POLI: MOBILE AR BY HEARING *PO*SITION FROM *LI*GHT

Anonymous submission

ABSTRACT

Connecting digital information to physical objects can enrich their content and make them more vivid. Traditional augmented reality techniques reach this goal by augmenting physical objects or their surroundings with various markers and typically require end users to wear additional devices to explore the augmented content. In this paper, we propose POLI, which allows a system administrator to author digital content with his/her mobile device while allows end-users to explore the authored content with their mobile devices. POLI provides three novel interactive approaches for authoring digital content. It does not change the nature appearances of physical objects and does not require users to wear any additional hardware on their bodies.

***Index Terms*—** Mobile authoring tool, coded light, interactive environment, augmented physical objects.

1. INTRODUCTION

Augmenting physical objects with digital content may enrich user experience for human-object-interaction. For example, when a visitor sees an artifact created by the Maya peoples in a museum, a video or text introducing related background about the Maya civilization might help the visitor better understand the artifact. Traditional Augmented Reality (AR) systems and applications [3] achieve this goal by leveraging various visual codes for defining interactive regions and using head-mounted or wearable displays for presenting augmented content. For instance, Rekimoto *et. al.* [5] proposed an "Augment-able Reality" system for registering information in the real world using wearable computers. The system augmented the physical space and objects with infrared beacons and printed 3D barcode paper cards and let the user's wearable computer discover those IDs. The user was able to author her voice notes and still images using the wearable equipment (*e.g.,* using a microphone to record the voice, using head-mounted display and handheld mouse to choose the still images). However, visual markers/tags/RFIDs change the natural appearance of physical artifacts. Moreover, users have to wear additional devices in order to receive the augmented contents.

In this paper, we propose POLI, an interactive multimedia authoring and retrieval system using mobile devices. The POLI system provides a system administrator a mobile app to author digital content to physical objects using a mobile device. It also enables end-users to discover the authored digital content using their own mobile devices. The POLI system does not require users to equip with any addition devices except their mobile phone with a light sensor. The POLI system uses coded light for discovering physical objects' locations in real time, which does not change the nature appearances of physical objects.

Fan *et. al.* [1] proposed a system named "HiFi", which used a projector and a camera for defining interactive regions of physical objects and authoring digital content. The authoring process was done using a desktop PC. HiFi allowed users to discover the content of defined hotspots using their phone which is mounted with an external a light sensor on its top front side. However, the light sensor was physically connected to a PC through wires, where a decoding algorithm decoded the phone's location from the sequence of light intensity values and sent the decoded phone's location to the mobile app running on the user's phone. The app then calculated whether the user's phone was inside of a hotspot. The limitations of Fan *et. al.*'s work are: 1) the authoring process is not flexible. The system administrator has to author hotspots from a local PC and more importantly the system administrator can only author objects within the camera's view. Anything that outside of the camera's view cannot be authored. 2) The movability of end users are constrained by the length of the physical wire that connects the light sensor and the PC where the decoding algorithm is running.

The POLI system solves the aforementioned limitations of "HiFi". The POLI system allows administrators to author digital content to physical objects using their mobile phones without requiring a central camera. They can physically stand beside the objects and use natural hand gestures to define hotspots. POLI system also allows an end users to retrieval the authored digital content using their own mobile devices without restricting their physically movability, because the light sensor is essential plugged into the audio jack of their phone and is self-contained without requiring wires that connect to a local PC.

2. POLI SYSTEM DESIGN AND IMPLEMENTATION

POLI is an interactive multimedia authoring and retrieval system using mobile devices. For instance, a staff member of a painting art gallery can use the Mobile Authoring App to attach a video to a physical painting. Visitors can use the Mobile Client App to retrieve the authored digital content

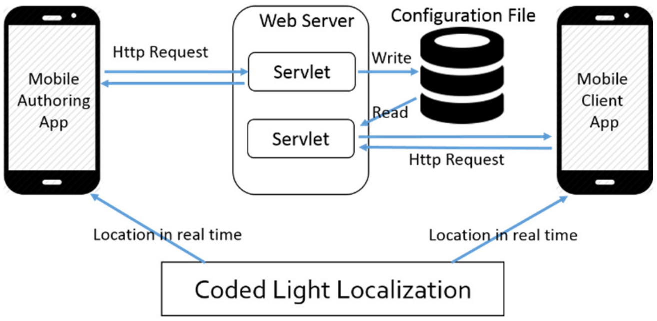

Fig. 1. POLI interactive multimedia authoring and retrieval system.

once their mobile devices entering the physical boundary of a physical painting.

The POLI system is illustrated in *Fig. 1*. It consists of a Mobile Authoring App, a Mobile Client App, a Web Server and a Coded Light based Localization System. The Mobile Authoring App allows a system administrator to define the location and the effective area of a hotspot and also author digital content to the hotspot. The Mobile Client App enables a client user to view the authored digital content of a hotspot on a phone if the phone enters the effective area of that hotspot. The Web Server interacts with the Mobile Authoring App via Http Requests to receive the configurations of all hotspots definitions, which include their locations, sizes and the associated digital content. The Web Server also interacts with the Mobile Client App to retrieve and send the related authored digital content to the client user. The Coded Light Localization projects a sequence of coded light to the physical space periodically such that each pixel of the projected area has a unique coordinate. In order to get its location, a tiny light sensor needs to be plugged into the phone's audio jack. A location decoding algorithm samples the light intensity from the phone audio jack and decodes the phone's location in real time. The location decoding algorithm is a part of both the Mobile Authoring App and Mobile Client App. Different parts of the POLI will be explained in the following sections.

2.1 Hotspot in Physical Space

We define that a hotspot is a region or an object in physical space where digital content have been assigned. A user can interact with digital content when the user's mobile device enters the physical boundary of a hotspot. For instance, a hotspot can be an art painting hanging on a wall where a video is authored within its physical boundary. When a user's mobile phone enters the painting's physical region, the authored video will be retrieved and played on the mobile phone. In addition, a hotspot may also be a region where a device control is assigned. For instance, when a user moves a phone into the physical region of a lamp whose light intensity control is authored, the phone may retrieve the lamp light intensity control and show an interactive GUI for tuning the light intensity.

In summary, a hotspot in physical space is an object or a region where a digital content or an interaction is authored to enhance its user experience. A user can interact with a hotspot with a mobile device in the hotspot physical boundary.

2.2. Defining a Hotspot

To facilitate the procedure of defining hotspot, POLI allows a system administrator to define a hotspot simply using a mobile phone. One property of characterizing a hotspot (*e.g.,* a physical object or a region) is its effective area in the physical space. Therefore, defining a hotspot for a physical object or a region is a procedure of associating digital content with a physical area. The POLI system leverages a projector-based localization method to compute a user's phone position in real time. The details of the localization technique will be explained in Section 2.2.1.

The phone's real time location then can be used to define the physical region of a hotspot. We have designed three novel interactive ways to define the region of a hotspot, which will be introduced in Section 2.2.2.

2.2.1. Real-time Localization on Mobile Phone using Coded Light

The localization system uses a DLP projector which periodically projects "Coded Light". Coded light contains a sequence of gray code images. If these images are projected sequentially, each pixel of the projected area will receive a different and unique sequence of bright/dark light intensity values. If bright and dark are coded as 1 and 0, then the sequence forms a gray code. Lee *et. al.* [4] leveraged similar technology to calibrate the projection onto any arbitrary size and oriented surface. Jones *et. al.* [2] used gray code images to figure out the layout around a TV in order to project game content beyond the TV's screen. Schmidt *et. al.* [6]used gray code images to infer a light sensor's coordinate inside a projection so as to active different GUI elements.

Fig. 2 shows a sequence of 6 gray code images. When they are projected to a surface in sequence periodically, the projected surface can be distinguished with 8 * 8 different coordinates. The first three images are used to encode the horizontal coordinate, while the last three images are used for encoding the vertical coordinate. If we divide each image equally into 8 * 8 grids, then each grid has a unique coordinate. For instance, if we use black for 0 and white for 1, then the gray code for the top left corner is ((0, 0, 0), (0, 0, 0)), which is (0, 0). In the same way, the gray code of the bottom right corner is ((1, 0, 0), (1, 0, 0)), which is (7, 7). Generally speaking, it requires to project $[logW] + [\, logH]$ gray code images sequentially to distinguish each pixel in a $W * H$ projection area. By projecting such a sequence of gray code images periodically, each pixel in the projected area can be located many times per second which can be used for real-time applications.

Fig. 2 Six Gray code images (the first three to distinguish the horizontal coordinate and the last three for the vertical coordinate) which can distinguish 64 positions (8 in horizontal * 8 in vertical).

A light sensor can sense the light intensity change with a high sensitivity. However, to capture the fast change in the readings of a light sensor, we need to sample the readings of a light sensor in high speed. We finally leverage the audio channel of a mobile phone and come up with the design shown in *Fig. 3*. The light sensor unit is attached to the top front part of the phone and plugged into its audio jack. When the mobile device with the light sensor plugged in the audio jack is placed into the "Coded Light" projection area, the POLI Mobile Authoring App and POLI Mobile Client App can sample the reading of the light sensor from the audio channel with high sampling rate and decode the location represented by the gray codes. For the area projected by a projector with the resolution of $W * H$, the decoding algorithm reads in $[logW] + [\,logH]$ light intensity values each time, and thresholds each value into zero or one so as to compute the gray codes which represent the location(x, y), $0 \le x < W, 0 \le y < H$.

2.2.2. Interactive Ways of Defining Region of a Hotspot

As described in section 2.2.1, the location decoding algorithm on mobile phone calculates the phone's location in real time by decoding a sequence of light intensity values of Coded Light received from a light sensor that is plugged in the phone's audio jack. To further define the region of a hotspot, we need an interactive approach to define a region where digital information or interaction could happen using point locations decoded by the decoding algorithm. We have explored the following three possible authoring approaches (Figure 3) for the mobile authoring app.

- *Center + Radius Authoring*

"*Center + Radius Authoring*" defines the region of a hotspot by its center and an effective radius R (*Fig. 4 (a)*). The actual POLI Mobile authoring UI for this method is shown in *Fig. 5 (a)*. When a user moves a phone to the center of a hotspot and clicks the "center" button on the GUI,

Fig. 3 To localize a phone, a light sensor is plugged into the audio jack of an android phone (Galaxy Nexus) for sampling the coded light in real time, which will be feed to a decoding algorithm.

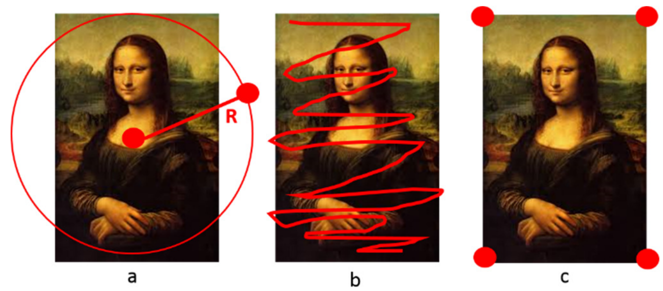

Fig. 4. Illustration of three authoring mechanism with a painting: (a) Center + Radius authoring; (b) Scribble authoring; (c) Four-Corner authoring

the position decoded at that moment will be stored as the center's coordinate. When the user moves the phone to the edge of the imaginary circle for the hotspot and clicks the "edge" button, the position of the edge point will be decoded and then the distance between edge point and the center point is calculated and used as the effective radius R of the hotspot. When a client user's mobile device is inside the circle area of the painting, the authored digital content (*e.g.*, a video) will be retrieved and played by the POLI mobile client app in real time.

- *Scribble Authoring*

"*Scribble Authoring*" is illustrated in *Fig. 4 (b)*, which defines the region of a hotspot by a sequence of points on the trace scribbled by a user. The actual POLI mobile authoring UI is shown in *Fig. 5 (b)*. With this app, a user can define the area of a hotspot by pressing down the "Scribble" button, and scribbling over the hotspot area before releasing the button. The app then records a sequence of decoded positions that represent the scribbled trace. We define the effective area of a hotspot as the minimum bounding box of the scribbled trace. Let's assume the coordinate of the sequence of positions from time *i* to time *j* are represented as (t, x_t, y_t), t = *i*,…,*j*. the minimum bounding box of the scribbled trace is computed and represented by four lines, two horizontal and two vertical lines: $x_{min} = \min x_t$, $x_{max} = \max x_t$, $y_{min} = \min y_t$, $y_{max} = \max y_t$, t = *i*,…,*j*. If a client user's mobile phone's

Fig. 5. Three POLI mobile authoring UIs: (a) Center + Radius authoring; (b) Scribble authoring; (c) Four-Corner Authoring

position (x, y) satisfies the constraints: $x_{min} \leq x \leq x_{max}$ && $y_{min} \leq y \leq y_{max}$, the authored digital content will be retrieved and played on the phone.

- *Four-Corner Authoring*

The idea of "Four-Corner Authoring" is illustrated in *Fig. 4 (c)*, which defines the area of a hotspot by the four corners of a rectangle. The POLI mobile authoring app has four buttons to let a user press and define the coordinate of the four corners (*Fig.5 (c)*). If a client user's mobile phone is in the rectangle area defined by the four corners, the digital content will be retrieved and played on the phone. Let's define the coordinate of the top-left, the top-right, the bottom-left and the bottom-right corners as the following $(x_{tl}, y_{tl}), (x_{tr}, y_{tr}), (x_{bl}, y_{bl}), (x_{br}, y_{br})$ and the coordinate of the client user's mobile phone as (x, y). If (x , y) satisfies the following two constraints:

$$\big((y_{tr} - y_{tl}) * (\mathrm{x} - x_{tl}) - (x_{tr} - x_{tl}) * (y - y_{tl})\big) * \big((y_{br} - y_{bl}) * (\mathrm{x} - x_{bl}) - (x_{br} - x_{bl}) * (y - y_{bl})\big) \leq 0$$

and

$$\big((y_{bl} - y_{tl}) * (\mathrm{x} - x_{tl}) - (x_{bl} - x_{tl}) * (y - y_{bl})\big) * \big((y_{br} - y_{tr}) * (\mathrm{x} - x_{br}) - (x_{br} - x_{tr}) * (y - y_{br})\big) \leq 0$$

, the phone is inside the hotspot and the authored digital content will be retrieved and played on the phone.

2.2.3. Creating & Uploading the Configurations of Hotspots

After defining the effective area of a hotspot, the next step is to choose the type of digital content for the hotspot. POLI Mobile Authoring app enables a system administrator to choose a multimedia file (*e.g.,* video, audio or a text) from the authoring phone. Once the digital content is selected, a configuration of the hotspot is created as a JSON string. The format of the JSON string using "*Four-Corner Authoring*" method is defined in Fig. 6. The generated JSON string and the actual file indicated in the "*mmname*" field of the JSON string form the configuration for a hotspot. POLI supports a user to define as many hotspots as possible. The JSON string for each hotspot will be combined together as a JSON Array. This JSON Array and all the multimedia content files described in each JSON string's "*mmname*" files are the configurations for all the hotspots. They are uploaded to the POLI Web Server via HttpRequests.

{"tl_x": *, "tl_y": *, "tr_x": *, "tr_y": *, "bl_x": *, "bl_y":*,"br_x":*,"br_y":*, "type": *, "mmname":*}

Fig. 6 The configuration of a hotspot in JSON format: The first eight parameters are used to define the coordinates of four corners. The "*type*" parameter defines the type of the digital content, which could be one of the four tags: "VIDEO","IMAGE" "AUDIO" and "TEXT". The "*mmname*" stores the file name of the digital content.

2.3. Interacting with a Hotspot

The ultimate goal for authoring digital content in hotspots is to let visitors explore the digital content and interact with the hotspots. For explorations and interactions, a client user needs a mobile phone with a light sensor (Fig. 3) and the POLI Mobile Client App. When the phone enters a hotspots area, the POLI Mobile Client App will send a HTTP request to the POLI Web Server to fetch the configuration file, which describes all the existing hotspots (*e.g.,* their locations, effective areas and the authored digital content). The mobile app constantly decodes the phone's current location based on the light intensity changes sampled from the phone's audio jack, and computes whether it is inside one of the hotspots. Once the mobile phone is in a hotspot, the Mobile App will send another HttpRequest to the Web Server to fetch the corresponding digital content. For instance, if the digital content is a video, then the video will be streamed from the server to the mobile phone, where the user can view it.

3. APPLICATION

In the previous section, we have explored the system architecture of POLI, how a system administrator defines hotspots and how a client user interacts with hotspots. In

(a) (b)

Fig. 7. Simulated mini Museum Environment Set-up in a room: (a) A projector is mounted on the top side of a wall and is projecting the coded light to the opposite side wall. (b) Two paintings are hang in the projection area and the museum registration paper is placed on a table.

this section, we introduce a concrete implementation of the POLI, a *Museum Assistant* system.

While visiting a museum where many artifacts are displayed, visitors might get a better understanding of the physical artifacts if relevant descriptive information is presented along with them. Currently such information is usually presented on a printed paper/board besides the displayed artifacts. The disadvantage of printed text/image is that the content shown are static. They cannot be used to serve audio/video data to visitors. Moreover, if the existing content need to be updated or new content need to be added, a new paper/board has to be re-printed. Museum Assistant system is a combination of both hardware and software solution that allows museum administrator to author more vivid and vibrant digital content to the displayed artifacts, and meanwhile enables visitors to view those digital content on their personal mobile devices.

To demonstrate the idea of *Museum Assistant* system, we simulated a mini-museum in a room with two paintings on a wall (Fig. 7). The mini Museum includes a projector which is mounted on the top of a vertical wall and projects the coded light to the opposite side vertical wall with two paintings. The periodically projected coded light as a localization infrastructure, which assigns each pixel in the projection area a different location. Any mobile device with our light sensor and decoding algorithm can identify its own location in the projection area in real time.

With this setup, a museum staff member can use a mobile phone to author hotspots in the projection area. When starting the *MuseumAuthoring* app, the staff needs to first identify the Web Server, where all the hotspots' configuration information will be stored. In our current implementation, we use IP address to identify the Web Server (Fig. 8 (a)). Second, after entering the IP of the Web Server, the staff will have a chance to choose a favorite way to define the region of the hotspot from the three options (Fig. 8 (b)) that we have described in Section 2.2.2 and use that method to define the physical region of a hotspot. Third, the staff can choose a multimedia file from the phone by pressing "*Click to choose a video to author*" button (Fig. 8 (c)). For instance, if it is a video, the staff gets a chance to preview the video content in the "*Video View*". When the staff is satisfied with one video, he/she can press "Next" button to proceed to the next screen (Fig. 8 (d)). Meanwhile, the video file will be uploaded to the Web Server via an

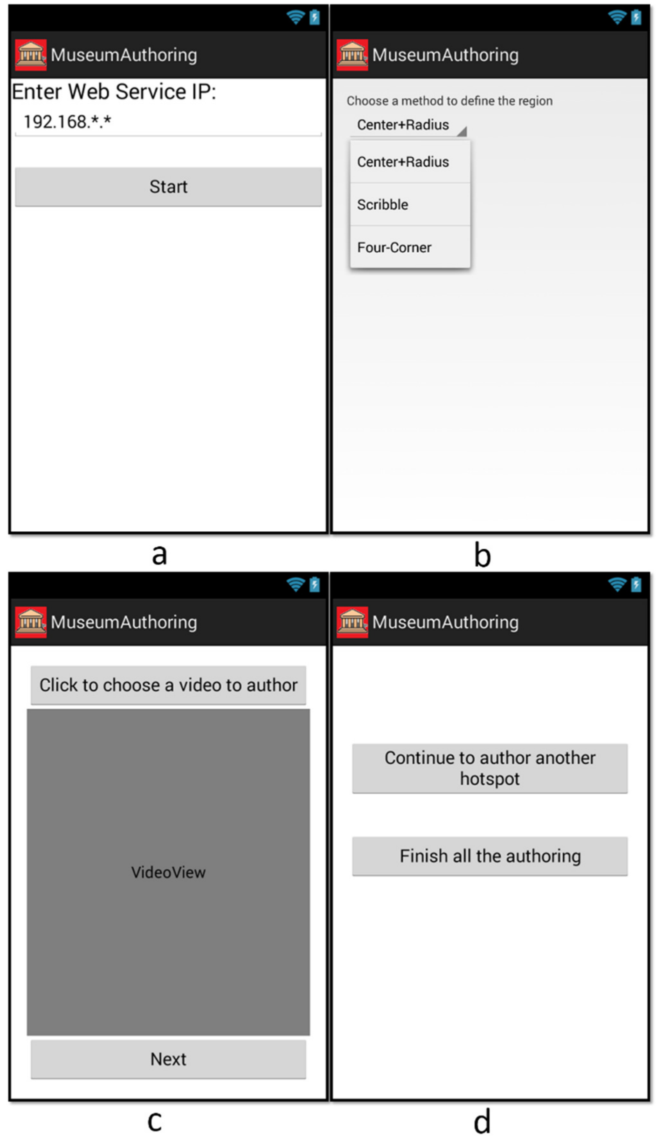

Fig. 8. Mobile Authoring Steps for the Museum Assistant System.

(a) **(b)** **(c)**

Fig. 9. A visitor registers to the mini-museum (a), and discovers the authored digital content for two paintings using a mobile phone (b, c).

HttpRequest. After that, the staff can continue to author another hotspot by pressing the "*Continue to author another hotspot*" button, which will take her to the screen Fig.8 (b). At the same time, the configuration of this hotspot is created as a JSON string as described in Section 2.2.3 and the JSON string will be added into the JSON array. In other words, the JSON array contains the configuration for all hotspots that have been defined so far. If the staff has finished all the authoring, he/she can simple press the "*Finish all the authoring*" button to upload the JSON array configuration for all the authored hotspots on the Web Server.

On the other hand, a visitor can use a mobile phone with POLI setup to retrieve and view the digital content from hotspots. To prevent any unauthorized access to our system, a visitor needs to first register a phone to the system. To do so, the visitor needs to start the client app and places the phone in the area of the registration paper, which is placed on the surface of a table (Fig. 7 (b) and Fig. 9 (a)). The client app will send the phone's location to the web Server for verification. Once the Web Server confirms that the phone is placed in the registration paper area, it will give this mobile phone access to discover all the hotspots in the *Museum Assistant* system and sends the configuration of all hotspots to the client app. When the client app gets the configuration file, it knows all the locations, effective areas and the authored digital content. When the visitor moves around in the museum, the client app keeps tracking of the phone's position and checks if the position is inside one of the hotspot area. If yes, then the client app will send the Web Server another HttpRequest to fetch the digital content. In this case, a related video is streamed to the phone and shown on the screen (Fig. 9, b and c). Note that the videos are different in Fig. 9 b and c, because the visor's phone is in different painting's hotspot area.

4. DISCUSSION

4.1. Multi-users support

The POLI system allows multiple users to interact with hotspots simultaneously, because the POLI Mobile client app running on each user's phone computes the phone's location independently and also interacts with the POLI Web Server independently in terms of fetching the digital content authored in hotspots. The only case that a user might be interfered by another user is when another user physically occlude the line-of-sight between the light sensor on the user's phone and the projector's light beam, in which case the light sensor cannot receive the light from the projector.

4.2. System Performance

The coded light localization uses gray code which makes it robust to accidentally errors in sensor readings. Gray codes have the advantage that each adjacent gray code only has one bit different. Therefore, even if the decoding algorithm detects one bit incorrect (*e.g.,* interpreting a low light intensity (0) as a high intensity (1)), the decoded location would still be just one pixel away from the true location, which will have very little effect on detecting hotspots.

5. CONCLUSION AND FUTURE WORK

In this paper, we present the POLI system, which allows a system administrator to author digital content to physical objects using a mobile phone and allows client users to explore the hotspots on their own phones. The POLI system leverages a coded light based fast localization system to discover a user's mobile device's location in real time. We have designed and implemented three novel interactive ways of defining hotspots. For future work, we plan to extend the authoring and retrieval into 3D by leveraging multiple projectors (*e.g.*, one on the ceiling projecting downward and one on the side wall projecting horizontally).

6. REFERENCES

[1] M. Fan, Q. Liu, H. Tang, P. Chiu. (2014). HiFi: hide and find digital content associated with physical objects via coded light. In *Proceedings of the 15th Workshop on Mobile Computing Systems and Applications (p.6). ACM.*

[2] B. Jones, H. Benko, E. Ofek, A. Wilson. (2013) IllumiRoom: Peripheral Projected Illusions for Interactive Experiences. In *Proc. CHI'13*. 869-878.

[3] D.W.F. Krevelen, R. Poelman. A Survey of Augmented Reality Technologies, Applications and Limitations. *International Journal of Virtual Reality*, 2010, 9(2):1-20.

[4] J. Lee, P. Dietz, D. Aminzade, S. Hudson. (2004) Automatic Projector Calibration using Embedded Light Sensors, In *Proc. UIST'04*. 123-126.

[5] J. Rekimoto, Y. Ayatsuka, K. Hayashi (1998). Augment-able reality: Situated communication through physical and digital spaces. In *Wearable Computers, 1998. Digest of Papers. Second International Symposium on (pp. 68-75). IEEE.*

[6] D. Schmidt, D. Molyneaux, X. Cao. (2012). PICOntrol: using a handheld projector for direct control of physical devices through visible light. In *Proc. UIST'12*. 379-388.