 13. Example of cross-group participation. Three different school groups, each wearing shirts of different colors, collaborate on the game play, and remained intermingled until the end.

Collaborative and playful interaction

Out of the 240 rounds observed, only a few were solo play, and approximately 85% of these in the presence of others. If there was more than one person present, then the game was engaged by multiple people. We observed playfulness in the form of mimicry of one another's movements and poses, as well as in the wild and free nature of the gestures and movement patterns (Figure 10). We frequently observed players 'photo bombing,' or to drop into someone

Figure 14. Players from separate groups remain intermingled with one another for digital post card photo.

else's photo unexpectedly, of other's win snapshots (Figure 11).

It was often the case that when one or more visitors were standing to watch the display, their presence would tend to encourage other visitors, related or otherwise, to follow suit. We called this 'chaining' behavior. It is similar to the 'honeypot' effect [12], except in this case the observers became immediate participants instead of lingering as observers. We saw 131 instances of this.

Unlike previous research [12], visitors tended to observe from a location that was close to the screen and *within* the camera's gaze, even if they were not actively engaging in the game play (Figure 12). We believe this is due to the motion flow software not trying to track individuals. Observers realized quickly they were *not* hampering gameplay by being in the field of view. In fact they were helping to the extent they moved about at all. The game mechanic created a fluid boundary between player and observer that seemed to reduce the barrier to participation. We also observed frequent instances of observers offering advice to players (gesturing to illustrate how to play or where to go, for example).

Urbanist William H. Whyte notes that part of what makes a plaza or other public gathering place lively is the extent to which people tend to be attracted to other people, and thus opt towards "self-congestion" [16]. Whyte notes that this form of congregation tends to occur in places of the most activity and flow. In the case of our game, observed chaining demonstrated that the game had the appeal and allure for visitors to evoke this behavior, and to transform the hallway into a place of greater engagement and appeal.

Most of the observed collaborative play was between people who we could discern were already part of a social grouping. However, there were 105 observable occasions were members of separate groups would co-mingle to finish a round (Figure 13). In the win photos, members of these separate groups did not reform themselves spatially into their separate units, but rather remained distributed (Figure 14). There were only 7 occasions when players of mixed groups actively segregated themselves for the win state photo. This is noteworthy, as observers of other public display experiences have noted the preservation of group clusterings throughout play cycles [12].

The 'chaining' behavior mentioned in the previous section applied across as well as within groups. The game seems to have provided what Whyte calls 'triangulation' [24]. Triangulation is a situation in which "external stimulus provides a linkage between people and prompts strangers to talk to each other as though they were not." People could play the game alongside each other and thus form a sort of informal connection without being formally introduced.

Appropriation of Images

The overall number of images shared via email and Twitter was far lower than the number of rounds played and won, during our sample period of 240 game rounds in 2 hours and 49 minutes. The photo-sharing kiosk was located on a separate wall from the game screen, and many visitors did not find their way over to it and discover that they could share the snapshot postcards.

However, we observed a spontaneous image appropriation behavior: 28 times during the sampled rounds, players and spectators used their own photo-taking devices (usually smartphones) to snap a picture of the winners (Figure 15). This is a remarkable figure considering that a total of 38 of the win-state post card images were sent from the kiosk provided. In retrospect, this makes sense, as we are accustomed and encouraged (through social media) to memorialize noteworthy moments using our own image capture strategies. The fact that this many players/spectators chose to capture the moment was encouraging, and took the appropriation a step beyond what we had envisioned.

Figure 15. Players would photograph themselves in front of the display.

Discussion of Results

We designed this collaborative interactive public experience with a surveillance camera to:

- heighten accessibility and awareness of the surveillance camera stream,
- encourage collaborative and playful engagement, and
- allow for ready appropriation of the image stream.

The results of our two-week analysis of gameplay in the museum support the conclusion that we were able to achieve these aims. We saw plenty of evidence of within group collaboration, and some evidence of across-group collaboration (something that is rare in public display interactions). The spontaneous use of personal imaging devices to capture the results of the game added another

layer to our understanding of how we might encourage appropriation and re-use of public camera streams in future.

CONCLUSIONS AND FUTURE WORK

The installation of the Pixel Motion game at the museum gave us a lot of valuable insights into creating games that use security cameras. We believe the work presented in this paper may be relevant to movement-based game design, public game design, cooperative play research, and also to those interested in using cameras as game input mechanisms. The motion flow software we used also allowed us to do interesting analysis of game play behavior that may be of interest to researchers interested in game metrics.

In our on-site research in the museum prior to the deployment of Pixel Motion, we found interaction patterns, such as turn taking and the honey pot effect, that have been observed by others creating public interactions [12]. In contrast, Pixel Motion was able to begin to mix players from different groups, encouraging collaboration in a way that we would like to use to build toward public interactions that help to build mutual awareness and interest across groups.

The results of our research suggest that, with the right augmentation strategies, surveillance camera streams could be repurposed to engage the public in interactions that encourage playful and collaborative engagement, and that allow for appropriation of the image stream for purposes other than those of the owner/operator.

ACKNOWLEDGEMENTS

Thanks to Bell Labs for their generous support in the creation of this project, especially Lawrence O'Gorman, our project sponsor. Thanks to the Liberty Science Center, Russell deMoose and Luke Brown Gold in particular who were invaluable throughout the installation and data collection at the museum. Thanks to Hendrik Strobelt for the beautiful visualization. Thanks to Lei (Raybit) Tang, Xinzhao Liang, and Hanzhou (Joe) Li for their hard work on various iterations of the game experience.

REFERENCES

1. Alt, F., Schneegass, S., Schmidt, A., Mueller, J., and Memarovic, N. How to Evaluate Public Displays. Proc. of PerDis 2012.
2. Converse. Prostreets. http://www.prostreets.com/.
3. Eriksson, E., Hansen, T.R., and Lykke-Olesen, A. Reclaiming Public Space—Desiging for Public Interaction with Private Devices. Proc. TEI 2007.
4. Fischer, P.T. and Hornecker, E. Urban HCI: Spatial Aspects in the Design of Shared Encounters for Media Facades. Proc. of CHI 2012.
5. Glaser B.G., and Strauss A. Discovery of Grounded Theory. Strategies for Qualitative Research. Sociology Press, 1967.
6. Hall, E. The Hidden Dimension. Anchor Books DoubleDay, 1966. pp 101-129.
7. Lavric, D. 2013. Why Almost Everyone in Russia Has a Dash Cam. *Wired*. http://www.wired.com/autopia/2013/02/russian-dash-cams/
8. Levin, T.Y., Frohne, U., and Weibel, P. 2002. CTRL [SPACE]: Rhetorics of Surveillance from Bentham to Big Brother. MIT Press.
9. Lewis, P. 2011. You're being watched: there's one CCTV camera for every 32 people in UK. The Guardian. http://www.theguardian.com/uk/2011/mar/02/cctv-cameras-watching-surveillance
10. Lyon, D. 2007. Surveillance Studies: An Overview. Polity.
11. Mueller, J., Alt, F., Schmidt, A., and Michelis, D. Requirements and Design Space for Interactive Public Displays. Proc. of MultiMedia 2010, pp. 1285-1294.
12. Mueller, J., Walter, R., Bailly, G., Nischt, M., and Alt, F. Looking Glass: A Field Study on Noticing Interactivity of a Shop Window. Proc. of CHI 2012.
13. O'Gorman, L., Yin, Y., and Ho, T.K. Motion feature filtering for event detection in crowded scenes, Int. Workshop on Pattern Recognition and Crowd Analysis, Tokyo, Nov. 2012.
14. Rosales, G. 2012. Coca-Cola Security Cameras: Turning the Tables on Big Brother. http://www.coca-colablog.com/brands/coca-cola-security-cameras/
15. The Surveillance Camera Players. http://www.notbored.org/the-scp.html.
16. Whyte, W. H. 1980. *The Social Life of Small Urban Spaces*. Project for Public Spaces, 1980.