2017

Influences of visual and auditory displays on aimed movements using air gesture controls

Jason Sterkenburg
Michigan Technological University

Steven Landry
Michigan Technological University

Myounghoon Jeon
Michigan Technological University

Follow this and additional works at: https://digitalcommons.mtu.edu/cls-fp
Part of the Computer Sciences Commons

INFLUENCES OF VISUAL AND AUDITORY DISPLAYS ON AIMED MOVEMENTS USING AIR GESTURE CONTROLS

Jason Sterkenburg, Steven Landry, Myounghoon Jeon

Mind Music Machine Lab
Michigan Technological University,
1400 Townsend Ave., Houghton, MI USA
{jtsterke,sglandry,mjeon}@mtu.edu

ABSTRACT

With the proliferation of technologies operated via in-air hand movements, e.g. virtual/augmented reality, in-vehicle infotainment systems, and large public information displays, there remains an open question about if/how auditory displays can be used effectively to facilitate eyes-free aimed movements. We conducted a within-subjects study, similar to a Fitts paradigm study, in which 24 participants completed simple aimed movements to acquire targets of varying sizes and distances. Participants completed these aimed movements for six conditions – each presenting a unique combination of visual and auditory displays. Results showed participants were generally faster to make selections when using visual displays compared to displays without visuals. However, selection accuracy was similar for auditory-only displays when compared to displays with visual components. These results highlight the potential for auditory displays to aid aimed movements using air gestures in conditions where visual displays are impractical, impossible, or unhelpful.

1. INTRODUCTION

Air gesture control – the operation of devices by in-air hand movements – has potential to empower users with a natural and rich level of control over their devices. Auditory displays, in combination with gesture controls, could improve technology accessibility for visually-impaired users and allow for eyes-free interaction for sighted users. However, it is unknown how auditory displays affect aimed movements using air gesture controls. Auditory display design is particularly interesting in application to air gesture controls because, unlike many other forms of technology, air gesture controls allow for continuous tracking of user hand positions. Currently, little is known about how different sonification strategies may affect aimed movement performance when using air gesture controls.

Some studies have investigated aimed movement performance using air gesture controls [e.g., 8-9] and even explored the concept of eyes-free aimed movements [9] using only kinesthetic information. Other studies have examined the impact of auditory displays on target acquisition performance [1-2]. However, to our knowledge there is little to no existing literature exploring the utility of auditory displays in conjunction with air gesture controls in aiding target acquisition tasks. Most existing literature surrounding the topic of auditory displays and air gestures have focused on target localization, i.e., finding the point of origin of a sound in space [e.g., 3-7].

We conducted an experiment to learn how auditory displays affect aimed movement performance using air gesture controls. We made comparisons between two sonification strategies: (1) a discrete auditory display – playing a sound whenever the user is on the target and (2) a continuous auditory display – playing sound continuously from the start of the movement until selection, and playing a discrete sound when the user is on target. We also made comparisons among auditory-only, visual-only, and visual-auditory displays, as well as a control condition for which there was no visual or auditory display.

2. METHODS

2.1. Design Guidelines

Soukoreff and MacKenzie [3] wrote a paper outlining several guidelines which supported the ISO 9241-9 standards for the evaluation of pointing devices in human-computer interaction. In keeping with standard evaluation of pointing devices, we followed each of those standards as much as possible. This is our justification for (1) our use of the Shannon formulation of index of difficulty, (2) our range of movement difficulties, (3) our adjustments for selection accuracy, (4) and our calculation of throughput.

2.2. Apparatus

We used a LEAP Motion as our hand-position tracking sensor and we used Pure Data – an open source graphical programming language – to develop our target selection task (Figure 1). As the participant moves their hand above the sensor, a cursor matches the position of the person's hand along the x-axis (no y-axis data were recorded) and makes corresponding movements on the screen. All cursor movements were mapped one-to-one to hand movements.

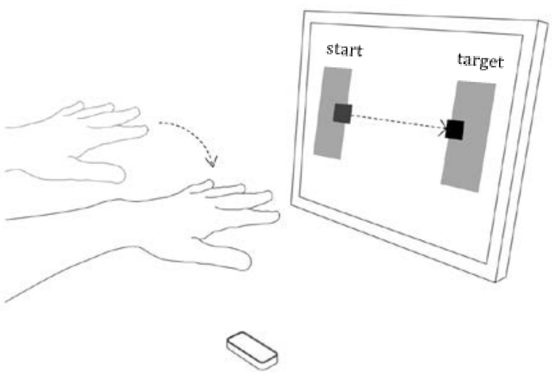

Figure 1: Illustration of experimental setup.

https://doi.org/10.21785/icad2017.065

2.3. Participants

A total of 24 undergraduate psychology students were recruited to complete our study (Table 1). All participants were given course credit as compensation for their participation. Only one person reported having experience using a LEAP Motion before.

Age (yrs)	Gender	Handedness
Mean=19.75	Males:14	Right:21
SD=1.96	Females:10	Left:3

Table 1: Demographic statistics for participants

2.4. Experimental design

We used full factorial within-subjects design with a total of six conditions:

AC – continuous auditory display
AD – discrete auditory display
VAC – visual plus continuous auditory
VAD – visual plus discrete auditory
V – only visual display
Control – visual removed upon start, no audio

2.5. Sound design

There were two different sound designs: a discrete auditory display and a continuous auditory display. The discrete auditory display consisted only of a pink noise that played as long as the cursor is within the target. The continuous auditory display constantly plays a sine wave that increases in frequency as the cursor gets closer to the target. The pitch increases as a function of the square (x^2) of current fraction of the total distance to the target that the cursor has traveled (Equation 1). The pitch increases one octave from the start to the target position. The continuous auditory display also played a pink noise when the cursor was within the target position.

$$Pitch_t = 440 + \left(\sqrt{440} * \left(1 - \left(\frac{|position_t - target\ position|}{target\ position}\right)\right)\right)^2 \quad (1)$$

2.6. Procedure

2.6.1. Practice

After providing informed consent and filling out a brief demographic survey, participants were first introduced to the general purpose of the experiment and given five minutes of guided practice during which they were exposed to each of the six different conditions. Participants were seated in a chair in front of a computer and a leap motion fixed at a 45 degree angle to the table. Participants were able to complete the task with their left or right hand but they were asked to not switch hands during the experiment. Participants were encouraged to take breaks between selections or conditions as needed.

2.6.2. Testing

After selecting the start button (open hand = select gesture), a target appeared somewhere to right on the screen. For visual conditions, participants can see the cursor and target, which changes color when the cursor enters it. For non-visual conditions the cursor and target are not visible. For each of the six conditions participants completed a total of 48 selections, 12 for each of 4 difficulty levels (ID = 2, 3, 4, 5). Each condition took about 6-8 minutes – the experiment lasted about an hour overall.

2.7. Statistics

Repeated-measures ANOVAs were conducted to identify differences between conditions. Two-tailed, paired-samples t-tests were conducted. A Holm-Bonferroni correction was used to decrease the number of Type-1 errors.

3. RESULTS

3.1. Selection time

Repeated measures ANOVA results indicate main effects for condition, $F(5,19) = 36.4$, $p < .001$, as well as difficulty $F(3,21) = 14.9$, $p < .001$. There was also a significant interaction, $F(15,545) = 3.83$, $p < .001$,which can be seen as a difference in slope of the lines in Figure 2. Paired comparisons (Table 2) showed participants were slower to make selections when using continuous (AC) and discrete (AD) auditory displays compared to conditions with visual displays and the control condition (Figure 2).

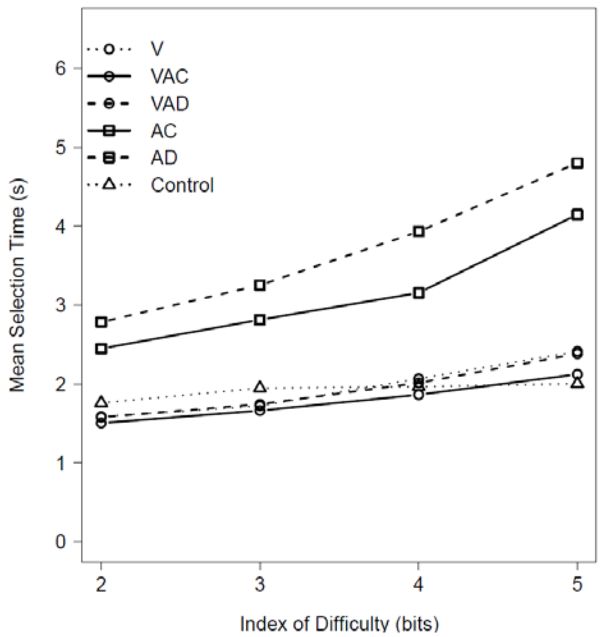

Figure 2: Average selection times for each condition across difficulty levels.

	AC	AD	Control	V	VAC
AD	<.001*	--	--	--	--
Control	<.001*	<.001*	--	--	--
V	<.001*	<.001*	1.00	--	--
VAC	<.001*	<.001*	.011	.057	--
VAD	<.001*	<.001*	1.00	1.00	.107

Table 2: P-values for pairwise comparisons of average selection times.

3.2. Selection accuracy

3.2.1. Error

Repeated measures ANOVA results show a main effect by condition, $F(5,19) = 34.4$, $p < .001$. Difficulty also showed main effects, $F(3,21) = 54.3$, $p < .001$, as well as an interaction with condition, $F(15,345) = 10.0$, $p < .001$, which can be seen by the difference in slopes of lines in Figure 3. Paired comparisons (Table 3) showed that participants' selection error, defined by the absolute value of the distance between the final cursor position and the closest edge of the target, was significantly higher for the control condition (Figure 3) compared to all other conditions. These tests also revealed that the discrete auditory display (AD) led to significantly higher error compared to all conditions other than the control. The AC condition led to significantly higher error compared to all conditions other than AD, VAD, and Control. All other conditions were statistically equivalent.

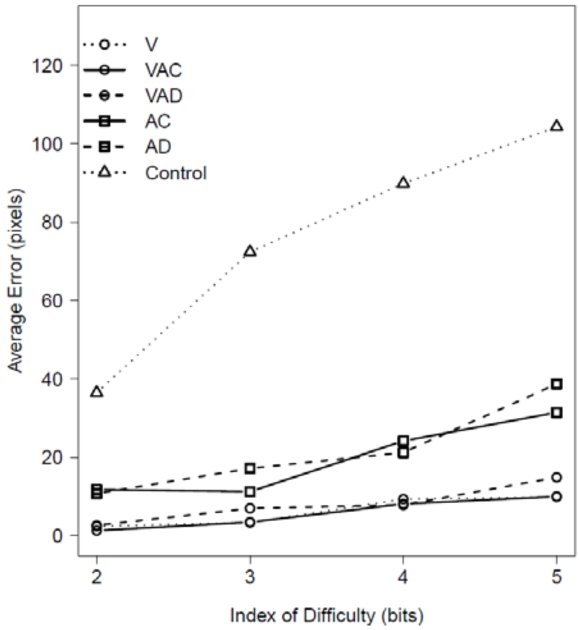

Figure 3: Average adjusted error for each condition across difficulty levels.

	AC	AD	Control	V	VAC
AD	1.00	--	--	--	--
Control	<.001*	<.001*	--	--	--
V	0.007*	0.002*	<.001*	--	--
VAC	0.006*	0.001*	<.001*	1.00	--
VAD	0.019	0.005*	<.001*	1.00	1.00

Table 3: P-value for pairwise two-tailed t-tests for selection error across conditions. * indicates a statistically significant difference.

3.2.2. Percent correct

ANOVA results showed a main effect for condition, $F(5,19) = 67.1$, $p < .001$. Figure 4 shows that conditions appear to largely be similar with the exception of the control condition which is significantly lower, which can be seen in Table 4. Difficulty also showed a main effect, $F(3,21) = 472$, $p < .001$, which is especially obvious in Figure 4. There was also an interaction between condition and difficulty, $F(15,345) = 2.99$, $p < .001$. There appears to be some separation between visual and non-visual displays at higher difficulties. Possible explanations will come in the discussion section.

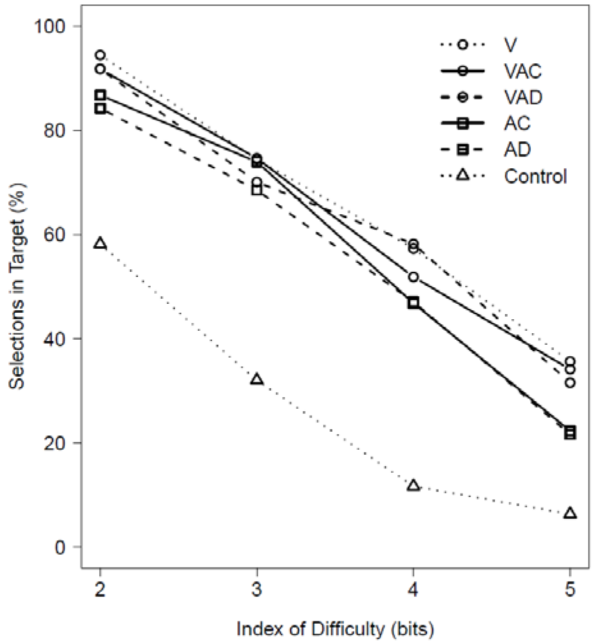

Figure 4: Average percent correct across conditions for each difficulty.

	AC	AD	Control	V	VAC
AD	1.00	--	--	--	--
Control	<.001*	<.001*	--	--	--
V	0.57	0.25	<.001*	--	--
VAC	0.76	0.43	<.001*	1.00	--
VAD	0.89	0.57	<.001*	1.00	1.00

Table 4: P-values for pairwise two-tailed t-tests for selection accuracy across conditions. * indicates a statistically significant difference.

3.3. Throughput

Throughput is a calculation that accounts for both the accuracy of the movement – difference between endpoint position and the center of the target – and movement time. This provides a measure of overall movement performance by information conveyed in bits per second. Repeated measures

ANOVA results showed significant differences between conditions $F(5,19) = 91.4$, $p < .001$ for throughput. Paired comparisons showed that participants had higher throughput with visual displays (VAC, VAD, V) compared to auditory-only conditions (AC, AD, Control) (Table 5). There was also a main effect for index of difficulty (ID), $F(3,21) = 35.0$, $p < .001$, and a statistical interaction, $F(15,345) = 2.04$, $p = .0122$.

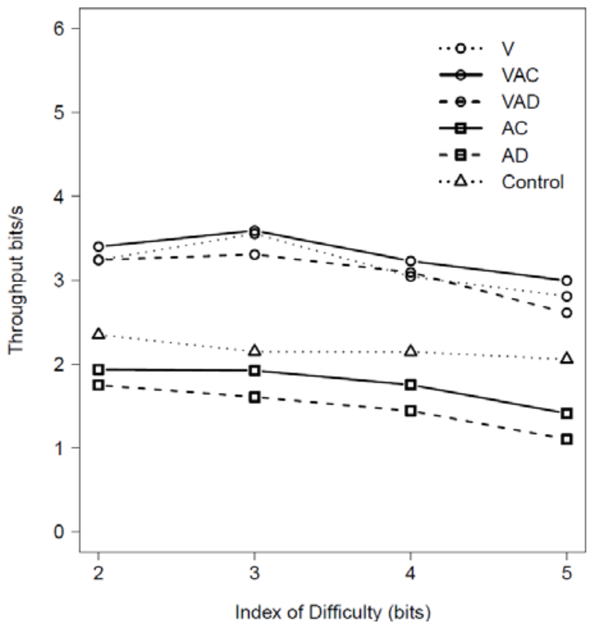

Figure 5: Average throughput across conditions for each difficulty.

	AC	AD	Control	V	VAC
AD	0.019	--	--	--	--
Control	<.001*	<.001*	--	--	--
V	<.001*	<.001*	<.001*	--	--
VAC	<.001*	<.001*	<.001*	.304	--
VAD	<.001*	<.001*	<.001*	.323	.047

Table 5: P-value results for pairwise two-tailed t-tests for throughput across conditions. * indicates a statistically significant difference.

4. DISCUSSION

These results convey a nuanced story about the influence of auditory displays on aimed movements using air gesture controls. As expected, visual displays resulted in faster movement times compared to auditory displays. Previous literature has shown that visual information is more readily integrated into trajectory corrections [10], suggesting that using auditory displays to convey information about movement trajectory is more effortful than with visual displays. One possible explanation for the difference in selection times between visual and auditory-only displays is that people are better able to accurately estimate the distance to the target and close the gap more quickly in the initial ballistic phase of movement when using visual displays, as opposed to the auditory-only displays which require more searching behavior. Regarding selection accuracy, however, auditory-only displays led to similar percentages of in-target selections compared to conditions with visual displays, especially at lower levels of difficulty. Interestingly, auditory-only displays consistently resulted in a statistical interaction, showing slower and less accurate movements, especially for much higher difficulty movements (ID = 4, 5). We suppose that the relatively poor performance for auditory-only displays for selection times may be because participants are receiving less information about the relative position of the cursor and the target, and as a result, need to make more fine motor corrections once they are close to the target.

Comparing between the continuous and discrete auditory-only displays, the continuous auditory display led to faster selection times and comparable accuracy, leading to overall higher throughput. The same pattern was not as clear when comparing continuous and discrete audio paired with visual displays (VAC and VAD), possibly as a result of participants deferring to visual information when it is available.

Overall, results indicate that auditory-only displays are not as effective as visual displays at guiding aimed movements in target acquisition tasks among sighted users. However, the data suggest that targets can be selected with similar levels of accuracy when using auditory-only displays, especially when movements are less difficult (ID = 2, 3). This suggests the potential for using auditory displays (continuous or discrete) for facilitating eyes-free target acquisitions using air gesture controls. For example, in vehicle contexts, auditory-only displays can result in the same accurate performance in the secondary gesture task, while maintaining visual attention on the road. Therefore, further applied research is required to identify the relationship among the task demand (e.g., level of difficulty), multi-modalities, and different types of auditory displays.

5. REFERENCES

[1] Hatfield, Brent C., William R. Wyatt, and John B. Shea. "Effects of auditory feedback on movement time in a Fitts task." *Journal of motor behavior* 42.5 (2010): 289-293.

[2] de Grosbois, John, Matthew Heath, and Luc Tremblay. "Augmented feedback influences upper limb reaching movement times but does not explain violations of Fitts' Law." *Frontiers in psychology* 6 (2015).

[3] Soukoreff, R. William, and I. Scott MacKenzie. "Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI." *International journal of human-computer studies* 61.6 (2004): 751-789.

[4] Pierno, Andrea C., Andrea Caria, and Umberto Castiello. "Comparing effects of 2-D and 3-D visual cues during aurally aided target acquisition." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 46.4 (2004): 728-737.

[5] Pierno, Andrea C., et al. "Effects of increasing visual load on aurally and visually guided target acquisition in a virtual environment." *Applied ergonomics* 36.3 (2005): 335-343.

[6] Zahariev, Mihaela A., and Christine L. Mackenzie. "Auditory contact cues improve performance when

grasping augmented and virtual objects with a tool." *Experimental brain research* 186.4 (2008): 619-627.
[7] Marentakis, Georgios, and Stephen A. Brewster. "A study on gestural interaction with a 3d audio display." *International Conference on Mobile Human-Computer Interaction*. Springer Berlin Heidelberg, 2004.
[8] Grossman, Tovi, and Ravin Balakrishnan. "Pointing at trivariate targets in 3D environments." *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2004.
[9] Cockburn, Andy, et al. "Air pointing: Design and evaluation of spatial target acquisition with and without visual feedback." *International Journal of Human-Computer Studies* 69.6 (2011): 401-414.
[10] Elliott, Digby, Werner F. Helsen, and Romeo Chua. "A century later: Woodworth's (1899) two-component model of goal-directed aiming." *Psychological bulletin* 127.3 (2001): 342.