AR Hand Gesture Capture for Interactive Data Analytics

Research Overview

Salang Kang | Robert Lam | Leon (Ang) Li

Also see the attached **web application** and **unity application** documentation.

Overview

Background

Prior to this research, there have been several studies on immersive data analytics and manipulating visualisations in VR or AR space. Maxime Cordeil, a researcher from Monash, has an example of this in his ImmersiveAxes project, which is largely a motivation for our own research project. As the field of immersive analytics grows, so does the need for more intuitive user control and fluid user experience within these platforms. Certainly, there exists prior research regarding elicitation studies and intuitive hand gestures for basic operations in VR/AR space; however there lacks research on more specific data visualisation techniques, such as filtering data.

Research Aims

The ultimate aim of this research project is to answer the research question, “what are the intuitive hand gestures that people would default to when interacting in AR/VR with data?” However, for our project, we were mostly focused on developing software that would assist with recording and analysing data for this purpose.

As such, our main objectives were to develop a **web application** to record data and a **separate, standalone application** to analyse the collected data and produce a measure of how similar the recorded gestures were. To accomplish this, we also needed to have a **gesture set** to collect data on and collect a pilot study to prove that our software works for the purposes of research. From a small-scale test, we should be able to confirm that we can measure the similarity of different gestures.

Significance

The development of new software to assist with recording and analysing gesture data will help researchers discover what the most intuitive gestures are for interacting with data. Data analytics is moving towards a more immersive field as AR and VR develop further, which creates a need for standard gestures that are comfortable and easy to use. Currently, there is a lack of research in what gestures can be used for more specific and niche data visualisation controls, such as filtering or exporting data. This research project hopes to bring further insight into which gestures users would be more comfortable with to perform various operations.

Gesture Set and Elicitation Studies

The gesture set was obtained after researching into other papers for various taxonomies and elicitation studies[1]. Three main categories were determined to be the most important for data visualisations:

TABLE 1: Taxonomy of interactive dynamics for visual analysis

Data & View Specification	**Visualize** data by choosing visual encodings. **Filter** out data to focus on relevant items. **Sort** items to expose patterns. **Derive** values or models from source data.
View Manipulation	**Select** items to highlight, filter, or manipulate them. **Navigate** to examine high-level patterns and low-level detail. **Coordinate** views for linked, multi-dimensional exploration. **Organize** multiple windows and workspaces.
Process & Provenance	**Record** analysis histories for revisitation, review and sharing. **Annotate** patterns to document findings. **Share** views and annotations to enable collaboration. **Guide** users through analysis tasks or stories.

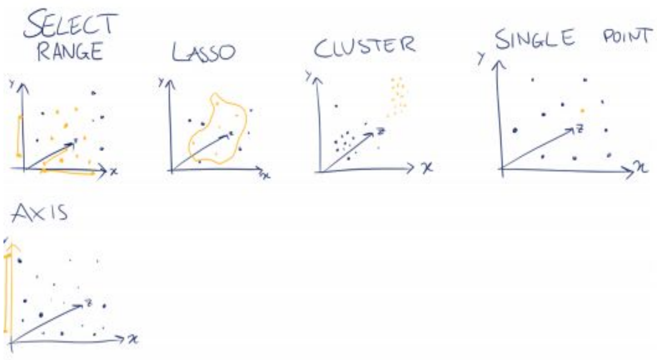
- Visualisation control
- Camera/view control
- Data annotation

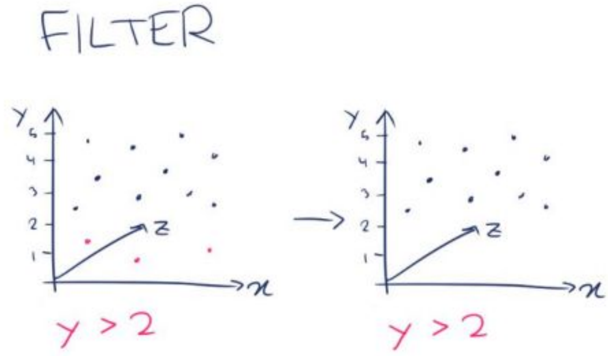
From this and various other references listed, we decided to come up with 13 different gestures to study. Similar studies conducted have around 20[4], with some having less and some having more. As the study is meant to gather data from a number of different volunteers, the focus is to be as less time-consuming as possible. Lowering the number of gestures would stop discouraging volunteers from conducting the test upon seeing the number of gestures available.

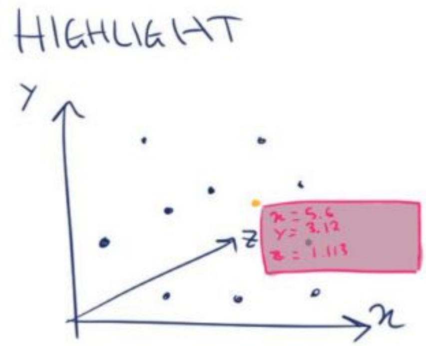
Rough sketches of each gesture was created, which were later converted into more meaningful animations through Unreal Engine 4, implemented into the Web Application.

These animations were modelled through Unreal Engine 4's built-in primitive shapes and programmed, before screenshots were taken frame-by-frame to generate an animated GIF. Some of the instructions did not need to be as detailed and were less animated compared to others.

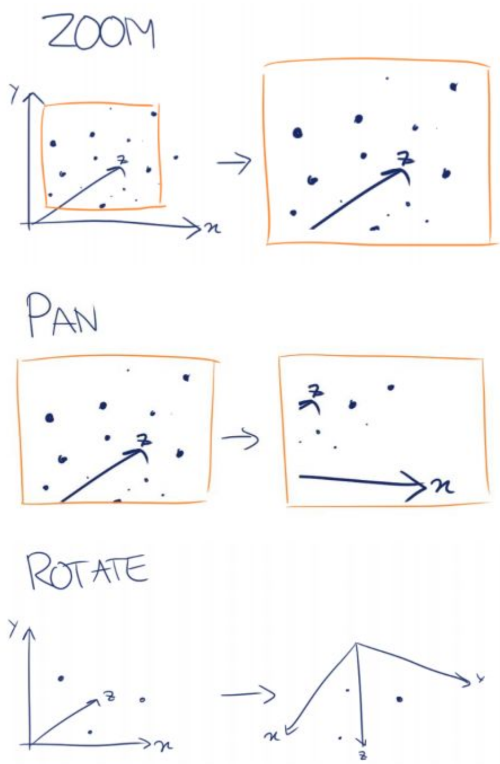
Visualisation control (selection and filter/highlight):



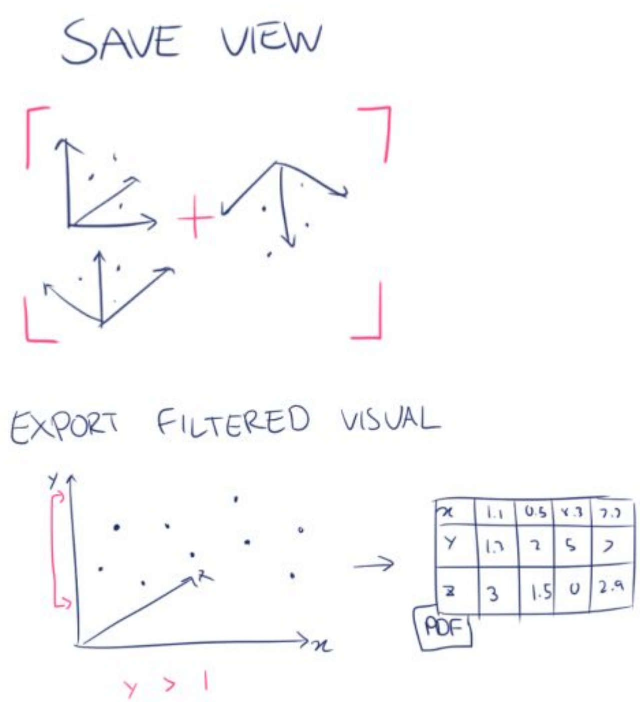




Camera/view control



Data annotation and export

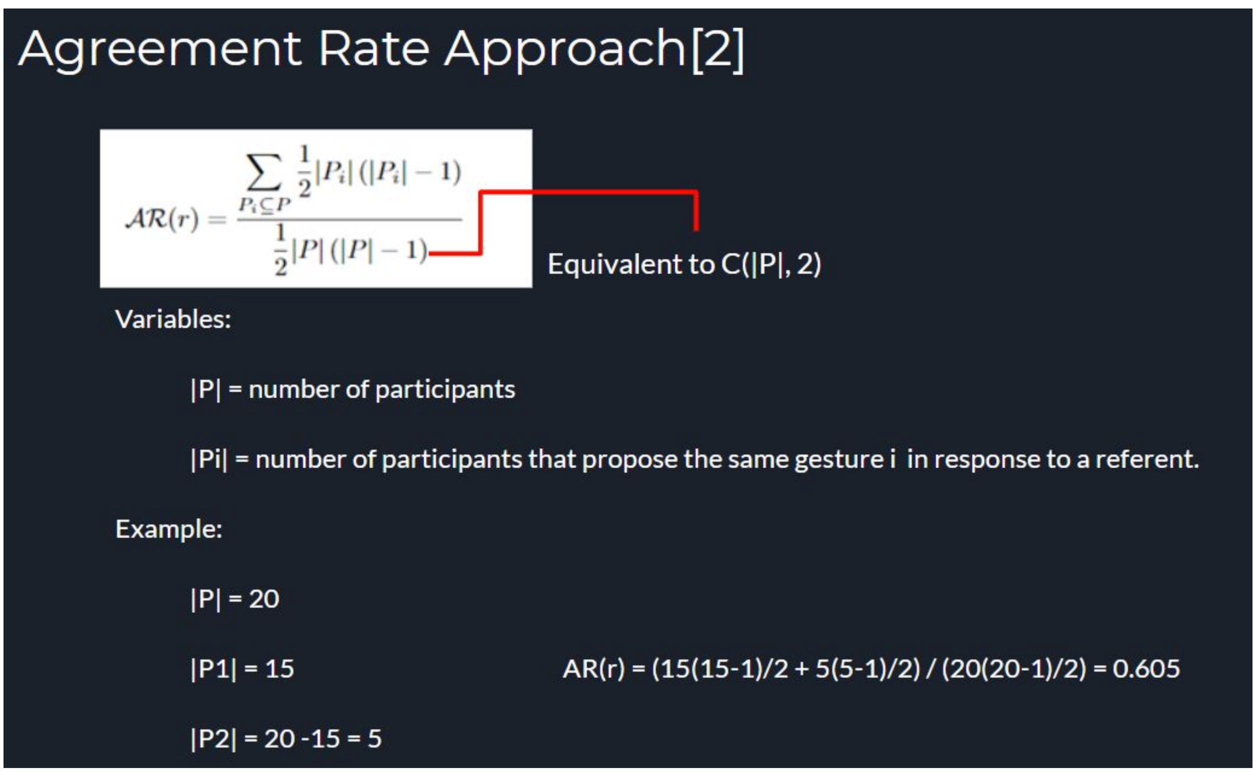


Similarity Measures and Agreement

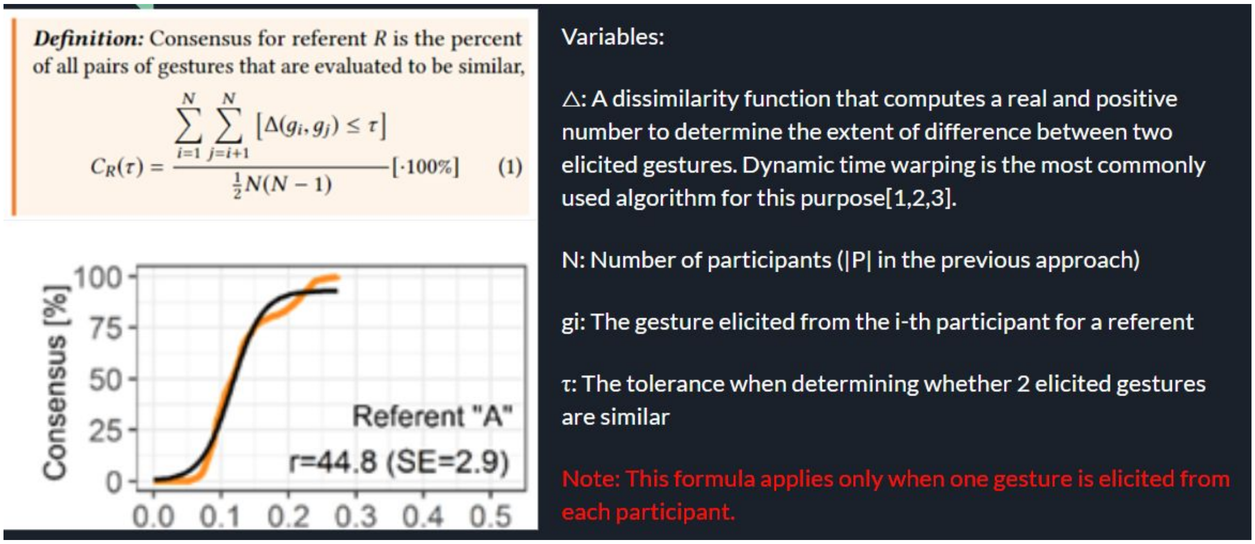
In order for the multitude of data collected to be useful, study on similarity measures and agreement were conducted.

Two main approaches were discovered - the **Agreement Rate** and the **Dissimilarity-Consensus Approach**. For the purposes of this study, we are utilising the latter, **Dissimilarity-Consensus Approach**.

For the agreement rate, the focus involved using a formula to determine the number of gestures that are similar and the ones that are different. However, the issue is that there lacks a unified function to measure similarity, which is required to remove bias for these gestures and provide a more quantitative analysis. Determining this would also take a significant amount of time due to the large number of participants required combined with the need to manually overview each gesture.



Instead, we look to the Dissimilarity-Consensus Approach, which focuses completely on computer-based analysis. In this method, we use a distinct similarity function to calculate how similar each of the recorded gestures are, based on their positioning and movement. Due to differences in time, dynamic time warping is a common algorithm that we use to normalise the motions. The analysis prioritises on ultimately calculating the Euclidean distance between each gesture to determine their similarity rating.



See also the attached presentation and write-up regarding the research conducted on similarity measures.

Results and Pilot Study

Throughout the course of our research, we developed a **Web Application** and a **Unity Application.**

The **Web Application** is purposed to be able to record and output data to a server in the form of .csv files. These files contain information about each individual joint over a number of frames, which would then be normalised for further analysis in the Unity application. Further information can be found in the Web application documentation.

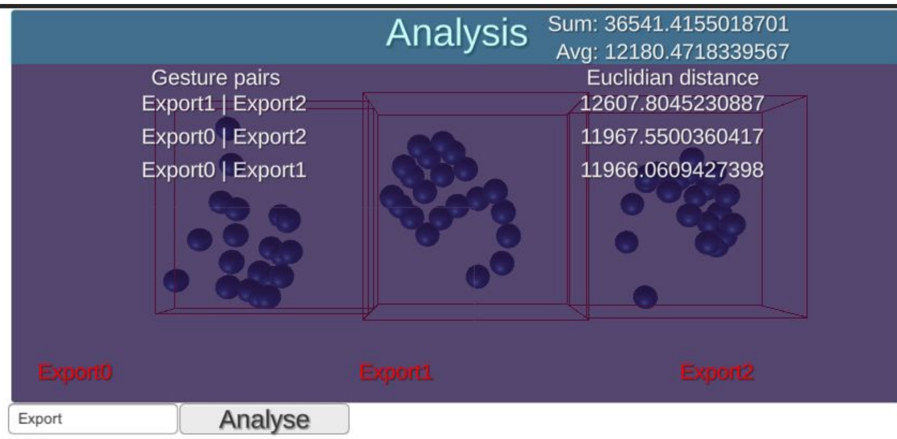
The **Unity Application** receives recorded data and is able to both reproduce the visualisation as well as analyse the similarity measures between them. Using the Unity application, the researcher is able to discern which is the most intuitive gesture for any particular operation. Further information can be found in the Unity application documentation.

A brief pilot study was conducted with the three of us involved in the research. We performed various gestures to achieve some results, analysed in the Unity application.

The sample data retrieved from the Web Application is in a csv format. It involves the time in milliseconds representing each frame as one column, with each subsequent column representing X, Y and Z coordinates of each joint as shown below.

	A	B	C	D	E	F	G	H	I	J
1	TIME	JOINT_0_X	JOINT_0_Y	JOINT_0_Z	JOINT_1_X	JOINT_1_Y	JOINT_1_Z	JOINT_2_X	JOINT_2_Y	JOINT_2_Z
2	26	254.4888931	240.1158243	0.000396691	266.3976667	247.1003134	-24.52472496	258.2266142	257.0852384	-39.43568039
3	39	254.4888931	240.1158243	0.000396691	266.3976667	247.1003134	-24.52472496	258.2266142	257.0852384	-39.43568039
4	53	264.2181854	247.2253002	6.05807E-05	265.07464	241.9124666	-21.21443367	245.9120601	243.9997965	-32.67797089
5	68	264.2181854	247.2253002	6.05807E-05	265.07464	241.9124666	-21.21443367	245.9120601	243.9997965	-32.67797089
6	85	267.2472666	251.1128689	0.000205688	256.4520082	242.3615727	-13.68378639	234.3376724	238.9132209	-20.13501167
7	102	267.2472666	251.1128689	0.000205688	256.4520082	242.3615727	-13.68378639	234.3376724	238.9132209	-20.13501167
8	117	257.6686392	239.5341091	-0.000672556	225.5725061	238.4866781	19.40099907	192.7702307	239.4176834	28.31077385
9	135	257.6686392	239.5341091	-0.000672556	225.5725061	238.4866781	19.40099907	192.7702307	239.4176834	28.31077385
10	299	228.5856078	265.624856	-0.000464171	184.1729079	265.0068837	-16.19651604	137.5751218	260.4734416	-18.90924644
11	326	225.4846356	264.2463847	-0.000815675	185.2939086	269.9210337	-20.55550957	130.8327556	268.6641764	-24.94201851
12	352	231.4258063	280.3667605	-0.00123585	186.1669904	282.9287009	-15.50327969	125.9835637	277.9801636	-19.02223969
13	366	234.6441125	278.3234672	-0.001833573	183.4004218	282.3595315	-14.15267086	129.9177561	279.0102392	-16.39928055
14	418	197.3193926	289.0173272	-0.000674464	158.1760388	285.1163849	-15.66781902	113.0171366	282.6875084	-18.309412
15	445	206.3393557	295.529041	-0.001288146	161.2581933	295.2209096	8.627953529	124.2049343	292.5296335	14.23631287

This data is then imported into the Unity application for further analysis.



As we can see, we are able to measure the Euclidean distance and have a similarity rating for various different gestures. From this, we can analyse and discover what is the most intuitive gestures to use for various gesture operations, as we came up with in our initial gesture set.

Our implementation of two functional programs that can record, reproduce and analyse data, as well as our initial pilot study to confirm that it works, prove that we satisfied our initial aims of developing software to help researchers discover the most intuitive hand gestures in immersive data analytics.

References

Taxonomy for gesture set:

[1] Jeffrey Heer and Ben Shneiderman. 2012. Interactive dynamics for visual analysis. Commun. ACM 55, 4 (April 2012), 45–54. DOI:https://doi.org/10.1145/2133806.2133821

Further references for gesture sets:

[2] Maxime Cordeil, Andrew Cunningham, Tim Dwyer, Bruce H. Thomas, and Kim Marriott. 2017. ImAxes: Immersive Axes as Embodied Affordances for Interactive Multivariate Data Visualisation. In Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17). Association for Computing Machinery, New York, NY, USA, 71–83. DOI:https://doi.org/10.1145/3126594.3126613

[3] Christopher R. Austin, Barrett Ens, Kadek Ananta Satriadi, Bernhard Jenny (2020) Elicitation study investigating hand and foot gesture interaction for immersive maps in augmented reality, Cartography and Geographic Information Science, 47:3, 214-228, DOI: 10.1080/15230406.2019.1696232

[4] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-defined gestures for surface computing. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09). Association for Computing Machinery, New York, NY, USA, 1083–1092. DOI:https://doi.org/10.1145/1518701.1518866

Similarity measures and analysis:

[5] Villarreal Narvaez, Santiago; Vanderdonckt, Jean ; Vatavu, Radu-Daniel ; Wobbrock, Jacob O.. A Systematic Review of Gesture Elicitation Studies: What Can We Learn from 216 Studies?. ACM conference on Designing Interactive Systems (DIS '20) (Eindhoven, du 06/07/2020 au 10/07/2020). In: Proceedings of ACM conference on Designing Interactive Systems (DIS '20), ACM Press : New York 2020

[6] Vatavu, R.-D., & Wobbrock, J. (2015). Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 2015, 1325–1334. https://doi.org/10.1145/2702123.2702223

[7] Vatavu, R.-D. (2019). The Dissimilarity-Consensus Approach to Agreement Analysis in Gesture Elicitation Studies. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1–13. https://doi.org/10.1145/3290605.3300454