All Across the Circle: Using Auto-Ordering to Improve Object Transfer Between Mobile Devices

Chengzhao Li*
University of Saskatchewan

Carl Gutwin†
University of Saskatchewan

Kevin G. Stanley‡
University of Saskatchewan

Miguel Nacenta§
University of St. Andrews

ABSTRACT

People frequently form small groups in many social and professional situations: from conference attendees meeting at a coffee break, to siblings gathering at a family barbecue. These ad-hoc gatherings typically form into predictable geometries based on circles or circular arcs (called F-Formations). Because our lives are increasingly stored and represented by data on handheld devices, the desire to be able to share digital objects while in these groupings has increased. Using the relative position in these groups to facilitate file sharing can enable intuitive techniques such as passing or flicking. However, there is no reliable, lightweight, ad-hoc technology for detecting and representing relative locations around a circle. In this paper, we present two systems that can auto-order locations about a circle based on sensors that are standard on commodity smartphones. We tested these systems using an object-passing task in a laboratory environment against unordered and proximity-based systems, and show that our techniques are faster, are more accurate, and are preferred by users.

Keywords: Object transfer; ad-hoc sharing; auto-ordering.

Index Terms: [Human computer interaction]: Interaction techniques—Pointing; [Collaborative and social computing]: Computer supported cooperative work.

1 Introduction

People commonly need to transfer objects and files from one mobile device to another. For example, conference attendees might decide to share business cards or research papers; a family gathered in a living room might share photos of vacations or grandchildren; and colleagues sitting around a conference table might need to share files or data related to a project. Typically, individuals will arrange themselves into circular or semi-circular physically proximate locations know as F-Formations [20, 26].

Current file transfer techniques (such as e-mailing the object or creating a link to a location in the cloud) require time and information. Instead, transfer could be accomplished using techniques that allow people to move objects by bumping the devices [14], making parallel gestures [15, 16], or providing a list view of all connected devices [11]. However, these techniques have limitations – for example, bumping requires close physical proximity, parallel gestures can take time to decide on and execute, and lists require that people know the mapping between the list entry and the device in the real world.

Another class of technique uses onscreen targets for transferring objects from one device to another [9, 28]. Target-based transfer is lightweight and natural compared to traditional file-sending mechanisms, because people can simply direct the object of interest toward the target. There are two main aspects to a target-based transfer technique: the local gesture used to select and direct the object, and the underlying infrastructure that identifies the target of the gesture. In this paper, we are interested in the second of these issues – identifying target locations.

In previous work, the target-identification problem is usually solved with approaches that require considerable infrastructure, such as magnetic sensors, infra-red cameras, multiple fiducial markers, depth cameras, or custom hardware [21]. Although a few infrastructure-free approaches have been developed (e.g., Virtual Compass [2]), these have high positional error rates, which make transfers in small circular groups infeasible.

The first of our auto-ordering techniques uses the smartphone's camera to find a fiducial marker (on paper or displayed on one of the phones themselves) placed in the middle of the group, and then uses the relative orientation of the marker observed by each phone to infer relative location. The second technique uses the smartphone's compass: users orient their phones toward a location at the center of the circle, and the relative angle each device reports can be used to resolve their positions.

In this paper we describe the two techniques, evaluate the capabilities and limits of each method, and assess performance in a controlled study that compares our techniques to a proximity technique and an un-ordered portal technique. Our studies found:

- Auto-ordering was faster than the other techniques, and less error prone than the portal technique.
- While participants were able to reduce transfer time using portals as they learned mappings, they never achieved better performance than the auto-ordering techniques.
- Participants overwhelmingly preferred the auto-ordering techniques to the portal and proximity techniques.
- The smartphone sensors underlying the techniques are accurate enough for groups of up to twenty people – many more than will typically be encountered in ad-hoc groups.

Our techniques provide a simple, intuitive and reliable solution for a common transfer situation — a small group gathered in an approximate circle — with sufficient accuracy and precision to reliably localize people around a circle.

2 Related Work

2.1 Location Sensing for Mobile Devices

A number of technologies have been explored for sensing the locations of objects in mobile contexts. Early systems used small transmitters to locate people and objects in an augmented environment. For example, the Active Badge system [36] sends ID information to sensors located around a building. To improve precision, ultrasound locators such as the Active Bat calculated position [30].

Recent approaches have used existing infrastructure or active sensing to provide position information. For example, researchers have used trilateration with WiFi signals to determine location (e.g., RADAR [1]). Motion-capture systems using magnetic tracking (e.g., the commercial Polhemus system) or infrared cameras (e.g., the Vicon system) provide precise 3D positions, but

* chengzhao.li@usask.ca
† gutwin@cs.usask.ca
‡ kstanley@cs.usask.ca
§ mans@st-andrews.ac.uk

Graphics Interface Conference 2016
1-3 June, Victoria, British Columbia, Canada

only within the range of the cameras or antenna. Chen et al [7] used the phone's built-in sensors to detect the spatial relationship between user and mobile device. Marquardt et al [25] built the Proximity Toolkit, based on a Vicon system and a KINECT sensor, to help developers easily obtain proxemics information in a room-sized environment. Kray et al [22, 23] used external cameras to track markers on devices. Li et al [24] and Schwarz et al [34] both looked at ways of arranging devices using marker based techniques for distributing content. Researchers have also investigated vision-based systems that track fiducial markers [19] or common image features [8], that allow pointing at IR-based tags on objects [35]. The Relate system [21] uses custom-built ultrasonic USB dongles to calculate relative positions to other devices. While the accuracy and precision were sufficient for file transfer activities, the custom hardware makes this solution unavailable to most users. Tracko, developed by Jin et al [17], synthesized Bluetooth low energy signals and inaudible stereo sound to deduce 3D locations of nearby devices. However, this system only works well within 1m, which makes it inappropriate for F-Formations that typically have diameters of 1-3m [20], and also performs poorly in noisy environments. The Virtual Compass system [2] uses Bluetooth received signal strength indicators (RSSI) and WiFi signals to calculate distances between devices and position them in a 2D plane. The system works without external infrastructure, but has a low accuracy – experiments showed that Bluetooth RSSI alone had a mean positioning error of 3.4m, that WiFi alone had a mean error of 3.9m, and that a combined technique had a mean error of 1.4m, with error of 2.7m at the 90^{th} percentile [2]. An error of 1.4m precludes the correct ordering of people standing around a circle. For example, taking a 1.4 m error and a two-meter diameter circular arrangement leads to an angular error of ±90 degrees along the circumference of the circle, making it implausible to use reliably.

2.2 Working Around a Circle

Researchers have studied the ways in which people organize themselves when they come together as a group (e.g., [10, 20, 26]). In particular, the physical arrangements that people use have been examined by Kendon [20], who described circular clusters called Facing Formations (F-Formations).

F-Formations can occur in many different settings, may be physically larger or smaller depending on the situation (e.g., a conversational group may have a smaller arrangement than a work group around a table), and may be only approximately circular (e.g., corner-to-corner, side-by-side, or face-to-face arrangements are also possible). F-Formations typically comprise between two and five people [26], and gestures or objects within the space between these people can become the focus of the interaction.

HCI researchers have used F-Formations as the basis for interaction techniques. In particular, Marquardt and colleagues [26] developed techniques that allow easy object transfer when people are beside one another, that provide screen previews based on device tilt, and that allow full screen sharing when people are in an F-Formation. However, these techniques required an external tracking system (i.e., a Vicon system).

2.3 Object Transfer Techniques

Several researchers in HCI have considered the problem of how to move objects from one device to another in multi-display environments. Below, we summarize this research based on three organizing principles identified in a detailed survey by Nacenta and colleagues [28]: the referential domain, the display configuration, and the control paradigm. Other researchers such as Rädle et al. [32] have proposed extensions to this architecture.

The *referential domain* is the way in which users refer to different displays. Two main reference types are spatial arrangement and named displays. Several object transfer techniques have been developed for both types: for example, real-world spatial locations were used with the early Put-That-There technique [5] and the Pick-and-Drop [33] technique; on-screen representations of real-world locations are used with "world-in-miniature" views (e.g., ARIS [4]); and arbitrary spatial locations are used with many portal-based techniques. Named displays, in contrast, use non-spatial methods such as text, numbers, or colors to refer to other devices. Many techniques have used this approach, including lists of displays in Multibrowsing [18] and Mighty Mouse [6], color-coded icons for different displays [10], as well as contact lists and shared-folder icons in commercial systems. Finally, some techniques allow users to cycle through a set of displays [3] by pressing a key or button.

Second, the *display configuration* is the way in which the displays are organized in physical and digital space. This dimension affects techniques that use direct manipulation to transfer objects, because the arrangement of displays limits the kinds of transfer actions. Stitched displays allow object transfer by moving the object across the edge of the display (e.g., [26]). Stitching can cause problems when different users see the displays from different directions [28], and so other techniques use the perspective of the user to organize display locations (e.g., Perspective Cursor [29]). Finally, "literal" techniques can use the actual devices themselves rather than their locations to enable object transfer. These techniques use parallel gestures (e.g. [15]), shaking [16], proximity, or bumping [14] as the mechanism for indicating which device is the target. One main drawback of these literal techniques is that they are limited by the physical reach of the user [28].

Third, the *control paradigm* is the way in which people actually perform the transfer. Previous techniques have used both of these mechanisms. For example, open-loop transfer was used by the Flick [37] and Multi-Monitor Mouse [3] systems; closed-loop transfer is used by all techniques that have a visible representation of the target (including world-in-miniature systems [5], pantograph-style movement [13], and portal-based techniques).

In terms of these dimensions, the auto-ordering techniques we developed use a spatial referential domain (using real-world locations), a perspective-based display organization (i.e., targets are arranged correctly for each person's view), and either open-loop or closed-loop control (since the technique supports both flicking and portal-based transfer).

3 THE AUTO-ORDERING TECHNIQUES

Auto-ordering of people engaged in F-Formations can be viewed as a technical problem of determining the relative location of the users, and faithfully rendering the relative locations on each user's device. The general problem of determining relative location can be quite complex, as it requires determining the position and orientation of individuals with respect to a coordinate frame. General positioning technology using GPS, WiFi, or Bluetooth localization do not have the spatial fidelity to resolve the relative locations of individuals standing in a circle, and dedicated hardware can be cumbersome or difficult to install. However, our assumption of users in an F-Formation allows us to constrain the problem to the point where sensors commonly found on mobile devices can perform auto-ordering registration. Figure 1 shows a schematic representation of the problem. Given four individuals (three shown, one holding the phone), we need to share a file with only one of them in an ad-hoc network. Labeling can happen through tags (color) or through relative location, (ordering around the circle). If the ordering were arbitrary, it would be more difficult for the user to assign the tag to the on-screen location.

Our design goals are to provide rapid operation that facilitates sharing, to minimize user error, and to have little or no physical setup required. We have developed two solutions, employing different sensor suites common to today's smartphones. The first is a marker-based technique that uses a fiducial marker to provide a visual reference that each phone can use to calculate its relative position. The second technique leverages the orientation sensors on the phone (accelerometer, gyroscope and compass) to determine the relative orientation of each user, which is then mapped to a circle around which they are standing.

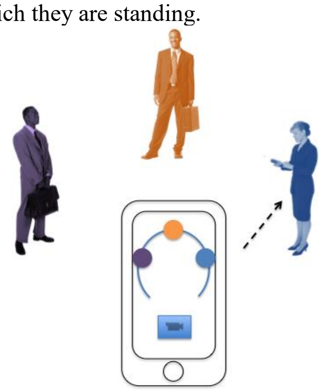

Figure 1: A typical transfer setting: four people in a circle, and the person holding the device must determine which on-screen portal corresponds to which person in the real world.

We assume that users roughly face each other (i.e., they form simple open or closed shapes where people face inwards) and are at roughly the same distance from the group's center, which is common in self-ordering behavior of small groups [20, 26]. These assumptions provide us with a critical insight that allows us to address the problem more accurately and reliably than previous attempts. Because we can assume that users are arranged around a circle, we can constrain the solution space to that manifold. To express ordering about a circle, one does not need relative location, but rather relative orientation – the polar coordinates of the person on the circle. Because the radius of the circle is fixed, or at least quasi-static with respect to the interactions, only the angular coordinate is required to determine relative position.

A key component of auto-ordering techniques is that they allow portal locations or flicking directions to be arranged to match the location of the actual people or devices in the real world. The general psychological principle of stimulus-response compatibility [28] predicts that digital arrangements that correspond to the physical world will be faster and produce less cognitive load because they allow people to use information provided by the real world instead of having to remember an arbitrary mapping (e.g., it may be easier to flick a document toward a real-world printer than selecting the printer from a list). Nacenta [27] showed mixed results when applying the idea of stimulus-response compatibility to transfer tasks – and no published experiments have assessed the use of world-to-interface correspondence for object transfer.

3.1 Marker-Based Ordering

Once users have organized themselves in an F-Formation, it is trivial to introduce a marker at the center of the circle, either as a piece of paper, or more likely as the screen of one of the participant phones. A visual or fiducial marker is a standardized shape or mark usually a heavy square, easily detected using image processing techniques. ARToolkit [19] is a software library used for building Augmented Reality applications, which users orientation sensors and image processing to detect the pose of a phone with respect to a marker, and optionally can render virtual objects over the marker. We use the Android AndAR implementation of ARToolkit.

AndAR returns a matrix representing the pose of the phone with respect to the marker. We used the matrix to calculate the relative rotation between marker and phone. The center of the circle is defined by the fiducial marker. Once each phone has its angle, it is trivial to resolve the position around the circle. Users on a display can then be rendered by calculating angular differences.

Typical implementations employing QR codes use paper markers, but we extended the technique to use a marker that is displayed on one of the phones in the circle (see Figure 2). In this variant of the technique, one person moves their device forward so that it is in view of the other phone's cameras, and this central phone displays a marker that is similar to the paper version.

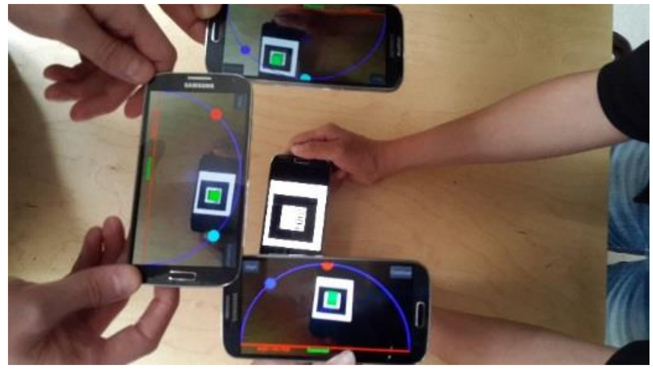

Figure 2: Marker technique using one of the phones to display the marker (no paper marker needed).

3.2 Compass-Based Ordering

Most mobile devices provide sensors which detect the orientation of the phone with respect to the Earth, with the direction of gravity providing the vertical axis and a compass heading providing orientation around that axis. Fusing data from the compass, gyroscope and accelerometers, Android provides an abstract sensor class *orientation* describing the phone's angular pose. Compass heading can be translated into relative position about a circle if the pose of the phone with respect to the circle is fixed.

In the compass-based registration system, users point their phones at the arbitrary center of a circle. If users are arranged in a closed shape such as a circle or square, this is trivial, as the arbitrary center is in the center of the shape they form. For more truncated shapes such as line segments or semi-circles, the alignment is only slightly more complex, and can be aided by selecting objects in the real world to serve as the center of the circle.

Once all users are pointed at the center of the circle, their positions around the circle can be inferred by the relative orientation they have with magnetic North. Assuming that magnetic North is always zero, then the relative location of each participant around the circle is simply the angular location of their reported orientation minus the angular location of the current user.

4 Technical Assessment

For the proposed systems to be functionally useful, the sensors must return sufficient angular resolution to localize a sufficient number of users in a circular arrangement. We consider three classic parameters: sensitivity, (the degree to which sensed values represent reality), precision (the degree to which a sensor returns the same value for the same stimulus), and span (the range over which sensor readings are valid), to characterize the sensors and to determine the number of people that can be reliably localized. We hold that at least five people should be reliably localizable around a semicircle, based on Bahl's work [1], therefore an angular resolution of at least 36° is required. We note that the Bluetooth-based Virtual Compass [2] is insufficient for practical use, as its mean angular error is almost triple this requirement.

The device used for testing was the Samsung Galaxy S4 (1.6 GHz processor, 5-inch 1080p display, Android 4.4.2).

4.1 Marker-Based Ordering

Fiducial markers can provide highly accurate (cm-scale) and stable pose estimates from camera images, ensuring that ordering will always be maintained. Because of the high-fidelity cameras on modern smartphones, the accuracy and precision of the marker-based localization are well below the 36° threshold specified. Typically, angular positions in of less than 3 degrees could be easily resolved. The primary limiting factor in the marker case is not the accuracy or precision, but the span, because the marker must always remain in view of the camera, constraining the number of possible angles. Two parameters of span must be considered: the distance and angle at which the marker can be resolved. The markers were mounted on a wall with certain distances to the ground, and also on a table. Marker size in pixels is a function of the physical size of the marker and its distance from the camera.

Table 1: Maximum ranges for different sized markers.

Marker Size (mm)	Maximum Range (mm)
50	1400
80	2300
100	2600
120	3200
150	3900

A 50 mm marker – a size which could be displayed on a large number of mobile devices if using our phone-based marker technique – can be reliably captured at a distance of 1.4 m (sd = 0.09), which is a reasonable maximum distance to the center of the circle for conversational arrangement such as those in [26].

The camera was moved at a constant distance of 1 m from the marker at varying rotation until the marker was no longer recognized. The maximum detectable angle with respect to vertical is approximately 75°, or almost vertical, meaning the camera should be able to resolve the angle from the marker in most comfortable-to-hold positions. Because of the highly accurate pose estimates, over an acceptable span and distance, the marker-based system is sufficient to provide automatic radial ordering.

4.2 Compass-Based Ordering

The abstract orientation sensor has a spans 360° but is often noisy, impacting the accuracy and precision of the position estimate. We recorded the reported orientation of eight phones arranged in a circular pattern for 15 minutes.

Approximately 9000 data points were recorded on each phone. The test result shows that the distribution of the angle for each direction follows the Gaussian distribution, with a mean of -3°, which establishes more than sufficient accuracy. Angular error for each phone was calculated to determine the precision. Figure 3 shows the histogram of measurement of angle errors of all measurement angles. The 99% confidence interval lies at ±6.67°, allowing up to 54 people to be placed around the circle in the limit. Practically, a much smaller number will need to be localized. Based on our criteria, the compass has sufficient accuracy and precision, and span to provide the quality of service required. In future work, we will confirm these experimental results with real-world groups.

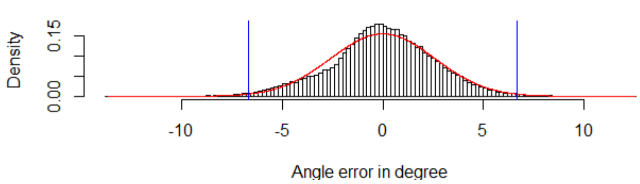

Figure 3: Histogram of angular error (red curve: best-fit Gaussian; blue bars: 99% confidence interval).

5 COMPARATIVE EVALUATION

We carried out a controlled experiment to compare the performance of auto-ordering techniques to two existing approaches – unordered portals (which provide an on-screen proxy for each person, but ordered arbitrarily), and a proximity technique which detects when two devices are physically close (using phones' NFC radios).

In all cases, a single phone was configured as the server, which maintained the known configuration of the phones and communications. Phones were dynamically added to the sharing representation as they logged into the app, and were connected to a Bluetooth star network via the server.

5.1 Experimental Conditions

The transfer techniques were implemented in a simple experimental system that asked participants to transfer objects to one of three other people (see Figure 4).

Wormhole. As seen in previous literature, our portal-based technique provided an on-screen proxy for each person; transfers were accomplished by dragging the object to the correct person's portal (Figure 4. Wormhole). Unlike the auto-ordered techniques, portals were ordered randomly.

Compass (Auto-Ordered). As described above, the Compass technique uses each device's compass reading to create a circular ordering for all devices (Figure 4. Compass). In order to make interactions as similar as possible across different conditions, the Compass technique used a proxy-based transfer action in which participants dragged objects to on-screen portals – although the portals were now ordered to match the locations of the other people around the circle. Our implementation can also use flick-based transfer, but only proxy-based transfer was used for the study.

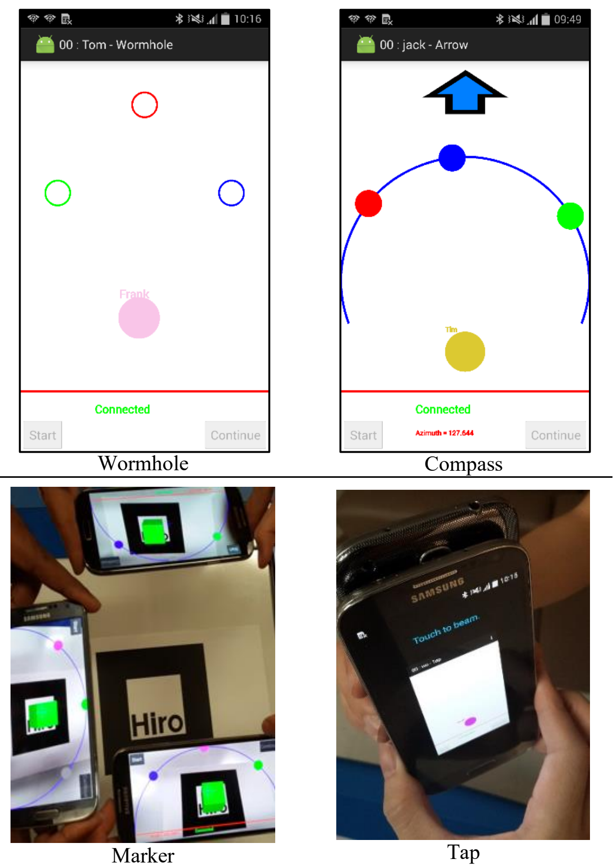

Figure 4: Transfer techniques used in the study.

Marker (Auto-Ordered). As described above, the Marker technique uses the device's camera to locate the fiducial marker and determine the orientation of the device to the marker; an ordering is then created using these relative orientations (Figure 4. Marker). As with Compass, the Marker technique used on-screen portals for transferring objects; the position of the portals was determined by the Marker-based ordering algorithm.

Tap. We developed a proximity-based transfer technique based on Hinckley's Bump system [14]. Tap uses Android Near-Field Communication (NFC) to control the object transfer. To use this technique, participants held their devices back to back; when the devices were close enough, the sender saw a popup message on their display, and tapped the screen to complete the transfer over Bluetooth (see Figure 4. Tap).

5.2 Participants and Apparatus

Thirty-two participants were recruited from a local university (ages 19-39, mean 26.6), in groups of four. All participants were frequent users of mobile devices (mean 23.6 hours/week). In seven of the eight groups, participants did not know one another.

The study used custom software developed with Android version 4.4.2, and was deployed on four Samsung Galaxy S4 devices (1.6 GHz processor, 5-inch 1080p display). The study was conducted in an open area of a research lab (approximately 8m by 8m). The floor was marked with a two-meter-diameter circle and four locations at the north, east, south, and west points of the circle.

5.3 Task

Participants were asked to transfer several objects to others in the group, using their mobile devices. We simulated a setting where people transferred objects to an ad-hoc group that would be together for only a short time – such as an impromptu meeting at a conference coffee break. To simulate this setting (in which people know that they want to transfer to a particular person standing in the circle), we had participants wear nametags with made-up names, and for each transfer, a circle (representing the object to be transferred) appeared on the participant's device with the name of one of the other people in the group. In the tap condition, they had to bring the phones in close proximity. For the Compass, Marker, and Wormhole techniques, the participant completed the transfer by dragging the circle to the correct on-screen portal (see Figure 4). In all portal conditions, the portals were labeled with the participant's username (not their real name). We did not show usernames on the portals because this would have allowed simple pattern-matching between the named transfer object and because, in a real-world situation, the transfer object would not show the intended recipient. To allow people to build a memory mapping between portals and people in the Wormhole condition, names could be shown by long-pressing anywhere on the screen (500ms).

5.4 Procedure, Study Design, and Hypotheses

Each group of four participants completed demographic questionnaires, and then were given an introduction to the four transfer techniques after completing informed consent consistent with our ethics approval. Groups worked with each of the four transfer techniques in an order balanced using a Latin square. For each technique, participants carried out five blocks of trials. In each block, participants completed three transfers (trials) to each of the other people in the group (in random order) for a total of nine transfers. To test the ability of participants to remap digital to physical locations when configurations change. After each block, the study simulated a new meeting of the four people – participants were moved to different physical locations around the circle, and the on-screen locations of the portals for all techniques except tap were reordered.

After each condition, participants completed an effort questionnaire based on the NASA Task Load Index (TLX) [12], asking about the technique they had just completed. After all four conditions, participants answered questions about their preferences.

The study used a 4×5x3 within-participant RM-ANOVA with factors *Transfer Technique* (Compass, Marker, Wormhole, Tap), *Block* (1-5), and *Repetition* to the same target (1,2,3 for each recipient). Dependent measures were transfer time and number of errors. Hypotheses were:

H1. Object transfer times for the auto-ordering techniques (*Compass* and *Marker*) will be faster than for either *Tap* or *Wormhole*;

H2. Error rate for *Compass* and *Marker* will be less than for *Wormhole*;

H3. Users will prefer the auto-ordering techniques over the other techniques.

5.5 Results

5.5.1 Transfer Time

As shown in Figure 5, mean transfer times ranged from about two seconds for the auto-ordering techniques to above twelve seconds for Tap. Figure 6 displays only the three faster techniques, and shows that Wormhole was slower in the first two blocks, and then the same speed as the auto-ordering techniques.

RM-ANOVA showed a significant main effect of *Technique* ($F_{3,93}$=973.6, p<.0001), and also a significant interaction between *Technique* and *Block* ($F_{12,372}$=7.35, p<.0001). Follow-up pairwise comparisons between conditions (using Bonferroni correction to maintain alpha of 0.05) showed that *Tap* was slower than all other conditions, and that both *Compass* and *Marker* were faster than *Wormhole*. We therefore accept hypothesis H1.

Observations during the trials suggested that the main reason that *Tap* was slower than the other techniques is that people needed to wait for the other person to be ready to carry out the technique – that is, *Tap* requires both sender and receiver to engage in the transfer, whereas the other techniques allow unilateral transfer.

Observations and participant comments also suggested that the reason for slower performance of the *Wormhole* technique was that participants needed to remember the mapping between the portals and the people in the real world. Our analysis of the number of times targets were repeated provides additional insight into this issue. Figure 7 shows the transfer times for the three portal-based techniques on each of the three repetitions per block. *Wormhole* was slower on the first trial (when the locations were unknown), and then similar in speed for the second and third trials. RM-ANOVA showed a significant interaction between Technique and Repetition ($F_{6,186}$=11.43, p<.0001).

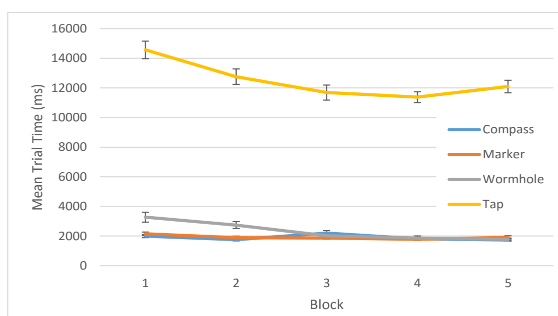

Figure 5: Mean transfer time (± s.e.), by technique and block.

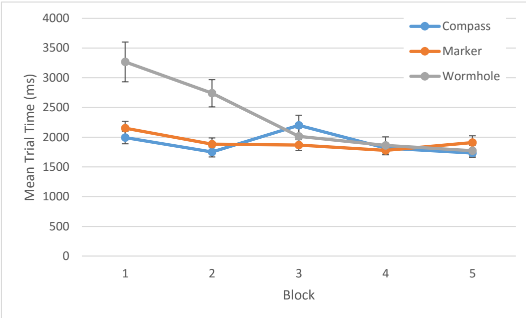

Figure 6: Mean transfer time (± s.e.) for portal techniques.

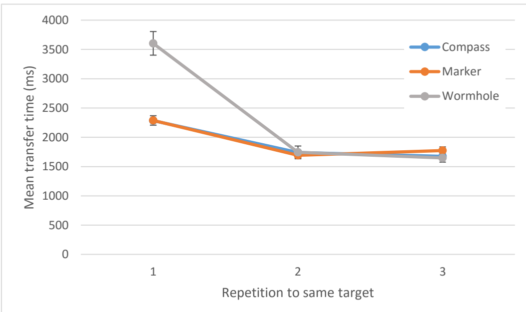

Figure 7: Mean transfer time by number of repetitions.

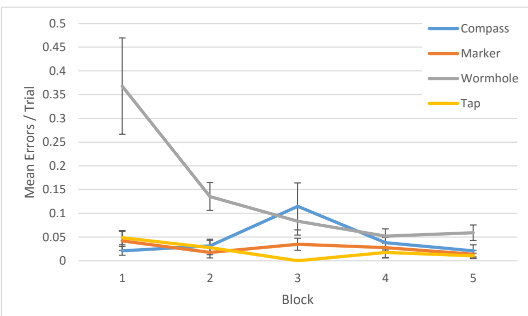

Figure 8: Error rate by technique and block.

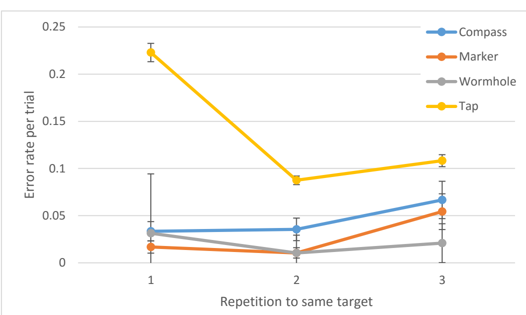

Figure 9: Error rate by repetition to same target.

5.5.2 Errors

We compared the number of errors made per transfer (an error was counted if the participant released the object on the wrong portal, or tapped devices with the wrong person). Figure 8 shows that the *Wormhole* technique had higher errors, particularly in the first two blocks. RM-ANOVA showed a significant main effect of Technique on errors ($F_{3,93}$=15.4, p<.0001), and also a significant interaction between Technique and Block ($F_{12,372}$=5.95, p<.0001). Follow-up pairwise comparisons show *Wormhole* had higher error rates than other techniques. We therefore accept hypothesis H2.

As shown in Figure 9, the number of errors for the *Wormhole* technique was highest on the first repetition (22%), falling to about 10% for the second and third trials. RM-ANOVA showed a significant interaction between Technique and Repetition ($F_{6,186}$=4.05, p<.005).

5.5.3 Subjective Responses: Effort and Preferences

Responses to the post-condition questionnaire (based on the NASA-TLX) are shown in Table 2. Friedman tests on the ratings showed that there were significant differences between the techniques in the amount of mental and physical effort, the amount of work required, and the level of frustration (all p<0.05). For all questions, ratings of the *Tap* condition were highest; the other three conditions (*Wormhole*, *Marker*, and *Compass*) were similar.

We also asked participants which technique they preferred in terms of several qualities – ease of use, speed, accuracy, and overall preference (see Table 3). Chi-squared tests showed that significantly more participants chose *Compass* for ease of use, speed, and overall preference. Most participants also chose *Compass* for accuracy, but the test was not significant. We therefore accept hypothesis H3.

Table 2: Mean (s.d.) of effort scores (1-7 scale, low to high).

	Compass	Marker	Worm	Tap	χ^2	*p*
Mental	2.2 (1.4)	2.6(1.4)	3.3(1.7)	3.1(1.7)	9.75	.021
Physical	2.1 (1.5)	2.8(1.8)	2.0(1.2)	4.3(1.9)	23.3	<.0001
Temporal	3.4 (1.9)	3.1(1.6)	3.4(1.8)	3.9(1.6)	4.43	0.21
Success	6.2 (1.1)	6.0(1.4)	5.7(1.2)	5.4(1.8)	5.87	0.12
Hard work	2.6 (1.9)	2.8(1.8)	3.4(1.8)	4.1(1.9)	10.9	.012
Frustration	2.0 (1.4)	2.6(1.9)	2.3(1.5)	3.6(2.1)	12.2	.0068

Table 3: Counts of participant preferences.

Which was:	Compass	Marker	Worm	Tap	χ^2	*p*
Easiest to use	21	5	3	3	28.5	.0001
Fastest	19	7	5	2	20.2	.0002
Most accurate	14	6	5	8	5.91	0.12
Overall preference	16	7	6	3	11.7	0.008

5.5.4 Participant Comments

We asked participants to provide written comments after each condition, and their remarks follow the performance and preference results provided above. First, several people commented on the speed of *Tap*, and particularly the need to wait for the other person to be ready. For example, one participant said "We have to wait until the person that I had to send the ball to is available to tap. Therefore, it is not that independent;" another stated "Very annoyed with having to wait for other people to transfer."

Second, participants recognized the correspondence between people in the real world and on the screen (for the two auto-ordering techniques) and the lack of correspondence for the *Wormhole* technique. Regarding *Wormhole*, one person said "I had to touch the circle to see who they were first before sending the ball." For *Marker*, one participant said "The exact positions were represented on the mobile device, hence it was easier to locate;" another stated that *Compass* was easy "because I didn't have to remember the positions;" and a third said (also about *Compass*) "It was just the most straightforward of them all, and I didn't have to remember any particular order to the positioning, as I could rely on both the screen or the person's name badge 'in real life'."

6 DISCUSSION

We developed and tested two new techniques for auto-ordering devices that are in an approximate circle (an F-Formation). We demonstrated the technical capabilities of the techniques (one based on the compass, and one based on a fiducial marker), and carried out a comparative study. The main findings of the study were:

- Auto-ordering techniques were faster than *Tap*;
- Auto-ordering techniques had lower error rates than *Wormhole*, particularly at first;
- *Compass* was strongly preferred by users.

6.1 Explanations for results

The findings from our evaluation generally match our expectations of the capabilities and limitations of the techniques and their underlying sensing technologies. First, the slow speed of the *Tap* technique appears to be caused by the requirement that both participants (sender and receiver) participate. Because participants were often engaged in a transfer to another person, this requirement meant that people spent considerable time waiting for the receiver to be available to match the tap gesture. Although this delay would not always occur (e.g., in a single-transfer scenario), there is a performance advantage for "one-sided" techniques.

The performance of auto-ordering compared to *Wormhole* appears to arise from these techniques' correspondence between the real world and the on-screen representation of targets. As noted several times by our participants, it was easier to carry out the transfer when the action was guided by the real world as well as the on-screen target. The higher error rate of *Wormhole* has a similar explanation – the arrangement of targets had to be memorized in each block, and when the arrangement did not match the real world, people had to deal with conflicting information.

We note that the correspondence problem is reduced for the *Wormhole* technique in some situations. For example, when people already know the names or IDs of the people around the circle, then the labelling of portals will provide enough information for people to carry out the transfer, without needing to build any memory mapping. Second, *Wormhole* can work as well as the auto-ordering techniques when positions remain stable. As shown in Figure 7, people's performance with *Wormhole* got closer to *Compass* in the second and third repetitions to each target, likely because people were able to memorize the mapping. In addition, Figure 6 shows that people were also faster with this technique in the later blocks of the study, suggesting that people learned how to best use the technique. However, research on Stimulus-Response compatibility has shown that there are performance advantages in using spatially compatible arrangements even after training [31].

However, even though *Wormhole* can work well in some situations, our study shows that this technique has poor performance when people are dependent on the real-world arrangement of the group – e.g., situations where people do not know one another. Our auto-ordering techniques do not suffer from this limitation, and perform no worse when people are familiar.

Finally, the preference for the *Compass* over *Marker* is likely due to the reduced constraints on how participants had to hold their devices – with *Marker*, people had to keep the fiducial marker in the camera's view while they carried out the transfers, whereas with *Compass* they had much more freedom to hold the device as they wished, as long as it generally pointed towards the centre.

6.2 Contributions and Generalization

We contribute new localization techniques, a comparative analysis of the techniques, and a further exploration of F-Formations.

- *New Localization Techniques*: Our primary contribution in this work is two new localization techniques for determining device ordering. Our techniques allow fast and accurate object transfer compared to two standard approaches. Our auto-ordering techniques are based on sensors and computational resources readily available on almost all smartphones, showing that our techniques are almost immediately usable. The technical evaluation demonstrated that the sensor accuracy, precision and span were more than sufficient for typical use. Our techniques were overwhelmingly preferred by participants, indicating a strong potential for uptake. Especially, the Compass could allow future research in ad-hoc message or file passing to be conducted simply, cheaply and reliably.
- *Comparative Evaluation*: Although our techniques were superior in many ways, they may not always be appropriate. Designers now have empirical evidence concerning the tradeoffs between the different techniques. Tap is slow, but had close to zero errors, and could be useful when security, and in particular, recipient selection is of paramount importance. Our analysis was the first to demonstrate Tap's inherent timing disadvantage anchored in the requirement for mutual action. Wormhole always performed the worst initially, but had its performance converge to that of our techniques within two repetitions. For tasks where repeated transfers have to take place wormhole might be superior as it can be performed without sensing. We were the first to demonstrate that in most of the cases discussed in the research of Marquardt et al [26], our techniques would be preferred.
- *Theoretical Grounding*: Our design draws heavily on the idea of F-Formations, which shows the spatial arrangements that people typically adopt in ad-hoc groups. The success of our techniques from both a performance and preference perspective provides additional support to the validity and utility of F-Formations as a construct for designing co-located collaborative systems. Furthermore, pointing to the circle is necessary for calibration but not necessary if people's locations are reasonably stable.

Beyond our immediate contributions, our work could have significant impact on other areas. The most obvious and immediate application of our work is as an interface widget in other file sharing studies. Given the apparent superiority of the technique, individuals studying other aspects of file sharing amongst collocated handheld devices (for example preview modes for received files) should adopt our technique for ordering to minimize the timing and learning confounds found in tap or wormhole techniques, respectively. While designed to solve the problem of auto-ordering for file sharing in ad-hoc groups, the technology has the potential for integration into larger collaborative systems; for example, facilitating file sharing amongst cliques of groups in a conference or work environment. Finally, our technology serves as a demonstration of how simple spatial sensors, now ubiquitously available on smartphones can be used in clever ways to facilitate collaborative actions. These kinds of interaction techniques might be interesting in co-located games, for example passing a virtual hot potato or as part of a live action game of "Simon."

6.3 Limitations

Although we have made novel contributions in this work, there remain some limitations to the study and a great deal of potential future work. Our work is heavily dependent on the use of sensors standard on commodity smartphones. While these devices are generally reliable, the sensors do have well known failure modes. Smartphone compasses can provide noisy or unreliable readings in ferrous environments. This limitation does not overly constrain the number of possible use cases, however, and the sensor has been robust in our tests. Furthermore, this issue being actively addressed by sensor scientists, and may be overcome in 3-5 years. In future work we plan to empirically test the robustness of the techniques in real-world settings and with real-world groups.

Our usage evaluation should also be followed up by further studies in a more natural, less controlled environment, and should include a greater diversity of experimental tasks. It would also be desirable to test the limitations of the systems for number of simultaneous users, speed of transfer and stability of spatial arrangement. The work here is an important step in leveraging new technology for new collaborative techniques. Finally, our implementations of portal-based techniques were intentionally limited (e.g., users could not rearrange portals to match the real world); in future work we will test whether these added capabilities could improve the overall performance of the Wormhole method.

7 CONCLUSIONS

By exploiting the common arrangements of individuals in small group gatherings, we were able to simplify a complex multi-agent spatial localization problem, to localizing agents around an approximate circle. We were able to leverage sensors available on commodity smartphones to localize individuals around a circle with a resolution sufficient for at least twenty-two people, which is more than what is currently supported for Bluetooth ad-hoc networks. We developed two techniques: a marker-based technique which can use either a paper or phone-based marker, and a compass-based interaction which had lower sensed precision but works without a calibration step. In a controlled experiment, auto-ordering techniques consistently outperformed unordered portals and a proximity-based technique, and were preferred by most participants. Because these techniques employ standard smartphones, they can be easily deployed, and can help facilitate digital object sharing in small group environments.

REFERENCES

[1] P. Bahl, and V. N. Padmanabhan. RADAR: An in-building RF-based user location and tracking system. *Proc. IEEE INFOCOM*, Vol. 2, 775-784, 2000.

[2] N. Banerjee, S. Agarwal, P. Bahl, R. Chandra, A. Wolman, and M. Corner. Virtual compass: relative positioning to sense mobile social interactions. *Pervasive computing*, 1-21, 2010.

[3] H. Benko, and S. Feiner. Multi-Monitor Mouse. *CHI '05*, 1208-1211.

[4] J. T. Biehl, and B. P. Bailey. ARIS: an interface for application relocation in an interactive space. *Proc. GI '04*, 107-116.

[5] R. A. Bolt. “Put-that-there”: Voice and gesture at the graphics interface. In *SIGGraph*, 262-270, 1980.

[6] K. S. Booth, B. D. Fisher, C. J. R. Lin, and R. Argue. The "mighty mouse" multi-screen collaboration tool. *Proc. UIST '02*, 209-212.

[7] X. A. Chen, J. Schwarz, C. Harrison, J. Mankoff, and S. Hudson. Around-body interaction: sensing & interaction techniques for proprioception-enhanced input with mobile devices. *Proc. Mobile HCI '14*, 287-290.

[8] D. Dearman, R. Guy, and K. Truong. Determining the orientation of proximate mobile devices using their back facing camera. *Proc. CHI '12*, 2231-2234.

[9] T. Grossman, and R. Balakrishnan. The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. *Proc. CHI '05*, 281-290.

[10] E. T. Hall. *The hidden dimension*. Doubleday, 1966.

[11] P. Hamilton, and D. J. Wigdor. Conductor: enabling and understanding cross-device interaction. *Proc. CHI '14*, 2773-2782.

[12] S.Hart, and L. Staveland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Advances in Psychology, 52, 139-183, 1988.

[13] M. Hascoët. Throwing models for large displays. *HCI '03*, 73-77

[14] K. Hinckley. Bumping Objects Together as a Semantically Rich Way of Forming Connections between Ubiquitous Devices. *Proc. UbiComp 2003*.

[15] K. Hinckley, G. Ramos, F. Guimbretiere, P. Baudisch, and M. Smith. Stitching: pen gestures that span multiple displays. *AVI '04,* 23-31.

[16] L. Holmquist, F. Mattern, B. Schiele, P. Alahuhta, M. Beigl, and H. Gellersen, Smart-Its Friends: A Technique to Easily Establish Connections between Smart Artefacts. *Ubicomp'01*, 116-121.

[17] H. Jin, C. Holz, and K. Hornbæk. Tracko: Ad-hoc Mobile 3D Tracking Using Bluetooth Low Energy and Inaudible Signals for Cross-Device Interaction. *Proc.UIST '15*, 147-156.

[18] B. Johanson, S. Ponnekanti, C. Sengupta, and A. Fox. Multibrowsing: Moving Web Content across Multiple Displays. *Ubicomp'01*, 346-353.

[19] H. Kato, and M. Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. *Proc. IWAR'99*, 85-94.

[20] A. Kendon. *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge University Press, 1990.

[21] G. Kortuem, C. Kray, and H. Gellersen. Sensing and visualizing spatial relations of mobile devices. *Proc. UIST '05*, 93-102.

[22] C. Kray, D. Nesbitt, J. Dawson, and M. Rohs. User-defined gestures for connecting mobile phones, public displays, and tabletops. In *Proc. Mobile HCI '10,* 239-248.

[23] C. Kray, M. Rohs, J. Hook, & S. Kratz. Bridging the gap between the Kodak and the Flickr generations: A novel interaction technique for collocated photo sharing. *IJHCS, 67(12)*, 1060-1072, 2009.

[24] M. Li, and L. Kobbelt. Dynamic tiling display: building an interactive display surface using multiple mobile devices. *Proc.MUM '12*, 24.

[25] N. Marquardt, R. Diaz-Marino, S.Boring, and S.Greenberg.The proximity toolkit: prototyping proxemic interactions in ubiquitous computing ecologies. *Proc. UIST '11*, 315-326.

[26] N. Marquardt, K. Hinckley, and S. Greenberg. Cross-device interaction via micro-mobility and f-formations. *UIST '12*, 13-22,

[27] M. A. Nacenta. Cross-display object movement in multi-display environments. Ph.D. Dissertation, University of Saskatchewan, 2009.

[28] M. A. Nacenta, C. Gutwin, D. Aliakseyeu, and S. Subramanian. There and Back Again: Cross-Display Object Movement in Multi-Display Environments. In *JHCI*, 24, 1, 170–229, 2009.

[29] M. A. Nacenta, S. Sallam, B. Champoux, S. Subramanian, and C. Gutwin. Perspective cursor: perspective-based interaction for multi-display environments. *Proc. CHI '06*, 289-298.

[30] N. B. Piyantha, A. Chakraborty, and H. Balakrishnan. The Cricket location-support system. *Proc. MobiCom '00*, 32-43.

[31] R. W. Proctor, and T. G. Reeve. Stimulus-response compatibility: An integrated perspective. Elsevier, 1989.

[32] R. Rädle, H. C. Jetter, M. Schreiner, Z. Lu, H Reiterer, and Y. Rogers. Spatially-aware or spatially-agnostic? Elicitation and Evaluation of User-Defined Cross-Device Interactions. *Proc. CHI '15*, 3913-3922.

[33] J. Rekimoto. Pick-and-Drop A Direct Manipulation Technique for Multiple Computer Environments. *Proc. UIST '97*, 31-39.

[34] J. Schwarz, D. Klionsky, C. Harrison, P. Dietz, and A. Wilson. Phone as a pixel: enabling ad-hoc, large-scale displays using mobile devices. *Proc. CHI '12*, 2235-2238.

[35] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg. Pose tracking from natural features on mobile phones. *Proc. ISMAR '08*, 125-134.

[36] R. Want, A. Hopper, V. Falcao, and J. Gibbons. The active badge location system. *Proc. TOIS '92*, 10, 1, 91-102.

[37] M. Wu, and R. Balakrishnan. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. *Proc. UIST '03*, 193-202.