AdHocProx: Sensing Mobile, Ad-Hoc Collaborative Device Formations using Dual Ultra-Wideband Radios

Richard Li
Microsoft Research, UW*
lichard@cs.washington.edu

Teddy Seyed
Microsoft Research
teddy.seyed@microsoft.com

Nicolai Marquardt
Microsoft Research, UCL
nicmarquardt@microsoft.com

Eyal Ofek
Microsoft Research
eyalofek@microsoft.com

Steve Hodges
Microsoft Research
Steve.Hodges@microsoft.com

Mike Sinclair
Microsoft Research
sinclair@microsoft.com

Hugo Romat
Microsoft
romathugo@microsoft.com

Michel Pahud
Microsoft Research
mpahud@microsoft.com

Jatin Sharma
Microsoft Research
Jatin.Sharma@microsoft.com

William A. S. Buxton
Microsoft Research
bill.buxton@microsoft.com

Ken Hinckley
Microsoft Research
kenh@microsoft.com

Nathalie Henry Riche
Microsoft Research
nath@microsoft.com

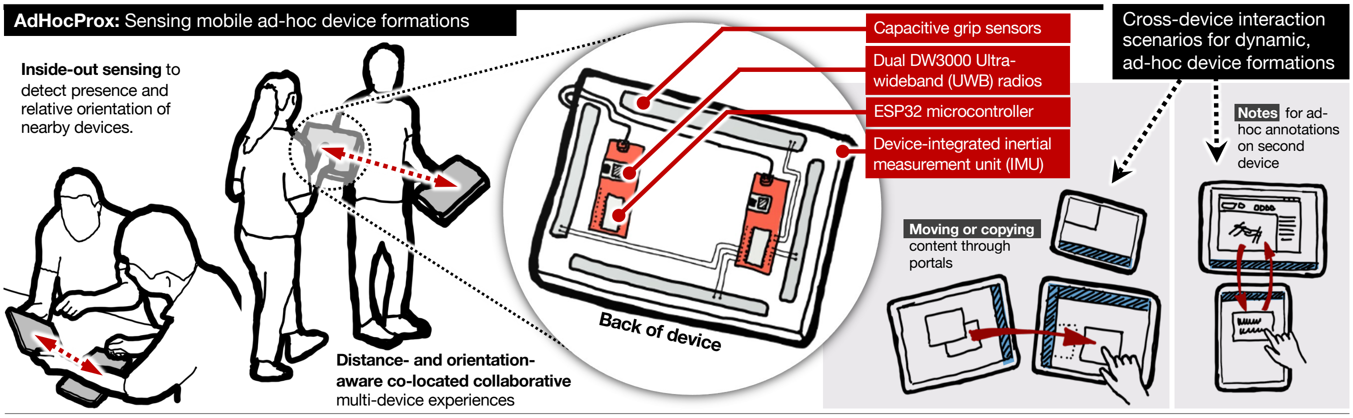


Figure 1: AdHocProx leverages multiple inside-out sensing modalities (including dual Ultra-Wideband radios, capacitive grip sensing, and fused inertial orientation) built into a device such as a tablet to detect presence, relative position, and device-motion gestures towards other nearby mobile devices. AdHocProx achieves this without recourse to externally-anchored beacons, central orchestration, or even reliance on local WiFi connectivity. This unlocks a number of cross-device interaction scenarios for highly dynamic, ad-hoc device formations outside of infrastructure-heavy "smart room" or lab settings.

ABSTRACT

We present AdHocProx, a system that uses device-relative, inside-out sensing to augment co-located collaboration across multiple devices, without recourse to externally-anchored beacons – or even reliance on WiFi connectivity.

AdHocProx achives this via sensors including dual ultra-wideband (UWB) radios for sensing distance and angle to other devices in dynamic, ad-hoc arrangements; plus capacitive grip to determine where the user's hands hold the device, and to partially correct for the resulting UWB signal attenuation. All spatial sensing and communication takes place via the side-channel capability of the UWB radios, suitable for small-group collaboration across up to four devices (eight UWB radios).

Together, these sensors detect proximity and natural, socially meaningful device movements to enable contextual interaction

---

*Richard is a Ph.D. student at the Paul G. Allen School of Computer Science & Engineering, University of Washington and completed this work on an internship at Microsoft Research.

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04.
https://doi.org/3544548.3581300

techniques. We find that AdHocProx can obtain 95% accuracy recognizing various ad-hoc device arrangements in an offline evaluation, with participants particularly appreciative of interaction techniques that automatically leverage proximity-awareness and relative orientation amongst multiple devices.

CCS CONCEPTS

• **Human-centered computing**; • **Ubiquitous and mobile computing**;

KEYWORDS

multi-device collaboration, ultra-wideband sensing, proximity, inside-out tracking

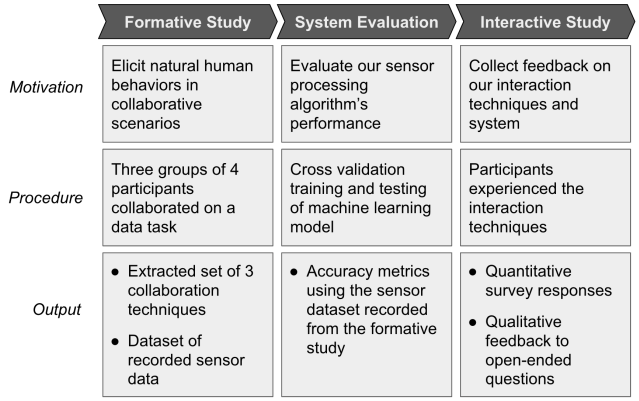
ACM Reference Format:
Richard Li, Teddy Seyed, Nicolai Marquardt, Eyal Ofek, Steve Hodges, Mike Sinclair, Hugo Romat, Michel Pahud, Jatin Sharma, William A. S. Buxton, Ken Hinckley, and Nathalie Henry Riche. 2023. AdHocProx: Sensing Mobile, Ad-Hoc Collaborative Device Formations using Dual Ultra-Wideband Radios. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 18 pages. https://doi.org/3544548.3581300

1 INTRODUCTION

Natural human-human communication leverages many nonverbal and highly contextual back-channels such as social cues, body posture, and nuanced movement of physical artifacts into and out of the focus of joint attention. As computing devices – particularly mobile ones – proliferate, they play an increasingly important role in our everyday interactions and collaborations. For instance, while conversing with a group, an individual may use their phone to look up relevant information, and even pass it around to share with others [49]. Alternatively, several users might concurrently work across shared and personal devices to contribute to a discussion, or to the outcome of a brainstorming session [25].

The problem is that connecting multiple devices together and configuring their use, particularly when owned by different persons, remains a tedious task that often disrupts people's workflow [86]. Configuring a communication channel between devices requires multiple steps, typically relying on IT-supported infrastructure such as WiFi, fixed Bluetooth beacons, or even round-trip communication with cloud services. And even when connected, devices typically do not self-reveal their relative proximity and corresponding interaction options (such as sharing files or maximizing application content across two screens) via "feed-forward" techniques or other graphical interaction affordances. These issues hinder multi-device usage in ad-hoc situations, when the time available to make devices work together is limited and might disrupt natural human social dynamics.

AdHocProx contributes a system founded on peer-to-peer networking with relative proximity via inside-out sensing between devices (Figure 1). This approach leverages context sensing to shift the burden of device discovery, spatial configuration, and (optional) cross-device connection from the user to the system; further, the resulting interaction techniques reveal multi-device usage opportunities as part of the user interface.



	Formative Study	System Evaluation	Interactive Study
Motivation	Elicit natural human behaviors in collaborative scenarios	Evaluate our sensor processing algorithm's performance	Collect feedback on our interaction techniques and system
Procedure	Three groups of 4 participants collaborated on a data task	Cross validation training and testing of machine learning model	Participants experienced the interaction techniques
Output	• Extracted set of 3 collaboration techniques • Dataset of recorded sensor data	• Accuracy metrics using the sensor dataset recorded from the formative study	• Quantitative survey responses • Qualitative feedback to open-ended questions

Figure 2: The studies presented in this paper, conducted to inform the design of and evaluate our system.

We use pairs of ultra-wideband (UWB) radios for low-bandwidth coordination between devices, as well as inside-out ranging to one another. Leveraging dual UWB radios enables AdHocProx to track not just proximity (*i.e.*, the absolute value of the one-dimensional distance between devices) but also the relative orientation (signed distance and angle) between devices. Our device prototypes also integrate capacitive grip sensing to detect when and where the user's hands – necessary to hold a mobile device up for use – may attenuate the UWB radio signal. Finally, each device's in-built inertial measurement unit (IMU) enables embodied tilt-based interactions in the context of the relative (device-to-device) orientation sensed via UWB.

AdHocProx can thus realize and expand upon cross-device interaction concepts presented in previous work such as GroupTogether [57], AirConstellations [55], or other related techniques [12] solely via on-device, inside-out sensing techniques. For example, when the user tilts an AdHocProx device toward that of another user, our system is aware of what device it is tilting towards. Thus, in contrast to prior cross-device tilt gestures in the literature, our approach does not require WiFi connectivity to establish spatial relationships between devices, and further does not rely on any extrinsic, environmentally-situated beacons or anchors – with the exception of Earth's gravity and magnetic field, of course, which feed into the fused IMU orientation.

To elicit and observe behaviors that arise during co-located collaboration with multiple devices, as well as evaluating AdHocProx's signal processing and machine learning pipeline, we conducted a formative study with three groups of 4 participants (total N=12). Participants worked in these groups to analyze and present data visualizations on paper taped to devices of different form-factors. Qualitative observations showed the importance of device arrangements (*i.e.,* devices side-by-side or across from each other) as part of collaborative work, and informed the design of four interaction techniques: Move, Copy, Pan, and finally "Note," which brings up a note-taking space when one (horizontally-oriented) device is brought into proximity of a second, vertically-oriented one.

An offline cross-validation evaluation in which we trained on sensor data from all four devices used in two of the three sessions, and tested on the remaining session, our processing pipeline obtained a 95% accuracy rate recognizing device arrangements. To gather feedback on four example interaction techniques enabled by AdHocProx, we conducted a follow-up study in which a subset of participants (N=6) from the first study. Participants experienced the techniques as implemented using the AdHocProx system in real-time. Participants generally responded favorably to the techniques, particularly the way devices could leverage awareness of one other.

Taken together, our work contributes the following:

1. Our implementation of AdHocProx, a system that recognizes formations of devices via dual UWB radios to coordinate devices and sense proximity as well as the angle between devices; capacitive grip sensing to reduce interference with UWB signals from the hand(s) holding a device; and IMU-based sensing of oriented device movements to correctly display corresponding graphical affordances.
2. Insights from a formative user study during which we observed user behaviors arising during co-located collaboration, from which we designed four interaction techniques.
3. A dataset of sensor signals for three groups of 4 participants (using a device each) in multiple arrangements.
4. Insights from a follow-up user study in which participants provided feedback on our designed and implemented interaction techniques.

After discussing Related Work, we detail the technical realization of the AdHocProx System. We then present a Formative Study that serves a dual purpose as a Sensor Data Collection activity for natural collaborative movements involving shared physical artifacts. Further, from this Formative study we distill four key behaviors: F1, Device Line-Up; F2, Device Roles; F3, User Proxy; and F4, Multi-device Territoriality, which collectively inform the set of four Designs for Interaction Techniques currently supported by our system. Finally, we present brief observations of an Interactive Study and Real-Time Evaluation of these techniques in action, provide a Discussion of some remaining points, and finally wrap up with a Conclusion and Future Work. Please also note that we share supplementary material including schematics, code, and data at the following repository: https://github.com/adhocprox/adhocprox

2 RELATED WORK

In this section, we motivate our work with a synthesis of work in *cross-device computing to support small-group collaboration* – with a focus on how *proxemics and micro-mobility* can inform the design of ad-hoc cross-device interaction techniques. We discuss technical approaches for sensing *device proximity, arrangement, and orientation* of nearby devices, and *grip sensing* to enable sensing of device micro-mobility that supports fine-grained nuances of sharing and interaction.

2.1 Cross-device Computing for Small-Group Collaboration

The field of cross-device computing explores how to design interfaces or applications spanning across multiple devices (*e.g.*, tablets, phones, laptops) to best support individual tasks or group collaboration (themes across this field are synthesized in [12]). While part of cross-device computing is about individuals interacting with device ecologies [53, 56, 86], a considerable focus is on multi-device applications for mediating collaboration in small groups. For instance, co-located group collaboration has been facilitated through digital tabletops [20, 72], shared usage of electronic whiteboards and vertically projected screen spaces [9, 10, 40, 41, 65], or the use of diverse ecologies of devices with different form factors [21, 73, 80].

Studies further investigated the use and value of multiple devices in knowledge work activities. For example, studies highlighted how access to both mobiles and additional synchronized interactive tabletops [75], or the use of a supplemental mobile overview device [11], can facilitate the collaborative decision- and sense-making process. And broadening beyond primarily screen-based devices (tablets, phones), other studies found that mobile-AR across devices (multiple viewpoints) mediates collaboration but also observed increased cognitive demand due to context switching. [78]. Later work also identified how different device configurations (e.g., form factors, formations) impact *"collaboration strategy, behaviour, and efficacy"* in co-located AR applications across multiple devices [79]. More generally, studies investigated the mechanics of mobile co-located experiences [51], including a framework mapping the social, technological, temporal, and spatial characteristics driving design. And more recently, Yuan et al. mapped out the current use and patterns of multi-device use [86].

2.2 Proxemics and Micro-Mobility

Towards facilitating groups' interactions across devices, related research proposed interaction techniques for easier access to and manipulation of digital content, complimentary functions of devices, or migration across devices (Table 3 in [12]). To inform the design of such interaction techniques and to better match interfaces and applications to people's expectations and current practices, work in cross-device computing has leveraged insights from seminal social theories. Two theories that closely relate to small-group collaboration are *proxemics* [28] and *micro-mobility* [50].

Hall's *proxemics* theory [28] correlates physical to social distance, where people move closer to and orient towards others for increased engagement. This seminal theory served as a fruitful inspiration in interaction design, where *Proxemic Interactions* [26] operationalized this theory, building devices and applications that react to the proxemics of people and other devices around them. Five main characteristics often drive these interactions: distance, orientation, movement, identity, and location. Examples include systems that adjust interactions with digital whiteboards [42] or tabletops [4] based on a person's proximity, mobile devices revealing sharing opportunities when moving closer [7], or environments where control of appliances is mediated and filtered through proxemics [48].

Related – and most relevant when in close proximity – *micro-mobility* [50] of physical artifacts is *"the fine-grained orientation and repositioning of objects so that they may be fully viewed, partially viewed, or concealed from other persons"* [57]. Examples of micro-mobility are subtle – and often fluid and ad-hoc – changes of position and orientation of documents to either suggest or inhibit shared access. Everitt *et al.* applied micro-mobility to the design of interactive tabletop workspaces for easy sharing, reorienting, and segmentation of documents [21]. GroupTogether later used micro-mobility to drive cross-device interaction [57], using subtle changes of position and tilting of devices as cues to provide fluid, less disruptive techniques for co-located collaboration. We incorporate notions of both proxemics and micro-mobility in our work, where we derive proxemic device arrangements to inform interaction techniques relevant to co-located collaboration – focusing on both proximity and relative orientation between devices, fused with micro-mobility cues of tilting and grip changes.

2.3 Sensing Proximity, Arrangements, and Relative Device Positioning

On a technical level, many multi-device collaborative systems need some form of sensing technology to recognize the location of devices, their distance from one another, or even fine-grained changes of proximity and orientation. We synthesize key technical approaches for the sensing of device arrangement, proximity, and orientation.

One common approach is to use camera-based tracking of device location [23, 37, 66], such as with RGB, infrared, or depth-sensing cameras. For instance, PolarTrack uses polarized filters to locate devices [64], EagleSense uses the segmentation of the depth-image to track people and devices [84], and ProximityToolkit uses a combination of infrared marker-based tracking and depth cameras [54]. These tracking approaches have in common that they are implemented as an *outside-in* sensing infrastructure, where a fixed setup needs to be deployed in the environment (in this case, one or multiple tracking cameras positioned at permanent locations in the environment). Increasingly, research moves towards inside-out tracking, where all tracking hardware is integrated into the device itself —- an example of this is Dearman et al.'s use of the back-facing cameras determining device formations [18].

Measuring acoustic ultrasound signals enables inferring approximate device positions. Systems used deployed networks of fixed infrastructure and mobile ultrasound transducers for tracking device location [1, 30]. Similarly, infrared outside-in setups can track devices, such as with the tracking of ActiveBadge tags [77], the ParcTab cell-based positioning approach [68], or in combination with visual and audio to sense people's devices [2]. Towards inside-out mobile setups, RELATE Gateways used mobile-only ultrasound transducers for sensing relative device proximity [29], and Wang et al. used ultrasound signals with commodity earphones sensing head position and orientation [76].

Existing internal sensors in off-the-shelf devices, such as magnetometers, can detect pairing of devices when touching [39], sense stacks of devices with magnet and reed-switch combinations [46], or recognize nearby docked devices with several magnets + sensor combinations [35]. Other embedded sensors, such as accelerometers or gyroscopes, can measure the tilting angle of a device, allowing users to automatically flip pages when the device is rotated [14], effectively navigate through long lists [22], or move content across devices [57].

Radio-based signals allow absolute or relative sensing of device location by measuring signal-strength estimates or time-of-flight. For example, Bluetooth signals in combination with audio data can support spatial location of devices [13, 38]. Ultra-Wideband (UWB), another wireless radio technology, has been used to localize objects in space [36], or to perform activity recognition [71]. Traditional UWB applications range from radar imaging to precise locating [88] and tracking [47, 62]. UWB radios incorporate a time-of-flight ranging method that estimates the distance between two radios. The UWB ranging method is used in many outside-in systems with fixed 'anchor radios' in the environment, allowing signal trilateration [17, 43], or combined in sensor fusion with camera-tracking and IMU sensors to increase tracking fidelity [67]. Because radio-based signals do not face some of the limitations of other tracking approaches – such as the visible line of sight needed with camera tracking, issues of attenuation or interference when using acoustic signals, or the limited ability of internal sensors to detect other devices – we use UWB radios for our inside-out tracking of device proximity. We further extend earlier UWB sensing approaches with a dual UWB radio setup and by fusing with signals from grip sensors and IMUs, allowing the angle of relative device orientations to be measured.

2.4 Detecting Human Grip and Grasping of Devices

Human grip and grasp represent a rich source of information to better understand user activities [52, 60]. It has been widely used to reveal users' intentions [8, 44, 81]. The way people shape their hands to interact with a physical artifact to handle objects or switch between them [8, 58] makes the interaction more explicit to understand [81]. Other research investigates how users grip digital devices such as tablets [74, 85, 87] or small devices [44]. Device bezel [45] and back-of-device [59, 83] remain the most common areas where the recognition happens. Grasp sensing allows detecting handedness [22, 82], avoiding unintended screen rotations [15], or calling up a graphical keyboard [16]. Other work explores front and back touch gestures [83] and bi-manual tablet interactions [74]. Capacitive sensing is a common approach to detecting grip [27]. Motion sensing has also been investigated in combination with touch to enable multi-modal gestures [24, 32, 34, 69]. Our work goes beyond these efforts by using grip sensing on the entire surface of the device, and its position relative to other devices in combination with inertial motion sensing. In our context, we use sets of grip sensors around the device to increase the reliability of our UWB-based inside-out tracking.

3 THE ADHOCPROX SYSTEM

We designed AdHocProx with the goal of leveraging proxemics to enable contextually-aware cross-device interaction, anywhere, any time. Many cross-device interaction techniques require a manual pairing or configuration process. For instance, an explicit process pairing process involving security codes is required for connecting over Bluetooth. Another example is the requirement of manually specifying the arrangement of monitors when setting up an additional monitor. These tedious steps frequently add so much friction that users often simply do not bother.

Previous efforts have been made to sense formations of devices. In the context of recognizing multi-person, multi-device formations, many of these works have relied on surveillance camera systems. Otherwise, signals from wearable data have been correlated to determine if two people are interacting. However, these both require either sensing in the infrastructure (*i.e.,* an outside-in camera) or

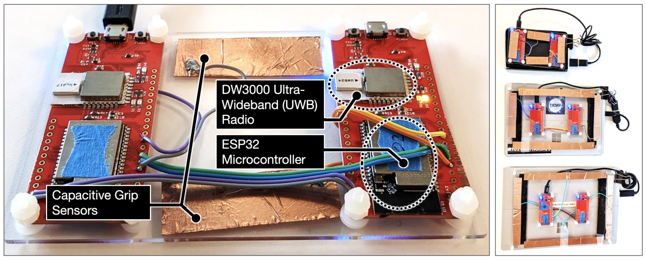


Figure 3: Left: The AdHocProx prototype hardware consists of two ESP32 microcontrollers, two DW3000 UWB radios, and four capacitive touch sensing electrodes. Right: The AdHocProx attachment on three devices with different form factors: Microsoft Surface Duo dual 5.6" screen smartphone, 10" Microsoft Surface Go and 12" Microsoft Surface Pro 7.

networking and computational processing in the infrastructure (*i.e.,* to facilitate synchronization of wearable sensor signals and recognize correlations between them). Infrastructure requirements introduce a setup time that is also often a significant deterrent in using such features, and largely prevents ad-hoc, mobile scenarios.

In our work, we designed AdHocProx to mitigate these issues. That is, we avoid both pairing and configuration steps: devices should automatically detect the presence of each other and automatically recognize their relative arrangement. We want to enable dynamic and natural contextually-aware interactions between and across devices with no manual configuration.

The AdHocProx system uses a pair of UWB radios attached to a host device to enable this pairing- and configuration-free sensing and networking (Figure 3). We empirically observed that the UWB ranging signal can be greatly affected when holding a device, especially when hands are close to the UWB antennas, so we include four capacitive touch sensing electrodes for detecting this and compensating. These two sensing modalities are part of the AdHocProx attachment, fixed to a host device with Velcro and connected via USB for power and transfer of data over a serial port. The host device’s onboard IMU is used in conjunction with these two sensing modalities.

3.1 Ultra-Wideband Radios

UWB radios transmit data in short pulses ($2-3$ *ns*) across a wide (> 500 *Hz*) bandwidth, enabling proximity ranging as well as communication. We build on UWB’s proximity ranging for relative orientation detection using by employing pairs of UWB radios per tracked device. Furthermore, we leverage UWB’s side-channel communication capabilities to transmit data for driving cross-device user interface updates in a peer-to-peer fashion.

3.1.1 Proximity Ranging and Relative Orientation Detection. UWB radios use time-of-flight measurements to measure proximity between two radios. We use a single-sided two-way ranging protocol, as shown in the left panel of Figure 4, sending an IEEE 802.15.4 standard data frame at each iteration. In order to measure the orientation of one device with respect to another, we included two UWB radios per device. Although the geometry of the system, as shown in the right panel of Figure 4, suggests that the angle between two devices can be inferred by simply comparing distances between each ranging measurement, we observed that the raw UWB signal was too noisy for this, both on its own and as induced by environmental factors. To handle these confounding factors, we opted to use a signal processing and machine learning approach instead.

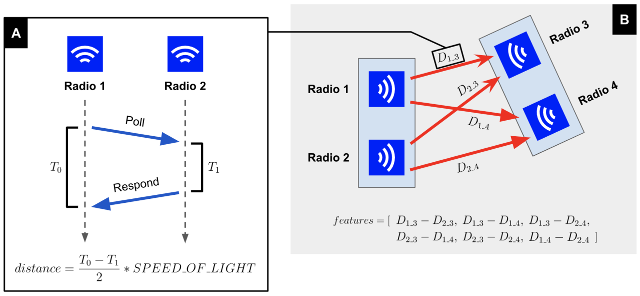


Figure 4: A: Time-of-flight calculation between a single pair of UWB radios to measure distance. B: Feature extraction based on pairwise differences of distances between two pairs of UWB radios to infer orientation of devices.

In order to avoid wireless cross-talk during ranging, we developed an ad-hoc peer-to-peer networking protocol that enables round-robin ranging by ensuring that only one radio is ranging at a given time. Our protocol also allows radios to dynamically enter and leave the ad-hoc network. Each radio identifies itself using a factory-programmed unique identifier. The protocol is implemented via the algorithm shown in Listing 1.

```
# get initial neighbors and order them based on ID
loop 5 seconds:
    known_nodes = discovery()
    known_nodes = sort(known_nodes)

# the device with the smallest ID starts the round robin process
if me == known_nodes[0]:
    my_turn = true

loop:
    # ranging packets include a flag to indicate whose turn it is
    my_turn = respond_to_ranging()

    if my_turn:
        # do discovery to find new neighbors
        known_nodes = discovery()
        known_nodes = sort(known_nodes)

        # range to every other node
        for node in known_nodes:
            if node != me:
                do_ranging(node)
        signal_next_node_to_range()
```

Listing 1: Ad-hoc peer-to-peer networking protocol algorithm.

We implemented this sensing system using a pair of ESP32-WROVER microcontrollers modules, each connected to a Decawave DW3000 UWB radio. The two microcontrollers connect to each other via a universal asynchronous receiver transmitter (UART), and the primary microcontroller connects to the host device using the universal serial bus (USB) protocol. We found that measuring proximity between two radios gave a sample rate of 500 Hz. In a network of four devices with eight radios in total, each device has a sample rate of 30 Hz with measurements to every other radio.

3.1.2 Peer-to-Peer Communication. We also use the UWB signal for peer-to-peer transmission of data for driving cross-device user interface updates. We do so by embedding the properties of user interface elements (*i.e.,* coordinate and scale), in the data frame transmitted for ranging. In particular, we found that we could reliably append up to 10 bytes to the end of the ranging data frame while maintaining the 500 Hz sample rate obtained by our ranging protocol between two devices (or 30 Hz across a network of eight radios), giving 5000 bytes-per-second between two devices (or 300 bytes-per-seconds across eight devices). This bandwidth is not sufficient for transmitting content such as images in an interactive fashion; for such resources, we send a URL for the device to download from or retrieve from cache. However, this bandwidth *is* sufficient for driving user interface updates across devices. For our user interface updates, we use 2 byte representations for the X-coordinate, Y-coordinate, and scaling of the user interface element being interacted with. The X-Y coordinate is on a global coordinate system that spans all of the devices in the network, negotiated based on the detected arrangement of the devices.

3.2 Capacitive Grip Sensing

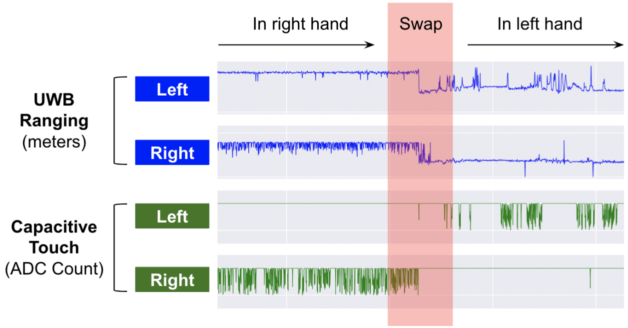


Figure 5: Example UWB ranging and capacitive touch sensor signals illustrating the interference introduced by a moving hand. During the period captured, the position of the device does not change, *i.e.,* the reported ranging should be constant.

We observed that human bodies can greatly affect the UWB ranging signal; a static hand between two radios can introduce up to 1 meter of error (Figure 5). In order to mitigate this interference, we included capacitive electrodes for coarse sensing of how the device is gripped. Four segments of copper tape were added to the AdHocProx enclosure to form the capacitive sensing electrodes that enable detection of grasp along any of the four edges. These electrodes were connected to the ESP32's touch sensing GPIO (general purpose input/output) pins. This information is used in the sensor processing pipeline to compensate for grip-related deviations in the ranging signal.

The UWB antennas in our prototype exhibit a 'butterfly' radiation pattern. We found that a hand moving along one edge of a device did not affect the signal as significantly as the hand being there or not. It might be possible to train a machine learning model to compensate for this based on the UWB signals alone, (*i.e.,* without capacitive sensors) but this process would either require explicitly designed and labeled training data, or a more sophisticated model that learns time-series features. Thus, we deemed the capacitive sensing approach to be simple yet effective, keeping the training and processing steps lightweight.

3.3 Inertial Sensing

In addition to the sensors mentioned above, the system uses the inertial measurement unit (IMU) onboard each host device. We fuse the raw accelerometer and gyroscope signals to determine the orientation of the device and detect tilt gestures. These tilt gestures, in combination with device arrangement information, are used to drive cross-device interaction techniques.

3.4 Sensor Processing Pipeline

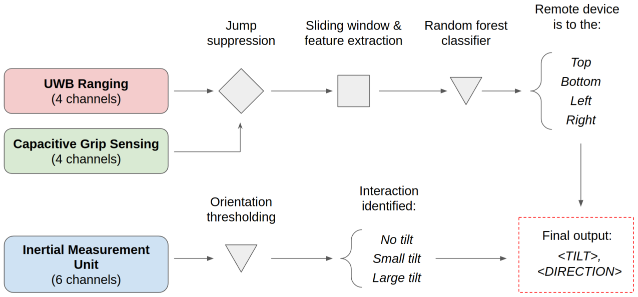


Figure 6: Sensor processing block diagram.

AdHocProx's sensor processing pipeline consists of two components, as shown in Figure 6: the device arrangement tracker, which outputs the orientation of a remote device; and the tilt gesture recognizer, which detects tilt events. The outputs from these two components are combined and provided to the user interface to render feedback.

To determine the arrangement of a network of devices, each device tracks the relative orientation of other remote devices using the UWB proximity ranging signals obtained by the ad-hoc, peer-to-peer networking protocol described previously. For each device, this protocol produces four signals to every remote device in the vicinity: distances are obtained for all permutations between the source device's two UWB radios and to the remote device's two UWB radios. The sensor processing pipeline, namely the top half of Figure 6, is then run on each set of signals, outputting an orientation for each remote device.

The UWB signal is extremely noisy. We found that one significant source of noise – inherent in designing a mobile, interactive device – is the presence of human hands. Figure 5 illustrates a device held with the right hand first, then swapped to the left hand in the highlighted region, producing an offset in the UWB signal of around 0.5 meters. We implement an algorithm that suppresses large changes in the UWB signal based on the capacitive grip signal: in a touch event, *i.e.,* when a hand touches or releases the device, the corresponding change in the UWB signal is recorded and compensated for.

After such large jumps in the UWB signals are suppressed, they are then passed into a sliding window of 0.2 seconds, with a 50% overlap (0.1 seconds), from which a set of features is extracted. These features are the differences in means, mins, and maxes for each pairwise combination of channels. On four channels, this step produces a feature vector of $3 * \binom{4}{2} = 18$ elements.

This feature vector is then fed into scikit-learn's default implementation of a random forest classifier with 100 estimators. The model outputs one of four cardinal directions; the remote device is determined to be to the left of, right of, above, or below the source device. Although it might be possible to increase the resolution of this output using more sophisticated modeling techniques, we found that the nearest 90 degrees was sufficient to effectively drive cross-device user interfaces. In practice, we observe that device arrangements on the same plane tend to be orthogonal.

In addition to tracking of neighboring devices' orientation, the sensor processing pipeline also uses the IMU to recognize tilt gestures of the local device. We fuse the accelerometer and gyroscope signals to produce an absolute orientation in space. The tilt gesture recognizer then simply applies thresholds to this orientation signal to determine if a tilt occurs.

We combine the outputs of these two processes to produce a contextually-aware tilt gesture recognizer. We use this output to drive a series of designed cross-device interaction techniques.

4 FORMATIVE STUDY AND SENSOR DATA COLLECTION

We conducted a formative study with a dual goal: (1) study natural behaviors of co-located work with analog material (pen and paper) to inform novel multi-device interaction techniques, and, (2) collect sensor data from devices handled by participants to be used for offline evaluation of AdHocProx's ability to recognize device arrangements.

Figure 7: Left: Photo of study environment. Right: A single sheet of paper with the same map printed three times, folded over itself to fit three different form factors.

4.1 Study Apparatus

Devices. We recruited three groups of four participants at a time. Each participant received a tablet device augmented with the prototype AdHocProx hardware. We used multiple device form-factors to observe a wider range of multi-device physical arrangements and activities. Thus, each group received two Microsoft 12" Surface Pro 7's (292 mm x 201 mm x 8.5 mm), a 10" Microsoft Surface Go (245 mm x 175 mm x 8.3 mm), and a Microsoft Surface Duo (folded, to present one 5.6" screen, 145.2 mm x 93.3 mm x 9.9 mm). Each device recorded UWB ranging between every tag to every other tag (7 channels), capacitive sensed grip points (4 channels), accelerometer values (3 axes), and gyroscope readings (3 axes). The AdHocProx attachment was connected to its host device over USB, and a JavaScript application enabling quick cross-platform deployment (Windows and Android) using the WebUSB API ran on the host device logging AdHocProx's sensor data. In addition to recording the inside-out sensor data, a video of each session was also recorded for ground truth, using an Azure Kinect.

Material. We used lightweight digital devices overlaid with paper "displays" for this activity in order to focus on the human behaviors of interest to us, such as device formations, the dynamics of these ad-hoc spatial layouts as focus of a collaborative activity shifts and evolves, as well as the interactions afforded by distinct arrangements. Further, devices overlaid with analog, paper-based information graphics avoid bias towards entrenched digital habits with status-quo touchscreen user interfaces. Our goal was to observe natural physical dynamics of interest, and eliminate factors unrelated to the co-located collaboration task, such as reorienting devices because of screen glare, or shifting grip to reach onto (and tap, touch, or scroll) the screen.

We printed a stack of charts related to COVID-19 trends generated on Our World in Data[1] and clipped them to each device. We printed each chart three times on a single piece of paper, in full, half, and quarter size, matching the 3 device form-factors of the group (Figure 7). It thus enabled participants to exchange a chart between their different devices simply by folding the paper to display the appropriately sized chart. Note that while we printed the exact same charts in different sizes, smaller form-factors were naturally harder to read, especially at a distance for example, with smaller printed text and labels for example. We also allowed participants with larger devices to place two half-size charts side by side.

To collect accurate sensor data from the devices (and each of their motions), we instructed participants to keep each sheet of paper clipped to a device, in any and all of their interactions. We only allowed participants to unclip charts to exchange them with another participant. They did so by unclipping the chart from their device, passing the sheet to another participant, this participant then (un)folding it to the correct form-factor and clipping it to their own device right away.

Finally, we provided analog pens to each participant and informed them of the possibility to annotate directly on paper charts or to clip a white sheet of paper (folded to the correct size) for notes or annotation if desired.

4.2 Task and Procedure

Task. We opted for a task requiring participants to deeply engage with the material provided, leading to natural interaction with their devices (to read and write on the charts clipped to these) and fostering discussion and information/data exchanges with other

---

[1] https://ourworldindata.org/coronavirus

participants. The goal of each group was to produce a poster depicting the difference in spread and measures taken during COVID between countries. To achieve this, they had to identify the most interesting countries to feature in the poster along with the most interesting set of charts.

To limit the session duration while maximizing the different types of co-located collaboration behaviors, we opted to structure the session into 4 phases. We selected such structure following insights from Olson and Olson [61], characterizing aspects of team work such as working in sub-groups and noting that most productive teams often transition from whole group meetings to sub-groups and back (Figure 8). We selected phases that encompass multiple activities and roles (reading data, writing notes, sharing insights, *etc.*) as well as transitions from group to sub-group structures to enable us to observe a large range of device arrangements and co-located collaboration behaviors.

1. *Data Review and Triage.* Participants individually review the data provided to them. Then they must work together as a group, to determine what data is available collectively and divide this data into two sets. Each set will be analyzed by a pair of participants in the next phase. Participants are free to divide the data as they see fit.
2. *Analysis and Comparisons.* Each pair of participants study the data in depth, comparing and contrasting charts, metrics, countries and coming up with an understanding of what the most interesting insights are in their dataset.
3. *Share and Report.* Each pair share their understanding of the dataset and point to what they consider the most interesting insights with the other pair.
4. *Synthesis and Production.* Participants design their poster as a whole group, determining a few high-level points, selecting the relevant set of charts, and laying out their content on the poster board.

Procedure. Participants reviewed and signed a consent form for their participation. The experimenter then explained the rules to manipulate devices and paper props enumerated in the previous section, followed by the presentation of the four different phases of the study. The experimenter reminded them of each milestone before moving on to the next phase, also giving time reminder when a phase lasted longer than estimated from our pilot session. At the end of the session, the experimenter provided a $40 USD gift card to compensate participants for their participation. This study protocol was approved by our Institutional Review Board.

4.3 Formative Study Results

We recruited three groups of four participants each for a total of twelve participants, with different backgrounds such as support engineers and administrative staff. Their self-reported familiarity with the other participants in a group ranged from being complete strangers to very familiar (*i.e.,* they work together on a regular basis). Each session lasted around 75 minutes.

After the session, two researchers independently made observations from the video recordings of the sessions. After discussing with the broader research team, we report the most salient and interesting behaviors observed (Figure 9):

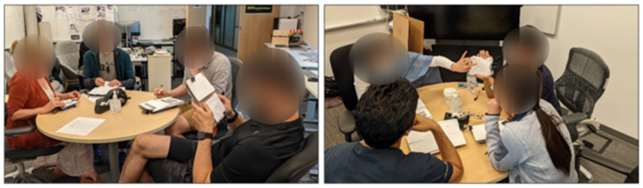
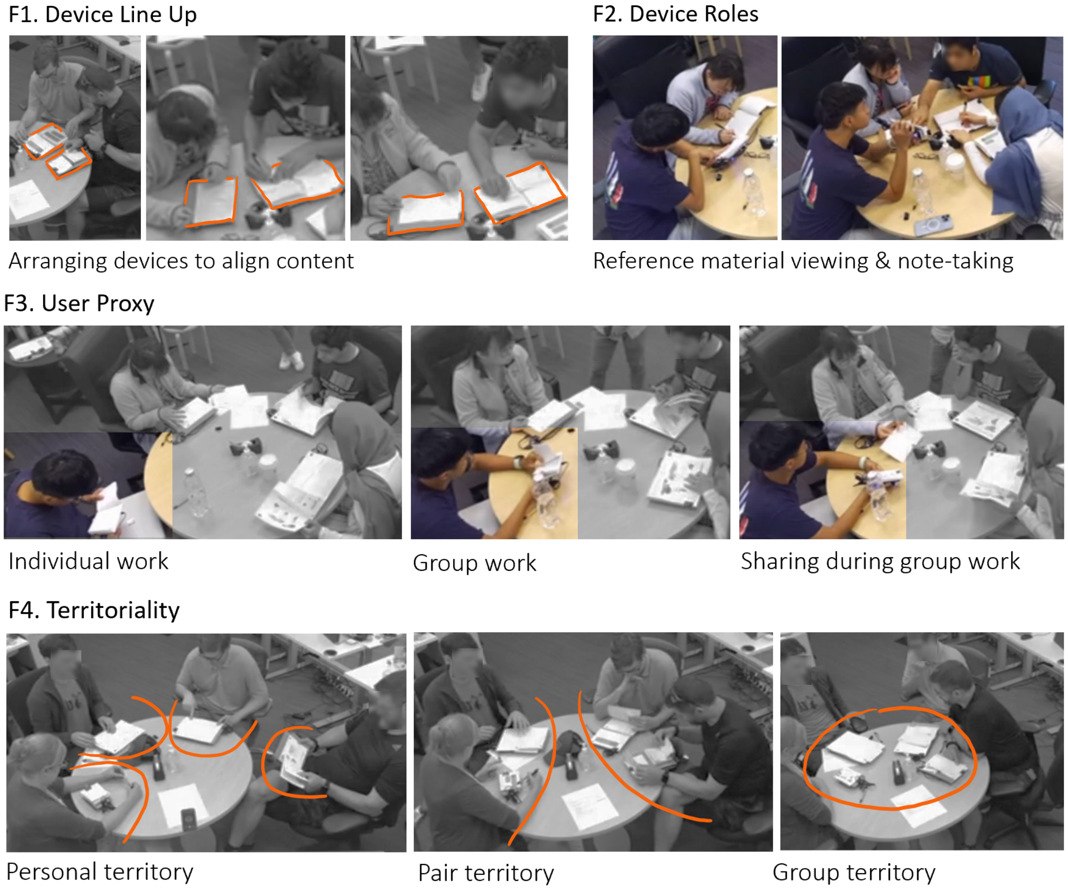


Figure 8: Example photos of study sessions. Left: Participants working alone or in pairs for deep analysis of the data. Right: Participants working together to develop a cohesive story.

(F1) Device Line-up. Participants frequently re-arranged their devices' position and orientation when interacting with others in the group. When discussing content on two devices, participants would frequently move their own device closer to the other person's device. Both would then re-arrange the devices to **line them up** to align content. Figure 9 also shows that participants changed the orientation of their devices (portrait or landscape) depending on the type of content to align. We observed occurrences of this behavior for all three groups in different phases and activities. For example, two participants aligned their devices next to each other to compare two different charts to triage them, or later line up visual features on each device (*i.e.,* to continue a line chart that started on one device and ended on another) during the analysis phase. Such device arrangement suggests the need to create one larger workspace with multiple devices.

(F2) Device Roles. We also observed participants adopt task-based division of labor at times, attributing **different roles** to their respective devices. We observed occurrences of the behavior in two of the groups. For example, one participant would tilt their device towards another participant to enable both of them to read the content, while the other participant would point at data on his own device to focus the discussion on a specific data point. In another instance, one participant held up a device vertically for an extended duration, while their collaborator noted relevant information on (the paper clipped to) another device using a pen. Figure 9 also illustrates that the roles devices play may be consistent across multiple group arrangements, for example the smaller device was used to show reference material both in pair and whole group work in Group 1. Such device arrangements suggest that people need ways to interact with and flexibly propagate content to other nearby devices.

(F3) User Proxy. We observed that devices often served as a proxy for the orientation and relative arrangement (formation) of the user's body relative to their collaborators during the sessions. This echoes findings from Marquardt *et al.* [57], with the distinction that participants were seated at a table, instead of standing (as studied by Marquardt *et al.*). Thus, even when devices remained resting on the table, participants still rearranged them (including sliding devices around) to afford the evolving collaborative activity. Figure 9 illustrates proximity and orientation of the device in response to the shifting focus of collaborative activity. These observations suggest that capturing device arrangements – and their dynamics – offers an insightful proxy for many aspects of individual and group activity around shared physical artifacts.



F1. Device Line Up

Arranging devices to align content

F2. Device Roles

Reference material viewing & note-taking

F3. User Proxy

Individual work

Group work

Sharing during group work

F4. Territoriality

Personal territory

Pair territory

Group territory

Figure 9: Examples of occurrences for our 4 salient findings.

(F4) Multi-device Territoriality. During individual work as well as dyadic collaboration, we observed instances of participants orienting their device toward themselves to signal unavailability to varying degrees – at times, even tilting the content away to partially conceal it from others in the group. By contrast, during the sharing phase, participants often placed their device in a more "public" orientation to afford viewing by others, including tilting it towards collaborators. During sub-group work, participants often reoriented their devices to form distinct clusters on the table (*i.e.* with one pair of devices proximal to one another, yet at a distance from the clustered pair of the other sub-group). Figure 9 shows an example of such "Pair territory" as well as other common device arrangements. This demonstrates that the principle of human *territoriality*, as previously observed in digital tabletop collaboration [70], also applies to multi-device co-located collaboration.

4.4 Offline System Evaluation

Our formative study doubled as a sensor data collection process. In particular, the sensor data reflected naturalistic behaviors of people interacting with the devices, albeit in a circumstance with a prescribed task for the group activity. From the sensor data collected, we compiled a dataset for offline evaluation. We annotated the orientations of the devices with respect to each other based on the videos recorded of the sessions. The granularity of our annotations was to the closest 90 degrees, in other words, one of the cardinal directions (*i.e.,* left, right, top, bottom) of a remote device with respect to the proximal one. These annotations consisted of labels and corresponding timestamp ranges, and they were paired with the time series sensor data for conducting our offline evaluation.

We sought to build a single generalized model independent of user or device form-factor. Thus our evaluation protocol, as shown in Figure 12, trained on data from two groups and tested on the third. Training and testing on the entire group's activity, across multiple persons using different sized devices, avoids overfitting our model to any individual user's idiosyncratic behaviors, or a particular device's characteristic properties. Therefore our evaluation protocol treats all devices equally, mixing data from all four devices per group for training, and evaluating on all four devices with the resulting generalized model.

To assess the impact of our capacitive grip sensing correction technique, we further conducted a separate ablation study. We ran the same evaluation protocol twice, once without any correction (Figure 10), and once with the previously described grip sensing correction (Figure 11). In the without-correction condition, we fed

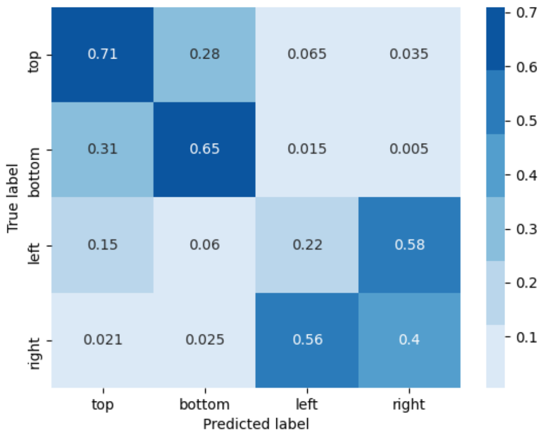


Figure 10: Without grip sensing correction, $accuracy = 49\%$

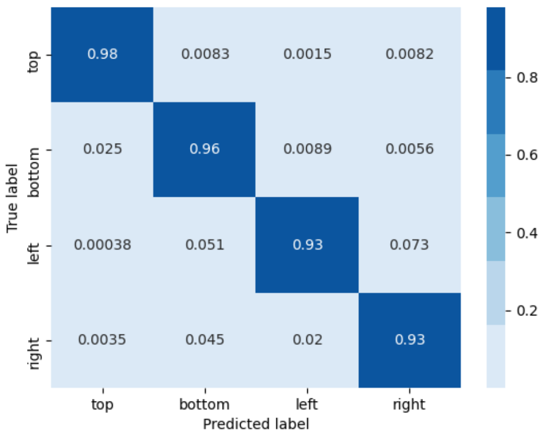


Figure 11: With grip sensing correction, $accuracy = 95\%$

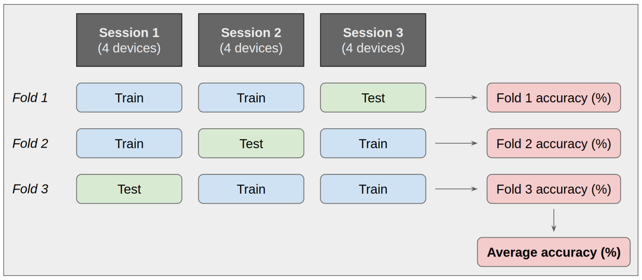


Figure 12: Cross-validation evaluation scheme for training and testing AdHocProx's machine learning models.

the raw UWB signals into our sensor signal processing pipeline. This baseline confirms that, without awareness of hand presence and placement, the system can confuse whether devices are to the left or to the right. But devices at the front produced more symmetric signals (because users most commonly tend to grip devices from the sides, rather than at the top or bottom).

In the capacitive-grip-corrected condition, we ran a *jump suppression* algorithm based on capacitive grip sensing, offsetting the UWB signals accordingly before feeding them into our signal processing pipeline. With the contextual awareness of what hand is holding the device, and when the grip starts and ends, the system was able to disambiguate devices to the left and right, resulting in an accuracy of 95% (Figure 11). We also noticed that most of the errors produced were temporally fragmented; by simply merging every 5 predictions using a majority vote, we obtained an accuracy of 99%. Since grip and device placements change on a scale of a few seconds (as opposed to multiple times per second at the level of individual sensor samples) this signal conditioning strategy offers a stable foundation for reliable interaction techniques.

5 DESIGN OF INTERACTION TECHNIQUES

We designed the AdHocProx system to serve as a spatial-sensing foundation for a range of possible cross-device interaction techniques. This foundation consists of not only the hardware attachment and sensor processing components described so far, but also a fundamental user interface layer to provide a cohesive experience across all AdHocProx-powered interaction techniques. This user interface layer realizes a visual representation and interactive behaviors that support the concept of a "portal" between devices.

5.1 AdHocProx Portals

An AdHocProx *portal* serves as a semi-public, limited shared space between two devices, much in the way that the porch of a home acts as a constrained social space to greet a possibly unfamiliar guest [33].

The sending user can only place a certain amount of content into a portal; the receiving user pulls the content out of the portal to complete the transaction. This design avoids a collaborator potentially interrupting and foisting content onto the main work-area of one's device at an untimely moment, as well as preventing (possibly unwanted) content from appearing on a receiving device, unless accepted by the receiver (via dragging the content from the local portal to complete the gesture). Note that the latter half of this handshake, *i.e.* explicitly dragging content out of the portal via touch, serves as a salient gesture that supports social awareness through its visibility to both sender and receiver (while also observable to other collaborators nearby, to support small-group awareness as well).

A portal automatically reveals itself along a screen edge when a remote device approaches from the corresponding direction. This design affords cross-device collaborative activity in accordance with observation **(F4)** of our formative study, where participants formed different territories with devices placed closer together to reflect collaborator's intent to work together. Moving an AdHocProx device within a threshold of approximately 0.8 meters triggers the

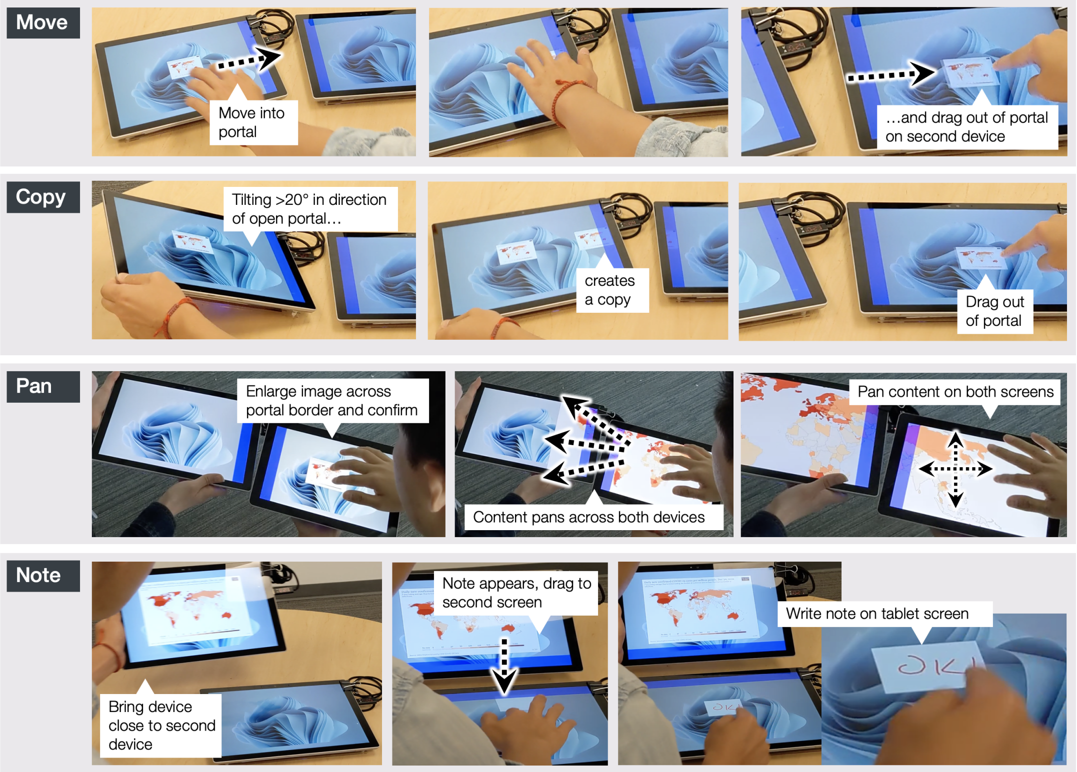


Figure 13: Illustrations of the interaction techniques: Move, Copy, Pan, and Note.

portal. We selected this threshold based on Hall's notion of proxemic zones, where an arm's-length of interpersonal distance defines an important perceptual and social boundary for one's personal space [28].

Our sensor processing pipeline runs locally on each device, recognizing which side of the remote device lies closest – and in what cardinal direction. The AdHocProx user interface layer then opens a portal along the corresponding edge of the local screen; the portal dynamically updates its on-screen placement if the remote device moves closer to a different edge of the screen, or disappears when the remote device moves away.

Due to the completely peer-to-peer nature of our round-robin UWB protocol, these portals open based solely on information available to each device locally. No external server or cloud service is necessary for spatial awareness, nor for portals to appear (or disappear) in response to devices coming and going, respectively.

The sharing of content is likewise peer-to-peer. Namely, AdHocProx shares references to files (rather than the content itself) by sending a URL over the spare bytes of the UWB ranging frames. For example, this can be a permissioned (read/write or read-only) link to an online document. Further, during the interactive drag-and-drop phase of a handshake gesture, AdHocProx keeps the coordinates of file icons synced between devices using these bytes as well. Due to UWB's limited bandwidth, we do not attempt to directly send media such as images and videos; rather, by sending a reference only, the other device can retrieve it when WiFi (or server access, if hosted in-cloud) becomes available at a later time.

5.2 Interaction Techniques

Based on the four phases and collaborative behaviors observed in our formative study (section 4.3), we designed four interaction techniques on top of the portals user interface layer, as detailed below (and in Figure 13.)

Note that we use the scenarios explored and behaviors revealed in our formative study as points of departure, rather than limiting the resulting interaction designs to a literal interpretation of passing pieces of paper around (for example). We also co-opt "digitally-authentic" behaviors such as copy or pan/zoom interactions that tablet and mobile device users have come to take for granted. Nonetheless, by having cross-device versions of such digital super-powers rooted in natural user behaviors, our hope is that the interface can embody these more abstract actions, making them more salient and memorable once discovered.

Move. Once a portal opens, a user can drag a file into it. This experience draws direct inspiration from phase 1 of the formative study, where the physical act of handing a sheet of paper to collaborators lets groups quickly triage and pass related data to one another.

Unlike email or using a cloud service, in which the sending of a file is an entirely disembodied digital experience, the physical metaphor behind this interaction remains consistent with the natural expectations of in-person collaboration. The interaction requires physical co-location and (unlike a digital sync progress bar) leverages the act of dragging the file as a salient physical indicator of the offer, exchange, and completion; a receiver can see the sender's hand approaching the portal, and likewise, the sender can observe if the receiver accepts their offering. The portal appears automatically, but the exchange of content is designed as a dyadic gesture with one portion completed by each user. (However, since touchscreens cannot detect which user makes contact, it remains possible for a single person to complete the gesture by reaching into the other user's personal device-space, but such a gesture is observable to both users, and would only be socially acceptable with tacit consent).

Copy. In addition to dragging a file into the portal, a tilt gesture (implemented as a tilt exceeding 20 degrees) triggers the duplication of a file that falls into the portal on its own. The tilt must correspond to the direction of an open portal; tilting in other directions has no effect. Hence, the gesture requires proximity as well as corresponding actions, while fluidly integrating three distinct elements – the verb (Copy), the selection source device, and the selection destination device – into a single, unified command phrase.

Since there is no physical metaphor for the duplication of a sheet of paper, we draw on a micro-mobility gesture [50] to initiate this action – namely the embodied, mutually visible act of tilting one's device towards a collaborator to afford sharing of content. This consideration also arose during the formative study, particularly during phase 2, where many participants asked the researcher conducting the study if some kind of copy operation was allowed in order for multiple people to look at the same item closely. Our observations with regards to lining-up devices **(F1)** and tilting the device as an indication of sharing information with others **(F3)** also informed the design of this technique.

Pan. Unlike the asynchronous *Move* and *Copy* operations, *Pan* enables users to zoom and pan around an image displayed across multiple devices. Respecting the limitations of the portal, to activate a Pan operation, the initiating user enlarges (*i.e.* zooms into) an image until it expands to reach the portal. In order to confirm the transaction, the receiving user can drag or zoom the image out of the portal on their own device. At this point, both users are able to perform zoom and pan operations on the image. The two screens functionally operate as one large screen, such that one user can touch their finger to one screen, and the other user drags their finger across the other screen, if desired, achieving a distributed zoom operation. Pan was inspired by multiple examples of participants arranging sheets of paper next to each other (mostly in phase 3 of the formative study, while sharing findings) in order to get a larger window into the data visualizations **(F1)**. For example, in many cases, two participants would have two different line graphs that were sequential along the X-axis, so they would arrange the two line graphs to be next to each other to give the feeling that it was one large graph; Pan attempts to create a digital version of this experience. Our observation of the different roles devices played **(F2)**, indicating that actions on one device can impact user expectations for another in a context-dependent way, also informed this interaction technique.

Note. A remote device held at a large tilt (greater than 70 degrees, as used by GroupTogether [57]) – much steeper than the 20-degree tilt that distinguishes *Copy* – creates a sticky-note icon in the portal that the user of the local device can conveniently drag out and write upon immediately. This interaction was inspired by instances of participants holding a sheet of paper up for another person to inspect or take notes on (for example, see Figure 8).

Since this facilitated one person more easily consuming content while writing on their own device, we assigned the role of capturing notes (writing) to the device flat on the table – with the devices thus serving different roles per **(F2)**. As a result, for the *Note* interaction the vertical device serves as the "remote" device, while the horizontal device flat on the table acts as the "local" device.

In phases 2 and 4 of the formative study, participants took notes or produced poster content (*i.e.,* headers for the posters) while in this pose, respectively.

6 INTERACTIVE STUDY AND REAL-TIME EVALUATION

We conducted a follow-up study to gather feedback from participants using our implementation of the interaction techniques described above. We implemented these using the same apparatus as used for sensor data collection in the formative study, namely a AdHocProx attachment and a host device – but with the screens exposed (rather than covered over with paper) and real-time sensor data driving live visual feedback.

6.1 Procedure

The goal of this follow-up study was to gather qualitative feedback on the usefulness of each techniques, in light of the task of the first study. For this reason, we invited participants from the first study to return, with six eventually able to do so. In individual sessions, each participant tried each the interaction techniques described above, implemented on top of the JavaScript application previously used for data collection, but with live sensor signal processing and machine learning models trained on the dataset collected in the formative study.

To elicit feedback on each technique, we used a semi-structure interview.

First, participants responded to each of the following statements on a 7-point Likert scale:

1. How useful would this technique have been if available during the formative study?
2. How easy was it to use this technique?
3. How likely would you use this technique if available in your everyday life?

For each question, the experimenter encouraged them to describe the reasons behind their rating.

Then, the experimenter asked them about the strengths and weaknesses of each technique. At the end of all four techniques, they were asked to rank them for overall preference as well as brainstorm any potential interaction techniques of their own.

Each participant received a $20 USD gift card for their participation.

6.2 Results and Feedback

Six participants from the formative study returned for the follow-up study. In general, participants responded favorably to the designed

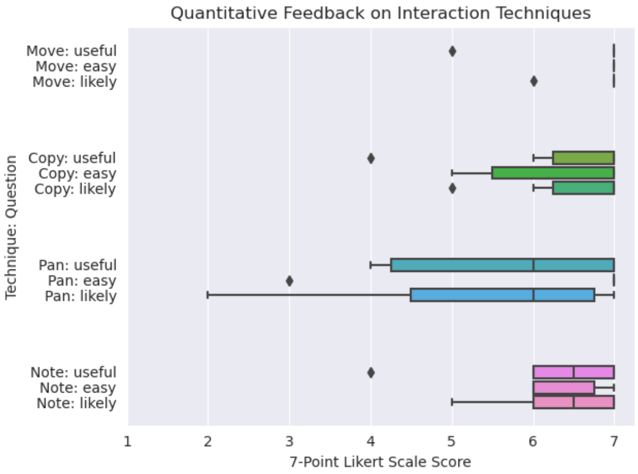


Figure 14: Likert scale responses to survey questions pertaining to the interaction techniques. Higher scores indicate greater positivity toward the interaction technique in question.

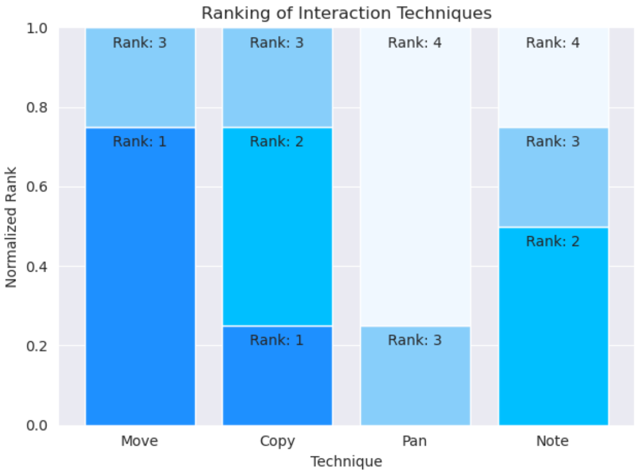


Figure 15: Ranking of interaction techniques. Lower rank value indicates greater preference.

interaction techniques, especially when keeping in mind how they would have applied to the formative study.

Figure 14 depicts the results from the Likert scale for these 6 participants, and Figure 15 depicts the final preference ranking of the four techniques. Note that, while these quantitative results are an indication of the potential usefulness of the techniques, our main goal with these questions was to elicit qualitative feedback to gauge whether our system could offer compelling solutions for ad-hoc, co-located collaboration. We report below the three major themes that emerged from this feedback.

6.2.1 Arrangement-Aware Portals. All of the participants understood the mental model of the portals, both in terms of how portals respond to proximity and arrangement, as well as providing a limited shared space between devices. This feedback suggests that portals can provide a solution to support multi-device territoriality **(F4)**.

In particular, they all appreciated that the portals allowed senders to visually confirm the content being sent while it is in the portal. Participants also all appreciated that the portals did not allow content to be forced onto a receiving device’s screen, and that the experience of receiving and accepting content by dragging it out felt more natural than clicking a digital button.

Outside of the design of the portals, some participants asked about the technical implementation of the portals. One recurring question was whether more than four portals (to more than four remote devices) was possible. Due to the bandwidth limitations of peer-to-peer ranging and data exchange exclusively over UWB, supporting portals for more than four proximal devices would be difficult to achieve at this time. However, there may be other ways to adapt our round-robin algorithm (*e.g.,* with lower sampling rate and response time) when more devices are nearby; likewise, backing off our strict use of peer-to-peer UWB to coordinate and sense proximity (*e.g.,* via a higher-bandwidth side channel once proximity is detected) might offer another approach. By contrast, having only one portal for each edge of the screen suits the small screens of mobile devices, while offering a side-by-side mental model of device proximity with a degree of simplicity (and device-to-device correspondence) that would be easy to undermine with added complexity.

Another question involved dismissing portals, whether due to a false positive sensor reading or simply to reclaim the screen real estate while near another device. We plan to add this straightforward capability to our prototype in the future, such as by swiping the portal away, or by automatically fading out if the user simply ignores the portal for several seconds.

6.2.2 Tilt-Based Interactions. For interactions involving tilt, such as *Copy* and *Note*, participants commented that they were generally easy to perform; some even noted that the experience of performing a physical action felt better than a purely digital action. This appears to echo our observations in the formative study **(F1) and (F2)**. Participants were concerned about the learning curve of such gestures, from the discovery of the gesture to finding the specific angle at which the gesture is recognized. In addition to the learning curve, some participants were also worried about the potential to falsely trigger the gesture. These concerns calls for more experimentation with the implementations of the threshold and technique itself. Some participants raised the issue that such tilt-based interactions might be cumbersome to perform with a larger form factor device such as the Surface Pro; on the other hand, some cited how pleasantly fluid the interaction was when performed with a mobile device such as the Surface Duo. This also echoes our observations in the formative study, in which participants with the more mobile device held the device in-hand a much higher proportion of the time, while larger devices tended to remain laying on the table.

6.2.3 Control and Agency of Shared Content. Before experiencing the Copy interaction, participants often asked about the ownership of the file transferred in the Move interaction. They were often concerned about whether the sender would be able to see changes made to the file by the receiver. Our design in fact transfers a URL to a file hosted on a cloud service, and a Move or Copy interaction just sends the URL to the receiving device. In this case, the distinction between the Move and Copy interactions depends on whether or not an instance is kept open on the sender's device. Once the receiving user accesses the document via WiFi and the cloud, if a write-enabled link is shared, changes from either user can sync per usual.

Outside of working on the file level, the Pan interaction enables users to simultaneously pan across a single image stretched across multiple devices. Multiple participants worried about another user having control of their own screen. This concern was so significant that Pan received the lowest scores on the Likert scales. This finding is particularly interesting as we observed many times in the formative study participants interacting with other's device and charts (See Figure 9, F2) from touching, browsing, holding, to even writing on others' data and devices. This could highlight a difference between how participants think of digital information vs analog ones, a question worth investigating in future work.

7 DISCUSSION

7.1 Sensing Techniques

We explored the use of UWB radios in this work due to their increasing availability in commercial devices, including smartphones [3, 6], and location trackers [5, 19]. However, current uses of UWB largely involve a number of static anchors (*i.e.,* UWB radios with a fixed, known location) that triangulate and track the location of a single moving UWB radio. In our work, we developed an ad-hoc round robin algorithm that not only allows devices to take turns ranging to each other without cross-talk, but also without any fixed, centrally coordinating device. This permits new devices to dynamically come and go, with their arrival or departure accounted for in the group round-robin.

AdHocProx required the use of two radios to estimate the angle of a remote device. While in principle UWB radios can measure phase-difference-on-arrival to infer angle [31], in practice this feature remains unavailable on current commercially-available devices or developer kits. We hope that this work and the AdHocProx prototype present compelling demonstrations of valuable functionality that make a strong case to include multiple radios in future mobile devices and tablets.

7.2 Devices as a Proxy

In this work, we used computing devices such as tablets and smartphones as proxies for the proxemics of their users. In our formative study, we observed users might hold up their device for another person to inspect for an extended amount of time **(F3)**, creating an opportunity for a digital intervention powered by AdHocProx. However, we did observe instances in which computing devices were less robust as proxies to their users. In particular, for larger device form factors laid out on the table. Participants moved their devices along (sliding them on the table) when engaged in deep discussions focused on the material. However, we also observed multiple instances of users leaving their device static on the table while orienting themselves differently in order to speak to someone else at the table.

In addition to the relationship between a person and their own device, it is interesting to look at the relationship between this person and other persons' devices. In our formative study, we found that participants, no matter how familiar they were with each other, felt no sense of invading private zones. While they did not grab other people's devices, participants frequently touched other people's devices to point out to specific content, sometimes reorienting it towards them. Figure 8 (right) and Figure 9 (top-right) depict such behaviors including a participant writing on another participant's device. Future work is needed to better understand the implication of this observation and if it would hold for multi-device interactions. The paper props used in the study, the somewhat "public" data (not owned by any of the participants), or the lack of access to other personal data on each device all may have contributed to such looser interpretation of private devices.

7.3 Pragmatic Considerations for Wider Adoption

The AdHocProx prototype described in this paper illustrates the potential of our approach. However, a number of additional considerations come into play when considering the suitability of AdHocProx for more widespread adoption.

Power consumption is a particularly important consideration for mobile devices. Decawave's DW3000 datasheet reports a typical current consumption of between 40 mA and 55 mA depending on the specific operation and configuration of the radio [63], and we empirically confirmed these values with our prototype. We used two DW3000 radios per device, but we would like to investigate the possibility of using a single radio with an RF switch to alternate between two UWB antennas as a possible cost and power-saving option. Our prototypes also include two ESP microcontrollers; we measured the combined current consumption of each ESP+UWB unit as 8 mA with both ESP and UWB radio idle increasing to 110 mA with both ESP and UWB fully active. While this power draw is not insignificant, we would expect a real-world implementation to use a single, low-power MCU instead of two ESP MCUs—and possibly even leverage an existing system MCU in the mobile device. This would of course reduce the power consumption, as would duty-cycling the AdHocProx hardware whenever possible. The full-active mode is only needed when a device is nearby several other devices, and until/unless this situation is detected it should be possible to run in a 'discovery' mode where a single UWB antenna is used at a much lower rate, such at 1 Hz. This could also offer a strategy to scale AdHocProx beyond its current limitation of four simultaneous devices. Although our prototype consumes extra power for signal processing and machine learning algorithms to be run on the host device, all of these processes are reasonably lightweight such that they too could be integrated into a single-MCU solution. The commercial success of UWB-based mobile device services (*e.g.,* the interaction between Apple iPhones and Airtags [5]) indicates that it is indeed feasible to incorporate the necessary hardware components in commercial mobile devices.

Security and privacy are also important considerations for wider adoption. At present, our UWB-based communication protocol is unencrypted. Security would thus require the development of new device attestation and encryption protocols to use in fully public, untrusted settings. With regards to privacy, since our system leverages proxemics, it inherently maintains a degree of privacy as physical proximity between devices is required for any information exchange to occur. Coupled with proximity, the interaction techniques we developed also offer users agency — a user that does not want a portal open to someone else can simply ignore it, or pull their device away.

7.4 Study Design Limitations

Our formative and follow-up studies also have a number of limitations. As with all qualitative studies, the generalizability of the results is limited. Insights from the formative study may have been impacted by the familiarity of members of the group and their personal styles of collaboration. The observations made during the formative study inspired a number of techniques, and the feedback gathered during the follow-up study shed some light at to the potential of these techniques (and the potential value of a dual UWB hardware setup) for co-located collaboration. However, additional rounds of studies are needed to formally evaluate the usability of the techniques and more systematically understand their usefulness in a co-located collaboration task.

In addition, both studies were conducted in the same space. It is important to note that the reception of UWB may be affected by reflection and absorbance in the environment, thus, testing this technology in different environments is important to better understand the robustness of the system.

7.5 Context-Aware Interaction Techniques

The primary goal of this work was to lay the groundwork for a system that reduces the effort necessary to work across dynamic formations of mobile devices in the context of *ad hoc* multi-user, multi-device collaboration scenarios. AdHocProx achieves a step in this direction by facilitating a peer-to-peer network connection between devices as well as automatically determining the formation of proximal devices. It is important to note that we only implemented a few interaction techniques inspired by our formative study on top of this platform. Our examples highlight how automatic spatial configuration of devices can make even simple interaction techniques more useful when applied across multiple devices. Furthermore, by combining AdHocProx's ad hoc networking capabilities with feedforward graphical affordances in the form of AdHocProx portals, we show a more natural "pairing" experience driven by physical proximity. As the basis for the system is laid out and demonstrated by the *ad hoc* dynamic realization of the interaction techniques described in this paper, we can envision future work focused on uncovering a suite of novel interaction techniques made possible by AdHocProx, for a variety of co-located collaboration tasks and contexts.

8 CONCLUSION AND FUTURE WORK

In this paper, we presented AdHocProx, a sensing system that uses UWB radios for peer-to-peer communication and inside-out tracking of the orientation of other nearby devices. We conducted a study in which participants took part in a group activity that required them to work together. From this study, we observed a number of collaboration behaviors, as well as collected passive sensor data for evaluating our system offline. Based on our observations, we designed a set of interaction techniques that build on the framework AdHocProx provides in terms of contextual awareness and the base layer user interface concept of portals. Finally, we conducted an interactive study to gather feedback on our implementation of these interaction techniques.

Our formative study was designed to elicit natural human behaviors in a collaborative setting based on proxemics and micromobility theories from sociology. Although we designed this study with the goal of leveraging findings from it as a basis for designing our interaction technique, our future work includes evaluating the utility of these techniques in context. The most straightforward example of such an evaluation would be to conduct another study with the same protocol as the formative study, but using digital devices with AdHocProx functionality instead of paper props. Participants would be prompted to compare the experience with their typical collaboration experiences. The reactions and feedback gleaned from our informal real-time interaction study indicate that participants are optimistic about finding utility for our interaction techniques in their everyday life.

Our paper specifically focused on augmenting co-located collaboration experiences using the AdHocProx prototype for proximity and orientation-aware interaction techniques. However, our system also has implications for hybrid and remote settings. We used our AdHocProx prototype for measuring the proximity and orientation of a device co-located with another. However, this technique could be extended to support awareness of devices in separate environments. For instance, a meeting room's projector screen could be instrumented with an AdHocProx attachment. Then, people in the meeting could perform the same interaction techniques described in this paper to interact with people both physically co-located in the meeting and those that are remote. AdHocProx's awareness of remote devices could be used in hybrid settings as well. We are also keen to explore UWB radios more tightly coupled to the body, such as on wrist-worn wearables or head-mounted displays, which could better resolve deictic pointing gestures and gaze direction, respectively.

In summary, we believe the ability to dynamically create peer-to-peer connections and automatically recognize formations that enable more natural interaction techniques anytime and anywhere lay the foundation for bringing many previously imagined cross-device interaction techniques to life in a practical way.

REFERENCES

[1] M. Addlesee, R. Curwen, S. Hodges, J. Newman, P. Steggles, A. Ward, and A. Hopper. 2001. Implementing a Sentient Computing System. *Computer* 34, 8 (Aug. 2001), 50–56. https://doi.org/10.1109/2.940013

[2] Xavier Alameda-Pineda, Yan Yan, Elisa Ricci, Oswald Lanz, and Nicu Sebe. 2015. Analyzing free-standing conversational groups: A multimodal approach. In *Proceedings of the 23rd ACM international conference on Multimedia*. 5–14.

[3] Andrew. 2022. How to enable ultra-wideband (UWB) on the pixel 6 pro. https://phandroid.com/2022/01/03/how-to-enable-ultra-wideband-uwb-on-the-pixel-6-pro/

[4] Michelle Annett, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. 2011. Medusa: A Proximity-Aware Multi-Touch Tabletop. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. ACM, New York, NY, USA, 337–346. https://doi.org/10.1145/2047196.2047240

[5] Apple. 2021. AirTag. https://www.apple.com/airtag/

[6] Apple. 2021. Ultra wideband availability. https://support.apple.com/en-us/HT212274

[7] Till Ballendat, Nicolai Marquardt, and Saul Greenberg. 2010. Proxemic Interaction: Designing for a Proximity and Orientation-Aware Environment. In *ACM International Conference on Interactive Tabletops and Surfaces (ITS '10)*. Association for Computing Machinery, New York, NY, USA, 121–130. https://doi.org/10.1145/1936652.1936676

[8] Cristina Becchio, Valeria Manera, Luisa Sartori, Andrea Cavallo, and Umberto Castiello. 2012. Grasping intentions: from thought experiments to empirical evidence. *Frontiers in human neuroscience* 6 (2012), 117.

[9] Jacob T Biehl and Brian P Bailey. 2004. ARIS: an interface for application relocation in an interactive space. In *Proceedings of Graphics interface 2004*. Citeseer, 107–116.

[10] Kellogg S Booth, Brian D Fisher, Chi Jui Raymond Lin, and Ritchie Argue. 2002. The" mighty mouse" multi-screen collaboration tool. In *Proceedings of the 15th annual ACM symposium on User interface software and technology*. 209–212.

[11] Frederik Brudy, Joshua Kevin Budiman, Steven Houben, and Nicolai Marquardt. 2018. Investigating the Role of an Overview Device in Multi-Device Collaboration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173874

[12] Frederik Brudy, Christian Holz, Roman Rädle, Chi-Jui Wu, Steven Houben, Clemens Nylandsted Klokmose, and Nicolai Marquardt. 2019. Cross-Device Taxonomy: Survey, Opportunities and Challenges of Interactions Spanning Across Multiple Devices. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, 562:1–562:28. https://doi.org/10.1145/3290605.3300792

[13] Daniel Chaffin, Ralph Heidl, John R Hollenbeck, Michael Howe, Andrew Yu, Clay Voorhees, and Roger Calantone. 2017. The promise and perils of wearable sensors in organizational research. *Organizational Research Methods* 20, 1 (2017), 3–31.

[14] Nicholas Chen, Francois Guimbretiere, Morgan Dixon, Cassandra Lewis, and Maneesh Agrawala. 2008. Navigation techniques for dual-display e-book readers. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1779–1788.

[15] Lung-Pan Cheng, Meng Han Lee, Che-Yang Wu, Fang-I Hsiao, Yen-Ting Liu, Hsiang-Sheng Liang, Yi-Ching Chiu, Ming-Sui Lee, and Mike Y Chen. 2013. iRotateGrasp: automatic screen rotation based on grasp of mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3051–3054.

[16] Lung-Pan Cheng, Hsiang-Sheng Liang, Che-Yang Wu, and Mike Y Chen. 2013. iGrasp: grasp-based adaptive keyboard for mobile devices. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 3037–3046.

[17] Roudy Dagher, Francois-Xavier Molina, Alexandre Abadie, Nathalie Mitton, and Emmanuel Baccelli. 2021. An Open Experimental Platform for Ranging, Proximity and Contact Event Tracking Using Ultra-Wide-Band and Bluetooth Low-Energy. In *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. 1–6. https://doi.org/10.1109/INFOCOMWKSHPS51825.2021.9484579

[18] David Dearman, Richard Guy, and Khai Truong. 2012. Determining the Orientation of Proximate Mobile Devices Using Their Back Facing Camera. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. Association for Computing Machinery, New York, NY, USA, 2231–2234. https://doi.org/10.1145/2207676.2208377

[19] Estimote. 2021. Estimote UWB Beacons. https://estimote.com/uwb-beacons

[20] Katherine Everitt, Clifton Forlines, Kathy Ryall, and Chia Shen. 2004. Observations of a shared tabletop user study. *Interactive Poster, CSCW* (2004).

[21] Katherine Everitt, Chia Shen, Kathy Ryall, and Clifton Forlines. 2006. MultiSpace: Enabling electronic document micro-mobility in table-centric, multi-device environments. In *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'06)*. IEEE, 8–pp.

[22] Kenneth P Fishkin, Anuj Gujar, Beverly L Harrison, Thomas P Moran, and Roy Want. 2000. Embodied user interfaces for really direct manipulation. *Commun. ACM* 43, 9 (2000), 74–80.

[23] Tian Gan, Yongkang Wong, Daqing Zhang, and Mohan S Kankanhalli. 2013. Temporal encoded F-formation system for social interaction detection. In *Proceedings of the 21st ACM international conference on Multimedia*. 937–946.

[24] Mayank Goel, Jacob Wobbrock, and Shwetak Patel. 2012. Gripsense: using built-in sensors to detect hand posture and pressure on commodity mobile phones. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*. 545–554.

[25] Saul Greenberg, Michael Boyle, and Jason Laberge. 1999. PDAs and Shared Public Displays: Making Personal Information Public, and Public Information Personal. *Personal Technologies* 3, 1 (March 1999), 54–64. https://doi.org/10.1007/BF01305320

[26] Saul Greenberg, Nicolai Marquardt, Till Ballendat, Rob Diaz-Marino, and Miaosen Wang. 2011. Proxemic Interactions: The New Ubicomp? *Interactions* 18, 1 (Jan. 2011), 42–50. https://doi.org/10.1145/1897239.1897250

[27] Tobias Grosse-Puppendahl, Christian Holz, Gabe Cohn, Raphael Wimmer, Oskar Bechtold, Steve Hodges, Matthew S Reynolds, and Joshua R Smith. 2017. Finding common ground: A survey of capacitive sensing in human-computer interaction. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 3293–3315.

[28] Edmund T Hall and Edward Twitchell Hall. 1966. *The hidden dimension*. Vol. 609. Anchor.

[29] Mike Hazas, Christian Kray, Hans Gellersen, Henoc Agbota, Gerd Kortuem, and Albert Krohn. 2005. A Relative Positioning System for Co-Located Mobile Devices. In *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services*. ACM, Seattle, Washington, 177–190. https://doi.org/10.1145/1067170.1067190

[30] Mike Hazas and Andy Ward. 2002. A Novel Broadband Ultrasonic Location System. In *UbiComp 2002: Ubiquitous Computing*, Gaetano Borriello and Lars Erik Holmquist (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 264–280.

[31] Milad Heydariaan, Hossein Dabirian, and Omprakash Gnawali. 2020. Anguloc: Concurrent angle of arrival estimation for indoor localization with uwb radios. In *2020 16th International Conference on Distributed Computing in Sensor Systems (DCOSS)*. IEEE, 112–119.

[32] Ken Hinckley, Michel Pahud, Hrvoje Benko, Pourang Irani, François Guimbretière, Marcel Gavriliu, Xiang'Anthony' Chen, Fabrice Matulic, William Buxton, and Andrew Wilson. 2014. Sensing techniques for tablet+ stylus interaction. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. 605–614.

[33] Ken Hinckley, Gonzalo Ramos, Francois Guimbretiere, Patrick Baudisch, and Marc Smith. 2004. Stitching: pen gestures that span multiple displays. In *Proceedings of the working conference on Advanced visual interfaces*. 23–31.

[34] Ken Hinckley and Hyunyoung Song. 2011. Sensor synaesthesia: touch in motion, and motion in touch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 801–810.

[35] Da-Yuan Huang, Chien-Pang Lin, Yi-Ping Hung, Tzu-Wen Chang, Neng-Hao Yu, Min-Lun Tsai, and Mike Y. Chen. 2012. MagMobile: Enhancing Social Interactions with Rapid View-Stitching Games of Mobile Devices. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia (MUM '12)*. Association for Computing Machinery, New York, NY, USA, 1–4. https://doi.org/10.1145/2406367.2406440

[36] Ke Huo, Yuanzhi Cao, Sang Ho Yoon, Zhuangying Xu, Guiming Chen, and Karthik Ramani. 2018. Scenariot: Spatially Mapping Smart Things Within Augmented Reality Scenes. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173793

[37] Shoichi Inaba and Yoshimitsu Aoki. 2016. Conversational group detection based on social context using graph clustering algorithm. In *2016 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*. IEEE, 526–531.

[38] Haojian Jin, Christian Holz, and Kasper Hornbæk. 2015. Tracko: Ad-hoc mobile 3d tracking using bluetooth low energy and inaudible signals for cross-device interaction. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. 147–156.

[39] Rong Jin, Liu Shi, Kai Zeng, Amit Pande, and Prasant Mohapatra. 2016. MagPairing: Pairing Smartphones in Close Proximity Using Magnetometers. *IEEE Transactions on Information Forensics and Security* 11, 6 (June 2016), 1306–1320. https://doi.org/10.1109/TIFS.2015.2505626

[40] Brad Johanson, Armando Fox, and Terry Winograd. 2002. The interactive workspaces project: Experiences with ubiquitous computing rooms. *IEEE pervasive computing* 1, 2 (2002), 67–74.

[41] Brad Johanson, Greg Hutchins, Terry Winograd, and Maureen Stone. 2002. PointRight: experience with flexible input redirection in interactive workspaces. In *Proceedings of the 15th annual ACM symposium on User interface software and technology*. 227–234.

[42] Wendy G. Ju, Brian A. Lee, and Scott R. Klemmer. 2007. Range: Exploring Proxemics in Collaborative Whiteboard Interaction. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems*. ACM, San Jose, CA, USA, 2483–2488. https://doi.org/10.1145/1240866.1241028

[43] R Karthikeyan, S Karthik, and S A Pasupathy. 2021. UWB Based Novel Approach for Orientation Estimation and Path Planning for Field Robots without IMU. In *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*. 1–6. https://doi.org/10.1109/ICAECA52838.2021.9675757

[44] Kee-Eung Kim, Wook Chang, Sung-Jung Cho, Junghyun Shim, Hyunjeong Lee, Joonah Park, Youngbeom Lee, Sangryoung Kim, et al. 2006. Hand grip pattern recognition for mobile user interfaces. In *AAAI*. 1789–1794.

[45] Sangtae Kim, Jaejeung Kim, and Soobin Lee. 2013. Bezel-flipper: design of a light-weight flipping interface for e-books. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. 1719–1724.

[46] Stefanie Klum, Petra Isenberg, Ricardo Langner, Jean-Daniel Fekete, and Raimund Dachselt. 2012. Stackables: Combining Tangibles for Faceted Browsing. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*

(AVI '12). Association for Computing Machinery, New York, NY, USA, 241–248. https://doi.org/10.1145/2254556.2254600

[47] Rakhesh Singh Kshetrimayum. 2009. An introduction to UWB communication systems. *Ieee Potentials* 28, 2 (2009), 9–13.

[48] David Ledo, Saul Greenberg, Nicolai Marquardt, and Sebastian Boring. 2015. Proxemic-Aware Controls: Designing Remote Controls for Ubiquitous Computing Ecologies. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 187–198. https://doi.org/10.1145/2785830.2785871

[49] Andrés Lucero, Jussi Holopainen, and Tero Jokela. 2011. Pass-Them-around: Collaborative Use of Mobile Phones for Photo Sharing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 1787–1796. https://doi.org/10.1145/1978942.1979201

[50] Paul Luff and Christian Heath. 1998. Mobility in collaboration. In *Proceedings of the 1998 ACM conference on Computer supported cooperative work*. 305–314.

[51] Sus Lundgren, Joel E. Fischer, Stuart Reeves, and Olof Torgersson. 2015. Designing Mobile Experiences for Collocated Interaction. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 496–507. https://doi.org/10.1145/2675133.2675171

[52] Christine L MacKenzie and Thea Iberall. 1994. *The grasping hand.* Elsevier.

[53] Nicolai Marquardt, Frederik Brudy, Can Liu, Ben Bengler, and Christian Holz. 2018. SurfaceConstellations: A Modular Hardware Platform for Ad-Hoc Reconfigurable Cross-Device Workspaces. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3173928

[54] Nicolai Marquardt, Robert Diaz-Marino, Sebastian Boring, and Saul Greenberg. 2011. The Proximity Toolkit: Prototyping Proxemic Interactions in Ubiquitous Computing Ecologies. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. Association for Computing Machinery, New York, NY, USA, 315–326. https://doi.org/10.1145/2047196.2047238

[55] Nicolai Marquardt, Nathalie Henry Riche, Christian Holz, Hugo Romat, Michel Pahud, Frederik Brudy, David Ledo, Chunjong Park, Molly Jane Nicholas, Teddy Seyed, et al. 2021. AirConstellations: In-Air Device Formations for Cross-Device Interaction via Multiple Spatially-Aware Armatures. In *The 34th Annual ACM Symposium on User Interface Software and Technology*. 1252–1268.

[56] Nicolai Marquardt, Nathalie Henry Riche, Christian Holz, Hugo Romat, Michel Pahud, Frederik Brudy, David Ledo, Chunjong Park, Molly Jane Nicholas, Teddy Seyed, Eyal Ofek, Bongshin Lee, William A.S. Buxton, and Ken Hinckley. 2021. AirConstellations: In-Air Device Formations for Cross-Device Interaction via Multiple Spatially-Aware Armatures. In *The 34th Annual ACM Symposium on User Interface Software and Technology (UIST '21)*. Association for Computing Machinery, New York, NY, USA, 1252–1268. https://doi.org/10.1145/3472749.3474820

[57] Nicolai Marquardt, Ken Hinckley, and Saul Greenberg. 2012. Cross-device interaction via micro-mobility and f-formations. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*. 13–22.

[58] Andrea H Mason and Christine L MacKenzie. 2005. Grip forces when passing an object to a partner. *Experimental brain research* 163, 2 (2005), 173–187.

[59] Mohammad Faizuddin Mohd Noor, Andrew Ramsay, Stephen Hughes, Simon Rogers, John Williamson, and Roderick Murray-Smith. 2014. 28 frames later: predicting screen touches from back-of-device grip changes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2005–2008.

[60] Halla B Olafsdottir, Theophanis Tsandilas, and Caroline Appert. 2014. Prospective motor control on tabletops: planning grasp for multitouch interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2893–2902.

[61] Gary M Olson and Judith S Olson. 2000. Distance matters. *Human–computer interaction* 15, 2-3 (2000), 139–178.

[62] James O'Malley. 2021. The imminent ultra-wideband revolution: After decades in the lab, big tech has finally found a use for a unique wireless technology. *Engineering & Technology* 16, 9 (2021), 1–4.

[63] Qorvo. 2021. Qorvo DWM3000 6.5 & 8.0 GHz Ultra-Wideband (UWB) Module. https://www.qorvo.com/products/p/DWM3000#documents

[64] Roman Rädle, Hans-Christian Jetter, Jonathan Fischer, Inti Gabriel, Clemens N. Klokmose, Harald Reiterer, and Christian Holz. 2018. PolarTrack: Optical Outside-In Device Tracking That Exploits Display Polarization. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3173574.3174071

[65] Jun Rekimoto and Masanori Saitoh. 1999. Augmented surfaces: a spatially continuous work space for hybrid computing environments. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 378–385.

[66] Elisa Ricci, Jagannadan Varadarajan, Ramanathan Subramanian, Samuel Rota Bulo, Narendra Ahuja, and Oswald Lanz. 2015. Uncovering interactions and interactors: Joint estimation of head, body orientation and f-formations from surveillance videos. In *Proceedings of the IEEE International Conference on Computer Vision*. 4660–4668.

[67] Hamza Sadruddin, Ahmed Mahmoud, and Mohamed Atia. 2019. An Indoor Navigation System Using Stereo Vision, IMU and UWB Sensor Fusion. In *2019 IEEE SENSORS*. 1–4. https://doi.org/10.1109/SENSORS43011.2019.8956942

[68] Bill N Schilit, Norman Adams, Rich Gold, Michael M Tso, and Roy Want. 1993. The ParcTab Mobile Computing System. In *In Proceedings of the Fourth Workshop on Workstation Operating Systems*. 34–39. https://doi.org/10.1.1.43.9880

[69] Dominik Schmidt, Fadi Chehimi, Enrico Rukzio, and Hans Gellersen. 2010. PhoneTouch: a technique for direct phone interaction on surfaces. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*. 13–16.

[70] Stacey D Scott, M Sheelagh T Carpendale, and Kori Inkpen. 2004. Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*. 294–303.

[71] Sarthak Sharma, Hessam Mohammadmoradi, Milad Heydariaan, and Omprakash Gnawali. 2019. Device-Free Activity Recognition Using Ultra-Wideband Radios. In *2019 International Conference on Computing, Networking and Communications (ICNC)*. 1029–1033. https://doi.org/10.1109/ICCNC.2019.8685504

[72] Chia Shen, Katherine Everitt, and Kathleen Ryall. 2003. UbiTable: Impromptu face-to-face collaboration on horizontal interactive surfaces. In *International Conference on Ubiquitous Computing*. Springer, 281–288.

[73] Norbert A Streitz, Jörg Geißler, Torsten Holmer, Shin'ichi Konomi, Christian Müller-Tomfelde, Wolfgang Reischl, Petra Rexroth, Peter Seitz, and Ralf Steinmetz. 1999. i-LAND: an interactive landscape for creativity and innovation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 120–127.

[74] Julie Wagner, Stéphane Huot, and Wendy Mackay. 2012. BiTouch and BiPad: designing bimanual interaction for hand-held tablets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2317–2326.

[75] James R. Wallace, Stacey D. Scott, and Carolyn G. MacGregor. 2013. Collaborative Sensemaking on a Digital Tabletop and Personal Tablets: Prioritization, Comparisons, and Tableaux. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 3345–3354. https://doi.org/10.1145/2470654.2466458

[76] Yuntao Wang, Jiexin Ding, Ishan Chatterjee, Farshid Salemi Parizi, Yuzhou Zhuang, Yukang Yan, Shwetak Patel, and Yuanchun Shi. 2022. FaceOri: Tracking Head Position and Orientation Using Ultrasonic Ranging on Earphones. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3491102.3517698

[77] Roy Want, Andy Hopper, Veronica Falcão, and Jonathan Gibbons. 1992. The Active Badge Location System. *ACM Transactions on Information Systems* 10, 1 (Jan. 1992), 91–102. https://doi.org/10.1145/128756.128759

[78] Thomas Wells and Steven Houben. 2020. CollabAR - Investigating the Mediating Role of Mobile AR Interfaces on Co-Located Group Collaboration. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376541

[79] Thomas Wells, Dominic Potts, and Steven Houben. 2022. A Study into the Effect of Mobile Device Configurations on Co-Located Collaboration Using AR. *Proceedings of the ACM on Human-Computer Interaction* 6, MHCI (Sept. 2022), 200:1–200:23. https://doi.org/10.1145/3546735

[80] Daniel Wigdor, Hao Jiang, Clifton Forlines, Michelle Borkin, and Chia Shen. 2009. WeSpace: The Design Development and Deployment of a Walk-up and Share Multi-surface Visual Collaboration System. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 1237–1246. https://doi.org/10.1145/1518701.1518886

[81] Raphael Wimmer. 2010. Grasp sensing for human-computer interaction. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*. 221–228.

[82] Raphael Wimmer and Sebastian Boring. 2009. HandSense: discriminating different ways of grasping and holding a tangible user interface. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*. 359–362.

[83] Katrin Wolf, Christian Müller-Tomfelde, Kelvin Cheng, and Ina Wechsung. 2012. PinchPad: performance of touch-based gestures while grasping devices. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*. 103–110.

[84] Chi-Jui Wu, Steven Houben, and Nicolai Marquardt. 2017. EagleSense: Tracking People and Devices in Interactive Spaces Using Real-Time Top-View Depth-Sensing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3929–3942. https://doi.org/10.1145/3025453.3025562

[85] Dongwook Yoon, Ken Hinckley, Hrvoje Benko, François Guimbretière, Pourang Irani, Michel Pahud, and Marcel Gavriliu. 2015. Sensing tablet grasp+ micro-mobility for active reading. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. 477–487.

[86] Ye Yuan, Nathalie Riche, Nicolai Marquardt, Molly Jane Nicholas, Teddy Seyed, Hugo Romat, Bongshin Lee, Michel Pahud, Jonathan Goldstein, Rojin Vishkaie,

et al. 2022. Understanding Multi-Device Usage Patterns: Physical Device Configurations and Fragmented Workflows. In *CHI Conference on Human Factors in Computing Systems*. 1–22.

[87] Yang Zhang, Michel Pahud, Christian Holz, Haijun Xia, Gierad Laput, Michael McGuffin, Xiao Tu, Andrew Mittereder, Fei Su, William Buxton, et al. 2019. Sensing posture-aware pen+ touch interaction on tablets. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.

[88] Yuan Zhou, Choi Look Law, and Jingjing Xia. 2012. Ultra low-power UWB-RFID system for precise location-aware applications. In *2012 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. IEEE, 154–158.