Chapter 5
Cognitive Issues in Mobile Augmented Reality: An Embodied Perspective

Nai Li and Henry Been-Lirn Duh

1 Introduction

Augmented reality (AR), which is characterized by overlaying virtual imagery onto a physical world scene, enables users to interact with virtual information situated in real space and real time. Driven by the advances of hardware and software, mobile AR emerges as a promising interface for supporting AR interaction. Wearable devices such as head-mounted displays (HMDs) and wrist-worn displays are important channels for people to get immediate access to AR contents while roaming the surrounding environment. Recent developments in mobile computing technologies also enable more handheld devices like mobile phones, PDAs and tablet PCs to be the platforms to implement AR. The convergence of AR and mobile devices delivers an innovative experience for users to explore the physical world.

Mobile AR reveals potential in a diversity of domains, including manufacturing, tourism, education, entertainment, and urban modeling [79, 114, 115]. Despite the discussion on technical issues of AR, the social influence of AR technology has received considerable attention in recent years [43, 63]. In order to develop AR applications with high effectiveness, the user-centred perspective should be incorporated, and more understanding on the impact of characteristics of AR systems on human activities is needed [126]. Human factors play a vital role in designing effective computing systems. To date, some researchers have explored human factors in

---

N. Li (✉)

Department of Communications and New Media, National University of Singapore,
Blk AS6, #03-41, 11 Computing Drive, Singapore
e-mail: g0900788@nus.edu.sg

H.B.-L. Duh
Department of Electrical and Computer Engineering, National University of Singapore,
21 Heng Mui Keng Terrace, #02-02-09, Singapore
e-mail: eledbl@nus.edu.sg

AR from different perspectives with attempts to propose design guidelines [24, 31, 67]. Cognitive issues, which relate to users' cognitive process for understanding an AR environment when interacting with the system, are identified as an important category of human factors in AR [24]. Duh, Ma, and Billinghurst [24] suggested that gaining insights into cognitive issues underlying the effectiveness of existing AR systems is of significance for guiding and improving future design.

Mobile AR introduces more possibilities to AR interaction than stationary AR interfaces fixed to certain locations, but also presents new issues for developing effective AR systems. At present, our knowledge about human factors in mobile AR is very limited, and therefore it is crucial to comprehensively identify the opportunities and challenges posed by mobile AR interaction on the cognitive process. In this chapter, we provide an overview of cognitive issues involved in mobile AR systems through the lens of mobile AR interaction, in order to both clarify how mobile AR interaction concerns human cognition and offer guidance for mobile AR design in the future.

In the following sections, we will introduce the embodied perspective and explain our rationale for selecting it as the theoretical approach to organize cognitive issues in mobile AR interaction. We will further discuss the cognitive issues affecting mobile AR in detail, based on the findings of existing literature.

2 Embodied Cognition in Mobile AR Interaction

The development of computing technologies has revolutionized human–computer interaction (HCI) by driving people to actively and naturally interact with technologies within a broad range of contexts. The interaction penetrates to activities in people's daily life beyond working tasks, and ubiquitous and mobile computing makes it possible to flexibly manipulate technologies without geographical constraints. Instead of viewing the human mind as solely an information processor and cognition being isolated with action, several new theoretical approaches can be applied to the study of cognitive functioning in the field of HCI.

Embodiment, referring to "the property of our engagement with the world that allows us to make it meaningful," has been extended to the HCI domain in more recent years [23]. The embodied perspective of cognition suggests that cognitive processes are grounded in the bodily interaction in real space and real time [119]. The bodily engagement situated in specific physical and social environments can shape human cognitive process. As people's physical interaction style with the digital world becomes increasingly direct and inseparable from physical and social contexts, the embodied perspective serves as a suitable approach for analyzing HCI and generating implications for developing computing devices [51]. According to the embodied perspective, HCI involves users' constructing meaning and understanding by using technologies in physical and social environments rather than using technologies simply to implement tasks and process information.

Mobile AR demonstrates great potential for incorporating embodiment into users' interaction with computing technologies. It shifts the presentation of virtual information from onscreen displays to directly overlaying information onto the physical world. In particular, it entails the characteristic of context sensitivity for building meaningful relations between virtual information and the physical environment, which leads to a seamless merge of virtual and physical worlds [121]. Indeed, the role of physical environments as an integral part of activities is perceived as an important aspect of embodied cognition [119]. Mobile AR also offers a growing number of possibilities for users to physically interact with AR contents. By taking advantage of the features of different mobile interfaces, manipulation of AR contents has become a vital part of mobile AR interaction. There is a trend toward developing natural interaction techniques based on human manipulation skills and perceptions [25, 110]. The enhanced bodily engagement with virtual information expands users' capability of directly interacting with computing technologies to construct understanding of the setting. Apart from supporting individual activities, mobile AR also facilitates the establishment of spaces for shared experiences in indoor and outdoor environments. It enables social dynamics of meaningful construction in real world activities among multiple users whilst adding unique information produced by computing technologies. The interdependent roles of multiple users such as collaborators or competitors supported by mobile AR are also important in the effectiveness of shared activities [63, 124]. Specifically, mobile AR can augment the physical world and integrate physical resources with the mechanics of joint activities [80]. The movement and location of users can trigger certain events in mobile AR-supported activities. Therefore, users' engagement is important to mobile AR interaction, and the interaction is bound to both physical and social surroundings. Mobile AR systems should seek to strengthen users' construction of meaning and understanding in AR environments. These features of mobile AR provide evidence for using the embodied perspective to better understand mobile AR interaction and explain how the interaction supports cognitive functioning in the process of building understanding of mobile AR contexts.

On the basis of the characteristics of mobile AR interaction, we identify three primary categories of cognitive issues in mobile AR interaction: information presentation, physical interaction, and shared experience. The cognitive aspects related to each issue will be examined in the following sections.

3 Design for Mobile AR Interaction

In this section, we outline three categories of cognitive issues in mobile AR interaction including information presentation, physical interaction and shared experience. For each category, relevant aspects affecting users' cognitive functioning are discussed, and examples of mobile AR applications are presented in order to exemplify the role of cognitive issues involved in the use of systems.

3.1   *Information Presentation in Mobile AR*

The display of virtual information upon the physical world is a fundamental property of AR interfaces. Compared to virtual reality (VR), AR creates opportunities for enhancing real world objects and environments instead of replacing them. The realization of worth of mobile AR largely depends on maximizing the relevance of virtual overlays with regard to the physical world [48]. Annotations, which provide information relevant to one's physical surroundings, have become a mainstream component in current mobile AR systems in order to facilitate the understanding of the real world [121]. The meaningful integration of physical environments with mobile AR-supported activities can shape cognitive process in an embodied way [119]. To highlight the significance of arranging AR contents in a well-structured way, view management refers to "decisions that determine the spatial layout of the projections of objects on the view plane" [8]. Virtual information display is an integral part of the view supported by AR. In this section, amount, representation, placement and views combination of information are taken into account to clarify the impact of information presentation on cognitive functioning in mobile AR.

3.1.1   Amount

The co-existence of informative virtual elements and real world scenes in mobile AR interfaces delivers a unique opportunity for users to construct meaning in the physical world. Planning the amount of information available to users becomes a critical issue in mobile AR in order to support the cognitive process of making sense of information [8]. This concern has become salient as diverse platforms characterized with different properties like display spaces and field of view are gradually applied to mobile AR.

   The synthetic scene may hinder cognitive functioning if visual complexity of real world background and virtual information is not well balanced in AR [102]. Presenting a large volume of information simultaneously can result in clutter display, which is characterized by information interference in a single display. Users may feel overwhelmed by the information and have difficulty focusing, thereby increasing their mental load when using mobile AR. By investigating the effect of cumulative cluster on human cognitive performances in AR, Stedmon et al. [102] contended that clutter display sets barriers for searching targets and the visual confusion has negative effects on users' understanding of information delivered by AR. The issue of information density displayed on handheld devices in navigation has also been investigated by Ganapathy et al. [33]. By showing labels of surrounding spots such as hotels, parks, and bridges, the scenario facilitated users' identification of points of interest around them. The authors found that users have certain preferences toward the amount of items presented on a single screen. Too many or too few items could increase the effort required to search and comprehend information. For military AR applications, the focused display was stressed in order to present

the most relevant information to users during military operations; insufficient information negatively affected the maintenance of situation awareness, while too much information generated cognitive overload [73].

The cognitive effects of the amount of information also need to be addressed when applying mobile AR to realize X-ray visualization [92]. AR can not only display additional instructions in the real world, but also generate X-ray vision by visualizing the structure of invisible objects. The presentation of meaningful hidden structures requires delivering an appropriate amount of depth cues to aid the understanding of spatial location of occluded information with respect to the occluding scene, and preserving important information within occluding structures [55]. The amount of depth cues conveyed by mobile AR is vital for users to recognize the spatial relationship between occluded and occluding objects and thereby capture an accurate model of the environment [4]. Also, the display of virtual information onto the physical background can bring ambiguities to understand the real-world information [55]. Additional virtual information might occlude important information in the real-world scene, which leads to increased mental loads to make sense of the environment.

3.1.2 Representation

The representation of virtual information in mobile AR needs to be considered in order to increase users' efficiency of recognizing and comprehending information. The context-dependent characteristic of information representation has been acknowledged by a body of research [29, 121].

AR serves as a new form of representation to attach location-based information to the physical world scene [120]. Conventionally, the separation between information and physical spaces produces "cognitive distance" for users since they have to switch across spaces to extract targeting spots from information displays and then apply the information to real-world situations [60]. The transition of attention in these processes increases users' cognitive loads. For example, when offering navigation guidance for driving, "divided attention" caused by moving attention between information display and the road view could affect drivers' information processing and driving performance [60]. By overlaying virtual information about road conditions onto the windshield, mobile AR had stronger capabilities for narrowing the gap between geo-referenced information display and the physical space than a 2D bird's eye view map display; also, with AR people could more easily concentrate on the view in front of the car while gaining information regarding current and upcoming road conditions [60].

Although AR shows potential for promoting cognitive process in displaying geospatial data, the degree of enhancement varies depending on the methods used. The forms of virtual imagery in mobile AR can vary, including points, textual annotations, 2D graphics and 3D graphics. The representation of information impacts users' efforts to link virtual data with real locations in order to understand the physical world [108]. For example, 3D representations were found to be more capable of

displaying spatial information than 2D representations; the high level of realism facilitated by 3D representations made the virtual scene more understandable [19]. For nearby buildings, plain and un-textured models were adequate to convey geospatial information [29]. Different AR presentation schemes of arrows for navigation guidance impacted users' interpretation of the distance [108]. Directly viewing spots in occluded areas in real time can allow people to identify their surroundings more easily compared to using a map with symbolic representations [111]. Investigations have also been carried out in order to obtain insights into how representations of information affect user experiences in urban planning activities. For instance, different representations including spheres, cylinders, and smoke were used to visualize the level of CO in outdoor environments with the support of a mobile AR system designed by White and Feiner [115]. They found that the type of representation affected users' cognitive and emotional reactions to the data. These users' reactions suggested that diverse types of representations should be designed to adapt to different contexts.

Additionally, the reference frame of information is critical to one's cognitive functioning in order to interpret surrounding environments supported with mobile AR. Egocentric and exocentric viewpoints serve as two primary reference frames in information presentation [76]. While the exocentric viewpoint provides a better overview of the surrounding context, the egocentric viewpoint presents information for local guidance based on the first-person perspective [2]. It is suggested that the egocentric viewpoint is more useful for inferring the spatial relationship between the user's current location and the plot of interest compared with the exocentric viewpoint [93]. To take advantage of both reference frames, Langlotz et al. [69] developed a mobile AR system that enabled users to first have a global view of a exocentric 2D map with nearby annotations, then when users got close to the annotated spot, they could switch the view to a first-person perspective in order to find a way to reach the actual spot.

3.1.3 Placement

Optimizing the placement of virtual information is an indispensible aspect to consider when managing the view in mobile AR. It is suggested the information layout can impact the understandability of information; appropriate arrangement of information helps users to connect the meaning of virtual information with the real-world view [48].

As an intuitive way of annotating physical objects, labeling plays a vital role in providing more information to support exploration in AR settings [125]. Label overlap and object occlusion are two crucial problems associated with placing annotations in mobile AR [125]. Visual clutter can cause meaning ambiguity in AR contexts and exert a negative impact on users' understanding of target objects [33]. In addition, the relative distance between the label and the target object could affect users' eye movement when reading information. In one study by Azuma and Furmanski [3], as the relative label distance increased, users needed to take longer time to read the label. Ganapathy et al. [33] provided empirical evidence that there

is an acceptable maximum distance between the annotation and the target object to ensure the readability of information. In addition to the placement issue in each frame, the information layout in frame-to-frame transition is also a factor that relates to cognitive functioning. Unnatural changes of the positions of virtual information between two frames might lead to visual discontinuities; minimizing the visual discontinuity serves as a goal when managing the view in AR [8].

The dynamic feature of mobile AR should also be taken into account when planning the placement of virtual information. Mobile AR presents greater challenges for managing labeling placements than regular static backgrounds. Compared with the stationary setting, it is more difficult for wearable devices to locate spatial features of target objects such as position and orientation. This in turn affects the system's ability to identify the visible part of the object within a users' visual field and may lead to inappropriate label placements [75]. Also, the location of virtual information is crucial to users' visual attention. For mobile wearable AR, there are three main types of coordinate systems including *head-stabilized* where "information is fixed to the user's viewpoint," *body-stabilized* where "information is fixed to the user's body position," and *world-stabilized* where "information is fixed to real world locations" [10]. World-stabilized views enabled users to access to context-dependent information registered in the real world of the far field. Head-stabilized and body-stabilized views are commonly adopted in augmentations in the near field. In comparison with head-stabilized display, body-stabilized display has shown advantages in helping to understand the location of information and thereby enhancing the speed of searching information [10]. The influence of different spatial layouts of the virtual menu on the efficiency of selecting 3D objects also suggests that the placement issue should be considered when presenting information in mobile AR [117]. Rather than constantly concentrating on the information display, users of mobile AR supported by handheld devices usually switch their attention back and forth between the handheld display and the surrounding setting. The layout of information could impact the effectiveness of understanding the information at a glance [9].

3.1.4   View Combination

A set of visualization techniques such as *zooming and panning*, *overview and detail*, and *focus and context* are employed to combine multiple views on a single display [1, 2]. These techniques attempt to make full use of the display space and increase the efficiency of searching information [1].

Zooming and panning separates focused and contextual information temporally, which allow users to continuously zoom and pan the view in order to target desired information. The integration of zooming and panning with mobile AR assists users in getting detailed information in a consistent way [2, 14]. However, if people need to associate local details with the surrounding context to understand the meaning of information, zooming in to obtain details could result in the loss of global context, which could then increase users' required cognitive efforts [17].

Overview and detail separates focused and contextual information spatially by using two linked windows in a single view. The separation of two spaces requires additional visual navigation which may introduce new cognitive issues. Focus and context, such as fisheye view, seamlessly keeps focused information within the global context in one view via image distortion. It seeks to make the focus salient in the context and simultaneously enable users to grasp the spatial relation between focus and context. This type of visualization is useful when users have to frequently switch between focused and contextual information [7], but recognition and interpretation of distorted views are two primary issues affecting cognitive functioning [17]. The trade-off between the amount of distortion and mental load should not be ignored when applying distortion-oriented views to present concurrent information in mobile AR interfaces.

Rather than simply deliver information of immediate surroundings, a body of systems have sought to visualize information regarding off-screen and occluded objects by taking advantage of these techniques of view combination [2, 50, 93]. In response to the information presentation limitations of handheld devices, such as small screen size and narrow camera field of view, expanded views have been explored in mobile AR with attempts to navigate physical environments [50]. For instance, Sandor et al. [93] provided evidence that utilizing egocentric space-distorting visualizations in mobile AR is a strategy for displaying off-screen and occluded points of interest. The provision of focused and contextual information in a single view makes users more easily able to predict the spatial relation between the current environment and the point of interest without frequently switching attention across separated views. However, the distorted space could introduce new cognitive burdens for users due to increased effort required to interpret the distortion. Cognitive dissonance may arise from reading the reconstructed model of the real world [50].

3.2   Physical Interaction in Mobile AR

In more recent years, the interactive aspect of visualization has received increasing attention, and user-centred visualization has been specifically emphasized [1, 28]. According to the embodied perspective, the cognition is grounded in bodily engagement with technologies [119]. As more interactive techniques are available for mobile AR, considerations should be given to cognitive issues of physical interaction in mobile AR in order to enhance the effectiveness of systems. Navigation, direct manipulation, and content creation represent three typical physical actions engaged by users in mobile AR, and are described in this section.

3.2.1   Navigation

The advancement of mobile AR brings about innovative experiences to navigational activities by narrowing the gap between physical surroundings and abstract virtual

representations. Spatial awareness, including "a person's knowledge of self-location within the environment, of surrounding objects, of spatial relationships among objects and between objects and self, as well as the anticipation of the future spatial status of the environment," is an important factor for evaluating the success of navigation [112]. Beyond just delivering straightforward information to users, the research on mobile AR has focused more attention on interaction dimensions in navigation in order to motivate users to do self-exploratory activities and enhance their spatial awareness of physical spaces [89].

Users' ability to change viewpoints to browse information is crucial in order to explore the environment in navigation. Some mobile AR applications make it possible to actively transit across different viewpoints in real time to get desired information [2, 69]. By taking advantage of handheld devices, users can easily zoom in and out to search plots of interest, which overcomes the limited size of handheld device and enables users to gain desired information efficiently. Since the states before and after zooming can be varied in the view perspective or the field of view, the smoothness of transition is important to avoid introducing extra cognitive loads [2]. Photo-based AR interaction was designed to capture different viewpoints by taking snapshots in navigation, which allowed users to review previous viewpoints without physically revisiting those locations [103]. This method is identified as an effective way to reduce effort and time when investigating the environment.

The modality of information display also impacts users' interaction in navigation. Users move their attention between information display and physical surroundings to cognitively map their present location and spatial relationships with targeting plots [112]. Mobile phones equipped with built-in projectors are applied to navigation in order to support information displayed on a mobile phone screen and projector [40]. A projector characterized by a big screen size and high resolution helps users effectively look for information. The combination of mobile phone screen size and projector display resolution reveals potential for increasing efficiency and comfortableness of navigation. Schöning, Rohs, Kratz, Löchtefeld and Krüger [97] examined the use of projector phones in augmenting information on a classical paper map. The projection display could overlay additional customized geographical information about plots of interest on the map, avoiding the split of attention between the screen display and the physical map in navigation. The user study indicated that the projection display was capable of enhancing the efficiency of navigation by reducing completion time and error rate [97].

Instead of simply aiding users in finding their destinations, mobile AR can convey rich location-based background information based on users' interests during the navigation. Presenting information at two stages lets users browse general information at first; then after selecting certain favorite spots on the display, they can read additional detailed information to understand the target and make further plans regarding their movements [33, 86]. Also, some mobile AR systems have provided filtering options for users to control the visibility of information based on their preference, which contributed to reducing display clutter and fostering information comprehension [86].

3.2.2   Direct Manipulation

Manipulation is an important theme in users' interaction with computing technologies. Moving beyond traditional interaction methods based on WIMP metaphors (Windows, Icons, Menus, Pointer), a range of innovative interaction techniques are developed to assist natural and effective manipulations in mobile AR [15, 90]. The increasing level of physical participation can shape user experiences and affect their understanding of the world [62]. Indeed, the trend toward direct manipulation provides an important area to investigate the impact of mobile AR on cognitive process. Tangible interaction, direct hand interaction and multimodal interaction, each of which reveals potential for enhancing user experience in direct manipulation, are presented in this section to address cognitive issues involved in interacting with virtual information supported by mobile AR.

Tangible Interaction

Integrating tangible interaction with mobile AR interfaces has emerged as a new interaction paradigm in recent years. Rather than relying on specific input devices, users can physically manipulate traditional tools to interact with virtual information. Tangible interaction adapts to people's natural behaviours in daily lives, which in turn contributes to the intuitiveness of manipulation and reduces cognitive loads. Also, it is useful for making full use of rich human physical skills to foster input capabilities in mobile AR.

The effectiveness of using physical tools to perform different functions has a high significance in tangible interaction [88]. It is essential to enhance the richness of functionalities embedded in a single tool to manipulate virtual information. For example, a series of tiles with different functions were designed to support tangible interaction in mobile wearable AR [84]. Users could easily pick up different data tiles and arrange them on a whiteboard to link the virtual information contained by the tiles. Also, they could manipulate operation tiles to implement certain functions on virtual information, such as deleting, copying and help. Through combining input and output functions in tiles, the system made the manipulation simple and natural [84]. Virtual vouchers representing different species and relevant information were developed to support the identification of specimens in field work; users could flip the handle of the voucher and the change of position and orientation of the voucher results in displays of different characteristics of the specimen [116]. Tangible interaction is also expanded to interactively displaying spatial information in outdoor navigation. One such application enabled users to rotate a cube in different directions to browse information and target desired locations, which was more intuitive than using a paper map [78].

The physical capability of handheld devices opens up possibilities for extending tangible interaction in mobile AR. By taking advantage of the camera as an input channel, users are capable of naturally interacting with virtual objects by manipulating and altering the orientation and position of handheld devices [41, 64]. In such

camera-based interaction, multiple functions are assigned to a single mobile phone. Two fundamental types of physical gestures are defined, including "static interaction primitives" and "dynamic interaction primitives" [90]. Static interaction primitives allow users to manipulate the object on the basis of different postures of mobile phone, such as pointing, rotation, tilting and adjusting distance, while dynamic interaction primitives depend on physical movement of mobile phone such as horizontal movement, vertical movement and diagonal movement. The combination and sequence of physical primitives relate to the ease and speed of manipulation. Empirical evidence has been provided to support the notion that tangible interfaces have facilitation effects on the speed of positioning virtual objects, but the advantage on rotating objects is not obvious compared with keypad input [44]. Although this type of tangible interaction could increase the speed of manipulation, the problem of interaction accuracy should not be ignored.

Direct Hand Interaction

In the human body, hands are conceived as the most natural and effective input device to perform direct human–computer interaction [27]. The ubiquitous feature of hands as input devices is able to assist the realization of AR within a wide range of environments. In recent years, there has been an endeavour to exploit the potential of users' two hands in supporting direct manipulation of virtual objects in mobile AR [61, 82].

Direct hand interaction has been adopted as a natural input technique in outdoor wearable AR. Wearing a HMD leaves users' two hands free, which makes it possible for users to use their hands as an intuitive input channel. By taking advantage of vision-tracked pinch gloves, interaction techniques were developed to enable users to directly manipulate virtual information through hand gestures in mobile AR [110]. Tinmith-Hand was known as a glove-based interface that allowed users to control virtual objects and create 3D models of buildings either within or out of the length of arm [82]. Users could rely on their hands to perform cursor operations such as selection, rotation, translation and scaling on 3D objects. For example, they could select certain actions in the menu by pinching their fingers and thumb, transit between different levels of the menu by pressing their finger against their palm and rotate/scale virtual objects by adjusting the distance or angle of both hands. The intuitiveness and ease of use of direct hand interaction in wearable AR have been recognized by a range of research [20, 82, 83].

Handheld devices provide unique possibilities for using hands to directly manipulate virtual objects without employing additional special equipment. When one hand holds the device, the functions of the free hand are needed to implement effective physical interaction [25]. One typical type of interaction is using the hand as a vehicle to communicate commands for manipulating 3D objects in mobile AR. Touch screen interfaces allow more intuitive and effective input modes than buttons, key pads and joysticks [52]. However, there are some limitations to touch screens when interacting with 3D objects. The small screen size makes it difficult to select

objects and causes fingers to occlude the display, and patterns of 2D interaction such as pointing and clicking are not well-suited for manipulating 3D objects [52]. Hence, attempts toward designing alternative interaction techniques based on touch screen interfaces are important in order to enhance the effectiveness of spatial inter- action. For example, in order to tackle the issue of occlusion, *SideSight* was designed to expand the interaction area by allowing users to manipulate virtual contents by multi-"touch" around the mobile phone [15]. It was capable of sensing the action of fingertips in the periphery around the mobile phone. Indeed, a combination of tradi- tional touch interaction and off-screen gesture interaction enriched the interaction experience with the support of mobile phones [15]. Back-of-device interaction was proposed to enhance capabilities of touch-based inputs of handheld devices in mobile AR [61]. It empowered users to engage in interaction with virtual objects by using one hand at the back of display while the other was holding the device, which simplified the spatial interaction in addition to solving the occlusion problem. The efficiency of an interaction technique may vary from task to task and there is a need to match the technique to the action involved in activities [45, 52]. Another type of interaction has been developed to augment virtual objects on the palm or finger of the free hand. Users can directly manipulate 3D objects by changing the motion of the hand anywhere at any time [71, 100]. This approach makes it more convenient for users to realize AR contents with their hands and flexibly choose the viewpoint with which to inspect the content within the length of arm.

Bimanual interaction is characterized by using two hands to simultaneously han- dle one object [34]. Inspired by the conception of bimanual interaction investigated by Guiard [34], two-handed interaction has been explored in mobile AR interaction in order to increase the efficiency of manipulation through coordinating the actions of two hands [25, 96]. Using two-handed interaction shows advantages in fostering the performance of manipulating 3D objects compared to one-handed interaction [44, 46]. However, bimanual interaction cannot ensure better performance, and it is important to optimally assign functions to two hands [47]. In mobile AR supported by handheld devices, the non-dominant hand is usually used to control the view- point, while the dominant hand is used to manipulate virtual objects [14]. The con- current action of two hands contributes to the reduction of shakiness of manipulation on the handheld device, which in turn increases the precision of interaction. Also, since users are empowered to merge two actions to engage in the task, the physical and cognitive efforts generated by alternating across different contexts to succes- sively perform two actions can be reduced [35]. Guimbretière et al. [35] also posited that the smoothness of combining multiple manipulations influences the effective- ness of a hybrid interaction style. Two-handed interaction is also adopted in wearable AR. It is suggested that the accuracy of spatial interaction with objects in 3D such as rotation and scaling can be well executed through specifying the relative position and orientation between two hands [82]. By taking advantage of physical skills of using pen and notebook in daily life, personal interaction panels were created to sup- port two-handed interaction to manipulate AR contents [89, 106]. The integration of a personal interaction panel with a HMD expanded the capability of spatial input in mobile AR by allowing users to select, rotate, drag and drop 3D objects floating in the physical world and alter the viewpoint to obtain desired information [106].

Multimodal Interaction

Multimodal interfaces engage users in interacting with virtual objects through multiple input modalities and/or output modalities in mobile AR. With the support of multimodal interfaces, users' have increased flexibility to choose interaction modes under different situations, which shows potential for increasing the efficiency of manipulation [22].

Integrating complementary multimodality is essential to the efficiency of inter-action [26]. Efforts have been made to combine a class of input modalities on the basis of channels of human perception and communication [101]. This effort has been directed at complementing different modalities well and realizing natural interaction with technologies [101]. Currently, hand gestures and speech are the two main input modes of a range of multimodal interfaces in wearable AR applications [42, 65]. Incorporating speech as a means of input could augment the capability of hand gestures in directly manipulating virtual objects [54]. Gestures serve as an effective medium for carrying spatial information regarding object manipulation (location, movement manner, size), while speech supports commands that are needed to manipulate an object based on the description of its properties, which is truly important when the object is not visible to users' view. Empirical studies have demonstrated that a hybrid use of gestures and speed can positively affect the efficiency of spatial interaction in AR compared with unimodal interaction, because the former addresses the problem of ambiguity when implementing a command [54]. Gaze is also utilized in multimodal interaction to assist in the natural position-ing of AR contents. Gaze input is valuable for promoting the effectiveness of hands-busy activities [5]. Gaze directions and the duration of fixation are assigned as commands to naturally position virtual objects, which can reduce cognitive loads since users do not need to engage in hand-eye coordination in hand-based manipula-tion [26]. The concept of multimodality is adopted in mobile AR supported by handheld devices as well. For example, alternative interaction techniques, beyond standard touch screen interaction, were designed to complement one another and promote input capabilities [15, 52]. So, it is critical to identify the strengths and weaknesses of each modality and appropriately define commands to support sub-tasks in an activity [26].

The combination of output modalities is also a critical aspect to support the interaction with virtual information in mobile AR. Incorporating multiple output modalities can adapt to users' preferences in different types of interaction involved in one activity. Kawsar et al. [59] adopted mobile phones and personal projectors to support manipulations of virtual contents during navigation, and the informa-tion could be displayed on both the mobile phone screen and projector. The per-sonal projection caused the separation of input and output spaces, which then required effortful hand-eye coordination to transition attention across the two spaces in navigation. However, after the user discovering the target in navigational process, the large projection display was well-suited for him/her to manipulate objects with two hands.

3.2.3   Content Creation

User-created content in mobile AR interaction has attracted growing interest in recent years [121]. The user not only receives and manipulates AR contents, but also plays the role of author to produce virtual information in mobile AR. Authoring AR content contributes to increased availability of information and in turn enriches user experiences in mobile AR.

Annotating environments for information sharing is a main type of content creation in mobile AR. Rekimoto et al. [87] brought forward the concept of "augmentable reality" to describe mobile AR applications that empower users to generate virtual information such as textual, graphical and voice annotations and attach them to surrounding environments. Users are also able to communicate situated information with other wearable users and those using normal computers. By taking advantage of content creation, Reitmayr and Schmalstieg [86] developed a wearable AR system to support information creation and sharing among users in tourism. People could annotate their surroundings by adding predefined icons of different shapes and colours, and then share those icons with other participants. Explorations were also carried out to expand content creation in situ for ordinary users in mobile AR [68]. AR 2.0 has been discussed in recent years to highlight the importance of user-generated virtual information within the context of mobile AR [99]. The mobility and low cost of mobile phones make them a suitable choice for being AR authoring platforms [68]. Every person can be an author of AR contents in place, and then publish his/her information to the audience with the support of a mobile phone.

Locating precise 3D positions of in-place objects presents challenges for creating annotations in mobile AR. Wither et al. [120] suggested that the combination of the aerial photograph and the first person perspective view in situ allows users to create annotations in an easy way; the features of corners, edges and regions in the aerial photograph were useful for precisely annotating the scene. The switch of users' attention between the screen display and the physical site to verify the annotation point is a cognitive issue involved in creating contents, especially when labeling small objects [69]. A panoramic image of the surroundings was displayed to make users annotate the environment from the first-person perspective, which promoted the efficiency of locating target position when touching the display [69]. When the annotated object is larger than the size of display, new interaction styles are needed to aid users in identifying the target and creating annotations. A pagination mechanism implemented on mobile phones was developed to help users effectively change visualized objects and target the object to add new comments [74]. When the scene model of the environment is not available, judging the distance of the target object from the user might introduce cognitive burdens [121]. A series of pictorial depth cues were designed to help users determine the distance to the target and accurately annotate the feature [122].

The mobility of users is a concern for designing interaction to create information in mobile AR. The interaction accuracy is more likely to be increased if users are walking around. Touch screen interaction characterized with high intuitiveness and ease of use is commonly adopted in mobile AR [14, 52]. But under the condition

that a user carries a handheld device in a mobile context, the unsteady status of view on the screen makes it hard to precisely interact with AR contents. In order to reduce the errors of annotation, a range of new interaction techniques have been developed [38, 70]. For example, freeze-set-go interaction allows users to freeze the real world scene first, and then add annotations once they are still and in a comfortable pose; users can then unfreeze the view when they finish authoring the content [70].

Sketch-based AR is another branch of applications designed to support content creation in an intuitive and flexible way. It allows users to create visual scenes for AR through sketching on the interface with tools such as a stylus. *Napkin sketch* was an example that assisted creative 3D image drawing on the tablet PC by taking advantage of sketch-based interaction [123]. The capability for supporting the transition to previous frames is essential for users to freely modify the content in the design process. In-place sketching is also applied to support 3D model creation in AR games [39, 53]. With in-place sketching, users are able to sketch game contents based on pre-defined sketching rules and symbols, and then play with those contents. Two users can also sketch game contents and manipulate them alternatively to engage in the game. Sketching itself can be considered playing if the aim of game is to design certain contents together.

3.3 Shared Experience in Mobile AR

In recent years, mobile AR has been applied to facilitate shared experience in multiple domains such as collaborative learning, urban planning, social gaming, and tourist guiding [37, 43, 80]. Mobile AR lets multiple people interact with virtual information while maintaining social dynamics in the real world. From the embodied perspective, the social experience is an essential aspect to influence cognitive processes of constructing meaning when people interacting with technologies [51]. It is necessary to analyze the affordance of mobile AR for supporting cognitive process in multi-user activities. Given the significance of social richness in facilitating human cognition shared experience, key components in social contexts of shared experience are discussed first in this section. Next, three fundamental issues, bodily configuration, artifact manipulation and display space, will be presented to yield insights into how they relate to the establishment of social contexts in shared experience supported by mobile AR.

3.3.1 Social Context

Mobile AR creates new opportunities for enriching collective activities. The social context in shared experience does not simply mean the co-presence of multiple people, but also includes people's social role and their awareness of shared experience in activities [63, 81]. Recently, the capacity of computing technologies to enhance social richness is emphasized when designing shared experience [18].

From this perspective, understanding the social context in shared activities is helpful for effectively applying mobile AR in multi-user experiences.

The presence of multiple social entities is a fundamental component in shared experience. With multiple users in shared activities, mobile AR supports two main types of physical presences: local co-presence and mediated co-presence. AR technology can augment shared experiences by enabling co-located users to view and manipulate virtual information in face-to-face situations indoors or outdoors [118]. Mobile AR can also expand distributed multi-user activities by establishing a virtually shared space. For example, multiple users situated in distributed outdoor contexts could solve a problem through interacting with shared virtual objects [89]. Another example allows an outdoor user to explore and annotate the environment with AR while an indoor user receives on-site information and exchanges ideas with the outdoor user through employing VR [49, 109]. Also, mobile AR enables users to share annotations created in situ with others over distance [69].

Rather than simply gather multiple users in co-located or mediated shared settings, mobile AR needs to be designed to support social roles of multiple users in shared activities [63, 124]. Mobile collaborative AR emerges as an important field in current mobile AR applications [85]. Being collaborators, users are required to continuously build mutual understanding toward a shared goal in joint activities [91]. Given the importance of building mutual understanding in collaboration, the dynamics of social interaction are identified as an indicator for assessing the effectiveness of technology in supporting collaborative activities [66]. The scale of collaboration can vary from context to context. To promote social interactivity in large-scale collaboration, Klopfer, Perry, Squire and Jan have assigned distinct roles to collaborators and delivered customized information to each role [63]. Collaborators, characterized by different roles, needed to share information with each other to jointly perform a task. The enhanced social interdependence strengthened their beliefs as a group and subsequently motivated their commitment to the social interaction. However, the degree of overlap among different roles has to be considered [63]. Too much overlap may weaken interdependence in collaboration, but too little overlap may negatively affect the amount of common ground among collaborators.

Mobile AR also possesses great capabilities to support multiple users as competitors in shared experience. Social aspects, such as types of co-presence of players, communication among players, relationship of players, are stressed in computing games in recent years [21]. It has been suggested that competition among co-located players shows potential for enhanced enjoyment compared to mediated and virtual co-play [32]. Co-located social gaming is a typical form of shared activity that can be facilitated by mobile AR technologies [124]. With respect to mobile AR, it does not only maintain the social dynamics of multi-player games in the real world, but also seamlessly integrates computer-generated contents with the physical setting. To enrich social experience, applications combining competition and collaboration in entertainment supported by mobile AR have been investigated [80]. Social interaction is perceived as a core component in the social context of this type of multiplayer games due to the fact that players in one team need to negotiate strategies to compete against the other team of players.

In addition to the role of multiple users, workspace awareness serves as a key element to address in the social context of shared activities. Workspace awareness refers to "the up-to-the-moment understanding of how other people are interacting with a shared workspace," and impacts the effectiveness of shared activities [36]. Specifically, workspace awareness is comprised of the awareness of other participants' presence, the interaction engaged in by other participants, and the happening of activities within the workspace. Given the interdependence among multiple users, it is important for computing technologies to convey workspace awareness when constructing a space for shared activities. By providing hands-on experience in a shared visual context, AR can present a situational picture and foster workspace awareness in multi-user activities [81].

As emphases are attached to social aspects of shared experience, there is a need to gain an insight into the mechanism underlying the effectiveness of constructing social contexts in shared experience. In the following sections, bodily configuration, artifact manipulation and display space in mobile AR are discussed to illustrate their roles in affecting the effectiveness of shared activities.

3.3.2   Bodily Configuration

Mobile AR facilitates the construction of shared spaces in order to engage multiple users in diverse collaborative and competitive activities in the real world. However, a shared space among multiple users does not guarantee enhancement of social richness [124]. The capacity of mobile AR to support multiple users' bodily configurations, such as location and movement, is a critical issue involved in the establishment of social contexts.

The mobility of users has been recognized as an important aspect in affecting social dynamics in activities [80]. Empirical evidence has shown that input devices enabling natural body movements can encourage communication among users in shared activities [72]. In mobile AR, the mobility around a game board has been greatly investigated [43, 113]. Multiple users sit or stand around the board and simultaneously manipulate AR contents with their own devices in the networked play. The arrangement of the game board influences users' physical movements in the shared space. Xu et al. [124] found that different game board configurations in a shared space enabled users to stimulate different physical and social behaviours during the play. Players adopted physical movements as a strategy to compete for a good position to track the board and perform the task. Also, they adjusted their locations based on the observation of the opponents' movements in the game. The involvement of social interaction and the awareness of one another's actions could be enhanced when the configuration of the game board is appropriately designed. Building a shared space where users can move independently can stimulate explorations in learning activities. For example, Kaufmann and Schmalstieg [57] contended that the physical setup of AR prompts users to walk around the 3D geometric model to obtain different viewpoints for understanding spatial relations and facilitating further construction.

In recent years, mobile AR research has increasingly examined users' mobility in a broad range of physical environments [18]. Rather than simply providing a shared space, the physical environment can serve as an integral element in multi-user activities. With the advantage of portability and mobility, handheld devices supported with AR technology introduce opportunities for users to physically explore the real world. The notion of "using the physical world as a game board" is proposed to highlight the importance of giving meaning to physical locations and movements in game contexts [13]. For example, Mulloni et al. [80] augmented physical locations with virtual characters and settings in game narratives, and players needed to move among different locations to collect and rearrange virtual information in order to solve the mission. Users got involved in social interaction with each other while carrying out physical explorations. Mobile AR introduces innovative components to collaborative field work in outdoor settings [63]. For example, collaborators with HMD were able to freely navigate the environment and exchange ideas in order to conduct investigations [89].

The physical setup of AR interfaces can have effects on organizing users' bodily configuration during activities, which relates to their engagement of social interaction to make decision or find out solutions. With the advantage of wireless connectivity, handheld devices are increasingly utilized as platforms to support collaborative activities. Morrison et al. [79] adopted handheld devices to augment a paper-based map to guide users' explorations in the physical world. The collaborators supported with a combination of AR systems and physical tools tended to gather around the device and discussed strategies together, compared to those supported with only a traditional 2D digital map. The enhanced joint attention positively affected the performance of problem solving within a group [6].

3.3.3 Artifact Manipulation

With the advance of interaction techniques in mobile AR, there are increasing numbers of endeavours to incorporate manipulation of AR contents into applications targeting multi-user activities [12, 89]. Designing physical interaction with virtual contents has become an important aspect in supporting the effectiveness of collective activities.

Cognitively, the manipulative artifact serves as a common ground for multiple users to negotiate meaning in shared activities. Direct manipulation of virtual objects, characterized by deep bodily involvement in creating and modifying shared artifacts, motivates users to jointly engage in explorations to complete the group task [58]. The hands-on experience contributes to the accumulation of common ground among multiple users in the ongoing process. Instead of artifacts being isolated from the interactional process, shared artifacts can shape social interaction patterns of users [104]. Additionally, interacting with artifacts offers an important source for promoting workspace awareness among participants [37]. Manipulative actions can publicly signal the behaviour performed by others, which in turn can exert influence on the effectiveness of group work.

Allowing users to manipulate virtual information independently is important to stimulate their participation in shared activities. Independence, characterized by users manipulating virtual objects and changing the viewpoint of objects individually, is recognized as an indispensible aspect in AR-supported collaboration [107]. Users are capable of interacting with virtual information based on their interests and knowledge, which stimulates their participation in activities. Also, enabling individual users to simultaneously interact with virtual objects shows promises for increasing the effectiveness of collaboration [111].

Some research has focused on exploiting interaction techniques to enhance shared experience in mobile AR [12, 77]. Tangible interaction is identified as an effective approach to assist in the manipulation of virtual objects in multi-user activities [12, 84]. Physical controllers with different representations provide a common ground that contributes to the establishment and sustainment of mutual understanding among multiple users. Tangible interaction characterized by high intuitiveness can minimize the distraction of action when participants are engaging in discussion [95]. Multimodal interaction has been investigated to strengthen the effectiveness of shared activities as well [5, 89]. For example, *MAGIC*, a mobile collaborative AR system for exploring archaeological sites, allows users to utilize both text and speech to post messages to their partners regarding actions or ideas; the messages can then be shared with co-located or distant users on the HMD [89]. Some researchers have examined gaze direction as a complimentary channel to coordinate collaboration with the support of AR technology [5]. They found that gaze-based interaction is especially useful for facilitating joint attentions in the interactional process to construct shared understanding in remote collaborative activities.

3.3.4 Display Space

The arrangement of output displays in mobile AR can also influence social interaction patterns and workspace awareness in multi-user activities. The manner of presenting information about users' interaction and its subsequent effects on group work should be addressed when designing shared experience supported by mobile AR [105].

Constructing a shared display space is essential to the effectiveness of multi-user activities. The shared display of information serves as a common focus for users, and has facilitation effects on problem solving by stimulating social interaction [94]. Sharing the interface among multiple users is widely applied in collaborative activities. For example, a shared on-site map was displayed on the HMD of individual users to coordinate the interaction among them in collaboration [89]. Projectors have been studied with regards to establishing a public display of information in AR-supported collaboration [11]. The small display screen on handheld devices is not a good option for presenting shared information. Recognizing projectors' advantage in expanding the public display space, some researchers have applied the projector phone to mobile AR, enabling users to flexibly view information on either their mobile phone screen or projection [40, 59]. People thought that

the projection was suitable for collaborative activities because it was convenient for them to share ideas with each other around the shared display [59].

Also, as more interaction techniques are incorporated into multi-user activities in mobile AR, publicly informing the occurrence and progress of one's action to others is vital to foster situational awareness among users. Other participants' behaviors help individuals to adjust their own competitive strategies or offer elaborations for jointly solving the problem [43]. One limitation of applying AR in collaboration is that the personal information display exerts some negative effects on users' awareness of others' actions, which can make it more difficult for them to achieve mutual understanding of each other's manipulations of virtual information [114]. In order to enhance situational awareness, some research has taken advantage of multiple sensory feedbacks to indicate individual's behaviours to group members [43, 57]. Empirical evidence has revealed that audio feedback is a useful medium to raise users' awareness of actions performed by others [43]. In *Construct 3D*, colour schemes were applied to distinguish the contributions made by different users [57]. Facilitating the visibility of each other's interactions is important to distributed collaboration. The lack of physical presence can result in reduced visual cues and nonverbal interaction, which can then reduce the awareness of each other's actions and affect the construction of shared understanding. Thomas and Piekarski [109] adopted VR as a channel to provide the representation of the outdoor user's environment to enhance indoor user's awareness of his/her outdoor partners' action and context. Also, the connection of outdoor AR and indoor VR allowed users to engage in interaction over distance simultaneously.

Considerations should be also given to the personal display in AR-supported shared experience. Privacy is a critical issue in configuring spaces for multi-user activities [16]. In the context of collaboration, although people are required to work together to solve the problem, they still need to maintain individuality and engage in personal activities [36]. People sometimes expect to keep certain work or reflection private rather than sharing all information with others. With respect to social gaming of a competitive nature, participants should possess personal spaces to engage in individual actions in order to win the game [107]. Hence, it is necessary to develop hybrid interfaces to support both public and personal displays for multi-user activities in mobile AR.

So far, there have been some explorations on combining public and personal displays to enhance the effectiveness of shared activities supported by AR technology [30, 56]. For example, in *STUDIERSTUBE*, users can customize the view of scientific visualization to meet their own needs and exert self-control on whether to publicly display it to their collaborators or not [30]. A check in/out model was proposed to enable users to better perform collaborative and strategic work while collaborating remotely [56]. In this application, an augmented workspace is divided into two spaces: a public space and a private space. The user can perform actions in the private space if he/she wants to hide it from others. Additional personal interaction panels are integrated to construct a personal space for individual interaction. PDAs have been applied as platforms to allow users to make personal notes in addition to controlling the virtual object in AR [19]. In one study, people were allowed

to view and manipulate virtual information privately through a personal interaction panel and choose when to display their work on a shared projection for further discussion [98]. Multi-player games have also used the personal interaction panel to keep individual player's actions invisible to their opponents [105].

4 Conclusions

Mobile AR is identified as a promising interface to support individual's direct interactions with technologies bound to physical and social environments. The close connection between computational resources and the real world creates many opportunities for users to actively explore their physical space. As mobile computing platforms advance the implementation of AR in diverse domains, designing effective mobile AR systems has become an important part in fulfilling the potential of AR technology to foster users' cognitive functioning.

Recognizing the significance of users' behavioural involvement in constructing meaning and understanding of mobile AR environments, this chapter approaches cognitive issues of mobile AR from an embodied perspective to examine how the involvement impacts cognitive functioning. Three primary cognitive issues are identified, which include information presentation, physical interaction and shared experience. A variety of issues involved in each aspect are addressed along with examples of existing mobile AR applications. Fostering a better understanding of cognitive issues is important in order to guide the design of mobile AR systems to enhance human cognitive functioning.

The above review suggests that cognitive issues are crucial to the effectiveness of mobile AR systems. Analyzing human factors from the lens of mobile AR interaction is helpful for yielding an insight into the opportunities and challenges for developing effective mobile AR systems. Furthermore, the influence of mobile AR on strengthening human cognitive process is context-dependent. Making a good match for the context should be taken into consideration when utilizing mobile AR technologies. In the future, more research is needed to evaluate the effectiveness of mobile AR systems from this embodied perspective and apply findings for improving the system design.

References

1. Aaltonen A, Lehikoinen J (2006) Exploring augmented reality visualization. In Proceedings of AVI 2006, Venezia, Italy, 23–25 May
2. Alessandro M, Dünser A, Schmalstieg D (2010) Zooming interfaces for augmented reality browsers. In Proceedings of MobileHCI 2010, Lisboa, Portugal, 7–10 September

3. Azuma R, Furmanski C (2003) Evaluating label placement for augmented reality view management. In Proceedings of ISMAR 2003, Tokyo, Japan, 7–10 October
4. Bane R, Höllerer T (2004) Interactive tools for virtual x-ray vision in mobile augmented reality. In Proceedings of ISMAR 2004, Arlington, VA, 2–5 November
5. Barakonyi I, Prendinger H, Schmalstieg D, Ishizuka M (2007) Cascading hand and eye movement for augmented reality videoconferencing. In Proceedings of 3DUI 2007, Charlotte, NC, USA, 14–17 March
6. Barron B (2003) When smart groups fail. JLS 12: 307–359. doi: 10.1207/S15327809JLS1203_1
7. Baudisch P, Good N, Stewart P (2001) Focus plus context screens: Combining display technology with visualization techniques. In Proceedings of UIST 2001, Orlando, FL, 11–14 November
8. Bell B, Feiner S, Höllerer T (2001) View management for virtual and augmented reality. In Proceedings of UIST 2001, 11–14 November
9. Bell B, Feiner S, Höllerer T (2002) Information at a glance. Computer Graphics and Applications 22: 6–9. doi: 10.1109/MCG.2002.1016691
10. Billinghurst M, Bowskill J, Dyer N, Morphett J (1998) An evaluation of wearable information spaces. In Proceedings of VRAIS 98, Atlanta, GA, 14–18 March
11. Billinghurst M, Kato H, Kiyokawa K, Belcher D, Poupyre I (2002) Experiments with face-to-face collaborative AR interfaces. Virtual Reality 6: 107–121
12. Billinghurst M, Poupyrev I, Kato H, May R (2000) Mixing realities in shared space: An augmented reality interface for collaborative computing. In Proceedings of ICME 2000, New York, NY, USA, 30 July −2 August
13. Björk S, Falk J, Hansson R, Ljungstrand P (2001) Pirates! Using the physical world as a game board. In Proceedings of Interact 2001, Tokyo, Japan, 9–13 July
14. Boring S, Baur D, Butz A, Gustafson S, Baudisch P (2010) Touch projector: Mobile interaction through video. In Proceedings of CHI 2010, Atlanta, Georgia, USA, 10–15 April
15. Butler A, Izadi S, Hodges S (2008) SideSight: Multi-"touch" interaction around small devices. In Proceedings of UIST 2008, Monterey, CA, 19–22 October
16. Butz A, Höllerer T, Beshers C, Feiner S, MacIntyre M (1999) An experimental hybrid user interface for collaboration. Technical Report CUCS-005-99, Columbia University, Department of Computer Science
17. Carpendale MST, Cowperthwaite DJ, Fracchia FD (1997) Making distortions comprehensible. In Proceedings of VL. 1997, Capri, Italy, 11–18 September
18. Cheok AD, Yang X, Ying ZZ, Billinghurst M, Kato H (2002) Touch-Space: Mixed reality game space based on ubiquitous, tangible, and social computing. Pers Ubiquit Comput 6: 430–442
19. Clark A (2004) Hybrid AR users interfaces in collaborative gaming. Honours Reports, Computer Science and Software Engineering, University of Canterbury
20. Costello A, Tang A (2007) An egocentric augmented reality interface for spatial information management in crisis response situations. In Proceedings of ICVR 2007, Beijing, China, 22–27 July
21. De Kort YAW, Ijsselsteijn WA (2008) People, places, and play: Player experience in a socio-spatial context. Computers in Entertainment 6: 1–11. doi: 10.1145/1371216.1371221
22. Dias MS, Bastos R, Fernandes J, Tavares J, Santos P (2009) Using hand gesture and speech in a multimodal augmented reality environment. Lecture Notes in Computer Science, Volume 5085/2009:175–180, doi: 10.1007/978-3-540-92865-2_18
23. Dourish P (2001) Where the action is. The MIT Press, Cambridge, MA, USA
24. Duh HBL, Ma J, Billinghurst M (2005). Human factors issues in augmented reality. In: W. Karwowski (eds.). International Encyclopedia of Ergonomics and Human Factors. London: Taylor & Francis.
25. Edge D, Blackwell A F (2009) Bimanual tangible interaction with mobile phone. In Proceedings of TEI'09, Cambridge, UK, 16–18 February
26. Elepfandt M, Sünderhauf M (2011) Multimodal, touchless interaction in spatial augmented reality environments. In Proceedings of the HCII, Orlando, FL, USA, 9–14 Jul

27. Erol A, Bebis G, Nicolescu M, Boyle RD, Twombly X (2007) Vision-based hand pose esti-
    mation: A review. Computer Vision and Image Understanding 108: 52–73. doi: 10.1016/j.
    cviu. 2006. 10. 012
28. Fikkert W, D'Ambros M, Bierz T, Jankun-Kelly TJ (2007) Interacting with visualizations. In
    Kerren et al. (eds). Human-Centered Visualization Environments (pp. 77–162), Springer-
    Verlag Berlin Heidelberg
29. Froehlich P, Obernberger G, Simon R, Reichl P (2008) Exploring the design space of smart
    horizons. In Proceedings of MobileHCI 2008, Amsterdam, the Netherlands, 2–5 September
30. Fuhrmann A, Löffelmann H, Schmalstieg D (1997) Collaborative augmented reality:
    Exploring dynamical systems. In Proceedings of IEEE Visualization '97, Phoenix, AZ, 19–24
    October
31. Furmanski C, Azuma R, Daily M (2002). Augmented-reality visualizations guided by cogni-
    tion: Perceptual heuristics for combining visible and obscured information. In Proceedings of
    ISMAR 2002, Darmstadt, Germany, 30 September-1 October
32. Gajadhar B, De Kort YAW, Ijsselsteijn WA (2008) Shared fun is doubled fun: Player enjoy-
    ment as a function of social setting. In: Markopoulos P, de Ruyter B, Ijsselsteijn W, Rowland
    D. (eds.), Fun and Games (pp. 106–117). New York: Springer.
33. Ganapathy S, Anderson GJ, Kozintsev IV (2011) Empirical evaluation of augmented presen-
    tation on small form factors-navigation assistant scenario. In Proceedings of ISVRI 2011,
    Singapore, 19–20 March
34. Guiard Y (1987) Asymmetric division of labor in human skilled bimanual action: The kine-
    matic chain as a model. J. Mot. Behav., 19: 486–517
35. Guimbretière F, Martin A, Winograd T (2005). Benefits of merging command selection and
    direct manipulation. ACM Transactions on Computer-Human Interaction 12: 460–476
36. Gutwin C, Greenberg S (1998) Design for individuals, design for groups: Tradeoffs between
    power and workspace awareness. In Proceedings of CSCW 1998, Seattle, WA, USA, 14–18
    November
37. Gutwin C, Greenberg S (2001) The importance of awareness for team cognition in distributed
    collaboration. Technical Report 2001-696-19, Dept Computer Science, University of Calgary,
    Alberta, Canada.
38. Güven S, Feiner S, Oda O (2006) Mobile augmented reality interaction techniques for author-
    ing situated media on-site. In Proceedings of ISMAR 2006, Santa Barbara, CA, USA, 22–25
    October
39. Hagbi N, Grasset R, Bergig O, Billinghurst M, El-Sana G (2010) In-place sketching for
    content authoring in augmented reality games. In Proceedings of VR' 2010, Waltham,
    Massachusetts, USA, 20–24 March
40. Hang A, Rukzio E, Greaves A (2008) Projector phone: A study of using mobile phones with
    integrated projector for interaction with maps. In Proceedings of MobileHCI 2008,
    Amsterdam, the Netherlands, 2–5 September
41. Harviainen T, Korkalo O, Woodward C (2009) Camera-based interactions for augmented
    reality. In Proceedings of ACE2009, Athens, Greece, 29–31 October
42. Heidemann G, Bax I, Bekel H (2004) Multimodal interaction in an augmented reality sce-
    nario. In Proceedings of ICMI. 2004, State College, PA, USA, 13–15 October
43. Henrysson A, Billinghurst M, Ollila M (2005a) Face to face collaborative AR on mobile
    phones. In Proceedings of ISMAR 2005, Vienna, Austria, 5–8 October
44. Henrysson A, Billinghurst M, Ollila M (2005b) Virtual object manipulation using a mobile
    phone. In Proceedings of ICAT 2005, Christchurch, New Zealand, 5–8 December
45. Henrysson A, Marshall J, Billinghurstz M (2007) Experiments in 3D interaction for mobile
    phone AR. In Proceedings of Graphite 2007, Perth, Western Australia, 1–4 December
46. Hinckley K, Pausch P, Goble JC, Kassell NF (1994) A survey of design issues in spatial input.
    In Proceedings of UIST '94, Marina del Ray, California, USA, 2–4 November
47. Hinckley K, Paysch R, Proffitt D, Kassell NF (1998) Two-handed virtual manipulation. ACM
    Transactions on Computer-Human Interaction 5: 260–302
48. Höllerer T, Feiner S, Hallaway D, Bell B (2001) User interface management techniques for
    collaborative mobile augmented reality. Computers & Graphics, 25: 799–810

49. Höllerer T, Feiner S, Terauchi T, Rashid G, Hallaway D (1999) Exploring MARS: Developing indoor and outdoor user interfaces to a mobile augmented reality system. Computers & Graphics 23: 779–785
50. Hwang S, Jo H, Ryu J.-H (2010) EXMAR: EXpanded view of mobile augmented reality. In Proceedings of ISMAR 2010, Seoul, South Korea, 13–16 October
51. Hurtienne J (2009) Cognition in HCI: An ongoing story. An Interdisciplinary Journal on Humans in ICT Environments 5: 12–28
52. Hürst W, van Wezel C (2011) Multimodal interaction concepts for mobile augmented reality applications. In Proceedings of MMM 2011, Taipei, Taiwan, 5–7 January
53. Huynh DT, Raveendran K, Xu Y, Spreen K, MacIntyre B. (2009) Art of defense: A collaborative handheld augmented reality board game. In Proceedings of ACM SIGGRAPH Symposium on Video Games, New York, NY, USA, 3–7 August
54. Irawati S, Green S, Billinghurst M, Duenser A, Ko H (2006) An evaluation of an augmented reality multimodal interface using speech and paddle gestures. In Proceedings of ICAT 2006, Hangzhou, China, 29 November-1 December
55. Kalkofen D, Mendez E, Schmalstieg D (2007) Interactive focus and context visualization for augmented reality. In Proceedings of ISMAR 2007, Nara, Japan, 13–16 November
56. Kame G, Matsuyama T, Okada K (2010) Augmentation of check in/out model for remote collaboration with mixed reality. In Proceedings of ISMAR 2010, Seoul, South Korea, 13–16 October
57. Kaufmann H, Schmalstieg D (2006) Designing immersive virtual reality for geometry education. In Proceedings of VR' 06, Alexandria, Virginia, USA, 25–29 March
58. Kaufmann H, Dünser A (2007) Summary of usability evaluations of an educational augmented reality application. HCI 14: 660–669
59. Kawsar F, Rukzio E, Kortuem G (2010) An explorative comparison of magic lens and personal projection for interacting with smart objects. In Proceedings of MobileHCI 2010, Lisboa, Portugal, 7–10 September
60. Kim SJ, Dey AK (2009) Simulated augmented reality windshield display as a cognitive mapping aid for elder driver navigation. In Proceedings of CHI 2009, Boston, Massachusetts, USA, 4–9 April
61. Kim SW, Treskunov A, Marti S (2011) Drive: Directly reaching into virtual environment with bare hand manipulation behind mobile display. In Proceedings of 3DUI 2011, Singapore, 19–20 March
62. Klemmer SR, Hartmann B, Takayama L (2006) How bodies matter: Five themes for interaction design. In Proceedings of DIS 2006, University of Park, Pennsylvania, USA, 26–28 June
63. Klopfer E, Perry J, Squire K, Jan M.-F (2005) Collaborative learning through augmented reality role playing. In Proceedings of CSCL 2005, Taipei, Taiwan, 30 May-4 June
64. Koh RKC, Duh HBL, Gu J (2010) An integrated design flow in user interface and interaction for enhancing mobile AR gaming experiences. In Proceedings of ISMAR 2010, Seoul, South Korea, 13–16 October
65. Kölsch M, Bane R, Höllerer T, Turk M (2006) Multimodal interaction with a wearable augmented reality system. IEEE Computer Graphics and Applications 3: 62–71
66. Kreijns K, Kirschner PA, Jochems W (2002) The sociability of computer-supported collaborative learning environments. Journal of Education Technology & Society 5: 8–25
67. Kruijff E, Swan JE, Feiner S (2010) Perceptual issues in augmented reality revisited. In Proceedings of ISMAR 2010, Seoul, South Korea, 13–16 October
68. Langlotz T, Mooslechner S, Zollmann S, Degendorfer C, Reitmayr G, Schmalstieg D (2011) Sketching up the world: in situ authoring for mobile augmented reality. Pers Ubiquit Comput 1–8. doi: 10.1007/s00779-011-0430-0
69. Langlotz T, Wagner D, Mulloni A, Schmalstieg D (2010) Online creation of panoramic Augmented reality annotations on mobile phones. IEEE Pervasive Computing 99. doi:10.1109/MPRV.2010.69
70. Lee GA, Yang U, Kim Y, Jo D, Kim K.-H, Kim JH, Choi JS (2009) Freeze-set-go Interaction method for handheld mobile augmented reality environments. In Proceedings of VRST 2009, Kyoto, Japan, 18–29 November

71. Lee T, Höllerer T (2007) Handy AR: Markerless inspection of augmented reality objects using fingertip tracking. In Proceedings of ISWC 2007, Busan, Korea, 11–15 November

72. Lindley S, Le Couteur J, Bianchi-Berthouze (2008) Stirring up experience through movement in game play: Effects on engagement and social behaviour. In Proceedings of CHI 2008, Florence, Italy, 5–10 April

73. Livingston M. A., Ai Z., Karsch K., Gibson G. O. (2011) User interface design for military AR application. Virtual Reality, 15: 175–184. doi: 10.1007/s10055-010-0179-1

74. Lucia A, Francese R, Passero I, Tortora G (2010) SmartBuilding: A people-to-people-to-geographical-places mobile system based on augmented reality. In Proceedings of UBICOMM 2010, Florence, Italy, 25–30 October

75. Makita K, Kanbara M, Yokoya N (2009) View management of annotations for wearable augmented reality. In Proceedings of ICME 2009, New York, NY, USA, 28 June-2 July

76. Milgram P, Kishino F (1994). A taxonomy of mixed reality visual displays. IEICE Transactions on Information and Systems, E77-D(12): 1321–1329

77. Mogilev D, Kiyokawa K, Billinghurst M, Pair J (2002) AR pad: An interface for face-to-face AR collaboration. In Proceedings of CHI 2002, Minneapolis, MN, USA, 20–25 April

78. Moore A, Regenbrecht H (2005) The tangible augmented street map. In Proceedings of ICAT 2005, Christchurch, New Zealand, 5–8 December

79. Morrison A, Oulasvirta A, Peltonen P, Lemmelä S, Jacucci G, Reitmayr G, Näsänen J, Juustila A (2009). Like bees around the hive: A comparative study of a mobile augmented reality map. In Proceedings of CHI 2009, Boston, MA, USA, 4–9 April

80. Mulloni A, Wagner D, Schmalstieg D (2008) Mobility and social interaction as core game-play elements in multi-player augmented reality. In Proceedings of DIMEA 2008, Athens, Greece, 10–12 September

81. Nilsson S, Johansson B, Jönsson A (2009) Using AR to support cross-organisational collaboration in dynamic tasks. In Proceedings of ISMAR 2009, Orlando, Florida, USA, 19–22 October

82. Piekarski W, Thomas BH (2001) Tinmith-Metro: New outdoor techniques for creating city models with an augmented reality wearable computer. In Proceedings of ISWC 2001, Zurich, Switzerland, 8–9 October

83. Piekarski W, Thomas BH (2003) ThumbsUp: Integrated command and pointer interactions for mobile outdoor augmented reality systems. In Proceedings of HCII 2003, Crete, Greece, 22–27 June

84. Poupyrev I, Tan DS, Billinghurst M, Kato H, Regenbrecht H, Tetsutani N (2002) Developing a generic augmented reality interface. Computer 35: 44–50. doi: 10.1109/2.989929

85. Reitmayr G, Schmalstieg D (2001) Mobile collaborative augmented reality. In Proceeding of ISAR 2001, NewYork, USA, 29–30 October

86. Reitmayr G, Schmalstieg D (2004) Scalable techniques for collaborative outdoor augmented reality. In Proceedings of ISMAR 2004, Arlington, VA, USA, 2–5 November

87. Rekimoto J, Ayatsuka Y, Hayashi K (1998) Augmentable reality: Situated communication through physical and digital spaces. In Proceedings of ISWC 1998, Pittsburgh, Pennsylvania, USA, 19–20 October

88. Rekimoto J, Sciammarella E (2000) ToolStone: Effective use of the physical manipulation vocabularies of input devices. In Proceedings of UIST 2000, San Diego, California, USA, 5–8 November

89. Renevier P, Nigay L (2001) Mobile collaborative augmented reality: the augmented stroll. In Proceedings of EHCI 2001, Toronto, Canada, 11–13 May

90. Rohs M, Zweifel P (2005) A conceptual framework for camera phone-based interaction techniques. In Proceeding of Pervasive 2005, Munich, Germany, 29–30 May

91. Roschelle J, Teasley SD (1995) The construction of shared knowledge in collaborative problem solving. In: O'Malley C. (ed.), Computer supported collaborative learning (pp. 67–97). Berlin, Germany: Springer

92. Sandor C, Cunningham A, Dey A, Mattila V.-V. (2002). An augmented reality X-ray system based on visual saliency. In Proceedings of ISMAR 2010, Seoul, Korea, 13–16 October

93. Sandor C, Cunningham A, Eck U, Urquhart D, Jarvis G, Dey A, Barbier S, Marner MR, Rhee S. (2009). Egocentric space-distorting visualizations for rapid environment exploration in mobile mixed reality. In Proceedings of VR 2009, Lafayette, Louisiana, USA, 14–18 March

94. Sareika M, Schmalstieg D (2007) Urban sketcher: Mixed reality on site for urban planning and architecture. In Proceedings of ISMAR 2007, Nara, Japan, 13–16 November

95. Sareika M, Schmalstieg D (2008) Urban sketcher: Mixing realities in the urban planning and design process. In Proceedings of CHI 2008. Florence, Italy, 5–10 April

96. Sareika M, Schmalstieg D (2010) Bimanual handheld mixed reality interfaces for urban planning. In Proceedings of AVI 2010, Rome, Italy, 26–28 May

97. Schmalstieg D, Fuhrmann A, Hesina G (2000) Bridging multiple user interface dimensions with augmented reality. In Proceedings of ISAR 2000, Munich, Germany, 5–6 October

98. Schmalstieg D, Langlotz T, Billinghurst M (2011) Augmented reality 2.0. In: Coquillart et al. (eds.), Virtual Realities (pp. 13–37), Springer-Verlag/Wien

99. Schöning J, Rohs M, Kratz S, Löchtefeld M, Krüger A (2009) Map torchlight: A mobile augmented reality camera projector unit. In Proceedings of CHI 2009, Boston, MA, USA, 4–9 April

100. Siltanen S, Hakkarainen M, Korkalo O, Salonen T, Sääski J, Woodward C, Kannetis T, Perakakis M, Potamianos A (2007). Multimodal user interface for augmented assembly. In Proceedings of MMSP 2007, Chania, Crete, Greece, 1–3 October

101. Seo B-K, Choi J, Han J-H, Park H, Park J-I (2008) One-handed interaction with augmented virtual objects on mobile devices. In Proceedings of VRCAI 2008, Singapore, 8–9 December

102. Stedmon AW, Kalawsky RS, Hill K, Cook CA (1999) Old theories, new technologies: Cumulative clutter effects using augmented reality. In Proceedings of IV 1999, San Francisco, USA, 24–29 October

103. Sukan M, Feiner S (2010) SnapAR: Storing snapshots for quick viewpoint switching in handheld augmented reality. In Proceedings of ISMAR 2010, Seoul, South Korea, 13–16 October

104. Suthers DD (2006) Technology affordances for intersubjective mean making: A research agenda for CSCL. Interactional Journal of Computer-supported Collaborative Learning 1: 315–337

105. Szalavári Z, Eckstein E, Gervautz M (1998) Collaborative Gaming in Augmented Reality. In Proceedings of VRST 1998, Taipei, Taiwan, 2–5 November

106. Szalavári Z, Michael M (1997) The personal interaction panel-a two-handed interface for augmented reality. In Proceedings of EUROGRAPHICS 1997, Budapest, Hungary, 4–8 September

107. Szalavári Z, Schmalstieg S, Fuhrmann A, Gervautz M (1998) Studierstube-An environment for collaboration in augmented reality. Virtual Reality 3: 37–48. doi: 10.1007/BF01409796

108. Tönnis M, Klein L, Dey A, Klinker G (2008). Perception thresholds for augmented reality navigation schemes in large distances. In Proceedings of ISMAR 2008, Cambridge, UK, 15–18 September

109. Thomas BH, Piekarski W (2002) Glove based user interaction techniques for augmented reality in an outdoor environment. Virtual Reality 6: 167–180

110. Thomas B H, Piekarski W (2009) Through-walls collaboration. Pervasive Computing 8: 42–49.doi:10.1109/MPRV.2009.59

111. Tsuda T, Yamamoto H, Kameda Y, Ohta Y (2005) Visualization methods for outdoor see-through vision. In Proceedings of ICAT 2005, Christchurch, New Zealand, 5–8 December

112. Veas E, Mulloni A, Kruijff E, Regenbrecht H, Schmalstieg D (2010) Techniques for view transition in multi-camera outdoor environments. In Proceedings of GI 2010, Ottawa, Canada, 31 May-2 June

113. Wagner D, Pintaric T, Ledermann F, Schmalstieg D (2005) Towards massively multi-user augmented reality on handheld devices. In Proceedings of PERVASIVE 2005, Munich, Germany, 8–13 May

114. Wagner D, Schmalstieg D, Billinghurst M (2006) Handheld AR for collaborative edutainment. In Proceedings of ICAT 2006, Hangzhou, China, 29 November-1 December
115. White S, Feiner S (2009) SiteLens: Situated visualization techniques for urban site visits. In Proceedings of CHI 2009, Boston, MA, USA, 4–9 April
116. White S, Feiner S, Kopylec J (2006) Virtual vouchers: Prototyping a mobile augmented reality user interface for botanical species identification. In Proceedings of 3DUI 2006, Alexandria, VA, USA, 25–26 March
117. White S, Feng D, Feiner S (2009) Interaction and presentation techniques for shake menus in tangible augmented reality. In Proceedings of ISMAR 2009, Orlando, Florida, USA, 19–22 October
118. Wichert R (2002) A mobile augmented reality environment for collaborative learning and training. In Proceedings of E-learn 2002, Montreal, Canada, 15–19 October
119. Wilson M (2002) Six views of embodied cognition. Psychonomic Bulletin & Review, 9: 625–636
120. Wither J, Diverdi S, Höllerer T (2006) Using aerial photographs for improved mobile AR annotation. In Proceedings of ISMAR 2006, Santa Barbara, CA, USA, 22–25 October
121. Wither J, Diverdi S, Höllerer T (2009) Annotation in outdoor augmented reality. Computers & Graphics, 33: 679–689. doi: 10.1016/j.cag.2009.06.001
122. Wither J, Höllerer T (2005) Pictorial depth cues for outdoor augmented reality. In Proceedings of ISWC 2005, Galway, Ireland, 6–10 November
123. Xin M, Sharlin E, Sousa MC (2008) Napkin sketch-handheld mixed reality 3D sketching. In Proceedings of VRST 2008, Bordeaux, France, 27–29 October
124. Xu Y, Gandy M, Deen S, Schrank B, Spreen K, Gorbsky M, White T, Barba E, Radu I, Bolter J, MacIntyre B (2008) BragFish: Exploring physical and social interaction in co-located handheld augmented reality games. In Proceedings of ACE 2008, Yokohama, Japan, 3–5 December
125. Zhang F, Sun H (2005) Dynamic labeling management in virtual and augmented environments. In Proceedings of CAD/CG 2005, Hong Kong, China, 7–10 December
126. Zhou F, Duh HBL, Billinghurst M (2008) Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. In Proceedings of ISMAR 2008, Cambridge, UK, 15–18 September