When Smart Devices Interact With Pervasive Screens: A Survey

PAI CHET NG, JAMES SHE, and KANG EUN JEON, HKUST-NIE Social Media Lab, Hong Kong University of Science and Technology
MATTHIAS BALDAUF, Institute for Information and Process Management, University of Applied Sciences St. Gallen, Switzerland

---

The meeting of pervasive screens and smart devices has witnessed the birth of screen-smart device interaction (SSI), a key enabler to many novel interactive use cases. Most current surveys focus on direct human-screen interaction, and to the best of our knowledge, none have studied state-of-the-art SSI. This survey identifies three core elements of SSI and delivers a timely discussion on SSI oriented around the screen, the smart device, and the interaction modality. Two evaluation metrics (i.e., interaction latency and accuracy) have been adopted and refined to match the evaluation criterion of SSI. The bottlenecks that hinder the further advancement of the current SSI in connection with this metrics are studied. Last, future research challenges and opportunities are highlighted in the hope of inspiring continuous research efforts to realize the next generation of SSI.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing design and evaluation methods**; *Smartphones*; *Gestural input*; *Mobile devices*; Displays and imagers;

Additional Key Words and Phrases: Pervasive screen, smart device, interactive technology

ACM Reference format:
Pai Chet Ng, James She, Kang Eun Jeon, and Matthias Baldauf. 2017. When Smart Devices Interact With Pervasive Screens: A survey. *ACM Trans. Multimedia Comput. Commun. Appl.* 13, 4, Article 55 (August 2017), 23 pages.
https://doi.org/10.1145/3115933

---

1 INTRODUCTION

Within the past few years, pervasive screens have become a striking omnipresent element of today's urban environments: from public screens at airports, train stations, and shopping malls, over semipublic screens in conference rooms, cinemas, and stadiums, to multiple personal screens in private settings such as living rooms, as depicted in Figure 1. Yet despite their increasing penetration, today's public screens are still mainly a passive medium acting as high-resolution billboards without or with limited interaction possibilities for interested passers-by.

This work was supported by the HKUST-NIE Social Media Lab, HKUST.
Authors' addresses: P. C. Ng, J. She, and K. E. Jeon, HKUST-NIE Social Media Lab, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong; emails: {pcng, eejames, kejeon}@ust.hk; M. Baldauf, FHS St. Gallen, University of Applied Sciences, Rosenbergstrasse 59, 9001, St. Gallen, Switzerland; email: iam@matthiasbaldauf.com.

https://doi.org/10.1145/3115933

Fig. 1. Screens are everywhere: eboard (a) and Wow Display (b) in a public space, multiple projector screens in a semipublic space (c), and multiple TV screens in a private space (d).

As suitable complementary devices for distant screens, academia has identified ubiquitous smart devices and started to explore the design space of screen-smart device interaction (SSI) to enable more attractive content and compelling interactive applications for public screens (Clinch 2013). These devices' rich and steadily growing features, including wireless high-speed connectivity, accelerometers, gyroscopes, flashlights, and cameras, provide the basic cornerstones for a broad variety of interactive SSI use cases. Examples include using a smart device as a versatile remote control (Baldauf et al. 2016) or for creating and submitting content for interactive applications on public screens (Alt et al. 2013).

Due to the vast number of SSI systems presented over the past few years, getting an overview of the state-of-the-art SSI, and thus advancing the field, has become difficult. Recent related surveys cover topics such as direct (touch or gesture-based) interaction with screens (Greimel 2011), interactive tabletops (Bellucci et al. 2014), interaction with very large screens (Ardito et al. 2015), or advertising as a specific SSI use case (She et al. 2014). However, a structured overview that surveys and categorizes the recent works on SSI is missing.

The present survey fills this gap and provides a structured timely analysis of the current body of SSI research. Its objective is to detect so far disregarded aspects of SSI and to uncover opportunities for future research into SSI. This article provides the following three key contributions:

(1) *State-of-the-art of SSI*: This work provides an extensive survey on SSI research, structured into three core elements of SSI: screen, smart device, and interaction. The latest development of SSI is investigated, and the factors that affect the performance of SSI are discussed.
(2) *Evaluation metrics*: Two generic evaluation metrics (i.e., interaction latency and accuracy) are adopted and refined to suit the SSI's evaluation criterion. We further identify the limitations of current SSI in connection to these metrics with hopes of inspiring continuous research efforts to advance the field of SSI.

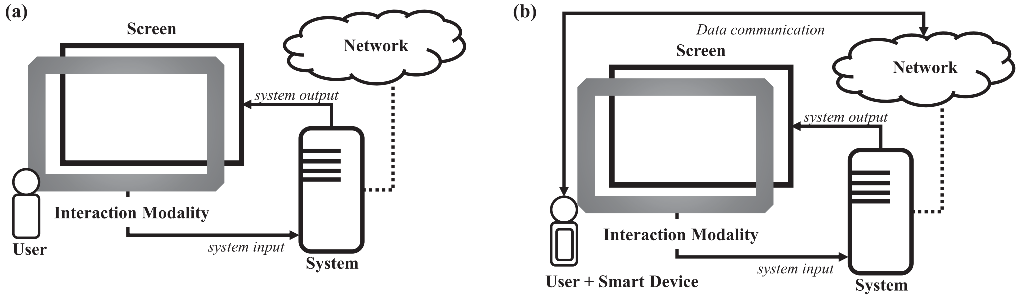

Fig. 2. Core elements of human-screen interaction (a) (She et al. 2014) and SSI (b).

(3) *Research challenges and opportunities*: The possible research challenges are presented according to the research road map span from short- to long term. Furthermore, future research opportunities for the next generation of SSI are identified.

The remainder of the article is organized as follows. In Section 2, SSI is defined according to its core elements, which provides the structure for the succeeding sections. Two evaluation metrics are also formulated as an assessing medium for the subsequent discussions in Sections 3 through 5. Section 3 gives an overview of the first core element of SSI: screen. Section 4 focuses on the role of the involved smart device, whereas interaction modalities between screens and smart devices are surveyed in Section 5. Finally, Section 6 highlights future research challenges and opportunities, and Section 7 concludes the article.

2 DEFINITION OF SCREEN-SMART DEVICE INTERACTION

In contrast to technical alternatives enabling direct human-screen interaction (Bellucci et al. 2010), such as touch-sensitive screens supporting direct touch (compare to CityWall (Peltonen et al. 2008)) or screens equipped with cameras spotting mid-air gestures (compare to WaveWindow (Perry et al. 2010)), SSI involves indirect interaction with a distant screen via a smart device. Figure 2 depicts the difference between human-screen interaction and SSI: instead of direct interaction with the screen, SSI makes use of a user's smart device to enable interaction with the screen or the system.

As its name implies, the notion of SSI comprises the following three core elements:

(1) *Screen and content*: Screen refers to any surface able to dynamically showcase multimedia content involving text, images, video, and sound, if the respective hardware is available. In a SSI scenario, the screen either comes with a built-in computing system or has been externally connected to a playback system that stores the contents and defines the scheduling policy for content playback.
(2) *Smart device*: The smart device is a personal computing device that supports remote interaction with a screen. These smart devices are not restricted to a predefined form factor as long as they satisfy the following fundamental core features, such as wireless connectivity (e.g., Bluetooth, WiFi, or cellular network) and advanced input capabilities (e.g., touch screen, camera, and built-in sensors such as accelerometers and gyroscopes) Even though the most popular smart devices today are smartphones and tablets, more recent respective devices, including wearables, are expected to embark on a new evolution in SSI.
(3) *Interaction modality*: The SSI modality refers to the style of interaction with the distant screen using the smart device. Modern smart devices with their rich features (high-quality displays, various built-in sensors, etc.) allow for a host of diverse interaction modalities.

> Fundamental for our notion of SSI is the (near) real-time interaction between the screen and smart device enabled by communicating with the playback system.

In comparison to the aforementioned human-screen interaction, SSI does not require any additional hardware but supports the reuse of already deployed off-the-shelf screens. However, users are required to use their smart device to interact with the screen. Human-screen interaction, on the other hand, offers spontaneous interaction with the screen since users do not have to carry or wear a smart device to interact with the screen. Whereas human-screen interaction limits the interaction range (i.e., users need to be in front of the screen to interact), SSI enables interaction from various distances, supporting screens situated in unreachable locations (Deller and Ebert 2011). Distant interaction further avoids hygiene issues, such as those associated with public touch screens (Dix and Sas 2010), and allows for the consumption of the entire content on a large screen, which is not possible in short distance interaction (Tomitsch et al. 2014).

Since both types of interaction systems have their own strengths, it is vital to take the design principle and target audiences into consideration before deciding which systems to use. For example, it is more appropriate to employ human-screen interaction when the target audience is elders, who do not always own a high-tech smart device. Furthermore, interaction accuracy in terms of the content being consumed is not a big issue with human-screen interaction, as users are always in front of the screen and conscious of the content with which they are interacting; however, it might be a problem with SSI, where content exchange between the screen and smart device is involved. The only possible latency with human-screen interaction is the process taken at the screen to translate users' touch commands, whereas SSI might need to deal with latency induced by the processes at the screen, smart device, and the server.

2.1 Research Methodology

As motivation and a starting point for our survey, we observe the gaps in related survey papers. As mentioned, none of the prior surveys have investigated the three core elements of SSI from a holistic approach. In view of this, we started to collect related works from various digital libraries (e.g., Google Scholar, ACM Digital Library, and IEEE Xplore Digital Library) and organized them according to the notion of SSI. To ensure that this survey does not overlap with previous related surveys, we also studied the latest surveys conducted by Greimel (2011), Bellucci et al. (2014), She et al. (2014), and Ardito et al. (2015), and filtered out the ones that have already been subject to similar discussion. Out of the total 287 papers downloaded, we have narrowed down our discussion to only the 69 most recent papers (i.e., the papers published within 10 years from the date of this survey) that address the related technologies according to the three core elements defined in SSI.

The cited works are organized in *csv* format and exported to Matlab for further analysis. Although the works cited in this survey might be insufficient to give an overall picture of current research due to page constraints, it should still capture meaningful insight regarding SSI, as depicted with the generated word cloud in Figure 3(a). Figure 3(b) shows the distribution of these 69 works according to their publication year. From this simple visualization, it is clear that being interactive is still the main concern in the field of SSI. This visualization also tells us that researchers have started to look into the interactive techniques involving mobile, smart device, multilarge displays, and so forth. This further assures us about the direction of this survey article, which is to investigate the interaction between a screen and smart device. We can see that this research field has received much attention lately, and ACM is the leading publisher that produces many high-quality works in this field. Among all of the works we have cited, the work of Boring et al. (2010) published in the ACM CHI series is the one with the highest citation count (i.e., 226 as of April 2,

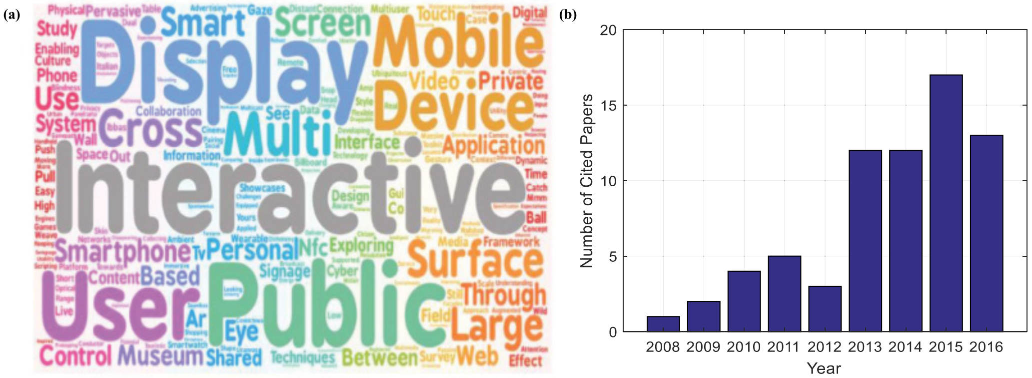

Fig. 3. (a) The word cloud visualizes the major focus of this survey corresponding to the collected paper. (b) The distribution of the collected papers with respect to publication year.

2017). The ACM CHI series has been the active venue for SSI-related research, with 21 works been cited in our survey, which is approximately one third of the total cited work.

2.2 Evaluation Metrics

This survey focuses on the interaction between the screen and smart device with a human in the loop and thus studies the effects of interaction latency on accuracy on the quality of experience (QoE). Rather than reinforcing the accuracy study in lower layers, such as the physical and network layers, this survey focuses on the latency and accuracy issue that is perceivable by humans, especially in the application layer involving multimedia exchange in diverse form factors. The following sections refine the scope of adopted evaluation metrics specific to the field of SSI.

2.2.1 Interaction Latency. Interaction latency is critical in today's society, especially in metropolitan areas where everyone is in a rush and everything needs to be done fast. Here we consider the latency contributed by a few interaction-related processes: at the screen, the smart device, and the server. Interaction latency can be induced due to the content scheduling policy on the screen or hardware specifications of the screen, such as frame-refreshing frequency and resolution.

First, considering the related processing over the screen as a black box, latency induced at screen t_s given the user triggering the interaction at time t_i can be described as follows:

$$t_s = f_s(p_s, C|t_i), \quad t_{sc} > 0, \tag{1}$$

where p_s is a set of design parameters/hardware specifications of a screen and C is the type of content scheduling policy adopted by the said SSI.

Second, a smart device always needs to handle multiple tasks (or threads) concurrently, and a certain degree of delay might be induced when a high-priority task is in execution. Note that the time taken to handle multithreading is subject to the hardware specifications and processing memory of the smart device. Another parameter that might affect the processing time is the size of the content—that is, denser data might take a longer time to process than sparse data. Similarly, by considering the related processing over the smart device as a black box, the latency contributed by the smart device t_d given t_i can be represented as follows:

$$t_d = f_d(p_d, \Lambda|t_i), \quad t_d > 0, \tag{2}$$

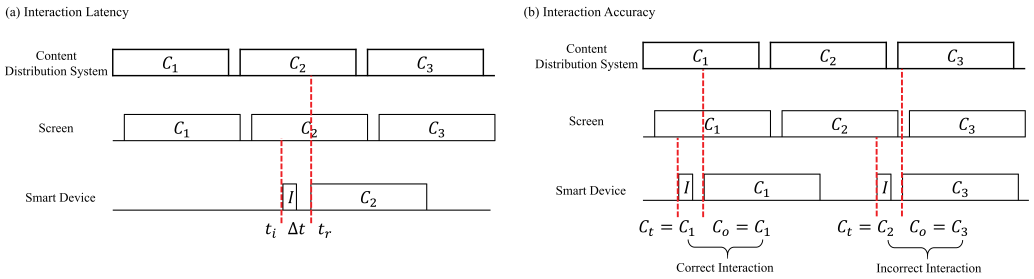

Fig. 4. The evaluation metrics: interaction latency (a) and interaction accuracy (b).

where p_d is a set of design parameters related to the hardware specifications of a smart device and Λ is the implemented algorithm that processes the collected data.

Today, most SSIs use a server to facilitate the interaction between the screen and smart device over a wireless channel; the possible parameters that might affect the data routing/exchange in the SSI are the allocated bandwidth, network capacity, traffic load, and so forth. In general, the interaction latency caused by the processes at server t_v conditioned on t_i can be described as follows:

$$t_v = f_v(p_v, \chi | t_i), \quad t_I > 0, \tag{3}$$

where p_v is a set of design parameters related to the specifications of a server and χ is the routing algorithm employed by the server.

Figure 4(a) illustrates the overall interaction latency Δt. By considering Equations (1) through (3), the overall interaction latency Δt can be estimated as follows:

$$\Delta t = t_r - t_i = t_s + t_d + t_I, \tag{4}$$

where t_r is the content receiving time and t_i is the interaction triggering time.

Over the years, several works have been proposed to reduce the induced latency. For example, Scherfgen et al. (2015) present a system that is able to select the best timing in capturing images such that the system is able to achieve a fast response rate. He and Ho (2016) propose an efficient multicast scheduling policy to achieve a real-time transmission with minimal response delay. Note that most works try to minimize the latency by tackling either the processing at the screen, smart device, or server individually, and so far no effort has been taken to address the latency issues contributed by these three elements concurrently.

2.2.2 *Interaction Accuracy.* Incorrect interaction might induce frustration, especially when the current content to be interacted with is perceivable by users. Since the content is the ultimate output consumed by users during an interaction, given the target content C_t and output content C_o, the interaction accuracy A_I in general can be formulated as follows:

$$A_I = \frac{\sum_{i=1}^{N} I_o^{(i)}}{N}, \tag{5}$$

where N is the total number of interactions performed by a user and I_o is the outcome of the interaction from $i = 1$ to $i = N$. The outcome of interaction I_o is defined as follows:

$$I_o = \begin{cases} 1, & \text{if } C_o = C_t \\ 0, & \text{if } C_o \neq C_t \end{cases}, \tag{6}$$

where C_t is the expected content to be obtained during the interaction. An error occurs when C_o is different from C_t; this scenario is depicted in Figure 4(b). However, Equation (5) is only valid when

C_o is perceivable, and this might not be the case for certain SSI applications. Hence, we further refine the metric by considering the accuracy from the following three cases: screen, smart device, and server.

An error occurs when the screen is not able to respond to the interaction commands, and such an error can be due to the hardware specifications of the screen. Another possible error is when the screen receives corrupted interaction commands, which can happen when more than two commands are sent at the same time and corrupt each other. Similarly, the output content $C_{o,s}$ conditioned on t_i can be formulated as follows:

$$C_{o,s} = g_s(p_s, \varsigma | t_i), \tag{7}$$

where $g_s(p_s, \varsigma)$ is a processing block that outputs the content subject to a set of design parameters p_s related to the hardware specifications of the screen and the received interaction command ς.

Most smart devices rely on certain intelligent algorithms to process the sensor/image/wireless signals and translate the corresponding interaction commands. Furthermore, some sensors might suffer from sensitivity in their data measurements due to their underlying hardware specifications. Considering the two scenarios discussed earlier, the output content $C_{o,d}$ perceived from the smart device perspective can be described as follows:

$$C_{o,d} = g_d(p_d, \Lambda | t_i), \tag{8}$$

where p_d is a set of design parameters related to the hardware specifications of the smart device and Λ is the corresponding signal processing algorithm.

As discussed, the server acts as a gateway to control the flow of interactions between multiple screens and multiple smart devices. In other words, the server processes the incoming interaction commands and routes the commands to the designated screens/smart devices accordingly. Errors might occur when the server forwards the commands wrongly in consequence to the incorrect routing process. In general, the output content subject to the processes at the server $C_{o,v}$ given t_i can be formulated as follows:

$$C_{o,v} = g_v(p_v, \chi | t_i), \tag{9}$$

where p_v is a set of design parameters related to the hardware specifications of the server and χ is the corresponding routing algorithm.

Last, even though the interaction commands are interpreted correctly, C_o could still be incorrect due to incorrect routing at the server. Supposing that the routing is correct, the C_o will depend on the processes at the screen/smart device. In general, the C_o corresponding to $C_{o,sc}$, $C_{o,sd}$, and $C_{o,sv}$ are summarized in Table 1.

3 SCREEN AND ITS CONTENT

Screens have become the major media channel for content presentation, including dynamic content such as videos and animations. These screens come in different sizes, ranging from large-scale screens, such as tiled video wall screens found on metropolitan streets and in shopping malls, to medium-scale screens, such as projector screens in cinemas and conference halls, to small-scale screens, such as 3D big screens in living rooms. However, not every form of content fits every size of screen, and not every size screen is suitable to be deployed in every type of location. In other words, we should always consider the underlying design purpose of the content, the message to be delivered, target audiences, and the deployment location before choosing the screen.

Moreover, in the SSI context particularly, the role of a screen is not only limited to content presentation but also to content that will attract passers-by and engage them in further interaction. User engagement is not a big deal in a private setting, as the user will always take the initiative

Table 1. Combination of Possible C_o

$e(g_{o,sv})$	$e(g_{o,sd})$	$e(g_{o,sc})$	C_o	Remark
0	0	0	$C_{o,sc} = C_{o,sd} = C_{o,sv}$	C_o are identical
0	0	1	$\{C_{o,sd}, C_{o,sv}\}$	C_o could be either $C_{o,sd}$ or $C_{o,sd}$ subject to the interaction command.
0	1	0	$\{C_{o,sd}, C_{o,sv}\}$	Same as above.
0	1	1	$\{C_{o,sd}, C_{o,sv}\}$	Same as above.
1	0	0	$C_{o,sc}$	C_o is always incorrect.
1	0	1	$C_{o,sc}$	C_o is always incorrect.
1	1	0	$C_{o,sc}$	C_o is always incorrect.
1	1	1	$C_{o,sc}$	C_o is always incorrect.

Note: $e(\cdot) \in \{0, 1\}$ measures the error with regard to C_o, where 0 means the correct output and 1 means incorrect output.

to interact with the screen, such as playing a game on the screen using her smartphone. However, in an urban environment, user engagement is always challenging, and one of the underlying challenges is the nature of the content displayed on the screen. Studies have observed that many people regard the content displayed on public screens as boring and therefore ignore them intentionally (Müller et al. 2009; Ribeiro and Metrôlho 2016). In view of the engagement issues with public screens, the following sections discuss the sources of content with the aim to identify the correlation between the content and deployment environment. We also examine the effect of the content scheduling policy on the interaction accuracy and latency.

3.1 Sources of Content

According to Clinch (2013), contents are important resources in SSI, and generally sources of content are generated from these two major types: system-defined content and user-generated content. System-defined content is attractive in terms of its professional design, whereas user-generated content encourages open creativity contributed by the general public. Both types of content have their own engagement power, and this article examines their engagement power subject to the environmental context.

3.1.1 System-Defined Content. System-defined content is usually paid content in which the content is professionally designed and created to maximize the benefit of the stakeholder. Many commercial SSI systems have been developed for advertising, as this is the sector that contributes to the highest revenue return. However, advertising content is deemed to be the most annoying content and consequently suffers from the engagement issue. To increase its engagement, many advertising companies have started to adopt SSI for their advertisements in urban environments (compare to Smart Signage (She et al. 2013) and WallShop (Muta et al. 2015)). Such SSI-based advertising systems leverage a smartphone to interact with the advertising content shown on the public screen. For instance, users can retrieve the content in which they are interested (review comments, product recommendations, etc.) into their smartphone for private browsing in addition to the main content shown on the screen. However, current interaction techniques are limited to system-defined content and unable to adjust their content to suit users' preferences. Considering the big data generated from social media networks and sensors, data analytics can be integrated into the current SSI systems for advertising such that the systems are able to predict users'

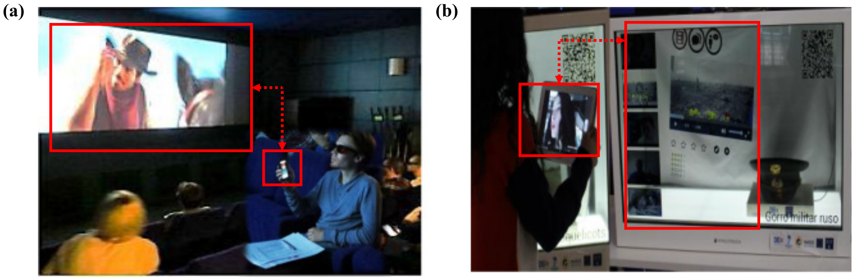

Fig. 5. Sources of content. (a) System-defined content (e.g., Interactive 3D Cinema (Häkkilä et al. 2014)). (b) User-generated content (e.g., Social Display (Bellucci et al. 2015)).

preferences and adapt their content accordingly. In addition to advertising, gaming (Dingler et al. 2015; Weißker et al. 2016) and edutainment (Hakvoort 2013; Ceipidor et al. 2013) are two other major domains that are highly dependent on system-defined content. Although it is nearly impossible for users to alter these types of system-defined content, SSI is able to turn a passive content viewing experience into an exciting interaction and thus increase user's engagement. Figure 5(a)) illustrates an interactive cinema scenario where users are allowed to catch the objects glowing out from the screen with their smartphones (compare to Interactive 3D Cinema (Häkkilä et al. 2014)).

Since system-defined content is centrally distributed, latency induced at the server can be greatly reduced since the network traffic is under monitoring. The only concern is the size and the resolution requirement of the content. Since these contents are professionally designed, screens with low specifications might not be able to display the received content correctly. Additionally, the size of the content might affect the refresh rate of the frame and result in obvious delay during the content transition time. Furthermore, interaction accuracy might also be affected when the request to change content is faster than the frame update rate. Taking MMM Ball by Weißker et al. (2016), for example, suppose that user A sends a command from his smartphone to move a ball to the left, and before the screen responds, user B joins in and asks the ball to move to the right. In this case, the screen might refresh its frame according to user A's command and switch halfway to entertain user B's command, or either users' command will be abandoned due to the slow frame update rate. Hence, it is very critical to choose the right screen to handle the requirements of the said SSI.

3.1.2 User-Generated Content. Shifting the content generation to users can possibly increase the engagement, as users might willingly involve themselves in an interaction when the content submitted by a user becomes the subject of hot discussion. In this case, the screen shows a mixture of predefined content that invites users to submit custom content in the form of text, images, and the like. For example, Figure 5(b) depicts how a user submits her commentary regarding a specific artifact displayed in a museum through her tablet (Bellucci et al. 2015). The submitted commentary, which is then made publicly accessible on the see-through display, has encouraged users to participate in further discussion. Multimedia sharing system is another respective SSI development that promotes user-generated content in which users can share their music videos or images on the screen through their personal smart devices (compare to Select&Place2Share (Seifert et al. 2014)).

Even though user-generated content can be very engaging, a certain level of administrative controls is deemed necessary to avoid content spamming. Such spamming, if not handled carefully, might have a negative impact and eventually drive people away from interaction. In terms of interaction latency, similar to system-defined content, it is very much subject to the physical attributes and specifications of the deployed screens. Note that interaction latency and accuracy are correlated in the screen context—that is, a significant latency will result in an interaction error. For example, the screen might not be able to update the content submitted by users when previously created content is still undergoing processing. Furthermore, a screen is always limited by its physical size, so parts of the user-generated content might not be displayed once the amount of submitted content exceeds a certain limit. To address this issue, one can consider making use of suitable visualizations to ensure the visibility of the entire content or to intelligently rearrange content and dedicate a larger display area for highly relevant content.

3.2 Support for Content Authoring

The ongoing penetration of smart devices has witnessed increasing interest in exploring development tools and frameworks that allow the easy creation of content and user interfaces spread over multiple screens. Modern Web technologies have often applied to realize cross-device user interfaces and interaction on off-the-shelf consumer devices. For example, Weave (Chi and Li 2015) is a scripting framework based on JavaScript and HTML to facilitate the creation of cross-device wearable interactions. Conductor (Hamilton and Wigdor 2014) is another respective framework based on cues that are broadcast between involved devices using Web sockets and JSON messages. Further related examples based on JavaScript APIs include Panelrama (Yang and Wigdor 2014) and Connichiwa (Schreiner et al. 2015). Frosini and Paternò (2014) present an API that can be exploited in both Web and Java applications and thus also supports native mobile applications written in Android.

An example of a middleware-based approach for developing multidisplay applications is Shared Substance (Gjerlufsen et al. 2011), which is a powerful generic data-oriented framework in a multi-device environment (displays, mouse pointers, user content, application states, etc.). The creators conclude that their data-oriented programming model with middleware support for sharing data and functionality provides a flexible, robust solution with low viscosity at both design time and runtime. In addition to frameworks and middleware for creating cross-device content, graphical approaches for content creation have also been explored. For example, XDStudio (Nebeling et al. 2014) is a builder tool for graphical user interfaces spread over multiple devices. XDStudio not only supports content authoring for several displays on a single device but also features on-device authoring.

3.3 Content Scheduler

A content scheduler defines how multiple content can be displayed either on a single screen or collaboratively over multiple screens. Here, we focus our discussion on two common types of content schedulers: the temporal scheduler and the spatial scheduler.

3.3.1 Temporal Scheduler. Figure 6(a) illustrates the concept of a temporal scheduler. A temporal scheduler is widely seen in most commercial screens to broadcast the advertisement sequentially according to the allocated time slot—for example, Smart Signage (She et al. 2013). The major issue with a traditional temporal scheduler is the inflexibility in content viewing. Since there is no way for users to rewind or replay the content, if a user misses a particular content, she needs to wait for the same content to be repeated, which can greatly degrade the QoE. To address this issue, Smart Signage allows users to drag the content shown on the screen to their smartphones such that users can still browse the content using their smartphones even though the content on the screen changes. However, such an approach suffers from low interaction accuracy, especially when the allocated display time for a content is too short.

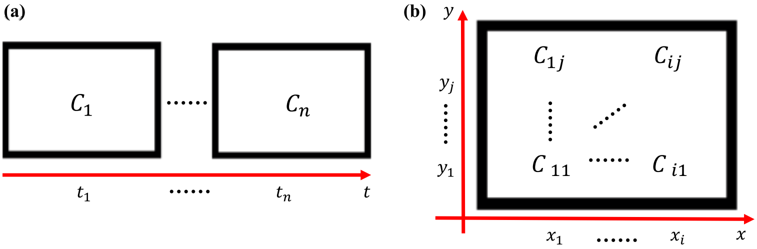

Fig. 6. Two types of content schedulers: the temporal scheduler (a) and the spatial scheduler (b).

3.3.2 *Spatial Scheduler.* In case of typical large public screens, multiple content can be displayed on a single screen at the same time by explicitly assigning each content to an available area on the screen, as illustrated in Figure 6(b). The Augmented Video Wall is one example that uses a spatial scheduler to display multiple content on the screen (Baldauf and Fröhlich 2013). The physical constraints of the screen size pose a great challenge to the spatial scheduler, where some content might need to be compromised so that it can all be squeezed onto a single screen. One solution is to use a tiled display wall with an array of multiple screens to provide a high-resolution display for multiple content. Other than the physical constraint imposed by the limited screen size, Dix and Sas (2010) also stated that the complexity of interaction increases proportionally with the amount of content being displayed on a single screen. Such complexity inherently increases the interaction latency and decreases the interaction accuracy.

4 SMART DEVICE AND ITS ROLES

This section discusses two elementary roles of a smart device: using the device as a second (miniature) screen and as a remote control. Finally, the interplay across multiple smart devices is discussed in Section 4.3.

4.1 Smart Device as Second Screen

A smart device, in most cases, features a private view in a much smaller form factor in addition to the current view shown on the screen. For example, Augmented Video Wall sends a user-selected video to the user's smartphone from a list of videos showing on the large screen (Baldauf and Fröhlich 2013). Most museum applications have also explored the use of a smartphone to grant users the accessibility to extra information regarding the displayed artifacts shown on the public screen (Hakvoort 2013).

The interaction latency in this case is very much dependent on the file to be transferred or the quality of the streamed video. Since the screen available on a smart device is typically smaller than the public deployed screen, we can reduce the possible latency by compressing the data to be transferred to the smart device. By minimizing the interaction latency, the interaction accuracy can also be improved. This is especially true for screens that showcase multiple content based on a time scheduler. Another factor that will affect the interaction accuracy is when two contents are closely spaced. Hence, careful placement of multiple content on the screen is needed when a spatial scheduler is employed to guarantee multiple users good QoE over the same screen without interfering with each other.

4.2 Smart Device as Coupling Device

Another widely exploited function of a smart device is to couple it with the screen so that a user can control or navigate the content on the screen using his own personal smart device. When a smart

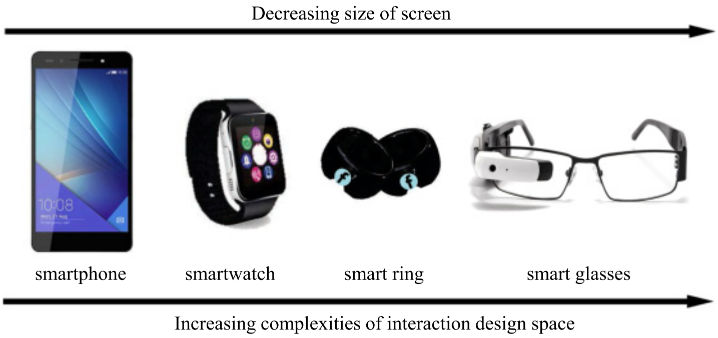

Fig. 7. Screen miniaturization limits the possible interaction design space (Ni and Baudisch 2009).

device is used as a coupling device, the dataflow between the screen and smart device is bidirectional in general. For example, uCanvas allows users to control the cursor on the screen through the smartphone's built-in magnetometer and accelerometer (Dingler et al. 2015). In addition to such sensors, computer vision techniques have also been exploited to enhance the interaction experience. For example, Jeon et al. (2010) use the smartphone's camera to compute the motion flow and inherently move the object shown on the screen, whereas Turner et al. (2013) introduce a gaze control that can be used to control the object on the screen through eye motion. Compared to the previous two roles, extra hardware is required in this system—that is, an eye tracking system to track the motion of the eye.

Using a smart device as a coupling device has been widely applied in gaming applications. For example, MMM Ball allows users to play the game shown on the distant screen via a smartphone-based gamepad (Weißker et al. 2016). However, this kind of smartphone-based gamepad often requires user to fix her eyes on the smartphone's screen and implicitly decreases the interaction accuracy, as users might miss the updated content shown on the screen. According to the investigation conducted by Baldauf et al. (2015), most smartphone-based gamepads have failed to substitute for a traditional joystick due to the lack of sensory touch feedback. The lack of sensory touch feedback has implicitly increased the interaction latency.

4.3 Interplay Across Multiple Smart Devices

With the emergence of wearable devices such as smart rings, smart glasses, and smart watches, smart device–based interaction is no longer limited to portable devices such as smartphones and tablets. The restriction of screen size due to the screen miniaturization (Holz et al. 2012), as shown in Figure 7, inevitably poses a great design challenge since some of the smartphone-based interaction modalities might not be applicable to wearable devices and reduce the degree of freedom in interaction design in connection to the limited design spaces. Nonetheless, such a limitation has indeed motivated research communities to seek other alternatives and inherently opened the door for greater opportunities, such as manipulating every part of the body as an input control to achieve hands-free and eyes-free interactions.

Since each smart device has its own pros and cons, for example, a portable smart device is more accessible, and a wearable smart device is more intuitive, the interplay across smart devices can create intuitive and accessible interaction by leveraging the strength of both. For example, Watchconnect (Houben and Marquardt 2015) introduced a smart watch to smartphone joint interaction by coordinating the smart watch's motion and the smartphone's touch input to extend their visual and tactile outputs to each other. Such interplay ensures users QoE by assigning the role of the smart watch and smartphone according to their own strengths. For instance, the smart watch is

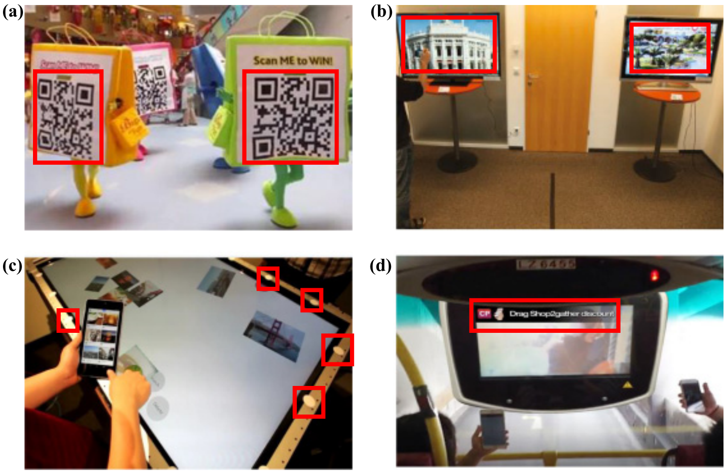

Fig. 8. Commonly used methods for device pairing: QR code (a), image recognition (Baldauf and Fröhlich 2013) (b), NFC tag (Fei et al. 2013) (c), and BLE beacon (d).

responsible for gesture motion, whereas the smartphone is responsible for storing the pull-down content.

5 INTERACTION AND ITS TECHNOLOGIES

To realize the *interaction* specified by SSI, three essential technologies are needed: device pairing, data communication channel, and interaction modalities. This section surveys the SSI system according to these three technologies and evaluates the interaction scalability in connection with the two adopted evaluation metrics.

5.1 Device Pairing

As discussed, a smart device must bind itself to a screen such that data can be exchanged over the dedicated communication channel. In addition to manual pairing (i.e., keying in the pairing password or typing in the URL manually), spontaneous device-pairing techniques have been developed that can be divided into two general types: vision based and radio based.

5.1.1 Vision Based. By leveraging a camera-enabled smartphone, together with computer vision techniques, users can bind their smart devices to the target screen. One of the more widely adopted vision-based techniques is the QR code, which binds the smart device to an accessible link for further interaction (Karaman et al. 2016). Figure 8(a) shows the use of a QR code to raise users' awareness regarding the events held in the shopping mall. However, many find that the QR code is not appealing enough to attract users, and researchers have started to exploit other image recognition techniques for device pairing. Referring to Figure 8(b), Baldauf et al. (2013) proposed spontaneous device pairing by scanning the image shown on the screen. This vision-based approach is a popular choice for device pairing due to zero deployment cost; however, it has a heavy computational load, which might increase its interaction latency. Furthermore, interaction accuracy might also be affected as a consequence of the lack of robustness of the implemented computer vision algorithm.

5.1.2 Radio Based. In earlier days, radio-based device pairing was used and was considered to be painfully tedious, as it involved several steps. First, the device must be initialized in the discovery mode; second, the target device (i.e., the screen in SSI context) must be selected; and third, the encrypted password must be keyed in to connect. Despite its tedious process, it is still widely applied in most SSI systems. For example, WiFi is used to associate the device to the local area network (compare to Smart Signage (She et al. 2013)) or Bluetooth to an ad hoc network (compare to IBBAS (Parker et al. 2016)). Lately, near field communication (NFC) tags and Bluetooth low energy (BLE) beacons, as shown in Figure 8(c) and (d), have emerged to be an alternative pairing mechanism. The drawback with the radio-based approach is the hardware deployment cost, as extra hardware is needed. Such extra hardware might induce extra latency subject to the processing memory of each piece of hardware. In contrast to the manual pairing, incorrect pairing might result since users are generally unaware, for example, of the target BLE device with which the smart device is communicating. Consequently, such incorrect pairing will lead to incorrect interaction.

5.1.3 Others. In addition to vision-based and radio-based device pairing techniques, some SSI systems use a hybrid approach to initiate the device pairing. For example, Yamaguchi et al. (2013) proposed a pairing technique by comparing the gesture motion obtained from the smartphone's accelerometer with the motion captured by the screen-side camera. Since each of the device pairing mechanisms has its own strengths, the developers need to be careful in deciding the right pairing mechanisms by considering the design principles, the purpose of the applications, and the target audiences. The deployment environment is another factor that developers need to consider. For example, it might be inapplicable to adopt a vision-based approach for some gallery-based SSI systems if the gallery setting is generally dark with a limited number of dim lights.

5.2 Data Communication

A data communication channel is established once the screen and smart device are paired. Most SSI systems are either connected to an ad hoc network or local area network via radio-based or optical-based technologies. Radio-based technologies, such as infrared (IR), are normally used for ad hoc communication, whereas Bluetooth can be used to support multiple users either in an ad hoc manner or within a personal area network (PAN). WiFi technology is widely used for data transfer, as it can support a higher data transmission rate and thus minimize the possible latency.

An optical-based communication channel is another wireless channel used in SSI systems. One recent example that used an optical communication channel for data transfer is by Lichtblick (Ostkamp et al. 2015). Lichtblick manipulates a sequence of optical signals emitted by the smartphone and translates them into a series of codes that can be recognized by the screen. The advantage of optical-based communication is that no Internet connection is needed, as both the screen and smart device can communicate directly over the optical communication link. The drawback is that the smartphone and screen must be within the line of sight for correct interaction. Despite all of the limitations, optical-based interaction is another best choice for future interaction development due to the low latency guarantee with a high data transmission rate.

5.3 Interaction Modalities

As discussed, interaction modalities are a set of commands for the users to interact with the system. The interaction modalities can be in the form of sensors, image processing, or optical signals. The commonly exploited interaction modalities for SSI are tangible touch, mid-air gesture motion, augmented reality (AR), and gaze control.

5.3.1 Tangible Touch. One common approach that provides tangible touch is by using an NFC-embedded screen with which users can interact using an NFC-enabled smartphone to select, pick up, and drop an object on the screen (Broll et al. 2013). In addition to NFC, another approach to

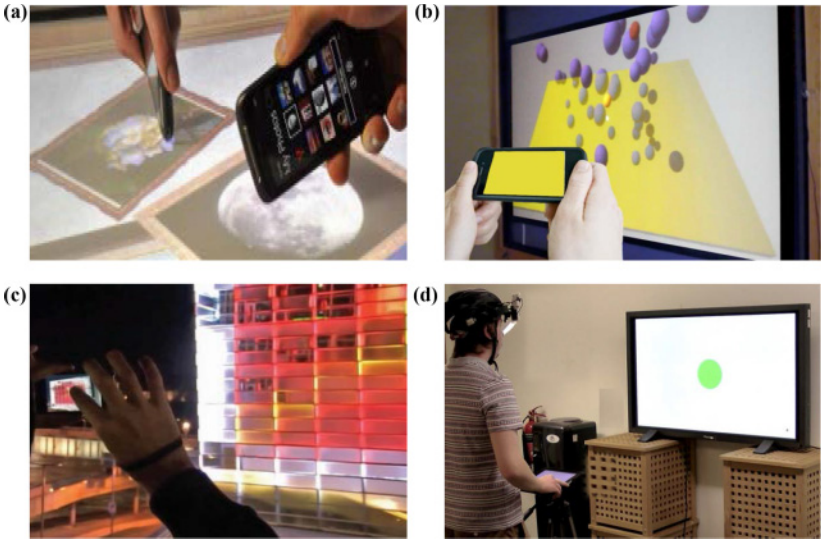

Fig. 9. Interaction modalities: tangible touch (Schmidt et al. 2012) (a), mid-air gesture motion (Pietroszek et al. 2015) (b), AR (Boring et al. 2011) (c), and gaze control (Turner et al. 2013) (d).

realize tangible touch interaction is to exploit the matching between the camera and touch events of both the screen and smart device. For example, a mobile device can be used as a stylus to select and pick up media content by touching the graphical representation on the screen (Schmidt et al. 2012) (Figure 9). Such interaction modality requires a robust image processing algorithm such that the system is able to identify the user's input correctly. Lack of robustness might eventually affect the interaction accuracy. Interaction latency might also be induced in consequence of heavy traffic load, especially when multiple users trigger interactions at the same time.

5.3.2 Mid-Air Gesture Motion. Various technologies have been exploited to achieve mid-air gesture motion interaction with the screen, such as using the data from a smart device's built-in sensors, and computing the motion through image detection or comparing the detected ray signal. Pears et al. (2009) propose a gesture interaction method by calculating the image marker position. In addition, a gesture interaction based on ray signals was recently proposed by Pietroszek et al. (2015). Such an interaction supports remote interaction over a 3D space. As opposed to the occlusion problem with tangible interaction, gesture interaction allows multiple users to interact simultaneously at a distance. However, mid-air gesture motion also suffers from a similar performance bottleneck as tangible touch since the implemented algorithm is sensitive to noise and unable to produce correct detection in a noisy environment.

5.3.3 Augmented Reality. Using AR techniques to interact with the screen is active research in SSI. For example, Figure 9(c) shows the AR interaction between a smartphone and media facade through live video (Boring et al. 2011). By calculating the spatial relationship between the video frames and the media facade, users can control the content shown on the media facade via their smartphones. According to Parker et al. (2015), AR interaction can promote multiple interactions while guaranteeing the privacy of each individual through private and personalized interactions. However, they also impose a high computational load that consequently affects overall interaction accuracy. Furthermore, the complexity of the related AR algorithm might also increase the challenge to achieve acceptable interaction accuracy.

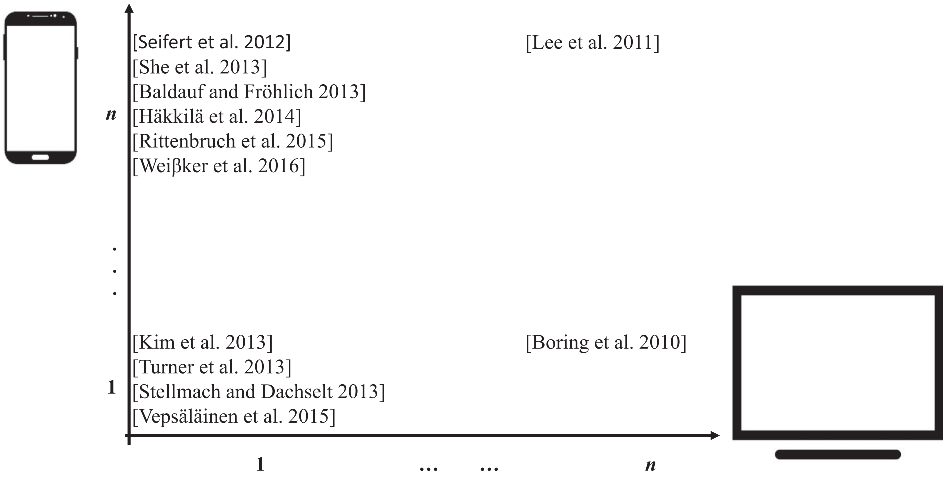

Fig. 10. Mapping of related works according to the scalability of their systems.

5.3.4 Gaze Control. In addition to using a smartphone-enabled camera, some systems have started to exploit the possibility of using a screen-side camera together with an eye-tracking system to achieve an eye-free interaction through gaze control. Examples of SSI systems that use gaze control are found in the work of Stellmach and Dachselt [2013] and Velloso et al. [2016]. Figure 9(d) illustrates the scenario where a user pulls and pushes a file between a smartphone and a screen through gaze control (Turner et al. 2013). Gaze control interaction is intuitive; however, such a mechanism is expensive in software computation in addition to the high hardware implementation cost. The interaction latency and accuracy are also subject to the specifications of the deployed hardware and the software algorithm.

5.4 Interaction Scalability

A scalable system which can support multiple interactions without compromising the interaction accuracy and latency is always desirable. Referring to the diagram shown in Figure 10, the interaction scalability is defined according to the ratio of screens to smart devices: one-to-one, one-to-many, many-to-one, and many-to-many.

5.4.1 One-to-One. Most of the earlier works in SSI are based on one-to-one interaction, as shown in Figure 10. One-to-one interaction is the most simple architecture in SSI development and has been widely applied in remote control applications, such as using a smartphone as a remote control to interact with a smart TV (Fu 2016; Vepsäläinen et al. 2015). Due to the limitations of eye tracking systems, most gaze control–based SSIs are mostly one-to-one interaction systems (Turner et al. 2013; Stellmach and Dachselt 2013). Since one-to-one interaction only involves one screen and one smart device, the major cause of latency is the processing time at the screen and smart device rather than the network traffic. Whereas for interaction accuracy, incorrect interaction occurs when the system fails to interpret the command sent by the user correctly.

5.4.2 One-to-Many. One-to-many interaction is widely applied in public scenarios, where multiple users are allowed to interact with a single screen simultaneously via their own smart devices. Multiple smart devices can interact with a single screen either independently or collaboratively. For example, Augmented Video Wall (Baldauf and Fröhlich 2013) enables multiple users to interact with the screen independently (i.e., the interaction event triggered by user A is independent of the interaction event triggered by user B). Collaborative interaction means that user A and B are working together to trigger the interaction. This kind of interaction is always seen in game applications where multiple users use their own smart devices to play the game shown on the screen (Weißker et al. 2016). Collaborative interaction is beneficial in facilitating group discussion, where multiple users can share their input from their smart devices to the screen for public discussion (Seifert et al. 2012; Rittenbruch 2015). Since one-to-many involves multiple smart devices interacting with a single screen, a server with a large capacity is needed to support the increasing traffic load. A low bandwidth server might result in undesirable latency and may affect the QoE. Furthermore, a high specification screen is desirable such that the screen can process multiple received content quickly and refresh their displays accordingly.

5.4.3 Many-to-One. Many-to-one interaction is less popular in a public scenario, as it is considered inefficient to deploy multiple screens yet only a single user can interact with them at any one time. This type of interaction is mostly applied in a workstation, where a user needs to work with multiple screens. To achieve a seamless interaction between multiple screens, Boring et al. (2010) proposed a touch projector system where users can move the content from one screen to another screen freely using their smartphones. One-to-many interaction is software computationally expansive as opposed to many-to-one interaction, which imposes heavy computation and processing on the hardware. Similarly, interaction latency and accuracy are still the ultimate design considerations considering the requirement of the hardware implementation.

5.4.4 Many-to-Many. One of the examples of many-to-many interactions is the dual interaction system (Lee et al. 2011) that allows users to share their designs from their smartphones to different types of screens, such as a projected screen and tabletop, simultaneously to facilitate collaborative discussion. Note that interaction latency and accuracy are critical in many-to-many interactions. Imagine a scenario where many people are talking at the same time; chances are that some conversations might be lost due to the noise and interruption. Similarly, when multiple smart devices are talking to multiple screens at the same time, interaction errors might happen when the system is not able to distinguish between the target and interference signals. Furthermore, many-to-many interactions can impose heavy network traffic and consequently increase the interaction latency. Hence, a scalable SSI system should be able to support an increasing interaction capacity—that is, many-to-many interactions without compromise in terms of interaction latency and accuracy. Such a scalable SSI system is desirable in a collaborative setting to support synchronous and asynchronous interactions across multiple screens and smart devices.

6 RESEARCH CHALLENGES AND OPPORTUNITIES

Having surveyed the existing work in the field of SSI extensively, we identify several future research opportunities. This section outlines such research challenges and opportunities organized on a basic research road map from short- to long-term research.

6.1 Short Term

Short-term research challenges and opportunities for the coming 2 years are described as follows.

6.1.1 Interoperability Between Multiple Screens. With ongoing advances in display technology, screens are no longer limited to a certain form factor, and any object can be augmented to become a screen. In addition to adding a screen to a vehicle or using a free-floating screen (Schneegass et al. 2014), as shown in Figure 11, recent research has also exploited the possibility of a projector screen being carried by a user and transforming it into a public screen for collaborative interaction (Wolf et al. 2016). Notwithstanding the emergence of various screens, most of them work independently.

Fig. 11. From a situated screen to a moving screen to a flying screen: large screen at Hong Kong Cyberport (a), moving screen on a truck (b), and free-floating screen (Schneegass et al. 2014) (c).

To address the issue related to interoperability, Müller et al. (2014) propose remote interaction with connected screens. Their experiment concludes that the cooperation of connected screens at different locations can increase user interaction and stir a remote honey-pot effect. Even though connected screens could be a solution to the interoperability issue, standardization is still needed to ensure interoperability between multiple screens.

6.1.2 Online to Offline Interaction. Despite the current development of smart technologies, many have raised concerns regarding the time people spend on the Internet. Since technology can take users from offline interaction to online interaction, many researchers believe that by reversing the use of such technology, it can coax users back to offline interaction. For example, by using a second screen to reveal users' activities with their smartphones, it creates an opportunity for serendipitous interaction (Jarusriboonchai et al. 2016). In other words, a second screen can be used as a social display to inform the surroundings of the current activity of a user. Through a second screen, passers-by can be made aware of the activity of a user and offline interaction can be stimulated when a passer-by is interested in the user's activity. Many research works have also started to manipulate wearable devices as a second screen to stimulate offline interaction. For example, Pearson et al. (2015) propose using a smart watch as a public display, which can deliver services like the weather forecast to people. A smart handbag has also been used as a public screen to stimulate social interaction (Colley et al. 2016). Obviously manipulating the current technology and reversing a private screen to a public screen encourages online interaction to revert back to social interaction. However, a second screen is not widely accepted in many communities due to privacy concerns, and further interdisciplinary research is needed to understand the public's acceptance as well as the reliability of the privacy protection adopted by the second screen.

6.2 Mid Term

The following two sections describe the mid-term research opportunities within a 5-year horizon.

6.2.1 Remote Communication and Collaboration. Networked public screens have been identified as a new communication medium that is able to connect people from different city districts, different parts of a country (see Gen. Schieck et al. (2014)), and even across country borders. Whereas for many applications the touch-enabled public screens are used, the role of additional smart devices for this usage scenario is unclear at this time. Smart devices might not only be used as input devices in this scenario, but their displays might additionally provide a private space to show personal and maybe privacy-critical information. Further, additional body-worn displays might enable even more complex communication and interaction modalities. Future research could investigate this interplay of public screens with public information and tiny private screens with private information for multimodal remote communication and collaboration scenarios in urban environments.

6.2.2 Interaction Beyond the Screen. Recently, several works have been proposed to extend interaction with real-world objects. For example, Snap-to-it (de Freitas et al. 2016) allows users to control any object using their smartphone. After using a vision-based approach to pair up the object, users can access the Web-based interface to control the target object using their smartphone. Similarly, an AR approach can be applied to interact with cultural artifacts exhibited inside a museum. Through an overlay video on the smartphone's camera, extra pieces of information regarding a particular artifact can be acquired and viewed in real time (Rattanarungrot et al. 2015). In addition to a vision-based approach, a radio-based approach has also been applied to interact with an Internet-connected object, which is generally known as the Internet of Things (IoT). One viable radio technology for IoT interaction is BLE. For example, a BLE-based smart ticketing system has been developed for public transportation (Narzt et al. 2016). With the current advancement of IoT, many opportunities have been created to extend smart device–based interaction with real-world objects.

6.3 Long Term

Challenges and opportunities in the long term (i.e., the next 10 years) are related to recent trends toward radical new display technologies and smart device classes beyond wearables.

6.3.1 Interaction with Shape-Changing Displays. Whereas screens have been becoming flatter and more fine grain over the past few years, the next generation of displays might dynamically change their shape and represent content in 3D. Recent respective research just started to explore potential applications for such shape-changing displays (Sturdee et al. 2015). We argue that the combination of such organic displays with today's and upcoming smart devices enables an entirely new set of interaction and use cases. For example, a smart device might be physically integrated into a dynamic display environment and become part of its novel shape and functionality.

6.3.2 Pervasive Screens and Future Smart Devices. HCI research has started to investigate novel classes of smart devices beyond today's wearables, such as smart glasses and smart watches. An example is AugmentedForearm (Olberding et al. 2013), a wearable display that covers the entire forearm and thus provides a large display surface. Other research focuses on (nearly) invisible devices implanted underneath the human skin. Researchers have begun to investigate the basic functionality of implanted devices such as sensing input through skin and communicating with external appliances (Holz et al. 2012). Related research investigated on-skin input for controlling mobile devices (Weigel et al. 2014). The combination of such future human-augmenting devices with nearby pervasive screens enables a variety of novel use cases and provides diverse research opportunities from seamless pairing to novel, maybe implicit, interaction with the screen.

7 CONCLUSION

SSI delivers multimedia content to the public through publicly deployed screens and engages users in interaction via diverse interaction modalities. In this survey, recent advancements in SSI have been presented and structured according to the notion of SSI. Two evaluation metrics (i.e., interaction latency and accuracy) have also been used to discuss the performance limitations of current works in SSI. The contents shown on the screen are the ultimate factor that draws users' attention before any interaction. Hence, it is important to understand how the content is created, distributed, and scheduled over multiple screens. Ubiquitous smart devices play various roles in our daily life. The interaction between screens and smart devices allows content to be exchanged from the public medium to the private medium, and vice versa. Contradictory to direct human-screen interaction, a data communication link needs to be established between the screens and smart devices before they can interact. The overhead induced during device pairing and its impact toward QoE is discussed. This work hopes to highlight the development of SSI in terms of interaction scalability (i.e., the capability to support multiple interactions without compromising the two evaluation metrics described earlier) and stimulate the next generation of SSI research with possible research opportunities. To conclude, SSI offers a means of interacting with pervasive screens through smart devices, reflects a smart lifestyle with context-aware services, and encourages social interaction in modern urban environments.

REFERENCES

Florian Alt, Alireza Sahami Shirazi, Thomas Kubitza, and Albrecht Schmidt. 2013. Interaction techniques for creating and exchanging content with public displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'13)*. 1709–1718.

Carmelo Ardito, Paolo Buono, Maria Francesca Costabile, and Giuseppe Desolda. 2015. Interaction with large displays: A survey. *ACM Computing Surveys* 47, 3, 46.

Matthias Baldauf, Florence Adegeye, Florian Alt, and Johannes Harms. 2016. Your browser is the controller: Advanced Web-based smartphone remote controls for public screens. In *Proceedings of the 5th ACM International Symposium on Pervasive Displays (PerDis'16)*. 175–181.

Matthias Baldauf and Peter Fröhlich. 2013. The Augmented Video Wall: Multi-user AR interaction with public displays. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, New York, NY, 3015–3018.

Matthias Baldauf, Peter Fröhlich, Florence Adegeye, and Stefan Suette. 2015. Investigating on-screen gamepad designs for smartphone-controlled video games. *ACM Transactions on Multimedia Computing, Communications, and Applications* 12, 1s, 22.

Matthias Baldauf, Markus Salo, Stefan Suette, and Peter Fröhlich. 2013. The screen is yours—comparing handheld pairing techniques for public displays. In *Proceedings of the International Joint Conference on Ambient Intelligence*. 32–47.

Andrea Bellucci, Paloma Diaz, and Ignacio Aedo. 2015. A see-through display for interactive museum showcases. In *Proceedings of the 2015 International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 301–306.

Andrea Bellucci, Alessio Malizia, and Ignacio Aedo. 2014. Light on horizontal interactive surfaces: Input space for tabletop computing. *ACM Computing Surveys* 46, 3, 32.

Andrea Bellucci, Alessio Malizia, Paloma Diaz, and Ignacio Aedo. 2010. Human-display interaction technology: Emerging remote interfaces for pervasive display environments. *IEEE Pervasive Computing* 9, 2, 72–76.

Sebastian Boring, Dominikus Baur, Andreas Butz, Sean Gustafson, and Patrick Baudisch. 2010. Touch projector: Mobile interaction through video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 2287–2296.

Sebastian Boring, Sven Gehring, Alexander Wiethoff, Anna Magdalena Blöckner, Johannes Schöning, and Andreas Butz. 2011. Multi-user interaction on media facades through live video on mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 2721–2724.

Gregor Broll, Eduard Vodicka, and Sebastian Boring. 2013. Exploring multi-user interactions with dynamic NFC-displays. *Pervasive and Mobile Computing* 9, 2, 242–257.

U. Biader Ceipidor, C. M. Medaglia, V. Volpi, A. Moroni, S. Sposato, M. Carboni, and A. Caridi. 2013. NFC technology applied to touristic-cultural field: A case study on an Italian museum. In *Proceedings of the 2013 5th International Workshop on Near Field Communication (NFC'13)*. IEEE, Los Alamitos, CA, 1–6.

Pei-Yu (Peggy) Chi and Yang Li. 2015. Weave: Scripting cross-device wearable interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI'15)*. 3923–3932.

Sarah Clinch. 2013. Smartphones and pervasive public displays. *IEEE Pervasive Computing* 12, 1, 92–95.

Ashley Colley, Minna Pakanen, Saara Koskinen, Kirsi Mikkonen, and Jonna Häkkilä. 2016. Smart handbag as a wearable public display-exploring concepts and user perceptions. In *Proceedings of the 7th Augmented Human International Conference*. ACM, New York, NY, 7.

Adrian de Freitas, Michael Nebeling, Xiang 'Anthony' Chen, Junrui Yang, Akshaye Shreenithi Kirupa Karthikeyan Ranithangam, and Anind K. Dey. 2016. Snap-to-it: A user-inspired platform for opportunistic device interactions. In *Proceedings of the 34th Annual ACM Conference on Human Factors in Computing Systems (CHI'16)*.

Matthias Deller and Achim Ebert. 2011. Modcontrol–mobile phones as a versatile interaction device for large screen applications. In *Proceedings of the IFIP Conference on Human-Computer Interaction*. 289–296.

Tilman Dingler, Tobias Bagg, Yves Grau, Niels Henze, and Albrecht Schmidt. 2015. uCanvas: A Web framework for spontaneous smartphone interaction with ubiquitous displays. In *Human-Computer Interaction*. Springer, 402–409.

Alan Dix and Corina Sas. 2010. Mobile personal devices meet situated public displays: Synergies and opportunities. *International Journal of Ubiquitous Computing* 1, 1, 11–28.

Shenfeng Fei, Andrew M. Webb, Andruid Kerne, Yin Qu, and Ajit Jain. 2013. Peripheral array of tangible NFC tags: Positioning portals for embodied trans-surface interaction. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces.* ACM, New York, NY, 33–36.

Luca Frosini and Fabio Paternò. 2014. User interface distribution in multi-device and multi-user environments with dynamically migrating engines. In *Proceedings of the 2014 ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS'14).* 55–64.

Jiuqiang Fu. 2016. Smart TV interface design and evaluation based on goal-oriented. In *Proceedings of the 2016 6th International Conference on Instrumentation and Measurement, Computer, Communication, and Control (IMCCC'16).* IEEE, Los Alamitos, CA, 327–330.

Ava Fatah Gen. Schieck, Holger Schnädelbach, Wallis Motta, Moritz Behrens, Steve North, Lei Ye, and Efstathia Kostopoulou. 2014. Screens in the wild: Exploring the potential of networked urban screens for communities and culture. In *Proceedings of the International Symposium on Pervasive Displays (PerDis'14).* Article No. 166.

Tony Gjerlufsen, Clemens Nylandsted Klokmose, James Eagan, Clément Pillias, and Michel Beaudouin-Lafon. 2011. Shared substance: Developing flexible multi-surface applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'11).* 3383–3392.

Felix Greimel. 2011. A survey of interaction techniques for public displays. *Advances in Media Technology.* Institute for Media Technology.

Jonna R. Häkkilä, Maaret Posti, Stefan Schneegass, Florian Alt, Kunter Gultekin, and Albrecht Schmidt. 2014. Let me catch this! Experiencing interactive 3D cinema through collecting content with a mobile phone. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems.* ACM, New York, NY, 1011–1020.

Gido Hakvoort. 2013. The immersive museum. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces.* ACM, New York, NY, 463–468.

Peter Hamilton and Daniel J. Wigdor. 2014. Conductor: Enabling and understanding cross-device interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'14).* 2773–2782.

Wei He and Pin-Han Ho. 2016. On achieving cyber-physical real-time snapshot acquisition in billboard/signage networks. *IEEE Internet of Things Journal* 3, 6, 1213–1221.

Christian Holz, Tovi Grossman, George Fitzmaurice, and Anne Agur. 2012. Implanted user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'12).* 503–512.

Steven Houben and Nicolai Marquardt. 2015. Watchconnect: A toolkit for prototyping smartwatch-centric cross-device applications. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.* ACM, New York, NY, 1247–1256.

Pradthana Jarusriboonchai, Aris Malapaschas, Thomas Olsson, and Kaisa Väänänen. 2016. Social display... We can see what you are doing on your mobile device. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion.* ACM, New York, NY, 53–56.

Seokhee Jeon, Jane Hwang, Gerard J. Kim, and Mark Billinghurst. 2010. Interaction with large ubiquitous displays using camera-equipped mobile phones. *Personal and Ubiquitous Computing* 14, 2, 83–94.

Svebor Karaman, Andrew D. Bagdanov, Lea Landucci, Gianpaolo DAmico, Andrea Ferracani, Daniele Pezzatini, and Alberto Del Bimbo. 2016. Personalized multimedia content delivery on an interactive table by passive observation of museum visitors. *Multimedia Tools and Applications* 75, 7, 3787–3811.

Jae Yeol Lee, Min Seok Kim, Dong Woo Seo, Chil-Woo Lee, Jae Sung Kim, and Sang Min Lee. 2011. Dual interactions between multi-display and smartphone for collaborative design and sharing. In *Proceedings of the 2011 IEEE Virtual Reality Conference.* IEEE, Los Alamitos, CA, 221–222.

Jörg Müller, Dieter Eberle, and Konrad Tollmar. 2014. Communiplay: A field study of a public display mediaspace. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, 1415–1424.

Jörg Müller, Dennis Wilmsmann, Juliane Exeler, Markus Buzeck, Albrecht Schmidt, Tim Jay, and Antonio Krüger. 2009. Display blindness: The effect of expectations on attention towards digital signage. In *Proceedings of the International Conference on Pervasive Computing.* 1–8.

Masafumi Muta, Soh Masuko, Keiji Shinzato, and Adiyan Mujibiya. 2015. Interactive study of wallSHOP: Multiuser connectivity between public digital advertising and private devices for personalized shopping. In *Proceedings of the 4th International Symposium on Pervasive Displays.* ACM, New York, NY, 187–193.

Wolfgang Narzt, Stefan Mayerhofer, Otto Weichselbaum, Stefan Haselb, et al. 2016. Bluetooth low energy as enabling technology for be-in/be-out systems. In *Proceedings of the 2016 13th IEEE Annual Consumer Communications and Networking Conference (CCNC'16).* IEEE, Los Alamitos, CA, 423–428.

Michael Nebeling, Theano Mintsi, Maria Husmann, and Moira Norrie. 2014. Interactive development of cross-device user interfaces. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems (CHI'14).* 2793–2802.

Tao Ni and Patrick Baudisch. 2009. Disappearing mobile devices. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 101–110.

Simon Olberding, Kian Peen Yeo, Suranga Nanayakkara, and Jurgen Steimle. 2013. AugmentedForearm: Exploring the design space of a display-enhanced forearm. In *Proceedings of the 4th Augmented Human International Conference (AH'13)*. 9–12.

Morin Ostkamp, Sven Heitmann, and Christian Kray. 2015. Short-range optical interaction between smartphones and public displays. In *Proceedings of the 4th International Symposium on Pervasive Displays*. ACM, New York, NY, 39–46.

Callum Parker, Judy Kay, Matthias Baldauf, and Martin Tomitsch. 2016. Design implications for interacting with personalised public displays through mobile augmented reality. In *Proceedings of the 5th ACM International Symposium on Pervasive Displays*. ACM, New York, NY, 52–58.

Callum Parker, Martin Tomitsch, Judy Kay, and Matthias Baldauf. 2015. Keeping it private: An augmented reality approach to citizen participation with public displays. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*. ACM, New York, NY, 807–812.

Nick Pears, Daniel Jackson, and Patrick Olivier. 2009. Smart phone interaction with registered displays. *IEEE Pervasive Computing* 8, 2, 14–21.

Jennifer Pearson, Simon Robinson, and Matt Jones. 2015. It's about time: Smartwatches as public displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1257–1266.

Peter Peltonen, Esko Kurvinen, Antti Salovaara, Giulio Jacucci, Tommi Ilmonen, John Evans, Antti Oulasvirta, and Petri Saarikko. 2008. It's mine, Don't touch! Interactions at a large multi-touch display in a city centre. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1285–1294.

Mark Perry, Steve Beckett, Kenton O'Hara, and Sriram Subramanian. 2010. WaveWindow: Public, performative gestural interaction. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 109–112.

Krzysztof Pietroszek, James R. Wallace, and Edward Lank. 2015. Tiltcasting: 3D interaction on large displays using a mobile device. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 57–62.

Sasithorn Rattanarungrot, Martin White, and Ben Jackson. 2015. The application of service orientation on a mobile ar platform??? A museum scenario. In *2015 Digital Heritage*, Vol. 1. IEEE, Los Alamitos, CA, 329–332.

Fernando Reinaldo Ribeiro and José Metrôlho. 2016. Context-aware information systems for public spaces: The public and private dichotomy. overview, challenges, and experiments. *International Journal of Multimedia and Ubiquitous Engineering* 11, 3, 11–22.

Markus Rittenbruch. 2015. Supporting collaboration on very large-scale interactive wall surfaces. *Computer Supported Cooperative Work* 24, 2–3, 121–147.

David Scherfgen, Rainer Herpers, and Timur Saitov. 2015. A robust inside-out approach for 3D interaction with large displays. In *Proceedings of the 2015 IEEE Symposium on 3D User Interfaces (3DUI'15)*. IEEE, Los Alamitos, CA, 137–140.

Dominik Schmidt, Julian Seifert, Enrico Rukzio, and Hans Gellersen. 2012. A cross-device interaction style for mobiles and surfaces. In *Proceedings of the Designing Interactive Systems Conference*. ACM, New York, NY, 318–327.

Stefan Schneegass, Florian Alt, Jürgen Scheible, Albrecht Schmidt, and Haifeng Su. 2014. Midair displays: Exploring the concept of free-floating public displays. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. ACM, New York, NY, 2035–2040.

Mario Schreiner, Roman Rädle, Hans-Christian Jetter, and Harald Reiterer. 2015. Connichiwa: A framework for cross-device Web applications. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA'15)*. 2163–2168.

Julian Seifert, David Dobbelstein, Dominik Schmidt, Paul Holleis, and Enrico Rukzio. 2014. From the private into the public: Privacy-respecting mobile interaction techniques for sharing data on surfaces. *Personal and Ubiquitous Computing* 18, 4, 1013–1026.

Julian Seifert, Adalberto Simeone, Dominik Schmidt, Paul Holleis, Christian Reinartz, Matthias Wagner, Hans Gellersen, and Enrico Rukzio. 2012. MobiSurf: Improving co-located collaboration through integrating mobile devices and interactive surfaces. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 51–60.

James She, Jon Crowcroft, Hao Fu, and Pin-Han Ho. 2013. Smart signage: A draggable cyber-physical broadcast/multicast media system. *IEEE Transactions on Emerging Topics in Computing* 1, 2, 232–243.

James She, Jon Crowcroft, Hao Fu, and Flora Li. 2014. Convergence of interactive displays with smart mobile devices for effective advertising: A survey. *ACM Transactions on Multimedia Computing, Communications, and Applications* 10, 2, 17.

Sophie Stellmach and Raimund Dachselt. 2013. Still looking: Investigating seamless gaze-supported selection, positioning, and manipulation of distant targets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 285–294.

Miriam Sturdee, John Hardy, Nick Dunn, and Jason Alexander. 2015. A public ideation of shape-changing applications. In *Proceedings of the 2015 International Conference on Interactive Tabletops and Surfaces*. 219–228.

Martin Tomitsch, Christopher Ackad, Oliver Dawson, Luke Hespanhol, and Judy Kay. 2014. Who cares about the content? An analysis of playful behaviour at a public display. In *Proceedings of the International Symposium on Pervasive Displays*. ACM, New York, NY, 160.

Jayson Turner, Jason Alexander, Andreas Bulling, Dominik Schmidt, and Hans Gellersen. 2013. Eye pull, eye push: Moving objects between large screens and personal devices with gaze and touch. In *Proceedings of the IFIP Conference on Human-Computer Interaction*. 170–186.

Eduardo Velloso, Markus Wirth, Christian Weichel, Augusto Esteves, and Hans Gellersen. 2016. AmbiGaze: Direct control of ambient devices by gaze. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. ACM, New York, NY, 812–817.

Jouni Vepsäläinen, Antonella Di Rienzo, Matti Nelimarkka, Jouni A. Ojala, Petri Savolainen, Kai Kuikkaniemi, Sasu Tarkoma, and Giulio Jacucci. 2015. Personal device as a controller for interactive surfaces: Usability and utility of different connection methods. In *Proceedings of the 2015 International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 201–204.

Martin Weigel, Vikram Mehta, and Jürgen Steimle. 2014. More than touch: Understanding how people use skin as an input surface for mobile computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'14)*. 179–188.

Tim Weißker, Andreas Berst, Johannes Hartmann, and Florian Echtler. 2016. MMM ball: Showcasing the massive mobile multiuser framework. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, New York, NY, 3796–3799.

Katrin Wolf, Markus Funk, Pascal Knierim, and Markus Löchtefeld. 2016. Survey of interactive displays through mobile projections. *International Journal of Mobile Human Computer Interaction* 8, 4, 29–41.

Tokuo Yamaguchi, Hiroyuki Fukushima, Shigeru Tatsuzawa, Masato Nonaka, Kazuki Takashima, and Yoshifumi Kitamura. 2013. SWINGNAGE: Gesture-based mobile interactions on distant public displays. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 329–332.

Jishuo Yang and Daniel Wigdor. 2014. Panelrama: Enabling easy specification of cross-device Web applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'14)*. 2783–2792.

Received August 2016; revised May 2017; accepted June 2017