We are now capable of interacting in various ways with data visualization on mobile devices. In this chapter, we characterize how interacting with visualization using a mobile phone or tablet differs from analogous experiences using a PC. We provide an overview of the topic organized by interaction modality, beginning with touch interaction and subsequently discussing instances of spatial interaction and voice interaction. As an outlook, we envision compelling opportunities for future
mobile data visualization research inspired by recent developments in the field of mobile human-computer interaction.

3.1 FOUNDATIONS

The high resolution displays of current mobile devices allow you to see minute levels of detail in visualization content. Many devices are also built with powerful processors, capable of not only representing thousands of data points simultaneously, but also of responding to changing data, a changing surrounding environment, and a changing stream of interactions with the device, as described in Chapter 2.

Mobile phones and tablets have screens of varying size and aspect ratio, as well as different sets of sensors, and these differences affect how you interact with these devices. Despite these differences, the current ecosystem of two mobile operating systems (iOS and Android) incorporate a consistent set of interactions across devices. On the one hand, the consistency can make interactions easier to discover, on the other hand, it may limit innovation in mobile interaction design.

In this chapter, we do not constrain ourselves to existing conventions, but turn to research activities that envision possible futures in which you might interact with visual representations of data in unprecedented ways. An overarching assumption of this chapter is that you want and often need to interact with data visualization via a mobile device: that there exist circumstances in which you are unsatisfied with static visual representations of data, and that you become frustrated when you cannot interact easily or at all, such as when a website tells you to revisit the page from a PC [14]. The data that you interact with on your mobile device may be highly personal and immediately relevant to your surrounding environment. Your location, your health and activity data, your personal finances, and the local weather forecast are all examples of personal data that you might already regularly interact with via your mobile device. We assume that there are various occasions in which you are genuinely curious about these data, for example, to make comparisons that inform your decision-making, or to examine anomalous values and possible trends.

The intent to engage further and interact stands in contrast to cases in which all that is required is a succinct representation of data designed for quick monitoring tasks and glanceability, cases that are discussed at greater length in Chapter 5. Mobile interaction also stands in contrast to cases of passive consumption in communicative visualization scenarios, such as when viewing news graphics. There was ample experimentation with interactive news graphics during the first half of the 2010s, however this enthusiasm tapered in the second half of the decade following reports of how little people interact with news graphics beyond scrolling [2, 118].

One response to this realization has been to expand the capabilities of scrolling to trigger various events in a visual representation of data, often referred to as scrolllytelling [119]. Another response has been to incorporate viewers’ personal information as a form of interaction, such as detecting their location or soliciting personal details from viewers, such as their education, income, or occupation; both approaches can be used to generate custom views of a dataset that people might be more inclined to engage with.
Another interesting invitation to interact has appeared in data journalism focusing on trends or correlations, where the viewers are prompted to first guess the trend by drawing it [3]. According to Nguyen et al. [82], this format allows the viewers to externalize and test their own beliefs. Graphic elicitation of viewers’ beliefs via drawing is particularly well-suited for mobile touchscreen devices. Similar approaches to engage a large mobile audience with news stories that incorporate visualization will continue to evolve in the coming years as media agencies strive to seek a balance between development cost and value for the viewers.

This chapter presents an overview of recent advances in interacting with visualization via mobile devices. We focus on mobile phone and tablet devices; we do not consider watches or other wearable devices such as bands, rings, or head-mounted augmented reality displays, though we do comment on their potential in our discussion of future opportunities later in Section 3.3. We do, however, mention a few compelling examples of handheld mobile augmented reality, as well as one example when a mobile device is used in conjunction with head-mounted augmented reality system called FieldView [124].

We organize our overview according to interaction modalities: touch interaction, spatial interaction, and voice interaction. We anticipate that people will continue to interact via these modalities in future devices, and accordingly we hope that this chapter be read as an overview of multimodal interaction possibilities for mobile data visualization. Due to the rapidly evolving and ephemeral nature of mobile software, of which there tends to be no archival description, we concentrate on mobile interaction with visualization as described in peer-reviewed archival research literature. Yet, where relevant we also refer to commercial mobile applications and compelling instances of mobile interaction with visualization designed by practitioners.

Before we proceed with our overview, it is necessary to establish context about interactive visualization and interacting with mobile devices. We also realize that the scope of this overview has foundations in several areas of research, each with an associated body of literature and a research community, and we refer interested readers to these communities where relevant.

3.1.1 Interacting with Visualization

The history of creating visual representations based on data and using them to perform analyses and communicate insights to others is a rich tapestry spanning millennia and cultures. Until a few decades ago, the act of interacting with visualization entailed drawing, engraving, or sculpting representations of data, and then examining and manipulating static physical media. With the advent of computers, interacting with visualization meant interacting with dynamic media. But what makes interactive visualization different from other interactive media, such as video games, illustration tools, or word processors? The answer is simultaneously a difference in one’s intents and a different set of interactions needed to satisfy these intents.

Looking back on the past three decades of visualization systems and research papers, many researchers have proposed typologies [21, 44, 102, 128] of intent and interaction that offer some consensus, at least at an abstract level, of what makes
interacting with visualization distinct from interacting with other media forms. At the level of intent, people want to analyze, monitor, and communicate aspects of data, and to ask questions or anticipate questions that their audience might ask. In many circumstances, the specific referents of these questions cannot be specified a priori. Similarly, the data may also have a dynamic nature, where new data may provoke new questions. As a result, a single static representation of data often does not suffice.

Therefore, many interactive visualization tools allow people to navigate across their data, to filter them and select subsets of interest, to sort them, to change the way they are represented, to adjust how these representations are arranged, and to augment these representations with their insights [117]. How these interactions manifest vary from one visualization tool or environment to another [116], and as Dimara & Perin [32] note, multiple interactions might redundantly support the same intent. As observed by Lee et al. [71], Jansen and Dragicevic [54], and Roberts et al. [93], computer-mediated interactions with visualization, until relatively recently, involved interacting with only a mouse and keyboard.

### 3.1.2 Interacting with Mobile Devices

You might recall a time where mobile devices came with a hardware keyboard and a trackball for interaction. Nowadays, mobile devices are typically equipped with a multi-touch display, perhaps one that can even detect the amount of force with which you press on the screen, and with a small number of hardware buttons around its periphery. In addition, there are various other ways by which you interact with your mobile device in contrast to how you might interact with your PC, and it can be helpful to discuss the gamut of possible interactions for what follows below.

First, many mobile devices are equipped with accelerometers, light sensors, pressure sensors, and multiple cameras. These sensors make it possible to detect events such as changes of position and orientation in space, your grip and hand posture, touch events with varying levels of pressure, or the appearance of a face. Second, microphones allow you to speak, record, and sample audio. Third, some mobile devices are compatible with a pen or stylus, allowing you to write, sketch, and make fine selections. Fourth, many mobile devices provide haptic or vibrotactile feedback, either as a means to provide notifications or as a means of indicating that an interaction was recognized by the device. Finally, many mobile devices have various levels of geolocation tracking via GPS, WiFi, and Bluetooth. While it is true that many laptop PCs ship with geolocation awareness, a camera, and a microphone, they tend to be stationary when in use; it is the mobility of phones and tablets that expands the interactive potential of these sensors.

After the introduction of the iPhone in 2007, several touchscreen interactions have become familiar to us, and their effects can often be predictable when using a new mobile application. Consider the many contexts in which you tap, pinch, swipe, and tap & hold. However, beyond common touch actions, there are fewer conventions among the other modalities of interaction that we have listed. As a consequence, designers must consider creative strategies for ensuring the discoverability and learnability of these interactions.
In differentiating the ways in which people interact with mobile devices relative to how they interact with PCs, we must also consider the potentially different contexts of interaction. When you interact with a PC, you are often sitting or standing still. You may be in a professional setting, in your home office, or in a classroom, and you typically commit to longer sustained periods of interaction. In contrast, the contexts in which one interacts with mobile devices are more heterogeneous in terms of their surrounding physical environments, the relationships between other people and devices, and the cadence of interaction. Consider, for instance, that work contexts are becoming increasingly situated and collaborative, where mobile phones and tablets are often sufficient to perform tasks. Regardless of whether one finds themselves in a professional or casual setting, interaction with a mobile device may be more intermittent and fleeting than with a PC.

Despite the different usage contexts for mobile devices and PCs, we note that the distinction between laptop PCs and tablets is beginning to blur. For instance, some Microsoft’s Surface devices [77] and others like it are equipped with touchscreens and can be converted between laptop and tablet modes. Meanwhile, tablets such as Apple’s iPad Pro [5] boast screens as large as laptops, powerful hardware capabilities as good as many PCs, and peripheral keyboard attachments. These hybrid devices provide affordances to combine bimanual touch- and gesture-based direct manipulation with conventional keyboard, mouse, and trackpad interaction in the context of both WIMP- (Windows, Icons, Menus, Pointer) and post-WIMP interfaces.

We also note that the distinction between larger smartphones and tablets is also blurring, as evident by the use of the phablet moniker for the former. The overview we present in this chapter is a retrospective on the past decade of research, where most of the examples that we consider are associated with a specific device type. As a collection, however, the examples we cite along with our commentary may inform interaction design for and future research involving these emerging classes of devices that blur the boundaries between laptop and tablet or tablet and phone.

With this brief initial description of interaction on mobile devices, let us next look at interaction for mobile data visualization in detail.

### 3.2 OVERVIEW

On the one hand, we have a foundational understanding of how people interact with visualization. On the other hand, we have a foundational understanding of how people interact with mobile devices. This overview examines the intersection of these interactions; and while the research pertaining to interacting with visualization on mobile devices predates multitouch-enabled phones and tablets (e.g., [57, 22, 43, 51, 30]), our overview focuses on research published since 2010. Table 3.1 summarizes our overview.

It is also worth considering what remains outside of this intersection. Are some interactions with visualization incompatible with mobile devices? In Chapter 2, we encountered several strategies and related challenges for responsive visualization design, which include modifying the interaction design for different device profiles. However, given the breadth of interactions in the research literature, there may be
Table 3.1: A chronologically-ordered summary of the specific projects that we reference in our overview and the main device form factors associated with these instances: □ = phone; □ = tablet; □ = large display; □ = head-mounted display. We also denote the modalities of interaction that each project incorporates: □ = touch (+□ = pen); □ = voice; □ = spatial (□ = using cameras, □ = using (geo)location)

<table>
<thead>
<tr>
<th>Project (Year)</th>
<th>Reference</th>
<th>Device(s)</th>
<th>Modalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tangible views (2010)</td>
<td>Spindler et al. [107]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>TouchWave (2012)</td>
<td>Baur et al. [13]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>TouchViz (2013)</td>
<td>Drucker et al. [34]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>Kinetica (2014)</td>
<td>Rzeszotarski &amp; Kittur [95]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>Tangere (2014-16)</td>
<td>Sadana et al. [96, 97, 98]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>GraphTiles (2015)</td>
<td>Bae et al. [8]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>Subspotting (2016)</td>
<td>Baur &amp; Goddemayer [12]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>TouchPivot (2017)</td>
<td>Jo et al. [56]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>GraSp (2017)</td>
<td>Kister et al. [61]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>VisTiles (2017)</td>
<td>Langner et al. [67]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>Ranges over time (2018)</td>
<td>Brehmer et al. [20]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>SmartCues (2018)</td>
<td>Subramonyam &amp; Adar [111]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>AffinityLens (2019)</td>
<td>Subramonyam et al. [112]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>Pan + zoom eval. (2019)</td>
<td>Schwab et al. [101]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>FieldView (2019)</td>
<td>Whitlock et al. [124]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>Pressure sensing (2019)</td>
<td>Wang et al. [123]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>MARVisT (2019)</td>
<td>Chen et al. [27]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>InChorus (2020)</td>
<td>Srinivasan et al. [108]</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>Orchard (2020)</td>
<td>Eichmann et al. [36]</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

some that are unlikely to apply in mobile contexts, and others may be difficult or tedious to perform, or they may be difficult to discover, having no precedent in other application contexts. For instance, visualization authoring often requires a series of interactions to select data, apply transformations to the data, and specify a visual representation; would people be willing to perform this series of authoring interactions via a mobile device? Another activity that often entails a series of interactions, photo editing, has until recently been reserved for PCs. However, mobile photo editing with apps like Instagram is now commonplace. Could visualization authoring similarly become an activity that people carry out using a mobile device?

There are also some interactions with mobile devices that are incompatible with visualization. For instance, there are interactions that do not involve looking at the display, such as silencing notifications, adjusting volume controls, or interacting via auditory or haptic channels (we do not discuss data sonification and its haptic analog in this chapter). Finally, we must acknowledge that the limitations of contemporary
devices may impose constraints on what interactions people can perform involving visualization, and we return to this topic at the end of this chapter.

As mentioned above, the overview skews heavily toward instances of interacting with visualization via mobile devices as described in archival research literature. We acknowledge that the marketplace of mobile applications incorporating interactive visualization features is growing and evolving. Major business intelligence software vendors such as Microsoft [76], Tableau [114], Qlik [91], MicroStrategy [79], and Thoughtspot [115] all have mobile versions of their visualization solutions as of the time of writing, and these take various approaches to interacting with visualization. There are also many mobile-first or responsive news graphics that feature interactivity to some extent; Ros [94] catalogued several throughout the mid-2010s. However, the ecosystem of applications and responsive interactive news graphics is highly ephemeral, and written accounts of their interaction design choices are uncommon.

We organize our overview according to interaction modality. At the same time, we acknowledge the additive nature of interaction modalities; spatial or voice interaction usually accompanies, rather than replaces, touch interaction. As suggested by Table 3.1, nearly all of the projects that we surveyed incorporate touch interaction. Spatial interaction receives the next most coverage, while voice is discussed to a lesser extent, this being a reflection of its prevalence in the research literature on visualization and mobile devices.

3.2.1 Touch Interaction

Research examining the potential of multi-finger touch interaction for visualization started to accumulate around the beginning of the 2010s, following the commercialization of new touchscreen technology and touchscreen tabletop displays in particular and early research by Isenberg et al. [52], Frisch et al. [38], and North et al. [84]. With the introduction of the iPhone in 2007 and the iPad in 2010 as well as the popularization of multi-touch mobile devices, visualization researchers and designers turned their attention to smaller devices.

There are several challenges associated with touch interaction for visualization via mobile devices. First, there is the fat finger problem, the mismatch between the size of graphical marks and human finger tips. This problem is particularly acute with small visual elements that are to be selectable on a mobile phone or tablet, as for example when picking individual points from a scatterplot or narrow segments from a stacked bar chart. Interaction designers therefore face a trade-off between the minimum visibility of marks and their selectability: if all marks are to be easily and directly selected via touch, they must be suitably large, akin to a button. Yet, using larger marks is infeasible in many situations, and thus designers must consider alternative approaches, such as a multi-step or hierarchical selection upon touching a region with the visualization, which may involve selection from a modal menu panel or a modal zoom lens of the touch area, such as in Kinetica [95] or Tangere [96]. Adding larger invisible touch targets or an invisible Voronoi tessellation around small marks is another approach to support interactive selection [14]. Zooming in as a prerequisite to individual mark selection is also possible as a last resort.
A related challenge is the simultaneous visibility and selectability of targets due to occlusion from the finger or hand. If selection has no visible consequences within the remaining unoccluded area of the display, such as via a lens widget, the result of the selection can only be known once the finger or hand is moved away.

Unintended touch is another challenge, particularly for touch surfaces whereupon people may rest their hands or at least the base of their palms, for example, when laying a mobile device flat on a table. Accidental touch may also occur when a mobile device is held in one’s hand or positioned at an inclination. The reduction of screen bezel widths across devices in recent years may exacerbate this problem. Also problematic is misinterpreted touch, in which one touch gesture is confused for another, such as confusing a two-finger pinch with a two-finger rotation. Both forms of touch recognition error continue to be problematic for touchscreen interfaces.

Given the limited vocabulary of touch and the restrictions on touch target size, interacting with visualization via mobile devices is often reliant upon menus and explicit modes of interaction, such as alternating between navigation and selection. However, sufficiently large menu interfaces often occlude content on small screens, and different interaction modes are difficult to discover. As a consequence, the number and variety of unique touch interactions tends to be small in most instances, which often forego the ability to select individual marks.

One of the major challenges presented to visualization designers with respect to touch-based interfaces is the relative lack of unique, differentiable gestures for initiating different operations. While it is possible to use more complex multi-finger or bimanual gestures, these are correspondingly more difficult to discover and remember, and are likely more difficult to perform, especially on handheld devices. For example, consider a simple swipe gesture. A designer may associate swiping along an axis to select visual marks in a view or swiping could pan the view to show other data ranges. Both of these mappings are possible, but the gesture must be uniquely assigned to one operation. Mapping tasks to interactions is a non-trivial problem in general; a deeper discussion is provided by Gladisch et al. [41]. While some gestures have become familiar over time, Isenberg & Hancock [53] caution that new gestures are difficult to perform, memorize, and discover. Instead, they advocate for postures that re-use and combine simpler interactions for direct parameter control, and that such postures are easier to remember and discover than a set of gestures.

This section on touch interaction is structured according to device class: tablet or mobile phone. For each class, we profile several notable projects with touch interaction of different complexity. At one end of the spectrum, we have tapping, holding, or double tapping, which do not involve any motion. As we increase the complexity, we encounter actions such as pulling to refresh or swiping, both of which involving motion along a single direction. Gestures such as pinching to zoom, lassoing to enclose, and dragging to reposition often involve motion along two dimensions. Finally, the projects in our list include single-handed as well as bimanual interaction, wherein the dominant hand may be holding a pen or stylus to either draw, write, point, or select.
Interacting with Visualization on Mobile Devices

Touch Interaction on Tablets

Since the release of the iPad in 2010, a number of visualization research projects have examined the potential of this form factor and the touch interaction it affords.

With TouchWave, Baur et al. [13] introduced a set of multi-touch interactions specifically tailored for directly manipulating stacked area charts on tablets without relying on widgets and mode switches. To keep the touch interactions simple, they started with established touch interactions (e.g., tap, tap & hold, pinch, swipe) as much as possible, and then expanded the set by incorporating multi-finger gestures, some involving motion and others not (see Figure 3.1). TouchWave also leverages contextual information where the interaction occurs to provide more appropriate response and feedback, such as a single tap on the background canvas to invoke a vertical ruler perpendicular to the horizontal time axis, superimposed with text annotations revealing the value corresponding to each band at the selected time point. The system reacts differently for vertical and horizontal scaling that uses the same pinch gesture: the context is compressed while the focus region is expanded in horizontal scaling, while vertical scaling magnifies the vertical axis uniformly to keep the relative sizes of the layers intact.

Figure 3.1: With TouchWave, Baur et al. [13] introduced a set of touch gestures for manipulating stacked area charts on tablets. Video stills courtesy of Dominikus Baur, used with permission; watch the full video at https://youtu.be/tZ1EJoY8Hck.

In the TouchViz project, Drucker et al. [34] designed and compared two interfaces for working with bar charts on tablet devices. This comparison is notable given the recent blurring between WIMP and post-WIMP interfaces for devices that straddle the boundary between laptop and tablet, as described above. To develop a set of post-WIMP gestures for their FLUID interface, Drucker et al. conducted a structured brainstorming session, where they chose to focus on the gestures that involved manipulation mapped directly onto objects on screen. Similar to TouchWave before, the post-WIMP FLUID interface strove to minimize the use of buttons and controls, ensuring that all gestures occur on the chart itself. In contrast, the alternative WIMP interface featured the same interactions accessed via buttons and menu commands.
(see Figure 3.2). Unsurprisingly, the FLUID approach was predominantly favored by study participants.

Figure 3.2: In the TouchViz project, Drucker et al. [34] compared a WIMP interface using menus of buttons (left) with a FLUID direct manipulation touch interface (right). Video stills courtesy of Ramik Sadana, used with permission; watch the full video at https://vimeo.com/57416758.

**Kinetica** [95] is a touch interface for unit charts, in which each data point is represented by a small circular mark. The system implemented a physics-based interface in which a person’s fingers and gestures acted on those marks, giving the impression that one can directly push or sweep marks around the display as if they were a set of colliding particles. Placing two fingers on the display could specify two control points for a histogram or a bounding box for a scatterplot. Overall, the system employed a number of custom gestures which applied to the unique physics-based view it provided. Not all of Kinetica’s gestures involved direct and continuous manipulation of marks. For example, a spiral-shaped swipe defined a spiral curve along which to position marks, though the marks would only move to fit the spiral curve after the gesture was completed. Moreover, common task such as changing the color, size, and position of marks still required a traditional control menu.

The **Tangere** system [96, 97, 98] was developed for interacting with different types of charts, including line charts, bar charts, parallel coordinates, and scatterplots (see Figure 3.3). The two primary goals were: (1) making touch gestures to invoke operations as simple as possible and (2) keeping them consistent across the different types of charts, the latter goal distinguishing it from the systems reviewed above.

In Tangere, a lasso around a set of marks selects them and swiping on an axis selects items in the spanned region, while tapping & holding, and dragging on an axis initiates a sort. Unlike PCs, modifier keys (e.g., shift, control) are not readily available on tablets to expand the set of operations. To expand the types of selection that could be performed, Tangere used bimanual interaction instead. A person typically holds a tablet with their non-dominant hand and performs touch gestures with the other dominant hand. In Tangere, the thumb of the non-dominant hand can touch the edge of the display while holding it (see Figure 3.3 right). This touch, called a “clutch,” acts as a type of modifier to change the functionality of the touch gesture being performed with the other hand. For example, normally a touch gesture on a

---

1. A video of Kinetica (© 2014 ACM) is available at https://youtu.be/7OYcGiKrmEg.
Figure 3.3: **Tangere** [96, 97, 98] featured actions compatible with multiple chart types, including linked highlighting or brushing across multiple views. In a later version of the system, one could “clutch” with the non-dominant hand to modulate the action performed by the dominant hand, akin to holding down a shift or control key with a physical keyboard (right). *Video stills courtesy of Ramik Sadana, used with permission; watch the full videos at [https://vimeo.com/195348951](https://vimeo.com/195348951) and [https://vimeo.com/195349037](https://vimeo.com/195349037).*

Interacting with Visualization on Mobile Devices

data item replaces the prior selection by this new item. With the clutch engaged, the new item can instead be added to the selection. Tangere employs the clutch operation with touch, drag, and pinch gestures to provide a broad set of different operations.

With the emergence of pen-enabled devices, researchers have started investigating the use of pen and touch in the context of data visualization. Although some smartphones (e.g., the Samsung Galaxy Note series) are equipped with a digital pen, existing research has thus far been conducted with tablets. We note that Frisch et al. [38] examined pen and touch interaction on stationary tabletop devices; while some tabletop interactions may be applicable to tablets and smaller devices, the focus of this overview remains on mobile devices. We refer to the union of both pen and touch interactions and not necessarily simultaneous pen and touch interaction.

In their **TouchPivot** system, Jo et al. [56] designed pen and touch interactions to support data exploration on tablet devices for novices (see Figure 3.4). Unlike other systems, TouchPivot deliberately incorporates WIMP interface components to leverage their familiarity and accessibility to novices. In addition, to facilitate understanding of data transformations such as pivoting and filtering, TouchPivot displays a data table and a chart together, keeping them in sync. To devise a gesture set that novices may easily understand and use, Jo et al. started from a survey of pen and touch gestures used in 13 previous studies to support data exploration. To keep their gesture set as small and simple as possible, they decided to use three touch gestures (tap, tap & hold, and drag) and four pen gestures (tap, simple stroke, lasso, and write).

Extending the concept of clutching, TouchPivot employs a fan menu to enable rapid exploration; dragging along the arc of the fan at the bottom left corner provides access to data columns, enabling people to pivot the data by the focused column and preview the distribution of values in the preview area in the bottom right part. Lifting
Figure 3.4: **TouchPivot** [56] featured a fan menu for the thumb of the non-dominant hand, and pen gestures for both tables and charts. *Video stills courtesy of Jinwook Seo, used with permission; watch the full video at https://youtu.be/Q6quofDi07I.*

Figure 3.5: In **SmartCues** [111], touch gestures annotate charts with text labels, color highlights, reference lines, and shaded reference bands. Video stills courtesy of H. Subramonyam; watch the full video at https://youtu.be/xQF0mFn5Q.

...a thumb from the fan menu confirms the pivot operation, updating the data table and moving the chart to the main chart view in the top right part.

Touch interaction with the table on the left side mimics the mouse interaction currently used on a PC. For example, a drag gesture pans the table view. In addition to writing with the pen, TouchPivot also employs a few pen interactions for manipulating the table. For example, drawing a vertical line stroke on the table sorts the data table and the chart based on the corresponding data column, while drawing a lasso on the table highlights the corresponding records in the scatterplot.

Subramonyam & Adar’s **SmartCues** [111] tablet application featured a touch interaction vocabulary for selecting and annotating charts with text labels, color highlights, reference lines, and shaded reference bands (see Figure 3.5). This vocabulary involves one- and two-finger gestures on axes, marks, and legends in bar charts, line charts, scatterplots, and tilemaps. Earlier in this section we encountered reference lines and text labels in response to a touch gesture in TouchWave; SmartCues takes this further, toward a more complete chart annotation system akin to desktop-based tools like Click2Annotate [26] or ChartAccent [92].

To summarize, we began this section with TouchWave and TouchViz, which considered touch interactions for a single type of chart. Kinetica introduced us to
physics-based interactions with particle-like marks that could be reconfigured into a wide variety of layouts. With Tangere and TouchPivot, we saw both an evolution of how to use the non-dominant hand via clutch interactions and fan menus, as well how to interact consistently across multiple chart types and tables. TouchPivot added pen gestures to our growing tablet interaction vocabulary. Finally, SmartCues focused our attention on annotation and the ability to easily identify and compare values via a small set of gestures. However, all of these projects considered a tablet form factor, leading to the question of whether the interactions that we have discussed will be compatible when we reduce the display dimensions to that of mobile phones.

**Touch Interaction on Phones**

Though much of the recent research pertaining to touch interaction with visualization on mobile devices has been carried out using tablets, we now review some of the research incorporating visualization and touch interaction on mobile phones.

Figure 3.6: (a) **GraphTiles** [8] & (b) **Orchard** [36] are mobile applications for navigating and exploring network data, such as the relationships between films, actors, and directors. GraphTiles image courtesy of B. Watson (from [9]). Orchard video stills courtesy of B. Lee; watch the full video at https://youtu.be/moCXzuotYyw. Hand gesture icons by GestureWorks ® [40] (© 2018).

Beyond considering a mobile phone form factor, **GraphTiles** [8] also stands out in that it considers graph data, whereas our previous examples visualized tabular data in bar charts, line charts, and scatterplots. With GraphTiles, one could swipe to navigate a tile-based node representation featuring superimposed link lines, and tapping & holding would select a node, thereby permitting a faceted search based on the selected node’s attributes (see Figure 3.6a). More recently, Eichmann et al. [36] envisioned another touch-based interface for navigating multivariate networks. With their **Orchard** application, one can scroll vertical lists of nodes and horizontally swipe to pivot a graph by link category, thereby building up a graph query trail (see
Figure 3.6b). Both applications avoid a conventional node-link representation in favor of designs that more easily support touch interaction on a small display.

Many of the projects discussed above involve interactions for selecting one or more individual marks. An alternative to selecting marks is selecting regions wherein marks appear. In Brehmer et al.’s visualization [20] of ranges over time on mobile devices, some configurations of the application involved the simultaneous display of dozens or hundreds of individual marks, and a single tap interaction would trigger a rectangle, wedge, or concentric band selection spanning or intersecting multiple marks, which could be repositioned via dragging (see Figure 3.7). The three geometric manifestations of the selection region remained a fixed size irrespective of the granularity of the data and the number of visible marks, and the size of this region was adequately large enough for single-digit touch selection.

Recent mobile devices are now capable of detecting touch pressure and can expose degrees of pressure to interaction designers. Apple refers to such interaction as 3D Touch, while Huawei refers to it as Force-Touch. Interaction design researchers such as Pelurson & Nigay [87] have begun exploring the potential of variable-pressure touch on mobile displays, leading to new interactions for navigation [29] and text selection [4, 25, 42]. The implications for interacting with visualization via mobile devices have not been fully examined, and we are unaware of visualization research or existing applications that incorporate tapping and pressing to varying degrees of pressure as an isolated gesture. Notably, Wang et al. [123] explored the use of pressure-based touch interaction for 3D visualization on mobile phones (see Figure 3.8), however the tap and press interaction preceded a drag gesture, which was mapped to continuous 3D navigation, where a light press/drag corresponded with X-Y rotation and a hard press/drag corresponded with X-Y translation. We revisit the topic of pressure sensing and opportunities with novel device capabilities below in Section 3.3.5.

Panning and zooming are fundamental operations when exploring time-oriented data [1]. This led Schwab et al. [101] to comparatively evaluate alternative gestures for panning and zooming under varying degrees of navigation difficulty (see Figure 3.9).
Consider, for example, how tapping twice in short succession has various repercussions across applications. In some cases, a double tap is unintended and is treated like a single tap. In others, it is ignored altogether. Interestingly, Schwab et al. found that while a continuous pinch to zoom is best in most cases, brushing along an axis and dragging orthogonally to an axes are effective in some circumstances, particularly when the difficulty of the navigation is high, where the index of difficulty of a navigation event can be computed according to the distance between the origin position and the target position and the specificity of the target.

Despite the many pan and zoom alternatives examined by Schwab et al. [101], the vocabulary of touch actions used by researchers for interacting with visualization on mobile phones is smaller than that of tablets. With mobile phones, we have yet to encounter instances of two-handed gestures for interacting with visualization such as Sadana & Stasko’s clutch action [98]; nor have we seen pen-based input akin to that of TouchPivot [56]. In our discussion of opportunities below in Section 3.3, we revisit the topic of enlarging an interaction vocabulary and we consider new possibilities for mobile phone interaction based on recent mobile HCI research, as well as possible interactions afforded by new and forthcoming mobile phones.

3.2.2 Spatial Interaction

In contrast to touch interaction, which involves interacting on the device, spatial interaction is about performing the interaction with the device by manipulating its position and orientation via movement. Spatial interaction with mobile devices also evokes the related concept of tangible interaction, which can be broadly interpreted...
Figure 3.9: Schwab et al. [101] compared alternative gestures for panning and zooming representations of time-oriented data on mobile phones. Video stills courtesy of Micha Schwab, used with permission (we added the icon annotations); watch the full video at https://multiscale-timelines.ccs.neu.edu.

to involve physicality and embodiment, according to Boy [19] and Maher & Lee [73]. The scope of tangible interaction encompasses far more than we are prepared to discuss here. For example, it even includes interacting with instrumented objects and environments, such as how Chan et al. [24] and Ebert et al. [35] detected the stacking, sliding, and dialing of acrylic discs and cubes across the surface of capacitive touchscreen tabletop displays. In both cases, the discs and cubes are directly linked to digital artifacts displayed on the table, and their movements map to functions in the tabletop application. In the context of visualization, discussion of tangible interaction brings data physicalization [55] to mind, however many physical renderings of data are not instrumented with sensors and thus are not capable of responding to changes in position or orientation.

As it relates to our current discussion of visualization on mobile devices, Spindler et al. [106] note that unlike the discs and cubes used by Chan et al. [24] and Ebert et al. [35], a mobile device is simultaneously a display for visual representations of data and the tangible object of interaction. This section will outline the interactions that are possible when mobile devices are used in this way, detecting changes in position and orientation via motion sensors, cameras, and (geo)location tracking.

Spatial Interaction using Motion Sensors

A basic requirement for spatial interaction is the ability to track a device’s position and orientation in space relative to its environment. This can be accomplished using sensors within the device itself or via external sensors, such as infrared trackers
installed in a room. In theory, 6 degrees of freedom (6-DOF) are used, where three spatial coordinates define the device location and three angles define its orientation. However, the practical use of these degrees of freedom is limited. On the one hand, the available sensor technology might limit the precision with which spatial position and orientation are measured. On the other hand, the human motor system naturally limits possible movements and the precision with which they are performed. In the following, we assume that tracking delivers reasonably good results and that spatial interaction design takes human factors into account.

Given the aforementioned assumptions, Spindler et al. [107] showed that mobile devices can be used for spatial interaction in various ways. Figure 3.10 shows three examples, where so-called **Tangible Views** are used to explore parallel coordinates, node-link diagrams, and space-time cubes. The two basic operations are movement and rotation of the device, which can in turn be combined to form gestures.

![Figure 3.10: Spindler et al.'s [107] Tangible Views applied to different visualizations.](Images from [117] licensed under cc by-nc-sa)

**Basic movement & rotation.** Before using device movement for interaction, it is necessary to define a reference space for movement. This involves determining whether movements are measured relative to the current device position or relative to absolute coordinates in a fixed reference space. Relative measurements are typically applied when a mobile device is used in large open environments, while absolute measurements are feasible in smaller spaces, such as in the GraSp project [61], where one moves a mobile device in front of a larger display (see Figure 3.11).

In principle, moving a device changes its position in 3D space, which would enable users to control three visualization parameters. However, adjusting a 3D position precisely is difficult given the properties of the human motor system. Therefore, it makes sense to constrain movement interactions, and the resolution with which positions are tracked can be reduced. Spindler et al. [105] observed that people can reach up to 44 different vertical positions reasonably well. Moreover, they can consider movements with respect to 2D reference planes, usually the horizontal and vertical planes in front of them. An example is the VisTiles project [67] (see Figure 3.12),
Figure 3.11: In the **GraSp** project [61], spatial movements with the tablet modulate what is visualized both on the tablet and on the larger display. Video stills courtesy of the Interactive Media Lab Dresden; watch the full video at [https://youtu.be/1LeBSZBLOQk](https://youtu.be/1LeBSZBLOQk).

In which several mobile devices placed on a table form an ensemble of visualization views that are dependent upon their distance and orientation relative to one another.

In general, the rotation of a mobile device is considered to take place around the device’s center of gravity. Again, while three independent dimensions are theoretically possible to control the visualization, it is common to apply constraints to make rotations practically feasible. For example, the rotation could be constrained to one or two dimensions. Moreover, the angle of rotation might be limited based on the situation; for example, rotating a handheld device around the axis of the forearm is limited to less than 180 degrees. 90 degree device rotation often toggles between portrait and landscape viewing modes in mobile applications. One example appeared in a previous version of the Apple’s iOS Stocks app; though this feature no longer exists in the current version of the app (at the time of writing), rotating from portrait to landscape increased the size of the line chart for the currently selected stock and hid the list of other stocks.

**Spatial gestures.** A new perspective on spatial interaction opens up when considering device movements and rotations as paths through space and time: they can be used within the limits of the tracking system and the human motor system to define gestures that correspond with adjustments to visual representations shown on a mobile device. That is, interaction is not based on a single vector with up to six dimensions (3D position and 3D rotation), but on a timed sequence of vectors, where the sequence contains the current position and rotation plus previous positions and rotations and their corresponding timestamps. While it is the users who perform the movements
Figure 3.12: VisTiles [67] are an ensemble of connected mobile devices that maintain a shared awareness of their relative positions and orientations, as changes to either modulate both what is shown as well as the interaction affordances on each device. Video stills courtesy Interactive Media Lab Dresden, used with permission; watch the full video at https://youtu.be/8MxPAMKmkSM.

and rotations, it is the task of the visualization designer to define a set of reference paths, so-called gestures, to be matched with the user input.

The space of possible gestures is immense, as any of the six degrees of freedom can be performed and combined with any timing. This has implications for the practical use of gestures, especially, on the discoverability of gesture-based interaction in visualization interfaces. While there are techniques for assisting users in drawing stroke gestures with touch or pen [11], no such techniques exist for spatial gestures. Ideally, spatial gestures are simple to perform, easy to remember, and do not induce any substantial fatigue. Detecting gestures is a non-trivial problem, which reaffirms the need to keep gestures simple. As Spindler et al. [107] indicate, there are two relatively simple spatial gestures for handheld mobile devices for interacting with visualization, namely tilting and shaking.

The tilt gesture corresponds to a brief rotation of the device around the forearm axis and a subsequent return to the default device orientation, as demonstrated by Dachselt & Buchholz [31]. It can have a positive or negative sign, depending on the direction of the rotation. Tilting can be used to navigate a visualization in a step-wise manner, where each tilt corresponds to a single step. This can be useful, for example, for switching between different pages of a visual representation. Additionally, the step size can bear meaning, whereby tilting with a larger angle could be mapped to bigger steps in the interaction. The shake gesture literally requires the device to be shaken. This corresponds to a change in the device position at high frequency, where the
direction of change switches frequently from a positive to a negative sign. The shake gesture is applicable for interactions that convey a “No, I don't want this” intent. For example, a shake could be used to dismiss the current visualization layout and request that the system generate a new one. Another example could be to reset a filter that has previously been applied to reduce the number of data items on the display. Again, the duration and energy of the shake could be used as an additional channel of control. The shake and tilt gestures are only two examples of what is possible with spatial interaction with the device displaying visualization. Depending on the type of device and data being visualized, different gestures can be used; Chapter 4 will touch upon aspects of spatial interaction in the context of 3D mobile data visualization.

The advantage of interacting spatially with the same device on which visualization content is displayed is a high degree of directness. However, device movements and rotations influence how well the user can see the visualized data. For example, by tilting the device, we lose the ideal perpendicular view on the device. When shaking a device, it is naturally hard to see any details in the visualization. This can be critical in cases where the interaction results in only subtle changes of the display, which might go unnoticed. It is therefore important to ensure that any feedback to spatial interaction addresses these issues.

Spatial Interaction using Cameras

Many contemporary phones and tablets feature high-resolution front-facing and rear-facing cameras; both can be used as additional inputs for spatial interaction.

One use of a mobile camera is position-based transfer of visualization content across devices. The **Visfer** project [7] envisions a set of networked displays, which might include mobile phones, tablets, PCs, and large wall-sized displays (see Figure 3.13). The large displays could be divided into multiple views, and each view is augmented with an animated QR code. Using a mobile phone or tablet, one can move the device to point the camera at a QR code and to transfer content from the large display to the smaller one. Several types of content transfer are possible, including a responsive adaptation of the large display view to the aspect ratio of the smaller device, a transfer of the view specification, or a transfer of summary-level data, thereby allowing the smaller device to display a different yet related view to what is shown on the large display.

Camera-based spatial interaction can also be used to navigate and manipulate virtual objects in mobile augmented reality. One example is **AffinityLens** [112], which is an application for affinity diagramming with physical sticky notes augmented with QR-like codes (see Figure 3.14). Detecting these tags and their related note content via the camera, the application can highlight note categories and text search results. It can also generate chart overlays that summarize note content, including word clouds, line charts, and bar charts. Another example is **MARVisT** (or Mobile Augmented Reality Visualization Tool) [27], which allows people to create and view unit charts of 3D glyphs distributed within a volume (see Figure 3.15). We return to the chart authoring aspect of MARVisT below in our discussion of future opportunities.

Finally, mobile cameras can be used to detect hand gestures, which could in
Interacting with Visualization on Mobile Devices

Figure 3.13: With **Visfer** [7], one can transfer visualization content from a large display to a tablet or phone via the use of QR codes. Video stills courtesy of Karthik Badam, used with permission; watch the full video at [https://youtu.be/KGIYqulePGA](https://youtu.be/KGIYqulePGA).

Figure 3.14: Affinity diagramming is augmented with **AffinityLens** [112], in which alternative visual summaries of sticky note content is shown on the tablet or phone via the use of QR-like codes detected by the device’s camera. Video stills courtesy of Hariharan Subramonyam, used with permission; watch the full video at [https://youtu.be/p9WNtB0rQEo](https://youtu.be/p9WNtB0rQEo).

turn modify the visualization, such as by triggering navigation or selection. In our discussion of future opportunities below in Section 3.3.5, we refer to two projects (ARPen [122] and Portalible [90]) that involve the capture of gestures performed with the hand that is not holding the device. Alternatively, these gestures could be performed by some other person who is visible to the camera.

**Spatial Interaction using (Geo)location**

Spatial interaction may also involve changing the position of a mobile device at a much larger scale, thereby necessitating the use of a geolocation sensing. Mapping applications such as Google Maps are canonical examples in which visual representations of alternative trajectories, arrival estimates, accidents, and traffic congestion are updated in response to your change in location. Similarly, fitness applications such as Runkeeper [6], Fitbit [37], or Strava [110] encode trajectory paths over the course of a run or cycle, overlaying these on a map. Another notable instance of location-based interaction with visualization on a mobile device is **Subspotting** [12], in which a representation of cellular connectivity along the New York City subway system is updated in response to one’s location along a subway line (see Figure 3.16).
Figure 3.15: **MARVisT** [27] is a mobile augmented reality visualization authoring tool, in which it is possible to encode the size of virtual marks based on real objects detected in the scene and subsequently place marks via spatial gestures. Video stills courtesy of Z. Chen; watch the full video at [https://youtu.be/cbtbJXwpwdk](https://youtu.be/cbtbJXwpwdk).

Figure 3.16: In **Subspotting** [12], one’s change in location along New York City’s subway lines updates an egocentric visual representation of cellular connectivity. Video stills © 2016 OFFC NY, used with permission from Dominikus Baur; watch the full video at [https://vimeo.com/153013236](https://vimeo.com/153013236).

Online news articles, mobile versions of websites, and mobile apps can also request geolocation information about the viewer as a means to provide personalized content (for example, consider the OECD Better Living Index [85]). Geolocation sensing could also be used to support *context-dependent interaction* or responsive interaction design for visualization on a mobile device, such as by disabling or simplifying interaction while the location of the device is changing rapidly, which is typically indicative of moving in a vehicle, as described earlier in Chapter 2.

Our final example of spatial interaction is **FieldView** [124], a research project that involves visualizing location-specific data on mobile device displays as well as in head-mounted augmented reality; currently, our focus is on the former, as we return to the latter in our discussion of future opportunities below. In the mobile instantiation of FieldView (see Figure 3.17), multiple visual representations of data aggregated over a spatial grid allow forest ecologists, wildfire fighters, search & rescue teams, and others working in similar roles to ensure coverage of a territory and to add or edit location-specific data, thereby updating any corresponding visualization. In a sense, this form of interaction is reminiscent of strategy video games whereby the
“fog of war” is lifted via change in location, whereupon more of the territory and its attributes become visible.

Figure 3.17: FieldView [124] combines data entry and mapping using a mobile phone with head-mounted augmented reality visualization. Video stills courtesy of the CU VisuaLab, used with permission; watch the full video at https://youtu.be/pHfdId4Gis.

3.2.3 Voice Interaction

Voice interaction has intrigued designers for decades, such as Bolt’s “Put-that-there” interface (1980) [18] or Bartlett et al.’s Itsy pocket computer (2000) [10]. It has the potential to address challenges in interaction design for mobile visualization use cases, where access to mouse and keyboard is missing and the display size is small [70]. Researchers have recently envisioned compelling scenarios that facilitate and augment data exploration on mobile devices by leveraging speech input. Srinivasan et al. [109] describe a novel tabular data manipulation scenario on tablet devices, discussing ways to complement direct manipulation via touch or pen with minimalistic speech input. Choe et al. [28] envisioned a novel way to help people explore their personal data on smartphones by incorporating speech interaction and Kim et al. recently realized this approach with their Data@Hand [60] application. Exploring self-tracking data often involves specifying date and time, or their ranges. While this is tedious to do on mobile devices with existing widgets, such as calendar and clock controls, people are already comfortable and familiar with specifying dates and times with speech.

Recently, Srinivasan et al. [108] developed InChorus, a multimodal interface that incorporates pen, touch, and speech to facilitate data exploration on tablet devices (see Figure 3.18). InChorus was designed to address two fundamental issues. First, most of the prior research of data visualization on tablet devices have been optimized for a specific visualization type, such as stacked graph, bar chart, and scatterplot. This could cause conflicts and inconsistencies when we need to design a system that supports multiple types of charts. Second, when constrained by only pen and or touch, systems face increased reliance either on menus and widgets or on complex gestures as the number and complexity of operations grow.

To design multimodal interactions that function consistently across multiple visualizations, InChorus brings speech interaction into pen and touch interaction: the directness and precision of pen and touch is complemented by the freedom of
Figure 3.18: **InChorus** [108] is a multimodal tablet-based visualization tool that incorporates pen, touch, and speech interaction. Video stills courtesy of Bongshin Lee, used with permission; watch the full video at https://youtu.be/cy0VSmUP_98.

each of the three input modalities can work individually for the operations that fit their inherent characteristics. However, what makes InChorus unique is that the three modalities can work together to provide a novel and more fluid interaction experience. For example, combining speech with touch in a meaningful way can help people perform a more powerful action with a simpler interaction because touch can provide a deictic reference to a speech command. Finally, for many operations, InChorus provides multiple ways to complete the operation using different input modalities. This flexibility helps to accommodate individuals’ personal preference. We revisit the topic of multimodal interaction in our discussion of future opportunities in the next section.

At this point, we end our overview of interaction for mobile data visualization. We have seen numerous examples of touch-based interaction on tablets and mobile phones, approaches to spatial interaction at local and global scale, and initial results of voice interaction. While the reviewed solutions already illustrate a wide range of possibilities, there is still more to investigate in the future.

### 3.3 FUTURE OPPORTUNITIES

We see several promising directions for future research and design with respect to interacting with visualization using mobile devices. We now reflect on the interaction vocabulary for mobile visualization and speculate on the future of multimodal interaction, multi-display interaction, mobile visualization authoring, and visualizing data in mobile augmented reality. While the latter two categories of opportunities are relatively nascent topics of discussion within the visualization community, we are not the first to discuss the former categories of opportunities; see Langner et al. (2015) [66]. Beyond visualization, we also look outward toward the broader human-computer interaction research community: to recently proposed mobile interactions and mobile technologies appearing at venues such as the CHI, UIST, and MobileHCI conferences.

For each new development with respect to interacting with visualization on mobile devices, one challenge will be evaluation, and as both Games & Joshi [39] and Blumenstein et al. [17] have observed, evaluation methodologies may need to
be adapted to consider mobile devices and contexts. The crowdsourced evaluations of zooming and panning on mobile devices by Schwab et al. [101] and of mobile-specific visual encodings for range data by Brehmer et al. [20] may serve as useful precedents in this regard, however some forms of interaction may require more controlled experimental environments and direct researcher supervision. The topic of evaluating visualization on mobile devices is addressed in greater detail in Chapter 6.

3.3.1 Consistency & Expressivity

Given the diversity of our overview, it appears as though we are in the early days with respect to interaction design for visualization via mobile devices, in that various interactions map to different intents and have different effects across applications. In other words, there is no standard set of consistent interactions with visualization on tablets and mobile phones. Part of this heterogeneity can be attributed to the variety of data types and visual encodings in use, which have their own set of affordances independent of display device. Further complicating matters is the ongoing technological evolution of the devices themselves. Our overview focused heavily on touchscreen devices introduced since the advent of the iPhone in 2007, and since that time, various multi-touch, pressure-based touch, and spatial interaction techniques have appeared. Some touch interactions have attained a more consistent meaning than others in this time, such as pinching or spreading two fingers to zoom content or pulling down to refresh content. However, as Schwab et al. [101] show in their comparison of zooming gestures, there are several other gestures associated with zooming. Likewise, one application’s panning gesture could be another application’s selection or brush highlighting. In applications with multiple possible interactions, designers may opt for multiple interaction modes, such as a selection mode and a navigation mode, or they might resort to a combination of gestures and a conventional interface of menus and buttons, akin to the TouchViz WIMP interface [34]. In these cases, the discoverability of individual interactions or interaction modes is an important concern for designers.

While acknowledging the challenges associated with the discoverability and consistency of interactions, it is also exciting to expand the vocabulary for interacting with visualization via mobile devices. We can look to the mobile human-computer interaction literature and to particular application domains for inspiration. For instance, consider a dialing touch gesture, one that evokes rotary phones or the classic iPod’s scroll wheel; Moscovich & Hughes [80] and Smith & Schraefel [104] showed that this gesture can be used to navigate text documents, and thus could be applied to navigating any continuous data dimension. Similarly, drawing a convex hull via multiple touch points could be useful to select content in scatterplots or node-link diagrams, particularly on larger tablet devices. Finally, mobile map and wayfinding applications have provided a rich set of interactions that could be applied to other forms of data, such as two-finger scrolling to tilt the viewing plane or two-finger rotation to toggle egocentric perspectives on content.
3.3.2 Multimodal Interaction

Many of the examples cited in our overview above feature multiple simultaneous modalities of interaction, though by examining each modality individually, this might not have been apparent. As Table 3.1 shows, instances of spatial interaction and voice interaction also tend to involve touch interaction. For instance, the InChorus system [108] incorporates touch input, pen input, and voice input. The additive nature of these modalities leads to the question of how designers should assign interactions to modalities, and whether interactions should be exclusive to one modality or should they be redundantly accessible via multiple modalities [41]. Once again, as this aspect of mobile visualization design matures, we may see more variety in the allocation of interactions across modalities and in turn a need for consistency, particularly as applications incorporating multimodal interaction move from the domain of research to commercial or publicly-available applications.

3.3.3 Multi-device Interaction

Designing for multi-device, multi-person environments leads to questions of how an interaction might differ across several heterogeneous devices, or how the repercussions of one person’s interaction with one device might manifest on other devices in the environment. If multiple views of a single dataset are distributed across displays, we arrive at the possibility of collaborative brushing and linking and other interactions that support mutual awareness and signalling. Linked navigation and selection across displays is another promising aspect of multi-device interaction with visualization. For instance, Voida et al. [120] envisioned a multi-device system involving an iPod touch and a tabletop display, in which the former is used both to view focused content in greater detail and to interact with the content using higher-fidelity multi-touch gestures afforded by the iPod touch. More recently, Berge et al. [15] demonstrated how a large wall-mounted screen could display an overview while a coordinated mobile phone could display a detail view of a subset of the overview, while Kister et al. [61] and Langner & Dachselt [65] showed how a detail view could be determined based upon the distance and relative orientation of a phone or tablet to the display. Langner et al. [68] would go on to demonstrate a multi-device system that supports both direct touch manipulation as well as remote interaction via a mobile phone for triggering details on demand views as well as highlight lenses and rulers for a wall-based visualization dashboard. Besançon et al. [16] demonstrated the use of spatial and touch interaction with a mobile phone to act as a remote for a paired large display, involving the visualization of 3D volume data, which called for 6-DOF navigation and the orientation of a 2D cutting plane across the volume. This intersection of 3D data visualization and mobile devices is discussed in greater detail in Chapter 4. Finally, Vistribute [49] takes us beyond tablets and mobile phones, being a framework for allocating interactive visualization elements across various types of displays, from mobile phones and tablets to PCs and wall displays.

Multi-person, multi-device environments are still an emerging area of visualization and human-computer interaction research. A study by Plank et al. [89] revealed that people are not accustomed to collaborating with visualization content distributed over
a set of coordinated tablet devices. To overcome a legacy bias of working with single
devices in isolation, visualization researchers and designers should identify a set of
discernible interactions for coordinating displays with corresponding attentional cues
for promoting collaboration. Frameworks such as VisTiles [67] and Vistribute [49] are
promising in that they may provide infrastructure for this research, which may in
turn reveal new compelling use cases for multi-device interaction.

3.3.4 Visualization Authoring

Mobile devices are capable of sensing and storing various types of data, including
motion, usage, location, sound, and images. Automatically-recorded data can also
be augmented or complemented with manually-recorded data. In addition to many
commercial self-tracking applications, some amateur “quantified-self” enthusiasts
develop mobile data collection processes and applications. Mobile applications such
as OmniTrack [59] allow people to track the data of their choosing, with options for
specifying the type and granularity of the data. However, while OmniTrack and other
self-tracking mobile applications allow people to record various forms of data, there
are few options with respect to configuring or authoring visualizations of the data; if
provided at all, visualization in these applications allow for little customization or
control over visual encoding design choices. While there may be some convenience in
being able to make visualization design decisions from the device that captured the
data, self-tracking enthusiasts seeking to perform in-depth data analysis are likely to
export their data from the device and visualize it using a PC.

There is evidently an opportunity to design effective mobile visualization authoring
interactions. Tableau’s Vizable iPad app [113] took an initial step toward mobile
visualization authoring, though it required connections to external data sources rather
than to data captured by the device itself. Mendez et al. [74] suggested a scenario
for mobile visualization authoring in which people would take photographs of charts
encountered in the physical world using a phone or tablet app. The app would infer its
data relations and allow for the manipulation of these relations as new or augmented
visual encodings.

Another opportunity is to bootstrap mobile visualization authoring via the capture
of autographic visualization [86], or visible material traces of real-world phenomena
such as air pollution or sea level change. One could imagine an interface wherein it
would be possible to define position, size, or color scales and encode image or video
content relative to these scales.

Lastly, we arrive at the intersection of mobile visualization authoring and mo-
bile augmented reality, instantiated in the MARVisT system [27] described above.
MARVisT allows people to specify visual encodings of unit charts via a mobile touch
interface, while individual units can be placed within a 3D volume via touch or spatial
interaction; the resulting unit charts can be navigated via spatial interaction.

3.3.5 Inspiration from Mobile HCI

By focusing primarily on visualization-related research projects in our overview, we
have only scratched the surface of mobile interaction design research and the work of
mobile interaction design practitioners. While our focus has until this point been on the intersection of mobile interaction design and data visualization, we now point to several recent developments in mobile HCI that can potentially be applied to future visualization research and design.

**Interaction with visualization on smartwatches.** Although we excluded smartwatch and other wearable devices from our overview, there are nevertheless opportunities here worth noting. Building off of earlier work by von Zadow et al. [121] that envisioned a wearable sleeve interface for interacting with a large wall display, Horak et al. [48] designed a system incorporating multiple smartwatches and a large wall display, wherein watches could store and display a subset of data based on touch and proxemic interaction, or they could act as a filter and remote control for the large display. To accomplish this, the system required a vocabulary of touch gestures for both display types, which are modulated by proximity to the display.

Smartwatch spatial gestures are steadily becoming more familiar, such as moving one’s wrist to face upwards, which reveals the watch’s home screen. Seyed et al. [103] have also considered new gestures for smartwatches, such as flip, slide, or detach. We also see opportunities to apply voice interaction to smartwatch visualization, as well as opportunities for pen or stylus interaction with smaller watch displays, perhaps via a finger-mounted stylus, such as the one demonstrated in the NanoStylus project [127]. Finally, there may be new interaction modalities to consider with smartwatches. For example, the MyoTilt project [64] uses electromyography allowing a smartwatch wearer to manipulate display content via a combination of arm tilt and forearm muscle engagement.

While smartwatch visualization interaction is still relatively uncommon, a review of this research area along the lines of our present overview may be worth undertaking in the near future.

**Gaze interaction.** Researchers have been examining the potential of eye gaze interaction on mobile devices for over a decade [33]. Initially, such interaction required specialized eye-tracking equipment, while contemporary front-facing cameras on mobile phones and tablets are now capable of eye-gaze tracking without any hardware modification, as demonstrated by Khamis et al. [58]. To our knowledge, there has yet to be a demonstration of eye gaze interaction with visualization for mobile devices.

**Pose and grip interaction.** Pfeuffer et al. [88] explored thumb + pen interaction on tablet devices: similar to the clutch and fan menu described above, the dominant hand’s pen actions are supported and augmented by thumb interaction with the non-dominant, device-holding hand. For example, to enable quick access to available options such as menu items, they employ thumb marking menus that can be operated by the thumb even while holding a device. To alleviate the issue caused by the thumb’s limited reach, they integrate indirect touch input with virtual handles, which were introduced in earlier work by Wolf & Henze [126]. Pfeuffer et al. demonstrated that thumb + pen techniques can be applied to manipulate and analyze data in spreadsheets on tablet devices: they focus on the common actions people perform with cells in spreadsheets, such as copy-paste, formatting, or data editing. In this project, the pen always writes (or draws), while touch always manipulates content,
following a design mantra advocated for by Hinckley et al. [46]. To our knowledge, we have yet to encounter a mobile application incorporating both visualization and bimanual pen+touch interaction.

Also promising are prototype devices that can detect changes in the positioning of one’s hands relative to the device, whether or not they are touching the device. Zhang et al. [129] demonstrated how to improve pose and grip sensing by augmenting devices with sensors placed along the bezel of the screen. Similarly, Hinckley et al. [45] demonstrated the ability to detect a finger hovering over the screen. It would be fascinating to experiment with how such techniques could be applied to visualization tools, such as Tangere [98], which remains to be a rare example of bimanual touch interaction with visualization on a tablet. For instance, grip and hover detection could be a way of eliciting tooltips for marks on a mobile display.

Mobile augmented reality. Recent developments in mobile augmented reality (AR) interaction also offer exciting prospects for visualization with mobile devices. ARPen [122] and Portal-ble [90] are instances in which the non-dominant hand is holding the mobile device as an AR lens, while the dominant hand interacts with virtual objects using a pen in the case of ARPen and using freehand gestures in the case of Portal-ble. Both cases require specialized hardware (a custom pen in the former case, a Leap motion sensor affixed to the phone in the latter case), as well as 3D marks distributed in space to interact with. Both projects demonstrate their respective techniques with abstract 3D volumes of virtual representations of real objects, though these volumes or objects could just as well be points in a 3D scatterplot, data glyphs in a 3D space-time cube, or other manifestations of volumetric data.

New mobile devices. Finally, we consider the potential of new and forthcoming mobile devices and their implications for interactive visualization design. This list includes wearable devices other than smartwatches: pendants, belts, e-textile garments, temporary tattoos featuring integrated circuits, and on-skin projection devices. Previously underutilized components of wearable devices may also be exploited, such as how Klamka et al. [63] created a smartwatch strap with a flexible e-ink display. These and other devices may have varying degrees of display resolution and sensing capabilities for detecting interactions.

Though we did not discuss head-mounted augmented reality devices in our overview, as Chapter 4 discusses augmented reality and 3D data in greater detail, it is nevertheless worth noting the capabilities of next-generation displays such as with the HoloLens 2 [75] and the Magic Leap One [72] and how mobile and wearable devices might be used in concert with them. For instance, Büschel et al. [23] demonstrated how a mobile phone instrumented with additional tracking sensors could serve as a precise panning and zooming tool for navigating 3D volumes. Work by Langner et al. [69] extended the VisTiles approach [67] by combining mobile devices with head-mounted augmented reality, thereby augmenting visualization views with additional 2D and 3D information around and above displays. Wearable input devices could also be used to manipulate content shown in head-mounted augmented reality, such as ARcord [62], in which a lapel-mounted strap serves as a way of specifying a value along a single continuous dimension by pinching and sliding up or down the strap.
Handheld devices with semi-transparent displays may also invite novel interactions for visualization; consider Lucid Touch [125], in which one interacts with the underside of the device and thereby avoids the problem of fingers occluding content. Finally, new flexible handheld display devices (e.g., [47]) may provide new input channels for interaction, such as squeezing, twisting, and stretching. Foldable mobile phones such as the Samsung Galaxy Fold [99], the Samsung Galaxy Z Flip [100], the Huawei Mate X [50], the Motorola RAZR [81] and the Microsoft Surface Duo [78] may provide opportunities for multi-view visualization applications, or they may be ideal for scrollytelling-based presentations that juxtapose text and visualization content.

### 3.4 SUMMARY

The world has been captivated by the prospect of mobile interactive technology for the better part of the last century. From science fiction films to expositions that envision the future of work and leisure [83], we have been captivated by the prospect of performing an increasingly diverse set of tasks from a mobile or wearable device. The emergence of interactive data visualization over the past three decades has both amplified and shaped this captivation, prompting us to question how we can interact with data in new contexts. In parallel, mobile device display sizes have grown and support an array of multi-point touch interactions, while both processor capabilities and display resolutions have improved to the point where it has become possible to visualize large and complex datasets from a mobile device.

There are still challenges with respect to responsive design (as summarized in Chapter 2), the legibility of text elements, and both the specificity and discoverability of interactions across modalities. However, we are nevertheless witnessing an increasing number of interactive visualization applications intended for mobile devices, as well as an increasing number of responsive and interactive visualizations embedded in websites. The breadth of application areas and data types, spanning both work and leisure, is both staggering and inspiring. Visualization and human-computer interaction researchers also continue to test the boundaries of interacting with visual representations of data on mobile devices.

In this chapter, we have put forward a summary of this research to date, one classified according to the possible interaction modalities, namely touch interaction, spatial interaction, and voice interaction. In focusing primarily on mobile phone and tablet devices, we acknowledge a need for further examination of interaction with visualization on watches and wearable devices. In the coming years, we expect innovation along these modalities to continue, along with a further blending of multimodal interaction for visualization on mobile devices. New modalities may yet emerge. We are also excited by the prospect of seeding mobile visualization research with recent advances in mobile human-computer interaction research, in which we will collectively ask how the addition of data and visual mappings modulate these interaction techniques. In particular, we are excited by the prospect of visualizing data in mobile and head-mounted augmented reality; much of this data is inherently spatial, situated and often three-dimensional, which is the topic of the next chapter.
References


[99] Sedig, K. and Parsons, P. “Interaction Design for Complex Cognitive Activities with Visual Representations: A Pattern-Based Approach”. In: *Transactions on Human-Computer Interaction (THCI)* 5.2 (June 2013), pp. 84–133. URL: http://aisel.aisnet.org/thci/vol15/iss2/1 (cited on page 69).


