Counterpoint: Exploring Mixed-Scale Gesture Interaction for AR Applications

Barrett Ens
University of South Australia
barrett.ens@unisa.edu.au

Aaron Quigley
University of St. Andrews
aquigley@st-andrews.ac.uk

Hui-Shyong Yeo
University of St. Andrews
hsy@st-andrews.ac.uk

Pourang Irani
University of Manitoba
pourang.irani@cs.umanitoba.ca

Thammathip Piumsomboon
University of South Australia
thammathip.piumsomboon
@unisa.edu.au

Mark Billinghurst
University of South Australia
mark.billinghurst@unisa.edu.au

Abstract
This paper presents ongoing work on a design exploration that interleaves microgestures with other types of gestures from the greater lexicon of gestures for computer interaction. We describe three prototype applications that show various facets of this multi-dimensional design space. These applications portray various tasks on a Hololens Augmented Reality display, using different combinations of wearable sensors. Future work toward expanding the design space and exploration is discussed, along with plans toward evaluation of multi-scale gesture design.

Author Keywords
Gesture interaction; wearable computing; Augmented reality

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous;

Introduction
Computer interfaces have often made successful use of gestures to allow rich and intuitive interaction. The strategy of mimicking our everyday interactions with real-world objects allows us to map high-level mental models onto appropriate sequences of physical actions [3]. Interaction designers have often sought to make

such interactions analogous to everyday interactions with real-world objects. With the mouse, for example, users apply a combination of motions and clicking gestures to select or move data objects on a virtual desktop. More recently, touch screens allow tapping and swiping gestures, which allow data objects and interface components to be manipulated using fingers, without an intermediary device. In next-generation Augmented Reality (AR) interfaces, gesture interaction will more closely resemble interactions with real-world objects, since users can point at [2] or grasp [16] virtual objects overlaid directly on the environment.

While such “natural” interactions can be highly appealing for AR applications, “in-air” gestures are currently prone to several disadvantages, such as fatigue, imprecision, and social awkwardness. Fatigue is caused by interactions that require the arms to be extended from the body for a prolonged period [9]. Precise interaction is known to be difficult without the aid of a haptic surface, limiting practical applications. Social awkwardness may arise in some social contexts, when gestures attract attention from observers who do not have a complete picture of what a user is doing [1].

To address these concerns, researchers have explored microgestures, minute gestures performed by the hands or fingers [22]. Microgestures rely on sophisticated sensing methods that are capable of detecting fine-scale hand motions. Such methods include computer vision techniques for articulated hand-tracking [18] and sub-millimeter radar [12].

Contribution

Microgestures have primarily been explored on their own, without considering the user’s long-term goals. In contrast, this work builds toward a design space for applying microgestures within the greater gestural lexicon [6]. This goal of this design space and ensuing exploration, which we term *Counterpoint*, is to elicit broad thinking about complex sequences of tasks, where interplay between complementary gestures is leveraged to achieve a sequence of user goals.

Our ongoing work will produce the following set of contributions:

- A complete design space for gesture interaction based on an exhaustive literature review.
- Introduction of the dimension of *scale* for mixed-scale gesture interaction
- A design space exploration that includes several example implementations of mixed-scale gestures
- The first evaluation of whether mixed-scale gesture interaction can mitigate issues of in-air gestures such as precision, fatigue and social awkwardness.

These contributions will help AR interaction designers support complex tasks and will hopefully encourage further exploration of the Counterpoint design space and mixed-scale interactions. This paper builds on our prior work that introduced the concept of mixed-scale gestures and the applications presented below [7]. Here we expand on our planned contributions, solidify the definition of our design space, discuss how our applications demonstrate the design dimensions, and outline our planned evaluation.

Counterpoint Design Space

The central idea of the Counterpoint design space is to facilitate the design of interactions leverage the advantages of different gesture types. For instance,

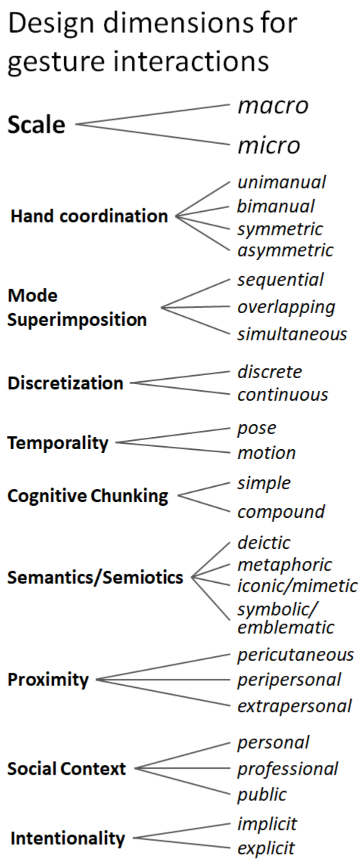

Figure 1: Design dimensions for gesture interaction, compiled from a literature review of gesture taxonomies and design spaces. The focus of Counterpoint is on designing gesture interactions that include combinations of microgestures and macrogestures.

large gestures such as pointing and 6-dof manipulation can be fast and convenient, but excessive reliance on such gestures can lead to arm fatigue or social awkwardness. Meanwhile, small finger movements can provide precision and subtlety, but may lack the expressivity of larger gestures.

While our main focus is on mixed-scale interaction, which integrates of microgestures and macrogestures into more complex interaction sequences, we need to view such designs within the context of a complete design space for gesture interaction. Therefore, one of the goals of this work is to provide an extensive gesture design space by consolidating the many previous taxonomies and design spaces for gesture interaction that have been introduced in prior work.

The course towards current thinking on gestural interaction design spans multiple research disciplines over several decades. Early taxonomies for gestures were introduced in the fields of linguistics and psychology, and were aimed at classifying different types and purposes of hand gesticulations commonly made during speech [4,5,10]. Later taxonomies introduced gestures made in conjunction with speech as a natural method of interacting with computers [17,20]. More recent design spaces have introduced new elements specifically geared toward computer interaction. Some of these are aimed at specific contexts for gesture interaction, such as multitouch input for tabletops [21], motion gestures for mobile phones [19], or wearable augmented reality [15].

In this work, we aim our design space toward designers of applications for wearable AR. In particular, we explore "in-air" gestures, which allow users to move freely and interact naturally in the environment where an application is situated, by eliminating the need for handheld devices. Moreover, we are interested in moving toward productive AR applications, which may include entertainment, but also situated analytics of environmental sensor data, interaction with networks of smart objects, and computer aided design (CAD). Such applications are likely to require complex sequences of nuanced commands.

Design Dimensions

A design space for gesture interaction is presented in Figure 1. These dimensions are compiled from a literature review of prior taxonomies and design spaces for gesture interaction, and is refined for our contextual focus on in-air gestures for wearable AR. Whereas this space encompasses many dimensions drawn from various sources, we introduce *scale* as the central dimension of this exploration. Microgesture and macrogestures have previously been explored independently, however to our knowledge, scale has not been included as a dimension in any previous gesture taxonomy or design space, and mixed-scale gestures have not been explicitly studied.

The scope of this paper does not allow us to include a detailed description of each dimension, however we are continuing to compile and refine a complete design space which will be fully presented in future work.

Design Space Exploration

To demonstrate the concept of mixed-scale interaction, we are implementing a series of applications within the Counterpoint design space. These interactions are not proposed as ideal workflows for the given scenarios, but are chosen to explore various facets of the design

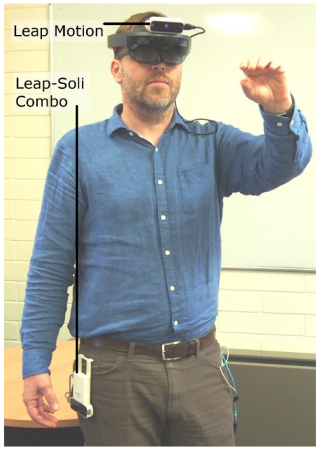

Figure 2: We use several wearable sensors to provide gesture input at multiple scales.

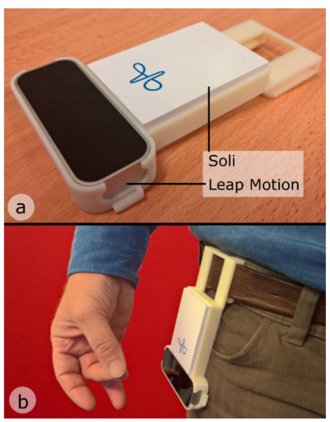

Figure 3: A belt-worn sensor configuration (a) allows "hands-down" [13] sensing (b) of microgestures.

space and to inspire further exploration of gesture combinations.

These implementations rely on various combinations of wearable sensors, as will likely become commonplace as wearable systems proliferate. Macro-scale hand gestures are tracked using a Leap Motion device mounted on a HoloLens [14] wearable AR display (Figure 2). For microgesture sensing, we use body-mounted Leap Motion [11] and Google Soli [8] sensors (Figure 3), which currently provide the best capabilities for detecting small hand movements. Below we describe three implementations, and explain how each leverages various dimensions of the Counterpoint design space (with keywords in *italics*). We continue to develop further examples as our work progresses.

Demo #1: Precise Object Manipulation
The first application (Figure 4a) demonstrates precise object manipulation, an operation that may be required for applications such as CAD or visual analytics. For demonstration purposes, we use a docking task. A head-worn Leap sensor (Figure 2) can detect *grasping and manipulation* gestures, which can be used to place an object near the specified position. The user can then fine tune the object's pose with the *arm down* in a relaxed posture [13]. A belt-worn sensor configuration (Figure 3) detects small movements of the thumb when placed against the sides or tips of the fingers. Six virtual sliders can be used individually to provide *precise* control over the object's three axes of translation and rotation (Figure 4).

This docking task solution provides an example of *unimanual*, *sequential* interaction. The initial 6-dof object manipulation is an example of *compound* *chunking*, the grouping of parameter manipulations to fit high-level mental models [3]. The subsequent fine-tuning is supported by isolating each axis into *simple* components. In this case, *pose* information provided by the Leap sensor determines which component to activate, while the Soli provides fine-scale microgesture sensing to control 1D sliders, which are mapped to either *discrete* or *continuous* input.

Demo #2: Virtual Puppetry
The second application explores virtual puppetry (Figure 5), which is inspired by marionette operators who commonly combine multiple gesture scales. As in the first application, a head-mounted Leap device senses large gestures made by the arm and hand, which in this case control the position and direction of a virtual puppet (Figure 5b). A Soli sensor, mounted under the wrist meanwhile senses subtle motions of the fingers below the occluded hand, which control various animations. For instance, while a pointed index finger indicates the puppet's running direction, the middle finger controls the speed of a running animation, which transitions to walking as the middle finger is extended. Conversely, lowering the index finger controls the speed and swinging direction of the puppet's toy sword.

This example demonstrates *unimanual* control of *simultaneous* modes. The index finger *pose* provides a *metaphoric* gesture to indicate the puppet's running direction while *continuous* microgestures control animation speeds. In this implementation, the limit of the arm's reach restricts the puppet to *personal* space.

Demo #3: In-Situ Video Editor
The third example explores a tool for in-situ cropping of short videos that might be taken with a head-worn AR

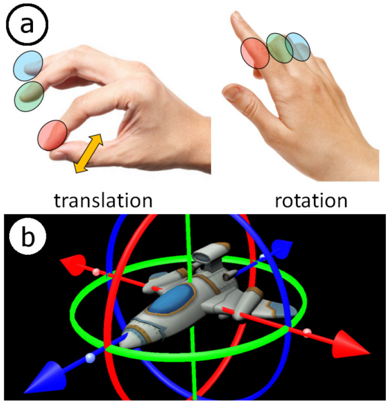

Figure 4: In demo #1, Six virtual sliders on the finger sides and tips (a) are used to control the pose of a virtual object (b) in a docking task.

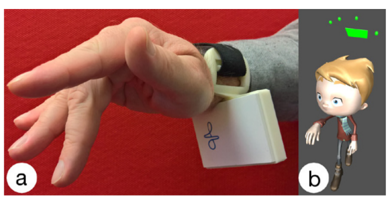

Figure 5: Demo#2 uses a wrist-worn sensor (a) to provide fine control of animations for a virtual puppet (b).

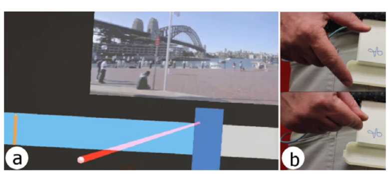

Figure 6: Demo #3 uses bimanual input to simultaneously specify a position on a film clip (a), and crop using a "snipping" microgesture (b).

display's camera. While one hand points to a position on a virtual scrubbing slider (Figure 6a), the second hand produces a "snipping" gesture, detected by a belt-worn Soli sensor (Figure 6b), to apply a cropping operation at the desired point.

This example demonstrates *bimanual* interaction; the head-mounted sensor detects large-scale *deictic* gestures on the video slider, while a pre-trained *mimetic* gesture triggers the cropping action. Other types of motion gestures could potentially be trained to provide a vocabulary of editing operations.

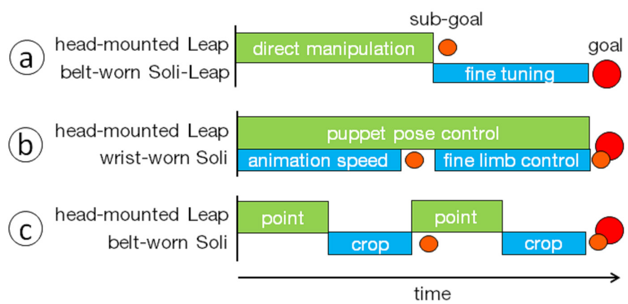

Figure 7: Our prototype implementations use interplay between macrogestures (shown in green) and microgestures (blue) to conduct a series of operations toward a user goal.

Figure 7 provides an overview of these three examples, emphasizing the multiple scales. A common feature across these examples is that the large-scale gestures (shown in green) tend to provide more *complex chunking*, whereas the microgestures (in blue) often control *simple* operations. At a higher level, the interplay between multiple scales allows an even greater chunking of operations into sequences that lead toward task sub-goals and goals.

Evaluation of Mixed-Scale Gestures

Our next steps include the design and execution of a user study to evaluate the benefits of multi-scale gesture design. We plan to implement a simple version of the demonstrated docking task, to compare multi-scale gesture interaction against individual use of microgestures and macrogestures. As metrics, we will measure task precision, user fatigue, and perceived social acceptability. Whereas prior studies of fatigue have used subjective ratings or macro-scale sensing [9], we will measure fatigue using electromyography to capture fatigue at multiple scales.

Conclusion

In summary, this paper outlines our ongoing progress in the development of the Counterpoint design space, aimed at facilitating design of multi-scale gesture interactions. We present three initial applications that explore this design space and demonstrate the interaction possibilities of multi-scale gesture design for wearable AR. Future work will continue these developments and introduce the first user study to evaluate multi-scale gestures by measuring fatigue.

References

1. David Ahlström, Khalad Hasan, and Pourang Irani. 2014. Are you comfortable doing that? In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services - MobileHCI '14*, 193–202.
2. Richard a Bolt. 1980. "Put-that-there": Voice and Gesture at the Graphics Interface. *Proceedings of the 7th annual conference on Computer graphics and interactive techniques - SIGGRAPH '80*: 262–270.
3. Bill Buxton. 1986. Chunking and phrasing and the design of human-computer dialogues. *Proceedings of the IFIP World Computer Congress*, 1986: 475–480.

4. D Efron. 1941. Gesture and environment. Retrieved January 15, 2018 from http://psycnet.apa.org/psycinfo/1942-00254-000
5. Paul Ekman and Wallace V. Friesen. 1969. The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding. *Semiotica* 1, 1.
6. Barrett Ens, Ahmad Byagowi, Teng Han, Juan David Hincapié-Ramos, and Pourang Irani. 2016. Combining Ring Input with Hand Tracking for Precise, Natural Interaction with Spatial Analytic Interfaces. In *Proceedings of the 2016 Symposium on Spatial User Interaction - SUI '16*, 99–102.
7. Barrett Ens, Aaron Quigley, Hui-Shyong Yeo, Pourang Irani, Thammathip Piumsomboon, and Mark Billinghurst. 2017. Exploring mixed-scale gesture interaction. In *SIGGRAPH Asia 2017 Posters on - SA '17*, 1–2.
8. Google. Project Soli. Retrieved September 14, 2017 from https://atap.google.com/soli/
9. Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed Endurance: A metric to quantify arm fatigue of mid-air interactions. *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*: 1063–1072.
10. Adam Kendon, Thomas A Sebeok, and Jean Umiker-Sebeok. 1981. *Nonverbal communication, interaction, and gesture: selections from Semiotica*. Walter de Gruyer.
11. Leap Motion. Leap Motion. Retrieved September 14, 2017 from https://www.leapmotion.com/
12. Jaime Lien, Nicholas Gillian, M. Emre Karagozler, Patrick Amihood, Carsten Schwesig, Erik Olson, Hakim Raja, and Ivan Poupyrev. 2016. Soli. *ACM Transactions on Graphics* 35, 4: 1–19.
13. Mingyu Liu, Mathieu Nancel, and Daniel Vogel. 2015. Gunslinger. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*, 63–71.
14. Microsoft. The leader in Mixed Reality Technology | HoloLens. Retrieved September 14, 2017 from https://www.microsoft.com/en-us/hololens
15. Thammathip Piumsomboon, Adrian Clark, Mark Billinghurst, and Andy Cockburn. 2013. User-defined gestures for augmented reality. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 8118 LNCS, PART 2: 282–299.
16. Ivan Poupyrev, Mark Billinghurst, Suzanne Weghorst, and Tadao Ichikawa. 1996. The go-go interaction technique. *Proceedings of the 9th annual ACM symposium on User interface software and technology - UIST '96*: 79–80.
17. Francis Quek, David McNeill, Robert Bryll, Susan Duncan, Xin-Feng Ma, Cemil Kirbas, Karl E. McCullough, and Rashid Ansari. 2002. Multimodal Human Discourse: Gesture and Speech. *ACM Trans. Comput.-Hum. Interact.* 9, 3: 171–193.
18. J M Rehg and T Kanade. 1994. Visual Tracking of High {DOF} Articulated Structures: An Application to Human Hand Tracking. *Proceedings of the European Conference on Computer Vision (ECCV)* 801.
19. Jaime Ruiz, Yang Li, and Edward Lank. 2011. User-defined motion gestures for mobile interaction. *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*: 197.
20. Alan Daniel Wexelblat. 1994. A Feature-Based Approach to Continuous-Gesture Analysis by.
21. Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-defined gestures for surface computing. *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*: 1083.
22. Katrin Wolf, Anja Naumann, Michael Rohs, and Jörg Müller. 2011. A taxonomy of microinteractions: Defining microgestures based on ergonomic and scenario-Dependent requirements. *Lecture Notes in Computer Science,* 6946 LNCS, PART 1: 559–575.