SideBySide: Ad-hoc Multi-user Interaction with Handheld Projectors

Karl D.D. Willis [1,2] *Ivan Poupyrev* [1] *Scott E. Hudson* [3,1] *Moshe Mahler* [1]

[1] Disney Research Pittsburgh
4720 Forbes Avenue
Pittsburgh, PA 15213
{karl, ivan.poupyrev, moshe.mahler}
@disneyresearch.com

[2] Computational Design Lab, [3] HCI Institute &
Heinz College Center for the Future of Work,
Carnegie Mellon University
5000 Forbes Avenue,
Pittsburgh, PA 15213
scott.hudson@cs.cmu.edu

ABSTRACT

We introduce *SideBySide*, a system designed for ad-hoc multi-user interaction with handheld projectors. *SideBySide* uses device-mounted cameras and hybrid visible/infrared light projectors to track multiple independent projected images in relation to one another. This is accomplished by projecting invisible fiducial markers in the near-infrared spectrum. Our system is completely self-contained and can be deployed as a handheld device without instrumentation of the environment. We present the design and implementation of our system including a hybrid handheld projector to project visible and infrared light, and techniques for tracking projected fiducial markers that move and overlap. We introduce a range of example applications that demonstrate the applicability of our system to real-world scenarios such as mobile content exchange, gaming, and education.

ACM Classification: H5.2 [Information interfaces and presentation]: User Interfaces.

General terms: Design, Human Factors, Experimentation

Keywords: Interaction techniques, handheld projector, pico projector, multi-user, ad hoc interaction, games.

INTRODUCTION

Handheld computing devices have become a ubiquitous part of modern life. They allow us to communicate, retrieve, and record information at any time, view digital content on the go, and play games in any location. Interaction with handheld computing devices, however, remains a largely solitary, single user experience. Today's devices do not typically provide interfaces and supporting technologies for co-located multi-user work, learning, and play – crucial elements of human interactions in the real world.

Our research is motivated by the vision of handheld computing devices that allow users to dynamically interact with each other in shared interactive spaces. Handheld projectors are an enabling technology that could realize this vision. They are sufficiently small to be grasped in a single hand, and light enough to be moved from place to place. Multiple users can project digital content directly into the physical environment. The relatively large size of projected images allows them to be easily seen by multiple users, making them a particularly natural fit for co-located multi-user interaction (e.g. Figure 1).

Figure 1: Interacting with the *SideBySide* system.

There has been rapidly growing interest in handheld projectors from both the industry and research communities [17,18], however, most work has focused on single user applications [e.g. 1]. Although there has been some work exploring the potential of handheld projectors for multi-user interaction [2,7], these systems require instrumentation of the surrounding environment with often complex sensing infrastructure. This significantly reduces the usefulness of multi-user handheld projection systems in any real world scenario where user mobility is required. We are unaware of any system allowing multiple handheld projectors to interact together without instrumenting the environment.

We present *SideBySide*, a system designed for ad-hoc multi-user interaction with handheld projectors. *SideBySide* does not require instrumentation of the environment and can be used almost anywhere. It uses a device-mounted infrared (IR) camera and a novel hybrid visible/IR light handheld projector. The *SideBySide* software platform tracks multiple independent projected images in relation to one another using invisible fiducial markers projected in the near-IR spectrum. The resulting system allows a broad range of new interaction scenarios where users are not tied to a fixed location. Exchanging digital content, such as personal contact information or media files, can be initiated without any infrastructure. Multiple users can play projected games together by creating ad-hoc gaming spaces. Interaction scenarios for education allow teachers and students to study together using a shared information space. Most importantly, these interaction scenarios can happen anywhere – in an elevator up to a business meeting, in a child's bedroom after lights out, or in the hallway outside a classroom.

We present the following contributions:

1. We outline the design and implementation of our hardware system, including a hybrid handheld projector unit that projects both visible and IR light.
2. We introduce our software system for tracking multiple projected images in relation to one another with invisible fiducial markers. In particular we present techniques for tracking overlapping fiducial markers and optical communication between devices.
3. We present a range of applications that demonstrate the exciting and engaging multi-user interaction scenarios possible with the *SideBySide* system.

RELATED WORK

Research related to the *SideBySide* system can be divided primarily into three categories: projector-based augmented spaces, multi-user handheld projector systems, and invisible marker tracking.

Projector-Based Augmented Spaces

An important aspect of multi-user interaction is extending interaction beyond a single user's screen and into the environment. Rekimoto's *Pick-and-Drop* system defined interaction techniques to support data transfer between multiple computers and handheld devices [15]. The *Augmented Surfaces* project later established a direct spatial relationship between laptop screen content and content projected onto nearby surfaces [16].

The use of arbitrary surfaces for content projection was explored with the *Everywhere Displays Projector* [14] to augment indoor spaces with projected content at any location using a steerable projector. The *Play Anywhere* system used a portable short-throw projector to create interactive tabletop experiences with handheld devices and tangible objects [27]. The *Bonfire* system also offered a portable form-factor to extend the desktop computing experience onto the tabletop using several handheld projectors mounted on the back of a notebook computer [8].

These projects highlight the potential for projected content to enhance the user experience by operating beyond the bounds of a single display. They are however, designed for use in a stationary position and are not suitable for use when moving between different locations. The recent emergence of handheld projectors offers a lightweight solution for unbounded mobile interaction.

Multi-user Handheld Projector Systems

Early pre-cinema use of handheld projectors stems from the development of the *Magic Lantern* in 17th century Europe and its later adaption in Japan for the *Utsushi-e* performance [25]. This performance involved small wooden slide projectors used from behind a large rice paper screen to act out a story for an audience. Multiple performers would coordinate their movement to change the size and location of projected images. This enabled the creation of relatively complex real-time animation in the pre-cinema era. This important pre-history of handheld projector-based interaction underlines the potential for multiple small projectors to interact together in a seamless way.

The relatively large public displays created by modern handheld projectors make them ideal for multi-user interaction scenarios. Sugimoto et.al. created a mockup system where two overlapping projection screens were used with a PDA touch screen to initialize file transfer between devices [21]. An exhaustive range of multi-user interaction techniques were developed by Cao et.al. using a motion capture system for location tracking [2]. These interactions focus on operations within a virtual workspace, such as content ownership, transfer, viewing, and docking.

Multi-user games have also been developed that allow users to work together to reach a goal. Hosoi et.al. developed a multi-user handheld projector game for guiding a small robot along a projected path [7]. Users line up pieces of track for the robot to follow and reach its goal. Cao et.al. developed a multi-user jigsaw game where users would pick up and place pieces of a puzzle together [3].

To enable interaction between multiple handheld projectors, these systems rely on infrastructure being added to the environment. This ranges from a fixed camera above the interaction area [7], to a professional motion capture system [2]. Relying on fixed infrastructure within the environment severely limits where handheld projectors can be used, substantially limiting their mobility. Our vision is to enable multiple users to interact side-by-side, anywhere, in any space. This is one of the main design considerations for this research, and one that strongly differentiates our technical solution from past work.

Invisible Marker Tracking

Fiducial markers have been used widely for location tracking due to their lightweight, robust performance. A well-known issue with structured, 2D barcode-style fiducial markers is their unnatural appearance that users cannot read or understand. Barcode style markers are difficult to integrate into the design of interactive systems due to their fixed aesthetic and form-factor that is intolerant to changes in color, shape, or material.

To address these concerns numerous systems have been developed to disguise or hide fiducial markers from the user. Custom marker patterns have been developed that are disguised to look like wallpaper [19], markers have been created with invisible inks for use with IR cameras [13], retro-reflective markers have been used together with IR cameras and lights [11], temporal sequencing of markers has been used with projectors and high speed cameras [6], structured pattern style markers have been projected with IR lasers [9,24], and several systems have been developed using IR projection from a fixed projector in either pure IR [5,20,23] or with hybrid IR/visible light [10].

Natural marker detection techniques have also been developed to detect imagery and objects based on their natural features without any structured marker pattern [12]. However, natural marker detection typically requires time consuming training for each object and is computationally expensive when compared to structured marker detection.

To track the movement of handheld projectors we developed a handheld projector that can project both invisible fiducial markers in the near-IR spectrum as well as content in the visible spectrum. As noted above, this has previously been accomplished using a fixed projector, but we are unaware of any system implementing IR projection in a handheld form-factor. The smaller form-factor and portability of handheld projectors introduces a number of new considerations for IR marker tracking with multiple users, where markers constantly move and can frequently overlap.

SIDEBYSIDE APPROACH

Our approach is guided by the vision of multiple users playing and interacting with each other using handheld projection devices (Figure 2). Key to this vision is interaction that can happen anywhere, at any time, in a fluid, spur-of-the-moment fashion. Users should be able to establish and break off interaction in a natural and transparent manner. Standing side-by-side and projecting images into the same space should be sufficient to establish immediate communication between devices and initiate interaction. Figure 3 shows the fundamental system design and technology solution we propose to realize this vision. Two requirements guide the development of our system: on-device sensing and lightweight communication.

On-device Sensing

We chose to embed all sensing on the projection device itself. Prior instrumentation of the environment, either with passive or active sensing infrastructure, should not be necessary. Each of our projection devices is equipped with a camera, an inertial measurement unit (IMU), a ranging sensor, and a push button (Figure 6).

A key requirement of any projector based multi-user interaction scenario is accurate registration of moving images projected from different devices. Each device must understand what other devices are projecting and where they are projecting it. We use on-board cameras to accurately track the position and orientation of moving projected images relative to each other.

Figure 2: The *SideBySide* concept – a self-contained, full color, handheld projection device allowing multi-user interaction in almost any space.

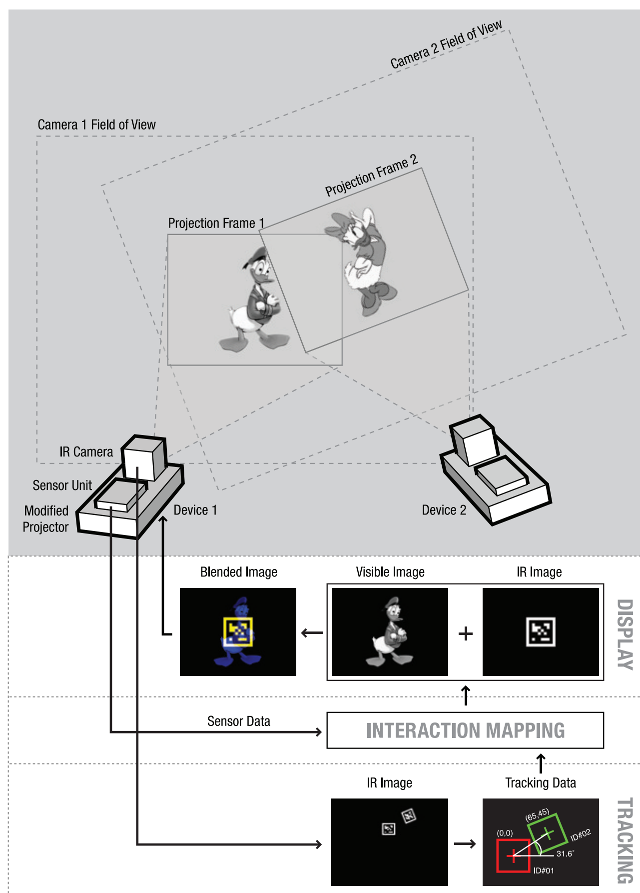

Figure 3: The *SideBySide* system overview.

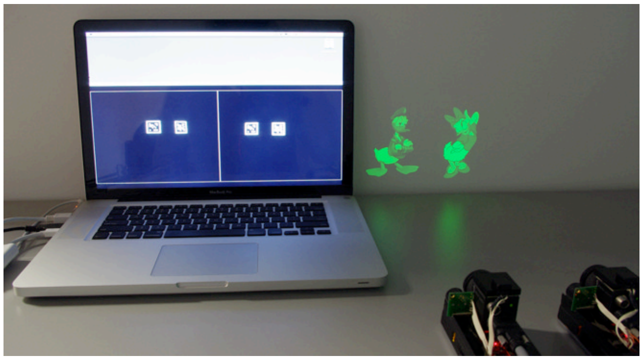

Figure 4: Projecting visible and IR images in a single stream. The characters are visible to the user, and the markers are visible to the IR camera.

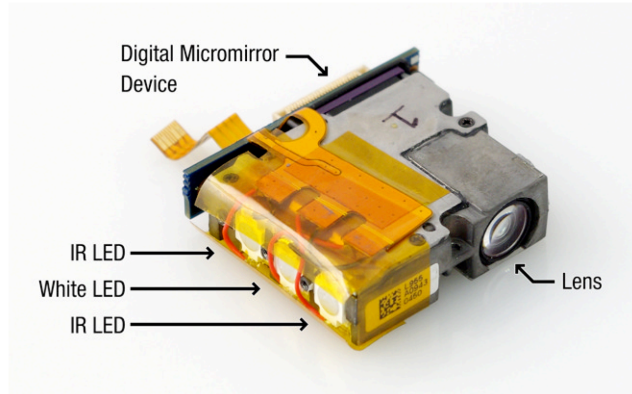

Figure 5: A DLP pico projector engine was modified to project both IR and visible light.

Because tracking visible projected images restricts the type of content that can be projected, we track invisible fiducial markers projected from the handheld projection device (Figure 3 and 4). To achieve this we modify the light sources of a handheld projector to add an IR projection channel. This allows independent visible and invisible content to be combined in a single projected light stream. An on-board IR sensitive camera constantly identifies and tracks the position and orientation of all fiducial markers projected. This data is then used to drive interactive applications. No tracking data is visible to the user, creating a fluid interaction experience.

Lightweight Communication

We enable instant and lightweight communication between devices. Other communication protocols, such as Wi-Fi or Bluetooth, typically require numerous steps to establish an explicit network connection. We use optical communication by projecting symbolic fiducial markers in the invisible IR spectrum. This can be used to communicate events such as button presses, or changes in the state of the application. Optical communication can also be used to initiate other forms of network communication as required.

SIDEBYSIDE HARDWARE PLATFORM

The *SideBySide* hardware platform consists of two core components: a hybrid IR/visible light projector and an on-board sensor assembly. We developed two prototype devices that are currently tethered to the same computer. This greatly simplified the development and evaluation of *SideBySide*. It is important to stress, however, that there is no communication and no data shared between devices – all interaction is based on locally sensed data. Our techniques and applications can therefore be implemented as a standalone handheld device in the future.

Hybrid IR/Visible Light Projector

We modified an off-the-shelf Optoma PK102 handheld projector that uses the Texas Instruments DLP pico chipset (dlp.com/pico), a commonly used light engine found in many commercial handheld projectors. Its optical assembly consists of three high power LED light sources emitting red, green and blue light. We replaced the red and green LEDs with equivalent IR LEDs (Osram SFH4232), allowing us to project an IR image without modifying the projector optics (Figure 5). Two IR LEDs create a sufficiently bright image to be used for marker tracking with the on-board IR camera.

The remaining green LED was replaced with an equivalent white LED (Osram LUWW5AM) to produce a brighter visible projected image. Due to the properties of the projector's optical assembly, undisclosed by the manufacturer, the projected visible image changes to a light green color. To provide the white LED with sufficient voltage (3.2V), we connect it to the blue LED power source that has the highest source voltage (~2.92V). The red and green power sources (~1.6V & ~2.5V respectively) are used with the IR LEDs that have lower voltage requirements (1.5V).

Although our prototype restricts the working color palette to monotone, our aim was not to engineer an entirely new projection device but to evaluate the feasibility of projected IR markers for tracking and registration. Designing a full RGBIR (Red, Green, Blue, Infrared) pico-projector can be achieved by numerous manufacturers, and is therefore beyond the scope of our current work.

On-board Sensor Assembly

We equipped our handheld projector with a PointGrey Flea3 (www.ptgrey.com) IR-sensitive black and white camera mounted above the projector (Figure 6). We use an IR pass filter to cut visible light from the camera image and avoid interference with fiducial marker tracking in the IR spectrum. The camera is fitted with a fixed focal length 3mm wide-angle lens that is mounted directly above the projector lens for optimal optical alignment.

We enable gestural interaction using an IMU with a three-axis accelerometer, gyroscope, and magnetometer. It enables us to measure absolute orientation of the device, as well as relative acceleration and angular velocity. An ultrasonic ranging sensor is used to determine the distance to the projection surface. Because the projected image is much brighter when the device is closer to the projection surface, the ranging sensor allows us to dynamically adjust the image threshold level for more robust marker recognition.

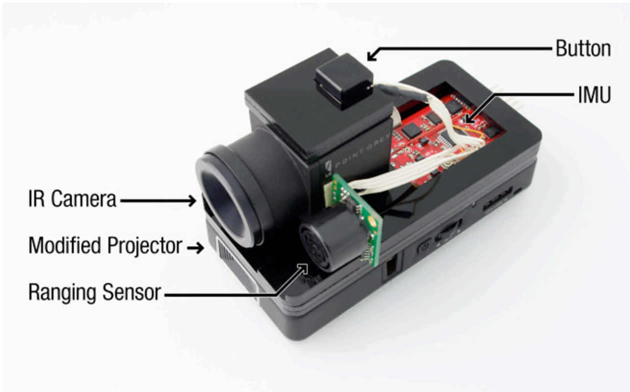

Figure 6: The *SideBySide* prototype with IR/visible light projector and an on-board sensor assembly.

The size of our prototype allows a user to hold and manipulate the device with a single hand. Because the system does not require any additional instrumentation of the environment, the device can be self-contained and reduced in size significantly. Many of the individual components used in our prototype can be found in current generation mobile devices. A compact and relatively inexpensive commercial implementation of our platform is very much possible.

SIDEBYSIDE SOFTWARE PLATFORM

The *SideBySide* software platform provides functionality to combine visible and IR imagery, track and register projected images using fiducial markers, resolve overlapping fiducial markers, and optically communicate arbitrary information between devices. In this section we discuss implementation details of the software platform.

Combining Visible and Infrared Imagery

To project an invisible image in the near-IR spectrum we load a monotone image into the red and green channels of an OpenGL frame buffer. The visible image is loaded from a separate monotone image into the blue channel (Figure 3) using OpenGL's additive blending mode. This combines the channels in real time using the graphics processing unit. Ghosting between channels, i.e. leaking from the IR channel into the visible channel (or vice versa), can be resolved by setting the system display color profile to 'generic RGB'. This allows the projection of clean, separate images in the visible and IR spectrums. Our approach can scale up to work with a full RGBIR image, where the IR channel is handled in a similar way to how alpha channels are used to specify transparency.

Tracking and Registration of Projected Images

To enable multi-user interaction with handheld projectors, each device must know the spatial relationship between its projected image and others nearby. To determine these spatial relationships we have developed tracking and registration techniques using invisible projected fiducial markers.

Each device projects a unique, static, 'reference' marker in the center of its projection frame. The on-board camera has a larger field of view than the projection frame, so it can observe markers from multiple devices (Figure 3). Each device firstly observes its own reference marker, then observes markers projected by other devices, next it determines the location and orientation of each marker in relation to its own marker, and finally estimates the size and orientation of the projection frames.

Our current implementation uses the *ARToolkitPlus* library [22] for marker tracking. A one-time calculation of the intrinsic camera parameters and homography matrix is performed to correct for lens distortion (Figure 7a). Once we have registered the markers in camera coordinates (Figure 7b), we remap them to the projector coordinate system. Remapping is based on the location of the projected marker in the scene and the scale ratio between the original pixel size of the projected marker and the size of the same marker observed by the camera. This process provides us with spatial information in the coordinate system used to design interactions and display graphics (Figure 7c).

Overlapping Markers

In multi-user interaction scenarios with handheld projectors, two projection frames often overlap. This can be accidental or an essential element of the interaction. For example, Cao et.al. used overlapping images for file transfer and magic-lens style navigation of large images [2]. Robust tracking of overlapping projected markers is therefore a key requirement for the *SideBySide* system.

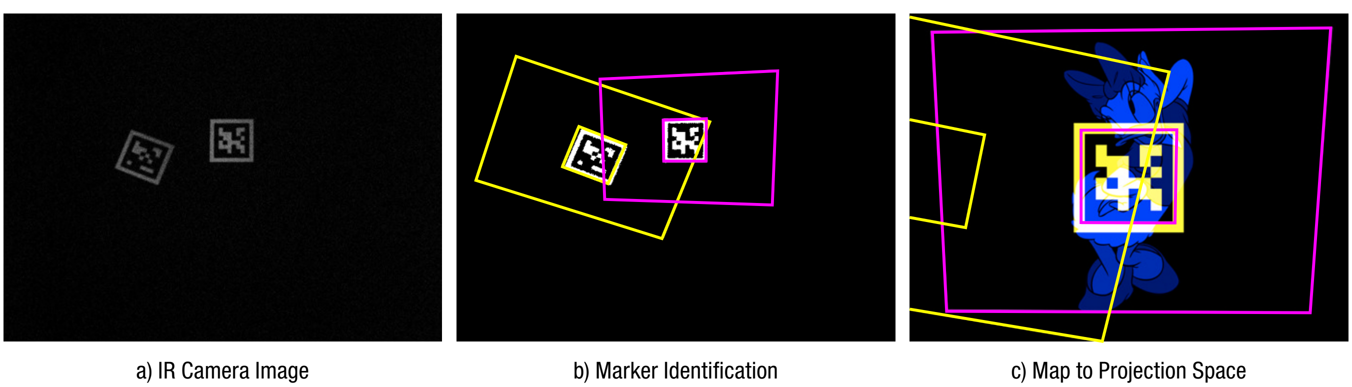

Figure 7: Invisible projected markers are perceived by an IR camera (a), located and identified (b) and then mapped from the camera coordinate space to the projection coordinate space (c).

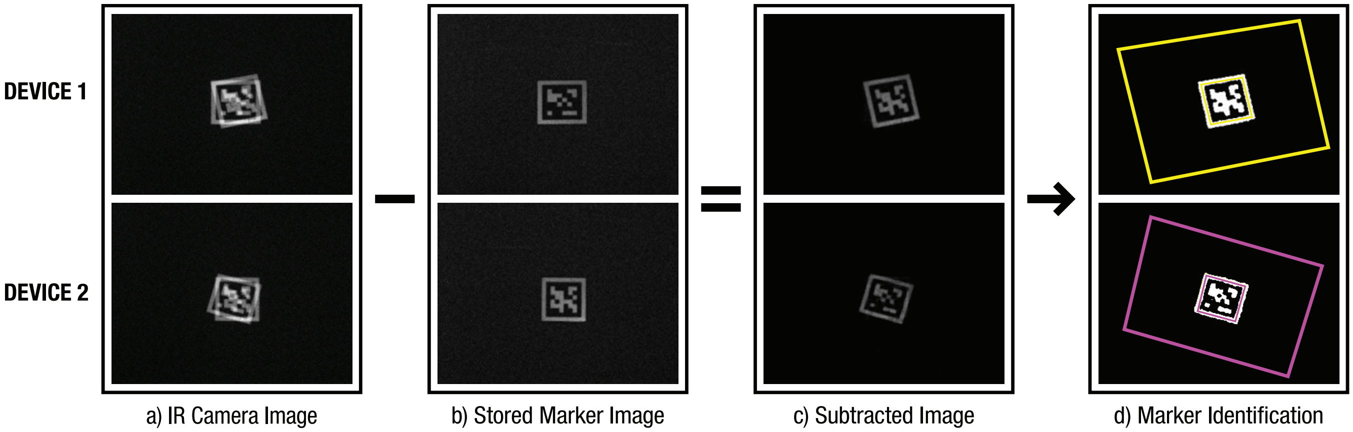

Figure 8: Recovery and tracking of overlapping projected markers.

Unlike printed markers that are typically opaque, projected markers can overlap and still leave adequate information to identify each marker individually. The overlapping region increases in brightness due to the additive nature of projected light (Figure 8a). A reference marker projected in the middle of the projector frame is known, unique and relatively static (Figure 8b). By subtracting it from the scene, we can recover the markers projected on top of it by other devices (Figure 8c). The recovered markers can then be identified and tracked (Figure 8d).

A naïve approach to marker subtraction would be to store a static binary representation of the reference marker in memory. However, environmental conditions such as lighting and properties of the projection surface constantly change as users move throughout the environment. Our marker recovery procedure continuously adapts to the changing environmental conditions by refreshing the stored image of the reference marker when there is no overlap, resulting in more robust performance.

Due to the close optical alignment of the camera and projector axes, projected images observed by the camera are relatively invariant to positional changes of the camera in relation to the projection surface. Even when the user points the device towards the projection surface at an acute angle, the projected markers appear relatively static and rectangular to the camera (Figure 9). Our adaptive marker subtraction technique is robust in situations when the angle of projection is constantly changing.

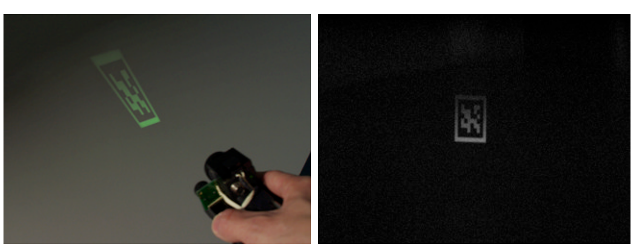

Figure 9: Projected markers that look distorted to the user (left) will appear relatively square when viewed from the device camera (right).

Optical Communication

To communicate events such as button presses, or changes in the state of the application we use optical communication by projecting invisible *event markers*. Event markers are standard fiducial markers used exclusively for symbolic data communication rather than location tracking (Figure 10). To project dynamic event markers, the sender device firstly identifies an empty region in the projection frame. It then projects a marker into this region for a given duration, in our implementation 1000ms. The event marker is projected towards the viewing area of the other device's camera, minimizing the possibility that it will fall outside the camera field of view. The camera on the receiving device observes the marker, recovers its ID, and executes the appropriate application-specific response. A total of 4096 marker IDs can be used with the *ARToolKitPlus* library to communicate between devices.

To measure the time taken for an event marker to be sent from one device and received by another device we performed a simple latency test. We recorded the transmission time of 100 event markers while both projectors were in a static position. All markers were successfully transmitted with a mean latency of 121.45 ms and standard deviation of 8.49 ms. This latency is roughly half the typical minimum human response time of 240 ms [4] and suitable for general-purpose interaction.

Communication bandwidth between handheld projection devices can easily be increased with the use of more so-

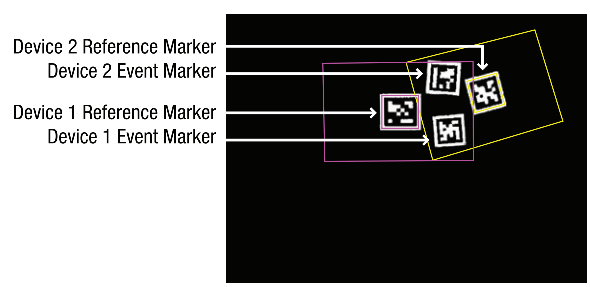

Figure 10: *Event markers* communicate event and state change information between devices.

phisticated spatial encoding techniques, such as QR codes, or temporal encoding techniques, such as modulated IR light. These techniques allow larger amounts of information to be transmitted between handheld projection devices such as URLs, geo locations, address data, status updates, calendar events, and email addresses. Optical communication can also be used to facilitate a 'heavyweight' network connection, such as WiFi or Bluetooth, for even larger files or secure communication between devices.

LIMITATIONS AND TRADEOFFS

Our current prototype has several limitations and tradeoffs.

Accuracy

We conducted an experiment to determine how precisely we could align two shapes projected from different projectors using our tracking system (Figure 11). We measured the distance from each corner of a square shape (Device 1) to each tip of an identically sized 'x' shape (Device 2). We took measurements in ten different locations distributed randomly across the projection area, for a total of 40 measurements. The mean misalignment between the two shapes was 5.28mm, with a standard deviation of 3.34mm. Although this level of accuracy is not ideal for perfect alignment of images, we found it to work well for general-purpose interaction where alignment is not the most critical requirement. Indeed, even professional level motion capture systems struggle to perform with a perfect level of precision [2]. Optimization of the underlying *ARToolKitPlus* tracking library can further improve the accuracy of the *SideBySide* system.

Projection Brightness

Images projected from handheld projectors have limited brightness and it is important for the projection device to remain in close proximity to the projection surface. In our current implementation, increasing the distance beyond 100 cm causes both the visual and IR projected images to dim. This makes it difficult for users to view projected content and decreases the reliability of marker tracking. Handheld projectors are typically used at a range of 60-70cm from the projection surface, where projected imagery is clear and easily seen [26]. Our current implementation functions reliably within this range.

Natural Light

A well-known problem for interactive systems using IR cameras and computer vision techniques is sunlight and other sources of IR light in the environment. This restricts usage of our system to settings without an abundance of natural light, such as indoor spaces with controlled lighting or outside spaces at night.

Camera Setup

A number of general camera decisions need to be made based on the application scenario. Motion blur can be an issue when the device is used for quick gestural interaction. Reducing the camera exposure time can help minimize motion blur; however, this comes at the cost of a dimmer image, decreasing tracking performance. The choice is

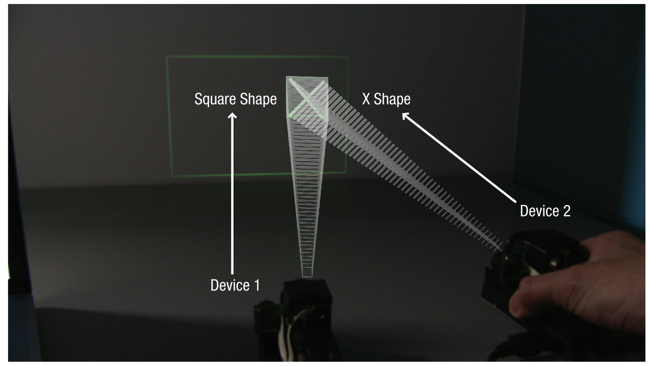

Figure 11: Testing tracking accuracy by aligning two identically sized shapes from different projectors.

therefore application dependent. When quick gestural movements are required, shorter exposure times should be used. When quick gestures are not required, longer exposure times will allow for better tracking performance.

Camera lens selection introduces another trade off. Wide-angle lenses capture a larger area well beyond the bounds of the projected image. This is useful to detect markers even when they are located at relatively large distances apart. However wide-angle lenses cause each marker to occupy fewer pixels of the camera image and result in coarser tracking and reduced accuracy.

Multi-user Interaction

Relative tracking limits the type of interactions that can currently be implemented with the *SideBySide* system. Camera-based systems for absolute localization, however, are a well researched area [11,19] and could easily be combined with our relative tracking approach.

Tracking overlapping markers is currently limited to either two overlapping markers at one time or to multiple markers overlapping different areas of the reference marker. It will fail when two markers overlap on top of the reference marker. This limitation can be overcome by using temporal sequencing of projected markers, so that each device projects a marker in sequence one after the other. Temporal sequencing can be also used to increase the number of projected event markers used at one time. With our current implementation, as the number of event markers increases it becomes increasingly difficult to locate free space for them to be projected. By projecting them one after the other this issue can also be resolved. We elected to leave the investigation of these techniques for future work.

Performance

Our current implementation runs two separate handheld devices on a single Apple MacBook Pro (2.89 GHz Intel Core 2 Duo) at a frame rate of ~45fps. Camera imagery is processed at 640x480 pixels and graphics are outputted at 800x600 pixels before being downscaled to the 480x320 resolution supported by the DLP projection engine. The underlying *ARToolKitPlus* tracking library is designed for use with mobile devices and [22] lists numerous optimizations that are applicable to our system. We believe an embedded version of *SideBySide* will run sufficiently on a current generation mobile device and within 2-4 years will match the performance of our current implementation.

APPLICATIONS

The *SideBySide* system supports a range of co-located multi-user interaction scenarios. By tracking the spatial relationship between multiple projected images, the system is ideally suited to interaction techniques utilizing the distance and angle between projections. Angle and distance can be used for continuous control, e.g. to zoom in/out, or as a binary control to trigger events, e.g. when the two projections overlap. Properties such as the alignment of the projections, the highest/lowest projection, or the relative size of the projections can be used to support further interaction techniques. With these basic building blocks *SideBySide* can support a diverse range of applications. We have developed example applications that focus on three areas: Mobile Content Exchange, Games, and Education.

Mobile Content Exchange

Handheld projectors can be used to support content exchange between devices in ad-hoc interaction scenarios.

Contact Exchange

The *Contact Exchange* application shows how common procedures such as exchanging contact information can be performed with the *SideBySide* system. One user scrolls through their list of contacts by tilting the device up and down. When she finds the contact she wants to exchange she presses a button and drags the contact on to the recipients projected address book (Figure 12, top). At that stage an invisible QR Code with embedded contact information is projected, and the recipient's device scans the code to complete the transfer. This technique can also be adapted for use with multiple devices by projecting the contact information code into open space and having multiple recipients scan the code at the same time.

File Transfer

The *File Transfer* application illustrates how users can transfer files from one device to another. The sender begins by scrolling through the files on her device with simple flick gestures. When she locates the file she wishes to transfer she drags it on to the receivers projected folder (Figure 12, bottom). As with the *Contact Exchange* example, a QR Code is then projected that contains a location where the file can be retrieved. Depending on the implementation this may be a URL or a Bluetooth address.

The *Contact Exchange* and *File Transfer* applications are two specific instances of content exchange. They greatly simplify the often-cumbersome task of locating a specific device to connect to, by providing a direct 'project-and-drop' style interaction. Any digital content, such as photos, videos, music, URLs, applications, and messages, can be transferred using variations of these techniques.

Education

Handheld projectors offer a novel, lightweight platform for learning applications.

Question & Answer

The *Question & Answer* application is designed to teach basic vocabulary to young children as either a first or second language (Figure 13, top). The teacher begins by projecting a question and several written answers from her device. At the same time the student's device depicts a pictorial version of the answer. The student answers the question by pointing to the correct written answer and pressing the button. This activity works both in a one-to-one context

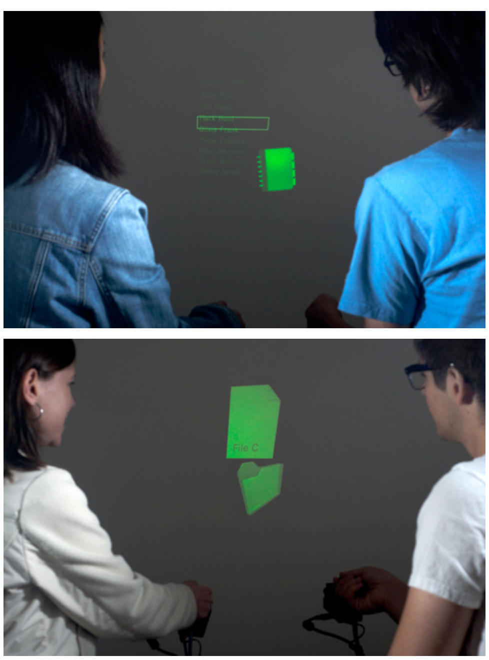

Figure 12: Mobile Content Exchange applications: *Contact Exchange* (top) and *File Transfer* (bottom).

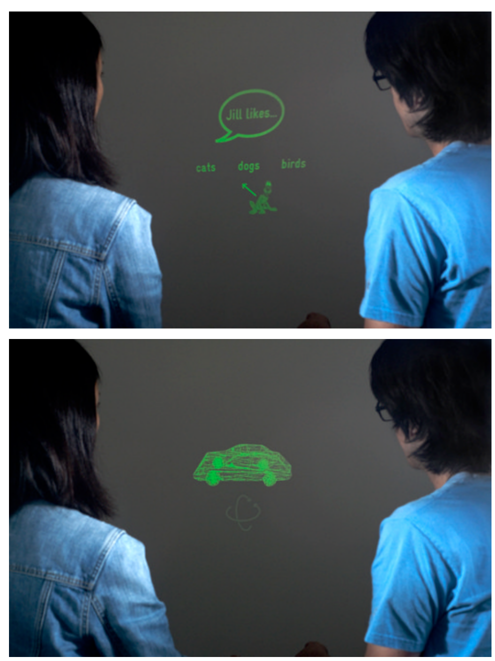

Figure 13: Education applications: *Question & Answer* (top), *3D Viewer* (bottom).

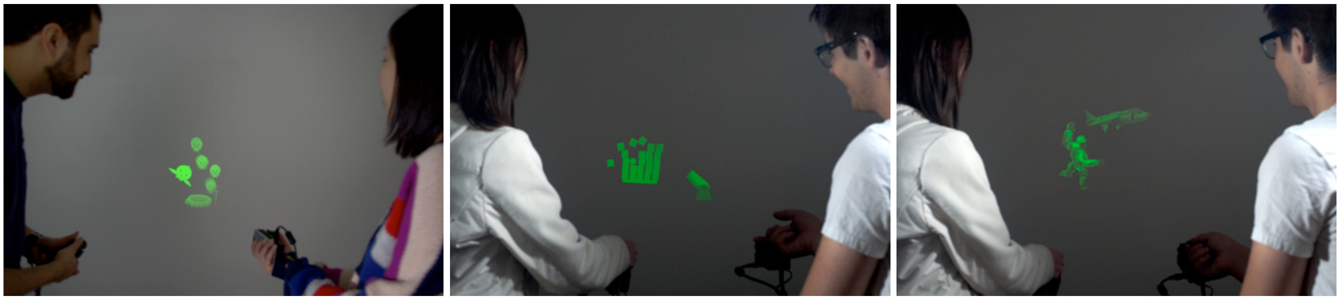

Figure 14: Game applications: *Falling* (left), *Cannon* (middle), *Gorilla* (right).

as well as with small groups of children taking turns to answer the questions. This allows for a more participatory style of interaction when compared to standard computers, but is more intimate than using a large overhead projector.

3D Viewer

The *3D Viewer* application lets two people control and view a 3D model together (Figure 13, bottom). One user projects an image of the 3D model and the other user projects a tool that controls the view. For example, the *orbit* tool maps the x,y difference between the two projection frames to the 3D rotation of the model. The *zoom* tool maps the distance between the two projection frames to the amount of zoom applied to the 3D model. These tools allow both users to actively control the vantage point or zoom level of the model. We envision the *3D Viewer* as a useful way to demonstrate 3D content in situations where a display or projector is not available. Models could range from mechanical objects, to architectural CAD renderings, to 3D molecular structures.

Games

Devices with embedded projectors offer an exciting new modality to play ad-hoc co-located games.

Boxing

The *Boxing* game is a playful rendition of a boxing match. Each player controls a boxer character by moving the projector (Figure 1). A single button on the projector is used to throw punches, but these land only if the player's character is within striking distance of the opponent. The game is played without a boxing ring, meaning the players are free to roam around and make full use of any available space. To communicate a punch being thrown we project event markers for the other device to read. If the distance between the two characters is within striking range, the other device responds to the punch. We found that despite the quick device movement during the game, the tracking system performs well.

Falling

In the *Falling* game two players work together to safely guide falling chicks into their nest. One player controls the nest location and direction of the falling chicks, and the other player controls the 'mama bird' to push the chicks into the nest (Figure 14, left). When one chick falls into the nest, the location of the nest shifts, and both players must re-adjust their positions to guide the next chick into the new nest location. We use a simple physics simulation to animate the falling chicks and IMU sensor data to make sure they are always falling downwards when the projector is rotated. We create an invisible physics object mapped to the location and orientation of the mama bird, so the chicks projected from one device appear to collide and interact with the mama bird projected from the other device.

Cannon

In the *Cannon* game two players work together to knock a stack of bricks off a platform by firing a projected cannon ball from one screen to another (Figure 14, middle). One user controls the location of the bricks by moving their projector, and the other user controls the location of the cannon in a similar manner. Pressing a button on the device fires the cannon ball from one projection frame to the next. Based on the relative location of the two projection frames the ball may miss the bricks or collide to knock them off the platform. To determine the location of the ball we project an event marker when the ball is fired. The device receiving the ball reads this marker and creates another ball with the same location and trajectory in its own projection space. When the ball is about to switch between projection frames, it is removed from one frame and made visible in the other. This gives the impression, due to persistence of vision, that a single ball has traveled across the two frames.

Gorilla

In the *Gorilla* game one player controls a gorilla character and the other controls a rescue plane (Figure 14, right). The rescue plane must try to capture the gorilla by moving in close and launching its net. The gorilla can resist being captured by punching at the rescue plane to destroy it. The button on each device triggers either the gorilla attacks or the rescue plane net. When the net is launched at the right angle and distance, the gorilla is captured and transfers from one projection screen to the other – temporarily hanging from the bottom of the rescue plane in the net.

CONCLUSIONS AND FUTURE WORK

We have presented *SideBySide*, a novel hardware and software platform for ad-hoc multi-user interaction with handheld projectors. Through use of a device-mounted camera and a hybrid visible/IR light handheld projector we have shown the viability of tracking projected content with invisible fiducial markers. Unlike previous systems that dramatically limit mobility by relying on infrastructure embedded in the environment, our system is completely self-contained and can be deployed as a handheld device.

Through the development of a range of example applications we have shown the wide and varied interaction scenarios where *SideBySide* can be put to real world use. Future work on the *SideBySide* project will continue to explore the possibilities of multi-user interaction by developing new applications and conducting formal user studies.

Enabling separate handheld projection devices to interact together is an important step towards truly seamless interaction across the greater ubiquitous computing landscape. We envision a day when digital content can traverse the boundaries of individual screens for fluid interaction between devices, people, and the physical environment.

ACKNOWLEDGMENTS

We thank Christopher Ioffreda, Yu Willis, Cheng Xu and Jo Ana Parker for assistance with photography, Sharon Hoosein for illustration assistance, and Mark D Gross for reviewing an early draft of this paper.

REFERENCES

1. Beardsley, P., van Baar, J., Raskar, R. and Forlines, C. Interaction Using a Handheld Projector. *IEEE Computer Graphics and Applications 25*, 1 (2005), 39-43.
2. Cao, X., Forlines, C. and Balakrishnan, R., Multi-user Interaction Using Handheld Projectors. In Proc. *UIST '07*, ACM (2007), 43-52
3. Cao, X., Massimi, M. and Balakrishnan, R., Flashlight Jigsaw: An Exploratory Study of an Ad-hoc Multi-player Game on Public Displays. In Proc. *CSCW '08*, ACM (2008), 77-86.
4. Card, S.K., Moran, T.P. and Newell, A. The Model Human Processor: An Engineering Model of Human Performance: Vol. 2 Cognitive Processes and Performance. in Boff, K.R., Kaufman, L. and Thomas, J.P. eds. *Handbook of Perception and Human Performance*, 1986, 1-35.
5. Chan, L.-W., Wu, H.-T., Kao, H.-S., Ko, J.-C., Lin, H.-R., Chen, M.Y., Hsu, J. and Hung, Y.-P., Enabling Beyond-surface Interactions for Interactive Surface with an Invisible Projection. *UIST '10*, ACM (2010), 263-272.
6. Grundhöfer, A., Seeger, M., Hantsch, F. and Bimber, O., Dynamic Adaptation of Projected Imperceptible Codes. In Proc. *ISMAR '07*, IEEE (2007), 1-10.
7. Hosoi, K., Dao, V.N., Mori, A. and Sugimoto, M., CoGAME: Manipulation Using a Handheld Projector. In Proc. *SIGGRAPH '07 Emerging Technologies*, ACM (2007), 2.
8. Kane, S.K., Avrahami, D., Wobbrock, J.O., Harrison, B., Rea, A.D., Philipose, M. and LaMarca, A., Bonfire: A Nomadic System for Hybrid Laptop-Tabletop Interaction. In Proc. *UIST '09*, ACM (2009), 129-138.
9. Köhler, M., Patel, S.N., Summet, J.W., Stuntebeck, E.P. and Abowd, G.D., TrackSense: infrastructure free precise indoor positioning using projected patterns. In Proc. *Pervasive '07*, Springer-Verlag (2007), 334-350.
10. Lee, J., Hudson, S. and Dietz, P., Hybrid Infrared and Visible Light Projection for Location Tracking. In Proc. *UIST '07*, ACM (2007), 57-60.
11. Nakazato, Y., Kanbara, M. and Yokoya, N., Localization System for Large Indoor Environments using Invisible Markers. In Proc. *VRST '08*, ACM (2008), 295-296.
12. Ozuysal, M., Calonder, M., Lepetit, V. and Fua, P. Fast Keypoint Recognition Using Random Ferns. *IEEE Transactions on Pattern Analysis and Machine Intelligence 32*, 3 (2010), 448-461.
13. Park, H. and Park, J.-I., Invisible Marker Based Augmented Reality System. In Proc. *VCIP '05*, SPIE (2005), 501-508.
14. Pinhanez, C.S., The Everywhere Displays Projector: A Device to Create Ubiquitous Graphical Interfaces. In Proc. *UbiComp '01*, Springer-Verlag (2001), 315-331.
15. Rekimoto, J., Pick-and-Drop: A Direct Manipulation Technique for Multiple Computer Environments. In Proc. *UIST '97*, ACM (1997), 31-39.
16. Rekimoto, J. and Saitoh, M., Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments. In Proc. *CHI '99*, ACM (1999), 378-385.
17. Rukzio, E., Holleis, P. and Gellersen, H. Personal Projectors for Pervasive Computing. *Pervasive Computing* (2011).
18. Rukzio, E., Schöning, J., Rohs, M., Häkkila, J. and Dachselt, R. Ubiprojection 2010: Workshop on Personal Projection. 2010, http://eis.comp.lancs.ac.uk/workshops/ubiproject2010/.
19. Saito, S., Hiyama, A., Tanikawa, T. and Hirose, M., Indoor Marker-based Localization Using Coded Seamless Pattern for Interior Decoration. In Proc. *VR '07*, IEEE (2007), 67-74.
20. Shirai, Y., Matsushita, M. and Ohguro, T., HIEI Projector: Augmenting a Real Environment with Invisible Information. In Proc. *WISS '03* (2003), 115—122.
21. Sugimoto, M., Miyahara, K., Inoue, H. and Tsunesada, Y., Hotaru: Intuitive Manipulation Techniques for Projected Displays of Mobile Devices. In Proc. *INTERACT '05*, Springer (2005), 57-68.
22. Wagner, D. and Schmalstieg, D., ARToolKitPlus for Pose Tracking on Mobile Devices. In Proc. *CVWW '07* (2007), 139-146.
23. Weng, D., Huang, Y., Liu, Y. and Wang, Y., Study on an Indoor Tracking System with Infrared Projected Markers for Large-Area Applications. In Proc. *VRCAI '09* ACM (2009), 239-245.
24. Wienss, C., Nikitin, I., Goebbels, G., Troche, K., Göbel, M., Nikitina, L. and Müller, S., Sceptre: an infrared laser tracking system for virtual environments. In Proc. *VRST '06*, ACM (2006), 45-50.
25. Willis, K.D.D. A Pre-History of Handheld Projector based Interaction. *Journal of Personal and Ubiquitous Computing*, Theme issue on Personal Projection (2011).
26. Willis, K.D.D., Poupyrev, I. and Shiratori, T., MotionBeam: A Metaphor for Character Interaction with Handheld Projectors. In Proc. *CHI '11*, ACM (2011), 1031-1040.
27. Wilson, A.D., PlayAnywhere: A Compact Interactive Tabletop Projection-Vision System. In Proc. *UIST '05*, ACM (2005), 83-92.