Human-Computer Interaction -- INTERACT'03
M. Rauterberg et al. (Eds.)
Published by IOS Press, (c) IFIP, 2003, pp. 487-494

Visually-tracked Flashlights as Interaction Devices

Ahmed Ghali, Steve Benford, Sahar Bayoumi, Johnathan Green, Tony Pridmore

The Mixed Reality Laboratory
The University of Nottingham
Nottingham, NG8 1BB, UK
{ahg, sdb, sxb, jzg, tpp} @ cs.nott.ac.uk

Abstract: We describe a technique for visually tracking flashlights so that they can be used as fun, cheap, intuitive and safe interaction devices with a wide range of surfaces. Our implementation includes a run-time system for tracking flashlight beams as they appear in a video image and an interface for dynamically configuring targets on the surface and associating them with sound files. We present two early applications of this approach: exploring a series of underground caves at a museum and interactive posters and wall displays. Early experience raises issues of dealing with wobbly flashlights, setting and conveying the extent of tracking range, and making users aware of system state. We emphasise that the defining feature of flashlights is that they throw a pool of light and suggest various ways in which this can be exploited to improve the flexibility of this approach.

Keywords: Augmented reality, visual tracking, visitor attractions, flashlights, laser pointers, wall displays

1 Introduction

Flashlights are cheap, robust, fun, familiar and safe and so make interesting interaction devices for the public. This paper describes a system for visually tracking flashlights that allows users to shine them onto a surface in order to trigger sounds that are associated with different features and target areas. Visual tracking involves extracting features of the image of the beam on the surface, including its centroid and area, from a video view in real time. Our technology also includes a tool for interactively defining targets on a video view of the surface.

This approach is primarily aimed at public settings such as museums, exploratoria, exhibitions, foyers and classrooms, where users might want an intuitive and engaging way of acquiring information by pointing at walls, ceilings, artifacts, tapestries, posters and other wall displays. It also has potential for use on the city streets at night.

This paper introduces the approach of interactive flashlights, describes its implementation, presents two applications, and discusses how early experience has led us to refine the technology.

2 Why use flashlights?

Visually tracked flashlights have previously been used for interacting with the graphical objects that were projected onto an immersive Storytent interface for young children (Green, 2002). This paper generalises and further develops the approach to work in more extensive open physical spaces, where there are multiple tracked surfaces whose boundaries are not always clearly delineated. This general approach can be compared to several other related interaction techniques.

First is the idea of tagging and tracking the flashlights themselves rather than tracking their beams. For example, (Regenbrecht *et al*, 2002) tracked fiducial markers attached to flashlights in order to control a virtual light source in an augmented reality environment. Here, the flashlights serve a more metaphorical purpose (indeed, other objects could easily be used in their place).

Second is the possibility of attaching light sensors to a surface and using flashlights to activate them. An interesting variant on this is modulating the beam of the flashlight to carry digital signals (e.g., the identity of the flashlight) that can then be will be active, reconfiguration may be difficult, and this approach may be infeasible if the surface is fragile, valuable or inaccessible. On the other hand it detected by the sensors (Hongshen and Paradiso, 2002). However, only certain points on the surface may be possible to attach the sensors to a wide range of objects that are moving around within the environment (e.g., badges worn by people) or to place them on surfaces where it would be difficult to obtain an overview from a camera (e.g., in a narrow tunnel or crawl space).

Third, there are a variety of mobile audio guides that use handheld computers rather than passive flashlights to deliver context sensitive audio information to museum visitors (see (Aoki et al) for examples). Some of these use position tracking to deliver location dependent audio and others allow proximity based interaction (e.g., through RFID) or allow localized pointing (e.g., through beaming infrared at special beacons that are attached to a surface). Like the light sensors, the latter involve attaching objects to the surface. They also tend to have a limited reading range and/or lack of high spatial resolution, making it difficult to accurately point at targets over a significant distance.

Fourth, is the use of laser pointers to remotely manipulate graphical objects on large shared displays (Davis et al, 2002; Olsen et al, 2001). Laser pointers potentially allow fine pointing and manipulation of objects and seem a highly appropriate technology for meeting rooms, lecture halls and similar environments. Given that visually tracking laser pointers is a closely related technique, it is worth clarifying some of the unique and interesting characteristics that emerge from video tracking flashlights that might make them especially suited to use in certain public settings

- First and foremost, the image of the beam on the surface takes the form of a pool of light, not a point of light. It is therefore possible to indicate an entire area of a surface at any moment in time, e.g., to sweep out several targets, or to work with targets of different sizes. On the other hand, the resolution of the beam is lower than with laser pointers.
- This pool of light also contains significant information that might be extracted by the tracking system. Its shape varies according to the flashlight's orientation to the wall. Its size varies with distance. Different flashlights' beams exhibit different patterns of light intensity, making them potentially identifiable.
- Working with pools of light opens up interesting possibilities for collaboration. The system might exploit the extent to which different beams overlap, or require large targets to be covered by several beams.
- They are readily available in a variety of physical forms (sizes, shapes weights, powers and designs) and mountings (including handheld, head-mounted, stand mounted and vehicle mounted). Some of them come with variable focus and color.
- Related to this, flashlights work at a variety of scales from mini flashlights that cast a pencil beam over a distance of just a few meters up to high powered searchlights that cast a beam over many hundreds of meters (opening up the possibility of pointing onto entire buildings or even low cloud cover).
- It is possible to cast shadows when using flashlights (e.g., by attaching a mask to the flashlight itself), opening up further possibilities for interaction.
- They are relatively safe, in terms of being shone into eyes and onto delicate surfaces.

3 The tracking system

Visually tracking flashlights involves the user directing a flashlight beam at a physical surface such as a wall; a nearby video camera detects and tracks the movement of the beam; image-processing software extracts key features of the projection of the beam as it hits the surface (e.g. its position and/or area); this in turn drives an application.

Each tracking system consists of a camera, a 500 Mhz PC, a sound display (external speakers or headphones), flashlights, and our VisionStation software. The latter consists of two key components: a target configuration system for interactively defining targets and associating them with sound files, and a run-time system that detects and tracks the beam of a flashlight across a surface and triggers the appropriate sound file whenever it hits one of these targets.

3.1 The target configuration system

Once the camera is in position, the interface shown in figure 1 is used to interactively create, remove and edit targets by clicking on the appropriate position in the video window (top of figure 1). The properties of a given target such as its height, width and its current sound-file, can then be manually edited via the dialog window (bottom of figure 1).

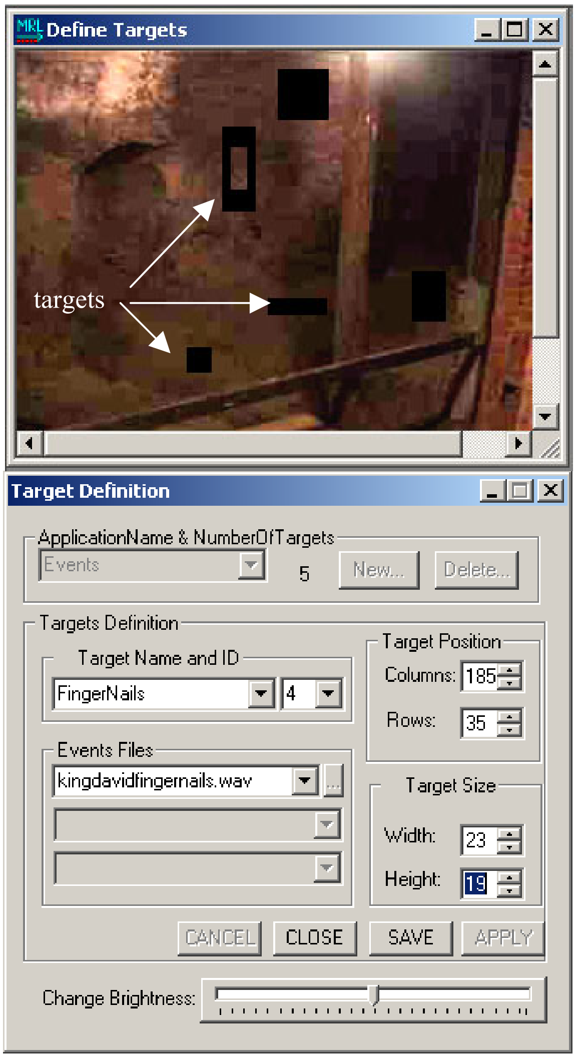

Figure 1: Target configuration system. Top: video window showing current targets, the selected target is unfilled. Bottom: editing target properties.

3.2 The run time system

The run time system monitors the area of the wall covered and then detects and tracks the beam of the flashlight across its surface. The current system reports the centroid and area of the beam. The system calculates the difference between the current image of the monitored area and a stored background image captured with no flashlight beam in view. The resulting difference image is filtered and used to extract the area of the surface illuminated by the flashlight beam, if any. The centre of gravity of the obtained area is then used to define the beam's position.

The system then looks for overlaps between the detected beam and the predefined targets. In the first version of the system this involved detecting when the centroid of the beam moved into the area of the target, which was taken to indicate that the user was pointing at the target. However, early experiences led us to develop an alternative approach that involved measuring the proportion of the target illuminated by the beam. We return to a comparison of these mechanisms later on. The run-time system can also be configured to play a background sound whenever the flashlight beam is detected in the image, but is not triggering a target. This is intended to confirm to the user that they are in tracking range, an anticipated requirements for when the technology is deployed on surfaces that cannot be fully covered by video cameras.

Deploying the system involves the following decisions. First, is the choice of the number of sections of wall to be tracked and the boundaries of each. Currently, it is necessary to deploy a separate tracking system for each section of wall, although it should be possible for each system to handle multiple cameras if desired. Second is the choice of targets and the design and recording of sound-files. Third, is the choice of camera and mounting position. Fourth is the choice of flashlight, covering size, weight, power, battery life, cost, focus and the shape of beam, whether it can be refocused, and whether it is handheld or head-mounted. Finally is the choice of the audio equipment, especially whether to use open speakers or headphones

4 Example applications

We now describe two example applications of interactive flashlights: exploring and learning about a series of underground caves at a public museum and creating interactive posters.

4.1 Exploring King David's Dungeon

We created a museum visiting experience in collaboration with a local castle museum, a historical monument that features a series of underground caves. These have served many purposes over the seven hundred years since they were first carved out, and many exciting stories are associated with them – they have been used as store rooms, larders, prisons (allegedly holding King David II of Scotland for eleven years) and secret passageways (including once kidnapping the Queen of England). The museum runs daily guided tours of this cave system.

We carried out public trials with our flashlight tracking system in King David's Dungeon in order to demonstrate its potential and to gain initial insights to help us refine the technology. We established three separate installations to cover three distinct walls of the caves, spanning two linked chambers. Our experience was open to the public for two days. During this time we hosted over one hundred and fifty visitors. Of these, roughly one hundred were adults and fifty were children. The majority of visits involved groups of two or more, although there were a few individuals.

Figure 2 shows the overall layout of the space and configuration of the system, including the locations and fields of view of the three cameras. We configured five targets for each section of wall. In order to get a broad range of experience we varied the size and physical nature of the targets, the content and duration of the sound-files and the flashlights and audio equipment used in each cave.

Cave 1 (figure 3, left): Three of the targets in this cave were associated with physical objects that we placed near ground level (two bricks and a fake spider). The remaining two were associated with existing holes in the wall. This cave featured a wooden barrier that defined a natural viewpoint and that prevented the public from approaching the cave wall by a distance of approximately five meters. Our camera was located beyond this barrier so that there was little chance of visitors occluding the image. Given the long distance to the wall, we armed our visitors with a relatively large and heavy flashlight (weighing 860g). The sounds in this cave described its use as a storeroom and larder and were played out through two nearby open speakers.

Cave 2 (figure 3, middle): This cave acted as a thoroughfare between the other two**.** The five targets were associated with man-made recesses in the walls. These were much closer together than in Cave 1 (their horizontal separation was roughly half the diameter of each – about 20cm). Given their proximity to the wall (maximum distance of two meters) visitors were given a smaller and lighter flashlight (150 g). Sound was played out through headphones that were wired to the computer.

Cave 3 (figure 3, right): Here, the five targets were on the opposite wall to Cave 2 and were associated with a variety of physical features – four recesses (not aligned, as in cave 2, and of different sizes) and a piece of carved graffiti high up on the wall. The sounds here told the story of King David. Visitors were armed with a medium size flashlight (460g) or a head-mounted flashlight (300g). Sound was played out through open speakers.

The sounds were less than five seconds long, with one exception that was fifteen seconds long (included as a test case). Each was intended to 'stand alone' (i.e., the sounds did not have to be heard in a particular order). The targets varied in size from 200 to 1500 cm^2.

4.2 Creating interactive posters

Our second application involved creating a series of interactive posters. These are paper based displays that are brought to life by the addition of interactive sound. Our aim here has been to demonstrate alternative potential uses of the interactive flashlights technology – wall displays for museums, foyers, classrooms exhibitions and possibly even the home – and also to test some refinements to the technology that were made in response to the caves experience (these are discussed in the final section of the paper).

Two examples are shown in figure 4. The first is an interactive Christmas decoration that was constructed and deployed in a family home during the Christmas break of 2002. Each snowflake is associated with its own Christmas carol. The decoration was mounted on the wall of the family living room with the camera, speakers and laptop PC being placed on top of a bookcase on the opposite wall.

The second is an interactive solar system poster that has been displayed in our laboratory over several weeks. Each planet is associated with a short recording that gives salient astronomical facts. Both posters were created by two children and an adult over a couple of hours, including assembling the poster, recording sounds and configuring the vision system. The primary use of both has been to show them to visitors to the home and lab respectively.

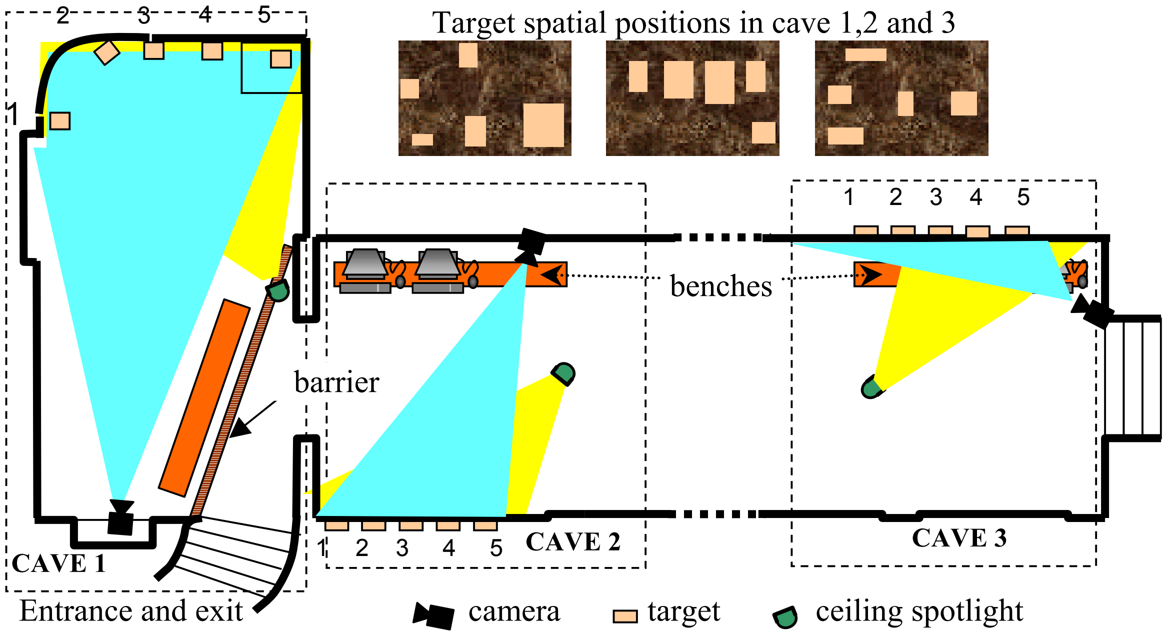

Figure 2: deployment in King David's dungeon: locations of cameras, lights, targets and barriers

Figure 3: Enhanced images of the three caves (also see video figure on the accompanying video tape)

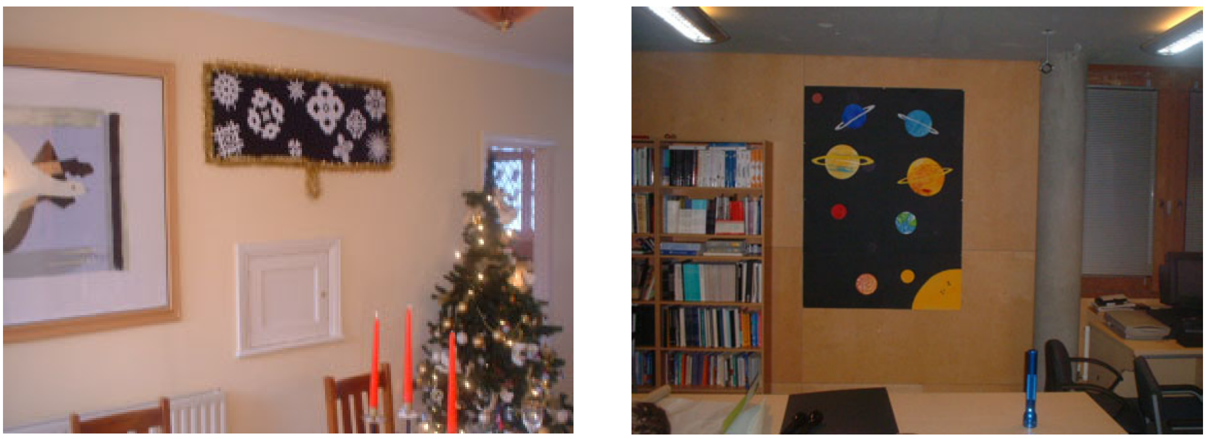

Figure 4: Interactive posters – Christmas decoration in a family home (left) and solar system poster in the laboratory (right)

5 Issues and refinements

Our observations of visitors using these initial applications have raised three key issues for refining the approach of visually tracked flashlights.

5.1 Wobbly flashlights

Previous studies of the use of laser pointers to interact with large computer displays have shown that jitter of the beam resulting from unsteady hands causes problems for fine grained interaction (Myers *et al*, 2002). We experienced similar problems with the use of the flashlights in the cave.

Wobbly flashlights would cause the beam to wander on and off of a target, which would cause the sound to repeatedly stop and begin again from the start of the recording. This was a particular problem for weaker visitors (e.g., young children), when using heavier flashlights over longer distances (when the same angular movements cause greater movement on the surface, an issue also observed from the use of laser pointers). Unfortunately, longer throws require more powerful and heavier flashlights. The effects of wobble also elicited giggles from younger children, which only made the wobble worse. This re-triggering was clearly annoying for visitors as it prevented them from hearing a complete sound file (one parent shouted out in frustration as their young child tried to hold a flashlight steady for long enough to hear a particular sound). It also produced a stuttering effect that painfully revealed the presence of the technology.

One way of reducing (but not removing) jitter with laser pointers is to filter position readings over time so that sudden jumps away and back to a relatively stable position are effectively ignored. However, we have adopted a different approach to dealing with wobbly flashlights.

The target triggering mechanism used in the caves would trigger a target when the centroid of the light beam entered the target area. The implicit assumption behind this is that the user is pointing with the flashlight (as if it were some degenerate form of laser pointer). We have subsequently revised our triggering mechanism to work on a different principle. This calculates the proportion of the target area that is overlapped by the beam area (the area of intersection of the beam with the target divided by the total area of the target). If this exceeds a configurable threshold then this target becomes a candidate to be triggered. If there is more than one such candidate, then it is for application software to decide how to respond. One option would be to trigger only the most illuminated target (the one with the highest proportion of overlap) so that only one sound is heard at a time. Another would be to trigger all candidates, allowing several sounds to be played back simultaneously.

We have used this revised mechanism with the interactive posters and our initial observations suggest that it is far more accommodating of wobble (although this remains to be tested experimentally). It is interesting to speculate why this might be so.

Figure 5 suggests one possibility. A small change in the 3D position or orientation of the flashlight can move the centroid a significant distance across the wall, maybe even taking it outside the target region. However, if the target is relatively small with respect to the width of the beam, most, or even all, of the target may remain illuminated.

It may be that users naturally view a flashlight as a device for illuminating, rather than pointing at, features of interest. Instead of keeping one point over the target they may seek to keep the target within the general pool of light, making the second triggering mechanism more effective. A further distinction should be drawn here between the size of the beam as detected by the system and its width as perceived by the user. To avoid false positives, the detection criteria used by the system are somewhat conservative, and typically mark a much smaller region than the one the user might consider to be illuminated by the flashlight. This user may therefore view a target as illuminated when the centroid is some way outside the target area and/or only a limited portion of the target area is overlapped by the system-detected beam area

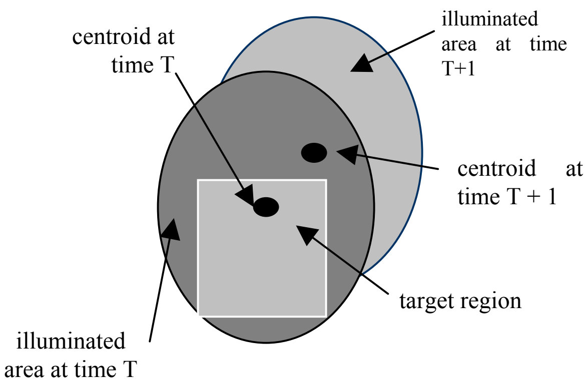

Figure 5: A small change in torch position or orientation can move the centroid outside the target region, but leave the target illuminated.

Our revised mechanism is also interesting because it exploits the fact that a flashlight throws a pool of light rather than a point of light. As noted in the introduction, we might potentially exploit this in other ways, for example, supporting collaborative triggering by allowing users to overlap or tile their beams, or extracting information about the distance and orientation of the flashlight to the surface or possibly even its identity. In short, a flashlight is not a laser pointer, and tracking technologies might make use of its unique visual properties.

As a further comment, another approach to this problem is to learn from visitors' own coping strategies. Some visitors were observed to rest the flashlight on nearby surfaces (such as the barrier in front of Cave 1) or to sit children on parents' knees on the benches (allowing the flashlights to be held close in to the body and parents to subtly help hold them). A physical solution to the problem of wobble might be to provide convenient supports such as a searchlight style swivel mounting for long throws.

5.2 Confirming tracking range

An important feature of the caves experience was that only limited areas of the total cave walls were interactive (due to limited camera field of view) and that the boundaries of these were not explicitly visible. Again, this contrasts with the previous uses of laser pointers to interact with computer displays, where the boundaries of surfaces are usually explicitly clear. Anticipating that this might confuse visitors, we had introduced the mechanism to confirm when a flashlight beam was in tracking range (playing a background sound – the wind).

We observed experiences where this mechanism was clearly understood and useful. In the caves, a young boy begins exploring the wall opposite the targets in Cave 3. After several minutes he moves over to the interactive side and then, just as he appears to be getting bored, triggers the wind sound for the first time. This captures his attention and he explores further, eventually triggering a target and then moving quickly on to find the others. In other examples we see visitors encountering the wind sound and exploring the boundary of the tracked region for a while, before moving on to find targets. In another case, an adult explains the meaning of the wind sound to a child.

On the other hand, there also appeared to be some confusion about the precise boundaries of the interactive area. In Cave 3, there was a mismatch between a clearly visible physical boundary where the wall met the ceiling and the field of view of the camera, which only roughly followed this.

The interactive posters are different again. In this case their physical boundaries are very clearly defined and the camera field of view covers the entire poster and also an area of the wall beyond. Our initial observations here suggest that the audible confirmation of tracking range may be unnecessary, as the clearly visible edges of the posters seem to naturally define a naturally expected tracking range. Further, some visitors appear to find it confusing to hear the sound when the flashlight is shining on an area of the wall that is outside the poster.

This suggests to us that a useful extension to the tracking system would be to be able to designate a subset of the video view to be the active tracking area when configuring the system. This would allow the tracking range to be set to match the natural physical features of the environment and hence users expectations. If the tracking algorithms were only applied to this selected area, then problems with occlusion of the view (e.g., by objects passing in front of the camera or shadows being cased onto the surface) would also be reduced (as these would have no effect if they were outside of the active area).

As a further comment, we saw several examples of the flashlight being fun to use on its own without the tracking system. Several groups of visitors happily wandered into other areas of the caves and explored them for a while. This reinforces the suggestion that it may be beneficial to have an interactive device that is useful in its own right (e.g., for when the technology is busy or if it fails).

5.3 Understanding system state

Another difference between the caves experience and our current posters is the extent to which the targets themselves are explicitly visible. With the posters, it soon becomes clear that the targets are the snowflakes and the planets. Once this is understood, then the consistency of the design means that it is immediately clear how many targets there are and where they are.

The caves experience was very different. Here, the targets were natural features of a rough cave wall. It was not always clear which features were targets. The most successful target was the graffiti carved near the ceiling in Cave 3, which we associated with the legend of King David carving the story of Christ in the walls with his fingernails. This feature commanded attention (on spotting it, one visitor grabbed her friend's flashlight to guide it there) and the story generated a strong reaction ("uugghhh … with his fingernails!"). This target was notable because it was an identifiable physical feature with interesting visual detail and actually required a flashlight to be able to read it clearly.

Our visitors adopted two broad strategies for finding targets. The first involved immediately trying out the most obvious physical features, and then applying the rule of consistency. The second (less popular) involved 'painting the wall', trying to systematically explore every spot. Some visitors seemed to switch over to painting towards the end of their turn, or on one occasion at the beginning of their turn following being a spectator for the previous visitor's turn. We suspect that these visitors were trying to find out if any targets had been missed. This raises a further issue. One of the characteristics of our approach (indeed, part of its appeal) is that the technology can be hidden from view; visitors point an everyday tool at existing physical features and there are no artificial computer graphics or other visual indicators of where the content is. However, this means that there are also no visible indications of system state, for example, showing how many targets have been found so far and how many remain to be found.

We also observed that people frequently revisited targets that they had already triggered, and we wondered whether they had forgotten them or perhaps were testing whether the system was still working (e.g., after an anticipated target hadn't delivered any sound). Again, it is difficult to convey this kind of persistent state information (that a given target has already been triggered) using audio alone. Perhaps some other indication of closure is necessary, telling visitors that they can safely move on from this section and hand over to someone else.

6 Conclusions

We believe that visually tracked flashlights offer an engaging and intuitive way of interacting with a range of surfaces in public environments such as museums, exploratia, galleries, foyers, exhibitions, schools, homes, or even the city streets at night.

Perhaps the single most important lesson from our work to date is to recognize the unique characteristic of flashlights – that they cast a pool of light onto a surface. This observation may be exploited to create more effective tracking techniques, for example that are less susceptible to the problem of wobble. It also suggests approaches to capturing richer information about the position and identity of the flashlight.

Our experiences have also raised issues about interacting with large extensive surfaces. In such cases it may be necessary to restrict tracking range to match natural physical features of the surface and to confirm this range to users. It is also necessary to rethink how users understand the state of the system and their interaction history.

7 Acknowledgements

This work has been carried out through the EPSRC funded Equator Interdisciplinary Research Collaboration (www.equator.ac.uk) and the European Union funded Shape project under the Disappearing Computer Initiative (www.shape-dc.org). We also gratefully acknowledge the support of Nottingham Castle Museum for access to the caves under Nottingham Castle.

8 References

Aoki, P., & Woodruff, A., Improving Electronic Guidebook Interfaces Using a Task-Oriented Design Approach, DIS 00, 319-315*, Brooklyn, New York*

Davis, J. & Chen, X., LumiPoint: multi-user location-based interaction on large tiled displays, Displays, *Vol. 23-5, Elsevier Science 2002.*

Green, J., Schnädelbach, H., Koleva, B., Benford S., Pridmore, T., Medina, K., Harris, E., Smith, H., Camping in the digital wilderness: tents and flashlights as interfaces to virtual worlds, *Chi 2002 Extended Abstracts*, *780–781, 2002.*

Hongshen, M. and Paradiso, J., The FindIT Flashlight: Responsive Tagging Based on Optically Triggered Microprocessor., *Proc. Ubicomp, 2002, Gothenberg, Setpember, 2002*

Myers, B., Bhatnager, R., Nichols, J., Peck, C., King D., Miller, R. & Long, C., Interacting at a distance: measuring the performance of laser pointers and other devices*, CHI 2002, 33-40, Minneapolis, USA, April 2002, ACM.*

Olsen D., Nielsen T., Laser Pointer Interaction, *CHI 2001Proc. 17-22, 2001*

Regenbrecht, H., Wagner, M., Interaction in a Collaborative Augmented Reality Environment, *CHI 2002 Extended Abstracts*, 504-505, 2002.