International Conference on Artificial Reality and Telexistence
Eurographics Symposium on Virtual Environments (2018)
G. Bruder, S. Cobb, and S. Yoshimoto (Editors)

# A Study on AR Authoring using Mobile Devices for Educators

Kinfung Chu[†1], Weiquan Lu[‡2], Kiyoshi Oka[§1], Kazuki Takashima[1] and Yoshifumi Kitamura[1]

[1]Tohoku University, Research Institute of Electrical Communication, Japan
[2]National University of Singapore, Singapore

**Abstract**

*Augmented Reality (AR) on consumer devices is now commonplace and it finds application in areas like online retail and gaming. Among which, school education can especially benefit from the interactivity and expressiveness provided by AR technology, facilitating the learning process of students. Although AR-enabled hardware and applications are becoming increasingly accessible to both students and teachers, the entry requirement for AR authoring is still prohibitively high for school teachers. Given the vast variation in the learning ability of students and school curricula, an AR authoring tool that allows the rapid and easy creation of educational content seems to be very desirable among teachers. This paper proposes a gesture-based control method that satisfies the need of educational AR authoring and presents prototypes that work well with smartphone VR head mounts. Through user studies we show that our proposed control method is simple but effective for basic authoring tasks. Our prototypes are also found to be useful in teaching different concepts that require a high degree of spatial comprehension.*

**CCS Concepts**

• ***Computing methodologies*** → ***Mixed / augmented reality;*** • ***Applied computing*** → *Interactive learning environments;*

## 1. Introduction

There is evidence that AR can boost learning across various school subjects. AR effectively conveys 3D information by superimposing virtual objects on the real world, helping students understand 3D geometry in engineering and mathematics courses [CCHK11, KS02]. Well-designed outdoor AR content encourages students to learn the social aspect of the scientific facts in textbooks [SK07]. In general, AR can motivate students to learn by drawing students' attention and enlivening the learning process [DSInK13]. In spite of the merits of AR in teaching, adoption by educators is slow according to our survey conducted at an education technology conference [NUS17]. Out of the 19 educators who answered our survey, only one was using AR or VR in teaching. One possible reason for the slow adoption is the lack of freely available AR teaching materials for the vast variety of school subjects. Even though educational AR content has the potential to flourish in the future, teachers will always need a way to tailor the content for their own unique teaching methods. In the existing education approach, teachers often tend to create their own teaching notes and slides to suit their teaching style in concert with the learning ability of their students. Likewise, it is typical that teachers will need to create AR content from scratch or customize teaching materials based on available AR content. Furthermore, AR has the flexibility to adapt content to the condition of the physical environment, empowering teachers to design specific AR teaching materials for field trips.

Unfortunately, many existing commercial AR authoring tools have been unable to meet the needs of teachers for rapid creation of educational content. Commercial AR applications are mostly developed using game engines or AR programming frameworks, both of which require a plethora of domain-specific knowledge ranging from programming to computer graphics as prerequisites. The time teachers can afford to learn AR authoring and to create content is limited. It is rarely practical for professional educators to invest time and effort into the learning of AR authoring, especially when there are other pressing academic and operational demands on their time. Although authoring tools targeting non-tech savvy users such as ZapWorks™and Blippbuilder™exist, they still have complicated interfaces and work processes and require the lengthy develop-build-deploy cycle for testing on mobile devices. This problem intensifies when users develop by trial-and-error, an approach often adopted by amateur programmers. Unlike ordinary mobile applications, debugging of AR applications cannot be effectively performed on emulators, since it may be unintuitive if the video with superimposed content appears on the computer screen but not on the screen of the device that captures the video. As AR authoring poses unique challenges, a novel development methodology that caters to the need of non-tech savvy teachers appears to be necessary.

Inspired by Reality-Based Interaction (RBI) [JGH*08], we created an AR authoring tool that addresses the aforementioned prob-

† kennychu@riec.tohoku.ac.jp
‡ lu.weiquan@gmail.com
§ oka.kiyoshi@icloud.com

© 2018 The Author(s)
Eurographics Proceedings © 2018 The Eurographics Association.

delivered by EUROGRAPHICS DIGITAL LIBRARY www.eg.org diglib.eg.org

DOI: 10.2312/egve.20181331

lems. In terms of the gulf of execution and evaluation [ND86], the translations between Task goals and System goals are non-trivial. For example, to create an object in AR, users have to go through the tedious process of coding and building on a desktop computer followed by testing on a smartphone to eventually verify that they have successfully attained the goal. The cognitive load for execution and evaluation can be significantly reduced when the development and testing platforms are unified into a homogeneous environment running on mobile devices. Users are aware of their surroundings and can enjoy the visualizations provided by AR even during the authoring process. Also, as gesture is a typical form of human expression, it seems that the interface can be made intuitive if common authoring commands are represented by related gestures.

This paper studies how mobile AR authoring can be designed for teachers by exploring the usability of gesture control under hand-held and head-mounted conditions. The functionality of the prototype is then expanded to collect feedback from high school teachers. The main contributions of this paper are:

- the insights into designing mobile gesture-enabled AR authoring tools through our prototypes;
- the study of the comparative merits of hand-held and head-mounted conditions in AR object manipulation;
- the exploration of the potential of our prototypes through several real-world application scenarios.

## 2. Related Work

AR authoring within an AR environment was first studied by Lee et al. [LNBK04]. Fiducial markers were used to represent different menu options and virtual objects. Users could import a 3D model from a predefined model set and associate the imported model with a marker. The orientation of the model could be changed by rotating a physical cube. The paper also proposed a novel approach to creating interactive content. Logical operators and other abstract concepts could be selected from the menu and assigned to a model. By compounding various operators and options, dynamic content responding to user interaction could be achieved. The system was implemented on a desktop. After that, there were numerous efforts in bringing AR authoring to mobile devices. Fully menu-based systems supporting general 3D transformations were found in [TDNL15, SW16]. The work of Rumiński and Walczak [RW13] enabled users to rotate and scale an object by swiping and pinching the smartphone screen respectively. Jeon et al. provided similar functions and supported deletion by shaking [JMY*16]. Yang et al. introduced to the menu-based authoring tool a selection method that could be controlled by tilting the smartphone [YSCH16]. The vast majority of existing mobile authoring tools appear to be menu-based and mainly utilize the touchscreen for menu operations and object transformation, leaving the effectiveness of other interaction techniques under-explored.

Nebeling et al. focused on simplifying 3D modeling and the creation of AR overlays using a smartphone in conjunction with a desktop [NNYR18]. 2D AR overlays were outlined using paper sketches and the sketches were captured using the camera of a smartphone. A desktop editor application was then used to perform foreground extraction. To create a 3D object, multiple images of a physical model made from Play-Doh were captured from different angles and the foreground of each image was extracted. A quasi-3D object could then be shown in AR environment by displaying the right image corresponding to the current viewing angle. The authors proposed a novel approach to 2D/3D modeling but most operations relied on the desktop editor, leaving the possibility of mobile AR authoring largely unexplored.

The use of gestures in mobile AR was studied by Hurst and Wezel [HvW13]. They argued that the small form factor of mobile devices restricted the extent of pointing and selection interaction via the touchscreen. Thus they investigated the potential of finger tracking for affine transformations of virtual objects in terms of task performance and engagement. The size and position of a virtual object could be changed by various gestures such as pinching, pushing and grabbing. The authors found that although finger control increased engagement, the accuracy was not high enough for concrete applications, concluding that the proposed finger control was only useful for entertainment purposes. It is therefore necessary to conceive an alternative gesture control method before gestures can be used in AR authoring.

## 3. Design Principle

Regarding the authoring of educational AR content on mobile devices, based on the related work, it seems reasonable that an ideal authoring tool can be expected to support a) import of 3D models, b) easy selection of objects, c) swift geometric transformation (translation, rotation and scaling) of objects and d) dynamic content with which students can interact. With these basic functionalities, a wide variety of teaching materials can be created. For instance, teachers can import 3D models downloaded from the Internet, arrange them as they desire and design simple quizzes using the interactivity provided by the tool. Previous work only utilized the touchscreen to realize the above features, but such control methods are insufficient according to Hurst and Wezel. Therefore, a new interaction paradigm that provides an intuitive interface to the above features is needed for efficiency and ease of use by educators.

### 3.1. Reality-Based Interaction

Our aim is to provide users with RBI-inspired interaction methods [JGH*08], so that users can create AR applications using their prior knowledge of the real world. This is achieved first by targeting our platform to mobile devices. Teachers can be assured that what they see on the mobile authoring environment is faithfully reproduced in the final product deployed on their students' mobile devices. Ideally, authoring should be performed solely on mobile devices to avoid switching between development and testing platforms, enabling quick verification of the intermediate result during the authoring progress.

### 3.2. Gesture Design

Interaction techniques that minimize the obstruction to the vision of users are preferable. Menus shown on the screen of mobile devices are sub-optimal designs because they utilize a large portion of the display area, hindering users' view of the environment behind the menus. Perhaps, menus should only be used when there is

© 2018 The Author(s)
Eurographics Proceedings © 2018 The Eurographics Association.

complicated information to be processed or selected by users. An approach that fulfills this requirement would be interaction through gesture recognition. As users execute commands by making gestures in front of the rear-facing camera of the device, they can see the effect of their commands without being blocked by other overlays on the screen.

As the processing power of a mobile device is limited, we need to design a set of gestures that can be recognized by the monocular rear-facing camera of a mobile device in real-time. Also, to avoid the need for expensive hardware, we target the minimum hardware requirement of our authoring tool at medium-end smartphones so that teachers can use their own devices for application development. Therefore, any depth camera-based approach would not be considered as most smartphones in the market lack such hardware, leaving the RGB camera as the most viable option. Although computer vision-based techniques for gesture recognition may be computationally expensive, we realized that the counting of fingers could be made computationally efficient to be used feasibly in our system. A wide variety of gestures can be supported by simply monitoring the change in finger count over time. In spite of its simplicity, it has been shown in previous research that finger count-based gestures have the advantages of less arm fatigue and reduced mental demand in menu selection tasks among various types of mid-air gestures [KL14].

In order to avoid the low accuracy problem of gestures in AR as shown by Hurst and Wezel [HvW13], we defined a set of gestures that worked in combination with a cursor and an AR marker for enhanced precision. The cursor was always at the center of the screen and was used for selection while the AR marker was a frame of reference for AR objects, as shown in Figure 1. The gestures were designed based on the principle of gestural congruence [STB14] so that they could be easily memorized. Leveraging the prior knowledge of users, we designed the gestures in Figure 2 that represented actions in reality.

- **Selection:** the selection gesture was analogous to mouse clicking in computer and it was represented by holding and then retracting the index finger. Button click, object selection and object placement could be performed using this gesture while pointing the cursor to the target object or location.
- **Menu:** The menu gesture invoked a menu for operations that could not be easily performed by gestures alone. It was represented by opening the palm, a gesture that was intuitively associated with the opening action.
- **Deletion:** Deletion was performed by closing the palm, an action analogous to crumpling a piece of unneeded paper to discard it.
- **Translation and rotation:** To translate or rotate a virtual object, the object needed to be frozen first. Freezing an object meant making the position and orientation of the object relative to the camera view instead of to the AR marker. When the object was frozen, the object appeared stationary with respect to the camera. Users could then move the AR marker to change the orientation and the position of the object relative to the marker. Once the frozen object appeared at the desired position and orientation, it could be unfrozen to re-establish its frame of reference to the marker. An example of translating and rotating an object is illustrated in Figure 3. Both freezing and unfreezing were represented by the "V" sign, a gesture similar to the pause button on common media players.
- **Scaling:** Enlarging and diminishing were represented by three and four fingers respectively. When users were holding any of the two gestures, the scale of the object changed. Once they were satisfied with the object size, they could hold the fist to stop further scaling.

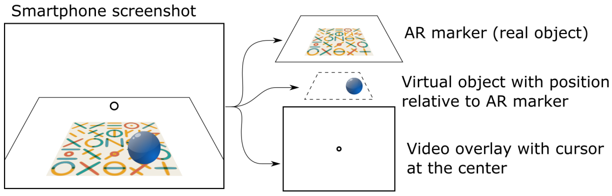



**Figure 1:** *AR marker provides a frame of reference for virtual objects. The cursor at the center of the screen is the target upon which various gestures act.*

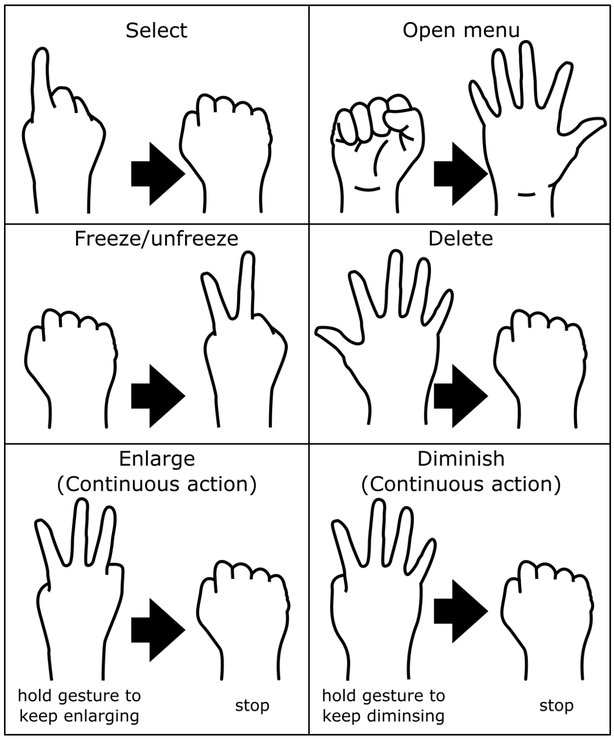



**Figure 2:** *Gestures are designed to convey meaning for ease of memorization. Gesture recognition is achieved by detecting the change in finger count.*

### 3.3. Modes of Operations

In addition to the usual hand-held operation mode found in previous work, we also studied the viability of head-mounted mode in mobile AR authoring. Gesture-based control enables users to interact with the system without touching the screen of mobile devices, thus it has the flexibility to support VR head mount. As pointed out by Kiyokawa [Kiy12], by using the head to control the view

© 2018 The Author(s)
Eurographics Proceedings © 2018 The Eurographics Association.

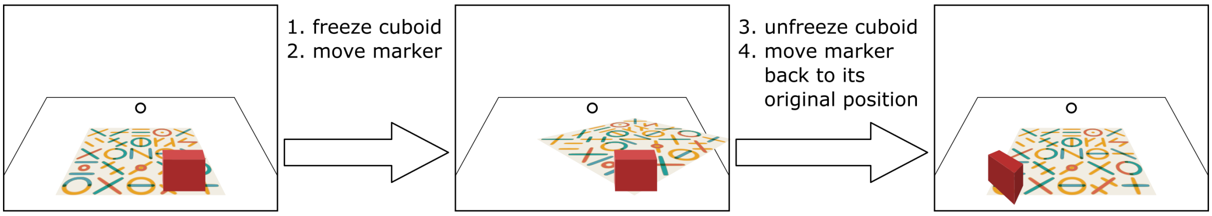
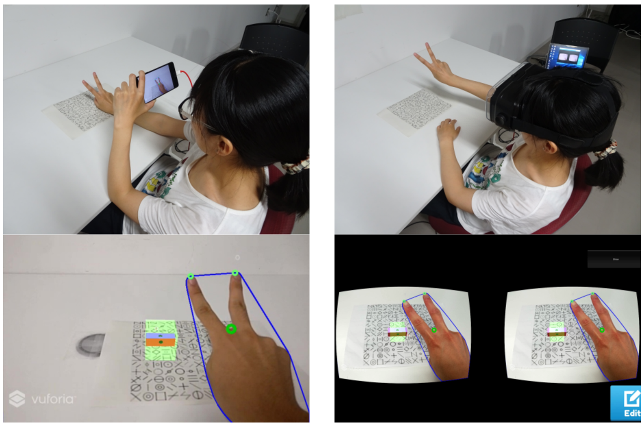



**Figure 3:** *To translate and rotate an object, first freeze the object and then move the marker. The frozen object will remain stationary with respect to the camera. After the object is translated and oriented as desired, unfreeze it to associate the object back with the marker.*

of the camera, both hands could be freed for interaction. Head-mounted mode also had the advantage of reduced arm fatigue and enhanced immersive experience. In our gesture-based design, one hand could make a gesture while the other hand could move the marker, thereby boosting the efficiency of authoring. A comparison of the postures between the two modes is shown in Figure 4.



**(a)** *Hand-held mode* **(b)** *Head-mounted mode*

**Figure 4:** *Making a gesture in front of the smartphone in the two modes. In head-mounted mode, the stereoscopic video can be viewed through a VR head mount.*

## 4. Prototyping

To know how suitable our proposed interaction method was for AR authoring, we implemented our first prototype on a smartphone and conducted an exploratory user study of the prototype. The prototype supported a) import of simple 3D objects, b) selection of objects using cursor and gestures and c) geometric transformation of objects. To simplify the experiment, only 10 predefined basic 3D models were prepared in the import menu. We were particularly interested in the relative merits and limits of both hand-held and head-mounted modes because the information was helpful for us to decide on which mode we could choose for implementing more complicated authoring commands such as d) dynamic content. We hypothesized that the level of physical strain was higher in the hand-held condition while it may be disorienting to view the surroundings through a head-mounted setup. To know the extent to which each condition affected the usability of the system and the cognitive workloads experienced by users, we performed the exploratory user study that required participants to perform a basic authoring task under hand-held and head-mounted modes.

### 4.1. System Architecture

The prototype was developed using Unity with Vuforia as AR plugin and OpenCV for gesture recognition, supporting both iOS and Android platforms. The video stream captured by the camera was analyzed to detect fingers and then augmented to add virtual objects as overlays in real-time. In head-mounted mode, the video was displayed stereoscopically to be viewed via typical smartphone-based VR head mount.

We adopted an approach similar to [SSP*14] for finger counting. Each video frame was first segmented using RGB values and then the contour of the hand within each frame was detected using Connected Component Analysis. By calculating the angles and distances among vertices of the contour and convexity points, the fingertips and thus the outstretched finger count could be found. For illustration purposes, the convex hull of the hand and the positions of the fingertips and the interdigital folds were indicated in the augmented video as shown in Figure 5, so that users were provided with feedback that their fingers were being correctly recognized by the system.

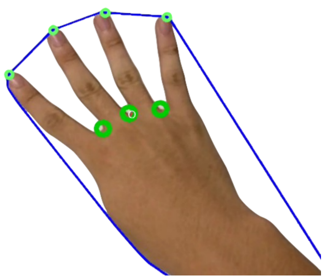


**Figure 5:** *The convex hull of the hand is indicated by the blue contour while the fingers and interdigital folds are identified with green dots as visual cues. These overlays are added to raise users' confidence during usage.*

## 5. Exploratory User Study

To evaluate the effectiveness of our gesture control in AR authoring and compare the performance under head-mounted and hand-held

© 2018 The Author(s)
Eurographics Proceedings © 2018 The Eurographics Association.

conditions, 32 university students in Japan and Singapore were recruited to test the system. Their fields of study ranged from communication to information content and interaction design. The average rating for their mobile AR experience was 1.4 on a scale from 1 to 3, with 1 being novice and 3 being expert. The mean age of the participants was 22.9. 15 participants were female.

All gestures were demonstrated by the experimenter to the participants before the experiment session started. After which, each participant was guided to perform a typical AR authoring task, consisting of the selection of a 3D model from a menu, placement, translation, rotation and scaling of the model to match a template. Each participant was then asked to perform the same sequence of actions without the guidance of the experimenter. Afterwards, participants answered the NASA-TLX [HS88] and SUS [Bro13] questions. They were then asked to express one challenging aspect about the task, one easy aspect about the task, as well as suggestions to make the system easier to use. All participants performed the experiment in both hand-held and head-mounted modes, with half of the participants starting with hand-held mode while the other half starting with head-mounted mode to counterbalance the conditions. After completing both modes, participants were asked to choose which mode they preferred and give reasons for their choice.

The participants used the Oneplus 3T smartphone and an Elecom plastic head mount. The smartphone was connected to the laptop via a 3 meter long USB cable for charging and monitoring of the smartphone screen by the experimenter. Calibration and eye checks were carried out using a standard Snellen chart before the commencement of the experiment to ensure that the participants could comfortably perceive the experimental stimuli.

### 5.1. Qualitative Result

In terms of user preference, 75% preferred head-mounted mode while the remaining 25% preferred hand-held mode. We analyzed these results through the lens of participants' opinions.

- **The viewing experience was better in hand-held mode:** Out of the 25% that preferred hand-held mode, 38% of them cited the superior visual quality as the main contributing factor for their decision. They "could not see things normally" and disliked the "narrower field-of-view" in head-mounted mode. Some mentioned that viewing through the head mount "induced motion sickness" and they were "forced to perform actions slowly due to the low FPS". Instead they could "see the hand more clearly" in the hand-held condition.
- **It was more comfortable and convenient to use hand-held mode:** Another 38% stated the lack of comfort and convenience in the head-mounted condition as the determining factor. It was "easier" and "more convenient" to use the hand-held mode because they could "just use the smartphone to complete the task without any other hardware", not having to "bother wearing the headset", which probably resulted in "less physical discomfort".
- **The physical strain was lower in head-mounted mode:** Out of the 75% that preferred head-mounted mode, about 50% of them commented that it was "tiring" and "exhausted" to hold the headset using their hands.
- **It was easier to get gestures recognized in head-mounted mode:** About 25% felt that their gestures were more readily recognized by the system in the head-mounted condition. They thought that it was "easier to control" and "position the fingers" in front of the camera.
- **It was easier to position the camera view and make gestures in head-mounted mode:** About 20% thought that they experienced less difficulty in performing the task under this condition. They "were able to use both hands" and found it "easier to select and interact with virtual objects". They could also "move the marker" conveniently. Furthermore, using the head to move the smartphone enabled them to "easily look around" and "position the cursor".

From the feedback of the participants we were able to conclude that our proposed gesture-based interaction techniques were useful for AR authoring. It was positively commented that "all operations can be performed using gestures without thinking about the touchscreen" and gesture-based interaction "is a direct way to control" eliminating the need to "go through several steps" in menu-based systems. For geometric transformations, "rotations can be done by just moving the marker" and it is "easy" and "intuitive" to move objects by moving the marker.

For the gesture vocabulary we defined, there were polarized comments about the ease of memorization of the gestures. Menu and deletion gestures received generally positive comments but it was hard for many participants to remember the gestures for enlarging and diminishing objects.

### 5.2. Quantitative Result

In terms of usability, the SUS score for the hand-held condition was significantly lower than that for the head-mounted condition based on a two-tailed t-test with $t[31] = -3.99$, $p < .001$. Regarding cognitive workload, the NASA-TLX score (lower is better) for the hand-held condition was significantly higher than that for the head-mounted condition, with $t[31] = 3.41$, $p < .01$. The workload of the hand-held condition was significantly greater in terms of physical demand, temporal demand, effort and frustration than the head-mounted condition.

### 5.3. Analysis and Implications

In the comparison between the head-mounted condition and hand-held condition, the former required less cognitive workload and had better usability. The higher workload for hand-held mode was probably due to the number of actions users needed to perform simultaneously. In the hand-held condition, the cognitive resources of users were split between holding the phone, positioning the cursor to targets and making gestures, whereas in the head-mounted condition users looked around and focused the center of the camera view to targets by moving their heads, a task that humans can perform more intuitively. The evidence from the NASA-TLX scores supports our suggestions. The mental demands were comparable in the two conditions, implying that the participants felt a similar level of difficulty. Likewise, the similarity of the performance scores suggested that the participants thought that they accomplished the tasks equally successfully. According to the higher effort and frustration scores in the hand-held condition, the participants felt that it was

© 2018 The Author(s)
Eurographics Proceedings © 2018 The Eurographics Association.

**Table 1:** *Results of NASA-TLX and SUS*

| | Mode | Mean | Std. dev. | t[31] | p |
|---|---|---|---|---|---|
| NASA-TLX | | | | | |
| Overall | hand | 58.7 | 15.1 | 3.41 | <.01 |
| | head | 46.5 | 19.3 | | |
| Mental demand | hand | 53.8 | 21.8 | 1.30 | 0.203 |
| | head | 47.0 | 22.4 | | |
| Physical demand | hand | 59.4 | 25.3 | 3.67 | <.001 |
| | head | 43.1 | 24.4 | | |
| Temporal demand | hand | 48.1 | 21.4 | 2.24 | <.05 |
| | head | 37.9 | 24.4 | | |
| Performance | hand | 49.2 | 20.4 | 1.77 | 0.0873 |
| | head | 42.3 | 22.2 | | |
| Effort | hand | 59.1 | 21.4 | 2.15 | <.05 |
| | head | 49.2 | 23.1 | | |
| Frustration | hand | 50.0 | 23.5 | 3.25 | <.01 |
| | head | 38.8 | 21.5 | | |
| SUS | | | | | |
| SUS | hand | 56.3 | 16.2 | -3.99 | <.001 |
| | head | 66.1 | 12.2 | | |

easier to perform the task in the head-mounted condition. Regarding the greater physical demand in the hand-held condition, one possible reason was that users needed to carefully adjust both hands to keep the marker within the view of the camera and to maintain a minimum distance between the camera and the gesture-making hand for proper gesture recognition.

Despite the numerous performance benefits afforded by the head-mounted condition, as suggested by the quantitative results, there were still a quarter of the participants who preferred the hand-held mode, a proportion that is surprisingly large. By looking closely into the opinion of the participants, it may be possible that the unsatisfactory display quality and unpleasant viewing experience under the head-mounted condition discouraged some participants from using this mode. The low screen refresh rate could have slowed down participants' actions and induced motion sickness and eye strain. The narrower field of view (FOV), low display resolution and the discomfort brought about by the head mount could also have been determining factors for the participants preferring the hand-held condition.

From this exploratory study, we learnt the areas in which the prototype should be improved. Overall, authoring through gestures on mobile devices seemed to be feasibly usable. Translation and rotation of virtual objects could be efficiently and intuitively performed using our proposed "freeze-and-unfreeze" approach. The result coincided with our expectation that gesture interaction could replace menus on the touchscreen for convenience and ease of control. Also, precise object selection was achieved through a cursor at the center of the screen. Although we designed some intuitive gestures such as those for menu invocation and deletion, we could make the gestures more intuitive by supporting more complicated gestures. Given that the convex hull of the hand and the position of fingertips were recognized by the system, it is very likely that the recognition system could be extended to support more gesture vocabulary in real-time. In terms of system performance, the accuracy of the gesture recognition should be improved to avoid unintended operations and unrecognized gestures, problems that occasionally occurred during the experiment. Furthermore, translation through freezing could be used in conjunction with other interaction techniques because it is not easy to accurately specify the final position through a perspective view.

From the perspective of user experience, we knew from the user study that display quality was important, especially in the case of head-mounted mode. A broad FOV, a high-resolution display and a responsive system were crucial for users to accept the head-mounted mode even if it provided evident advantages in terms of usability. Nonetheless, head-mounted mode was a viable option that worked well with gesture-based systems. As the display quality and the processing power of smartphones have been improving, head-mounted mode remains to be a promising approach for immersive AR authoring.

## 6. Usage Scenarios in Education

Based on the result of the exploratory study, we extended the functionality of the prototype and then interviewed a few teachers to study whether our prototype was useful from the perspective of practitioners. Also, we invited university students to show how our prototype could be applied in specific teaching scenarios.

### 6.1. Extension of Functionality

The prototype was extended in order to support 1) textual annotations, 2) event-driven change of content and 3) fine-tuning of virtual object size. Annotations and dynamic content are commonly used in pedagogy settings. Description text in AR can be used to show usage instruction and annotation for individual virtual objects while interaction within AR environment can promote active learning due to increased engagement of users. For instance, chemistry teachers may use the interactivity to create a simple quiz that requires students to show the bond breaking and bond formation that happen in the combustion of methane by reorganizing virtual molecules represented by the ball-and-stick model. If students correctly perform the task, a congratulation message will be shown. Otherwise, an animation showing the correct procedures will be displayed.

We decided that these new functionalities should be implemented in hand-held mode. For textual input, it seems more efficient to type on the screen keyboard than in mid-air. Also, the higher display resolution in the hand-held condition makes it easier to view text. For dynamic content, specifying the type of change and the triggering event requires multiple selections from menus, a task that seems to be more conveniently achievable on the touchscreen, as shown in Figure 6. Finally, the scaling of virtual objects can be fine-tuned by pinching the object on the touchscreen.

As our first expansion to the prototype, we added a simple interactive capability to the system. Users can specify either the change of surface color or the display of annotation upon the detection of a certain number of finger count. As the system now supports two operation modes, switching between hand-held and head-mounted modes is realized through a button on the touchscreen.

© 2018 The Author(s)
Eurographics Proceedings © 2018 The Eurographics Association.

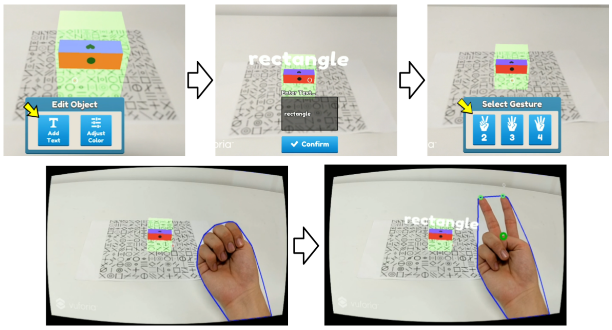



**Figure 6:** *An annotation can be typed on the touchscreen and the triggering gesture is then specified. The dynamic content can be immediately previewed.*

### 6.2. Interview with Teachers

We then studied the usefulness of our extended prototype in educational settings by interviewing 3 high-school teachers teaching biology, chemistry and physics. The mean age and teaching experience of the teachers were 23.3 and 1.3 years respectively. We gave them a brief overview of mobile AR, showed them the video demonstrating the functions of the extended prototype and collected feedback from them. All of them thought that it was easy to use the system and they would like to try using the system. Notable comments were expressed by a chemistry teacher and a biology teacher. The chemistry teacher expressed that he would like to use mobile AR to explain crystal structures and molecular structures to students. The biology teacher thought that the stages of human development could be explained using animated content in AR, which could attract students' attention. They further suggested that it would be helpful if students could share content among themselves. They also would like to "monitor students' learning activity" such as quiz answers through the application.

### 6.3. Application Design Workshop

We arranged a workshop to discover how our prototype could be used when applied to different teaching scenarios. After being presented with the functionality and principle of our prototype, 15 education-major students divided into 3 groups were asked to design and demonstrate example applications for any school subject. In all, the groups proposed 3 potential applications in the areas of anatomy, calculus and grammar, as illustrated in Figure 7. The group that developed the idea for anatomy pointed out that, in a university course of psychology, it was difficult to locate the basal ganglia in the brain. They proposed that our prototype could be used to vividly demonstrate to students the location of this brain structure. The left cerebral hemisphere could either be separated from the right one using the freezing operation or it could be made transparent using the palm-opening gesture, to reveal the internal structure of the brain. The brain could be viewed from different angles by rotating and tilting the marker. To return the brain parts to their original positions, the palm-closing gesture could be used. The application would enable students to easily interact with 3D anatomical structures that could not be easily taught using notes or slides. The other two groups of students developed solutions to visualize the solid of revolution in calculus and the different spatial prepositions in French. In the calculus example, a 2D curve was rotated around an axis to form a 3D volume which could be viewed easily in AR. In the French-teaching example, by looking at a virtual object from different angles and distances, the corresponding spatial prepositions used to describe the situations were shown on the screen. These 3 applications exemplified how mobile AR could promote active learning through simple interaction methods including gestures and markers.

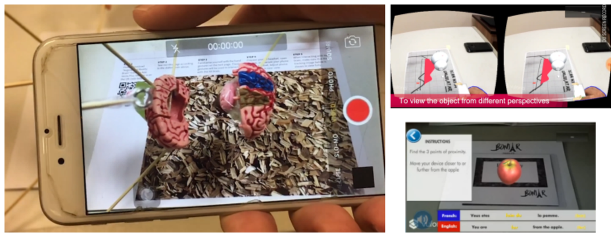


**Figure 7:** *The group on anatomy made a mock-up of the application using wooden sticks and plastic brain models. The other two groups made applications for calculus and French learning.*

## 7. Discussion and Limitations

Given the viability of our gesture-based interaction method in mobile AR authoring, it may be possible that the system can be used to develop a suite of mobile AR applications for education. Through our user study, we found that gestures can be used to efficiently accomplish tasks that are tedious on menu-based systems, especially 3D translation and rotation. However in our implementation, the occasional inaccuracy of the gesture recognition has been a source of confusion for users. First of all, the accuracy and responsiveness of the recognition should be enhanced to boost usability. Secondly, more comprehensive visual or audio feedback regarding the gesture recognition status should be provided to let the users know the type of operations they are performing. Thirdly, the recognition system should be improved to support a wider range of gestures instead of simply counting the number of fingers. For other functions that require complicated input, they should be realized using menus. There may be many ways in which the authoring tool can be extended and that will be a subject for future work.

## 8. Conclusion and Future work

We proposed the creation of prototypes with gesture-based interaction for mobile AR authoring to be used by teachers who need the fast and easy creation of AR teaching materials. The prototypes simultaneously supported head-mounted mode for intuitive operation and hand-held mode for enhanced display quality and touchscreen input. Our exploratory user study revealed that our gesture control method was effective for geometric transformation and object selection. It also showed that head-mounted mode and hand-held mode were complementary as the former had superior usability due to the reduced cognitive workload whereas the latter provided a

© 2018 The Author(s)
Eurographics Proceedings © 2018 The Eurographics Association.

more comfortable viewing experience. The prototype was then extended to support interactivity, an important element in educational contexts.

Through our interview and design workshop we obtained strong evidence of the usefulness of our prototypes in different pedagogical scenarios. It has been demonstrated by education-major students that our gestural metaphor was useful to not only educators but also learners who could now explore virtual objects with ease using gestures. Our prototype was especially beneficial for teaching difficult concepts or structures that demanded spatial understanding.

Future work should be dedicated to improving the gesture recognition system and the mobile application for students. The future authoring tool can be expected to support more gestures for more powerful authoring capabilities while remaining intuitive for non-technical users. The recognition accuracy also needs to be enhanced to avoid confusion. With an application variant specifically designed for learners, real-time interactions among the teacher and students can be made possible, resulting in a more engaging learning environment.

## 9. Acknowledgments

We gratefully acknowledge the support of Nvidia Corporation for supplying part of the hardware used in this research. This work was supported in part by the Cooperative Research Project Program of the Research Institute of Electrical Communication in Tohoku University, JSPS Bilateral Joint Research Projects/Seminars and JSPS KAKENHI (Grant Number 18H04103).

## References

[Bro13] BROOKE J.: Sus: A retrospective. *J. Usability Studies 8*, 2 (Feb. 2013), 29–40. URL: http://dl.acm.org/citation.cfm?id=2817912.2817913. 5

[CCHK11] CHEN Y.-C., CHI H.-L., HUNG W.-H., KANG S.-C.: Use of tangible and augmented reality models in engineering graphics courses. *Journal of Professional Issues in Engineering Education and Practice 137*, 4 (2011), 267–276. doi:10.1061/(ASCE)EI.1943-5541.0000078. 1

[DSInK13] DI SERIO A., IBÁÑEZ M. B., KLOOS C. D.: Impact of an augmented reality system on students' motivation for a visual art course. *Comput. Educ. 68* (Oct. 2013), 586–596. doi:10.1016/j.compedu.2012.03.002. 1

[HS88] HART S. G., STAVELAND L. E.: Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Human Mental Workload*, Hancock P. A., Meshkati N., (Eds.), vol. 52 of *Advances in Psychology*. North-Holland, 1988, pp. 139–183. doi:10.1016/S0166-4115(08)62386-9. 5

[HvW13] HÜRST W., VAN WEZEL C.: Gesture-based interaction via finger tracking for mobile augmented reality. *Multimedia Tools and Applications 62*, 1 (Jan 2013), 233–258. doi:10.1007/s11042-011-0983-y. 2, 3

[JGH*08] JACOB R. J., GIROUARD A., HIRSHFIELD L. M., HORN M. S., SHAER O., SOLOVEY E. T., ZIGELBAUM J.: Reality-based interaction: A framework for post-wimp interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2008), CHI '08, ACM, pp. 201–210. doi:10.1145/1357054.1357089. 1, 2

[JMY*16] JEON J., MIN H., YI M., CHUN J., KIM J.-S., CHO Y.-J.: Interactive authoring tool for mobile augmented reality content. *Journal of Information Processing Systems 12*, 4 (2016), 612–630. doi:10.3745/JIPS.02.0048. 2

[Kiy12] KIYOKAWA K.: Trends and vision of head mounted display in augmented reality. In *2012 International Symposium on Ubiquitous Virtual Reality* (Aug 2012), pp. 14–17. doi:10.1109/ISUVR.2012.11. 3

[KL14] KULSHRESHTH A., LAVIOLA JR. J. J.: Exploring the usefulness of finger-based 3d gesture menu selection. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2014), CHI '14, ACM, pp. 1093–1102. doi:10.1145/2556288.2557122. 3

[KS02] KAUFMANN H., SCHMALSTIEG D.: Mathematics and geometry education with collaborative augmented reality. In *ACM SIGGRAPH 2002 Conference Abstracts and Applications* (New York, NY, USA, 2002), SIGGRAPH '02, ACM, pp. 37–41. doi:10.1145/1242073.1242086. 1

[LNBK04] LEE G. A., NELLES C., BILLINGHURST M., KIM G. J.: Immersive authoring of tangible augmented reality applications. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality* (Washington, DC, USA, 2004), ISMAR '04, IEEE Computer Society, pp. 172–181. doi:10.1109/ISMAR.2004.34. 2

[ND86] NORMAN D. A., DRAPER S. W.: *User Centered System Design; New Perspectives on Human-Computer Interaction*. L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 1986. 2

[NNYR18] NEBELING M., NEBELING J., YU A., RUMBLE R.: Protoar: Rapid physical-digital prototyping of mobile augmented reality applications. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2018), CHI '18, ACM, pp. 353:1–353:12. doi:10.1145/3173574.3173927. 2

[NUS17] National TEL Conference, TEL2017. http://web.archive.org/web/20171216201040/nus.edu.sg/tel2017/, 2017. Accessed: 2018-07-10. 1

[RW13] RUMIŃSKI D., WALCZAK K.: Creation of interactive ar content on mobile devices. In *Business Information Systems Workshops* (Berlin, Heidelberg, 2013), Abramowicz W., (Ed.), Springer Berlin Heidelberg, pp. 258–269. 2

[SK07] SQUIRE K., KLOPFER E.: Augmented reality simulations on handheld computers. *Journal of the Learning Sciences 16*, 3 (2007), 371–413. doi:10.1080/10508400701413435. 1

[SSP*14] SONG J., SÖRÖS G., PECE F., FANELLO S. R., IZADI S., KESKIN C., HILLIGES O.: In-air gestures around unmodified mobile devices. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 2014), UIST '14, ACM, pp. 319–329. doi:10.1145/2642918.2647373. 4

[STB14] SEGAL A., TVERSKY B., BLACK J.: Conceptually congruent actions can promote thought. *Journal of Applied Research in Memory and Cognition 3*, 3 (2014), 124–130. Cognition and Education. doi:https://doi.org/10.1016/j.jarmac.2014.06.004. 3

[SW16] SAMBROOKS L., WILKINSON B.: Designing haratio: A novice ar authoring tool. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction* (New York, NY, USA, 2016), OzCHI '16, ACM, pp. 175–179. doi:10.1145/3010915.3011005. 2

[TDNL15] TANG J. K. T., DUONG T. Y. A., NG Y. W., LUK H. K.: Learning to create 3d models via an augmented reality smartphone interface. In *2015 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)* (Dec 2015), pp. 236–241. doi:10.1109/TALE.2015.7386050. 2

[YSCH16] YANG Y., SHIM J., CHAE S., HAN T. D.: Interactive augmented reality authoring system using mobile device as input method. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (Oct 2016), pp. 001429–001432. doi:10.1109/SMC.2016.7844437. 2

© 2018 The Author(s)
Eurographics Proceedings © 2018 The Eurographics Association.