EagleView: Design of People and Devices Visualizations for Analysing Interactions in Tracked Space

Suppachai Suwanwatcharachat
HCI MSc Final Project Report 2017
UCL Interaction Centre, University College London
Supervisor: Nicolai Marquardt, Frederik Brudy

Figure 1. EagleView receives video and high-level data of people in tracked space (left), then create visualizations on a web-based platform to quicken the analysis of recorded video (right)

ABSTRACT
Data related to physical space such as location of people, devices, and objects can be used to implement novel interactions in ubiquitous computing. It is common that the studies of these interactions involve a group of people doing different activities at the same time, requiring a video to be recorded for later analysis. However, video analysis process is time-consuming due to the need for repeated video playback to search for interesting activities. To aid researchers to quicken analysis of videos, we developed two types of web-based visualization tools. In particular, the real-time visualization showed information about people’s movement and actions based on five proxemic dimensions (distance, orientation, movement, identity, location) and another indicator for people’s attention. Overview visualization described activities over a time period in a single image. The visualizations were implemented using user-centred design method. An evaluation with seven expert users revealed that most visualization components were easy to understand and beneficial for video analysis. Real-time visualization provided a clean and objective representation of data, while overview visualization added a time dimension which transformed the data to reveal new meaningful information, for instance; a position heat map used a collection of people’s movement data to show the location where people stayed for the longest period; a movement trajectory created a continuous line which showed the patterns of how people navigated around the space. Through our study, we proposed an overview visualization on the timeline which would allow users to quickly search for interesting events without the need for video playback.

Author Keywords
Proxemic interaction; interaction analysis, tracked environments; visualization

ACM Classification Keywords
H.5.1 Multimedia Information Systems: Video; H.5.2 User Interfaces: Screen design, prototyping

MSC HCI FINAL PROJECT REPORT
Project report submitted in part fulfilment of the requirements for the degree of Master of Science (Human-Computer Interaction with Ergonomics) in the Faculty of Brain Sciences, University College London, 2017.

NOTE BY THE UNIVERSITY
This project report is submitted as an examination paper. No responsibility can be held by London University for the accuracy or completeness of the material therein.

Video Submission
https://youtu.be/1_NNMxu2qRE

Link to Source Code
https://github.com/strkfreedom/EagleViewVis
1. INTRODUCTION
The study of user interactions in a space is important for developing technologies in ubiquitous computing. Data related to physical space such as location of people, devices, and objects can help to implement novel interactions. Examples can be seen in context-aware systems, which use people’s location data to deliver information that is just-in-time and more relevant to user’s needs [27]. Further, the position data can enable a device to detect the presence of nearby devices and offer possible interactions to users [28]. Not only can this information be used to create new interactions, it can also be used to make an analysis on how people use applications and devices in private and public settings [37], or how supporting technologies are used in collaboration work environment such as surgery room [34].

The process to conduct such studies involves recording and analysing video. Using supporting tools, the recording of people’s movements and interactions can be done automatically. However, the analysis of this data remains a time-consuming manual process. In practice, video analysis involves multiple play through. The first play through would get an overview of the events, followed by repeated views on a specific part of the video. The repetitions are required to visually search for particular activities and identifying interesting behaviour. Because of these repetitions, the length of the video analysis process can take over two times the length of the video footage [20]. Existing video analysis tools do not help to reduce the need for repeated playbacks, but rather to facilitate the workflow by providing capabilities for video annotation and coding [10,12,17]. To summarize, video analysis was time-consuming due to the lack of support for visual search activities in video analysis tools. With the described problem, this research intended to address the following research question:

How can we design spatial visualizations that support visual search activities during video analysis of an interaction between multiple people and devices in a tracked space?

We created a web-based prototype showing visualizations in action. The application received high-level spatial data of people and devices from top-view Kinect cameras through EagleSense’s API (Figure 1, left) [47]. The left window showed real-time visualizations which translated five proxemic dimensions (distance, orientation, movement, identity, and location) along with their spatial relationship into visualizing components: head, body, sight, area around body, phone and tablet indicators, distance line between people or objects, spacing zones, future and past movement, colour grouping of people-to-people and people-to-device interactions (Figure 1, right) [12]. The right window showed overview visualizations, which provided a summary view of key data (Figure 1, right). Movement trajectory was an overview of movement pattern, while heat
The design was evaluated by expert users who had experiences in video analysis. Following an application training, participants tried out the prototype in a mock scenario-based task. Questionnaire and semi-structured interview were conducted to gather feedback on usability and usefulness of the tool.

The result showed that the application was easy to use for expert users. The meaning of each visualization components was clear and straightforward, except for two components. The first component was grouping visualization. When two people and a display were grouped together, a question arose whether both people were interacting with each other as well as interacting with a display, or both people were individually interacting with a display. A new design using two separate groups was proposed to make the meaning of the visualization clearer. The second component was distance visualization. There was a perspective distortion from a camera, which pushed the head of a person outward from the centre of the camera and from the person’s feet. This distortion created a reliability issue for an accurate distance measurement. To reduce the distortion, orthographic projections could be used.

Overall, the tool was found to be beneficial for video analysis. Real-time visualizations could be used as a clean representation of raw video that provided fundamental spatial information, while the overview visualization provided new insights that were difficult to extract from watching playbacks. For instance, it could reveal the different movement pattern of people who came to interact with a device in public settings, and that of the observing audiences around the first person. Driven by the concept of overview visualizations, a design of additional summary visualizations on the timeline was proposed. They would allow users to filter interesting parts of the video without video playback, which in turn, speed up the video analysis.

The contributions of this research are 1) creating a design which visualized key dimensions of proxemic interaction which translated data into easy to understand graphics. 2) By conducting user study, we described how users would integrate the visualizations into video analysis workflow. In particular, real-time visualizations could accompany traditional video playback as a clean presentation, and overview visualizations could transform the data by adding time dimension to create meaningful information. 3) A proposed design of overview visualization on the timeline could speed up video analysis by helping users to skip the first play through in the video analysis process.

2. LITERATURE REVIEW
This research is grounded in three areas of related literature (i) proxemic interactions (ii) proxemic-aware systems (iii) tools supporting interaction analysis.

2.1 Proxemic interaction
The study of interaction in a space involves the use of space, or in other words, proxemics interactions. The study of proxemics is described by Edward Hall as how people perceive the use of physical space when they are communicating or having social interactions with others [15]. People may use this understanding to adjust their position closer or further according to the underlying social meaning of different distances. Hall’s four proxemics zones proposed four different distance intervals and their corresponding social meaning (Figure 2). Intimate distance (0 to 50 cm) is used for close interaction such as touching or whispering. Personal distance (50 to 100 cm) is used for interactions with friends and family members. Social distance (1.0 to 4.0 m) is used for interactions with acquaintances. Public distance (4.0m or more) is used for public speaking. Although each zone had a specific description that provided specific distance, there might be some deviations to how individuals would perceive their use of space based on their own experiences and cultural background. Hall mentioned that the Americans and the Arabs did not share similar thoughts on how they interpret the distance during communication. Similarly, in a later work, Joosse et al. tried to measure a suitable distance to which a robot should maintain when engaging with people [21]. The study found Chinese people preferred a closer distance when compared to Americans people due to Chinese’s high-contact culture. To address the variety of people’s perceived use of distance, it is important that the zone distance in our visualization can be adjusted to match individual’s perception.

![Figure 2. Hall’s four proxemics zones [31]](image)

Distance is not the only measurable factor in proxemics interaction, rather, there are five key proxemics dimensions which are critical for the analysis of proxemics interaction (Figure 3) [12].

1. Distance is the measurement of space between people to people, people to interactive devices, and people to other objects. As described earlier, in Hall’s proxemics zones, distance is used to categorize four distinguish spaces used in social activities (Figure 2).
2. **Orientation** is the direction in which people and devices are facing. Low-level information is an angle a person is facing relative to a display. High-level information is whether a person is facing directly toward, partially toward, or away from the display. This measurement can be an indicator for a person’s center of attention at a moment. For example, a person directly facing a TV might indicate that the TV is currently watched by that person [26]. Conversely, a person facing away from a TV might indicate the lack of attention.

3. **Movement** is the combination of distance and orientation over time. The time dimension can be used to determine the speed of different motions such as walking or turning, which can indicate an attention. For example, a person who is walking pass a billboard screen and slowing down to a halt might indicate that the attention is caught by the billboard. In a similar way, a movement toward a large wall display indicates that the person intentionally approaches the display [7].

4. **Identity** is a unique indicator to distinguish one person or one digital device from another. This information is crucial in order to provide personalized service. A video application might resume the last movie you watched last session [28].

5. **Location** is the environmental context of the physical settings such as a door used to enter a room, a wall behind a display. Different designated areas can also be a type of location, which provide additional context for proxemics interaction, for instance, people enter activation space which allows them to see a large media façade before moving into interaction space which allows them to interact [10].

Around this area is the P-space, where people are located. The surrounding area around people is called r-space. People in R-space are unlikely to interact with the people in P-space. If other people wish to join the interaction, they need to relocate themselves into P-space. For instance, the O-space of two people facing directly to each other is the area between them. The O-space of 3 people standing in a circle is the inner circle formed by their positions. Other types of formation include side-by-side, or corner-to-corner orientations. The understanding of these formations can be used to determine whether people are interacting individually or in a group.

As an example of how f-formation can be used to analyze interactions, Paul et al. used the observation of f-formation to identify social interactions around a public tourist information center [33]. At the information counter, tourists who were in p-space were directly involved in the discussion with a helping staff, while people who unintentionally position themselves in r-space, because of the lack of space, did not directly participate in the conversation. It is important that people stay in f-formation to be included in the ongoing activity. This concept also applied to devices. When a group of people was accessing an information on a display, they included the device into the f-formation [Figure 5]. The analysis of f-formation showed how people spatially aligned themselves when interacting with technologies, which should be considered when designing new interactions.

Figure 3. Five proxemics dimensions [12]
These key proxemic dimensions provide a good coverage of information on an individual person or object, but not their interactions. F-formation, as explained by Kendon, describes how a group of people use and share physical spaces [22]. When forming a group, people arrange their positions in a way that their working spaces overlap, creating a joint shared space, which is called O-space.

Figure 4. F-formation consists of shared space (O-space), occupied space (P-space), and outer space (R-space) [22]

Although f-formation seems straightforward to manually identify when one occurs, computers rely on multiple factors to detect f-formation.
1. **Open or close formation** is determined by the position and orientation of each person. Open formation is seen in either side-by-side or corner-to-corner orientation, while close formation is seen in a complete circle.

2. **Shape of personal space** can be circular, ellipse, or other irregular shapes such as hour glass. Ellipse shape arguably reflects a realistic expectation of people that they maintained longer distance in the rear compared to the front and the side for safety [36]. Hour-glass shape extends the distance more to the front and the back than to the side [42].

3. **Social relationship** determines the distance they maintain during the interaction whether they are friends or strangers, as well as competing or collaborating.

Despite the multiple factors involved in f-formation, it is possible to accurately calculate the shape and size of f-formation in real-world settings. Setti et al. invented a Graphs-Cut for F-Formation (GCFF), an algorithm for detecting f-formation in public spaces [40]. GCFF simplified input data to only 2 variables: head position and its orientation. The algorithm then determines whether a person belongs to any group and calculate the area taken by the f-formation. With GCFF, the possibility is opened for researchers to apply the automatic detection of f-formation to help them in the interaction analysis.

### 2.2 Proxemic-aware systems

This section provides examples of technology which implemented the knowledge of how people use physical spaces. It was important that we knew how proxemics data could be used to create new interaction experiences.

**Attention grabber displays.** Many studies used people’s position data to create an interaction that seemed like people themselves are controlling what was appearing on the display. Then, people who noticed the interaction got their attention captured by the system. In a past study, a display placed in a museum detected body movement of people walking by and used them as a cursor for browsing through multiple pictures [26]. As a result, people noticed that their motion caused the screen to react accordingly, and became interested to interact with it. Another experiment tried not only to attract the attention but also to retain it as long as possible. Cheung and Scott [6,7] were able to attract and engage pedestrian in-situ walking past a display. They manipulated the direction and speed of floating pictures on the screen and used silhouette that mimicked the exact shape of user’s body. As the person got closer to the screen, the pictures moved slower and changed its direction, as well as increased the size of the silhouette.

**Proxemic-aware media player.** Other researchers used the proxemic information to introduce new methods to interact with displays. A proxemic-aware media player placed in a living room could detect the presence of a person who had just come into the room and turned itself on [1]. Further, the display recognized the person and showed the personal media collection. It was also able to recognize whether a person was giving attention to the display. If a person was receiving a phone call, the media playback would pause and wait until the call was done, then automatically resumed.

**Transfer of data between devices.** Marquardt et al. introduced a new method to transfer data between mobile devices [28]. It allowed users to see the presence of nearby devices which were likely needed for usage. It used a combination of proxemic dimensions as criteria to determine whether to present the visual indicator of nearby devices. The criteria filtered devices which were in the same room, in close distance, moving toward each other, and facing each other or standing side-by-side. This criterion helped to present the right devices to connect at the right moment.

**Possible proxemic interactions in a surgery room.** In an analysis of proxemic data to uncover key findings that could be used to solve issues faced by surgeons, Mentis et al. noticed the lack of accessibility to computers located out of reach during the operation where surgeon’s movement was restricted [34]. Mentis et al. suggested to use body movement as a controller and a pointer for the computer. A further suggestion was that the application should enlarge the images shown if the surgeon was detected to be far away.

**Proxemic-based game.** Delden et al. created a tagging and running game for children based on the location of each person in an indoor playground [8]. Distance, orientation, identity, movement, and location were used to analyze different strategies used by taggers and runners.

**Privacy protection.** Proxemics data could be used to detect privacy invasion and protected users from intrusive behavior. Brudy et al and Zhou et al. [3,49] considered the scenario of looking over a shoulder of a person who was using a large display or a tablet to see what was on the screen. They used the looker’s position, gaze direction, the user’s position and head direction to determine whether the behavior was deemed intrusive. If so, the tablet screen would adjust its brightness or hid part of its content to make it obscured and more difficult for the looker to see.

### 2.3 Tools supporting interaction analysis

Many applications were developed to assist the analysis of recorded video of interactions. Most of these tools provided basic functions such as video annotation and coding. VACA [5] offered a feature to select a portion of the video and insert short labels to create category coding directly on the timeline interface. VideoANT [18] provided a user-friendly interface for adding annotations to videos without obstructing the video playback windows. VCode and VData [14] gave graphical representations on its timeline for adding annotations of events which contained no duration or span over a period of time. VIPCAM [35] helped to analyze collaboration in multiple display environment by showing logs of applications a person used during a time period. ChronoViz [11] allowed researchers to integrate and
synchronize physical notes for video annotation. With the tools listed above, researchers have the necessary functions to make an analysis. However, these tools did not aim to speed up video analysis process.

To reduce the time required to watch playbacks of video, two research allowed users to search for specific parts of the video according to filtering conditions. 1) EXCITE provided dynamic query function on related data field [30]. The query result showed portions of video which matched the specified conditions. The process was done automatically by the application in a short duration, which helped reduce the time that researchers need to manually seek the video. Although this query technique was fast, it requires users to know in advance the specific criteria to search. This drawback would be overcome by our tool which presented the information first, then let users figure out the criteria later. 2) Glance [25] was a crowd sourced video analysis tool that provided a flexible query using natural language. The system coordinated multiple people to respond to query questions on different parts of the video. This way, it achieved both flexibility and speed in the video analysis process. The limitation was that each worker needed to be paid for the time they spent working. Nonetheless, it could be a viable option if researcher had the budget required.

Other researchers created tools for analyzing interaction in a specific context. VisTACO [36] specialized in analyzing spatial interactions involved in tabletop collaborations. It developed evaluation methods for tabletop settings according to the key characteristics of the technology such as orientation, spatiality, and synchronous inputs. GIAnT [48] presented a tool for analyzing user interaction with large wall display. Related data included user’s gazes, distance and position from the wall, movements, and the number of touch inputs. While these tools facilitated analysis of tabletop and large wall displays, no tools have been developed to analyze group interactions involving the use of mobile and fixed devices.

2.4 Potential improvements from previous literatures

By reviewing previous literatures, we found that several improvements could be made, all of which was later implemented in EagleView. 1) While proxemic dimension data could be beneficial for interaction analysis, it was difficult to gain these proxemic data by watching people in the video. Moreover, no analysis tools supported the use of proxemic dimensions. 2) While f-formation could benefit the analysis of people’s relationship in a collaboration space, there was no interaction analysis software that would automatically recognize the formation. 3) While general video analysis tools enabled annotation and coding on videos, it did not help to quicken the time consuming process of video analysis. 4) Recent video analysis tools aimed to speed up the video analysis by specifying filtering conditions, which was suitable for top-down analysis where users had already known what they were searching.

However, the need to know specific conditions could be a limitation. The tools might not be suitable for bottom-up analysis approaches. 5) No specialized analysis tools were made for analyzing group interactions with multiple devices that could collaborate together in ubiquitous computing.

3. DESIGN METHODOLOGY

This chapter discusses the design requirements, design principals, and the design methodology used to justify the design decisions in the Design & Prototype section. Later in the discussion chapter, we summarized how our tool met the design requirements.

3.1 Design requirements

In this section, we establish design requirements based on the study of current video analysis tools. Our tool aimed to present new visualization techniques for proxemic data and support video coding process by making information easier to access through visualizations. Our tool did not aim to be a complete, end-to-end tool for video analysis. In particular, it did not focus on annotation, video timeline, or statistical analysis features. The following design requirements were referred in the later chapters as R1 to R6.

R1. Playback visualizations in real-time

Researchers might conduct video analysis either by using continuous playback or by quickly jump to points which they found interesting [14]. Likewise, the visualizations should be shown in real-time during a playback, video seeking, and jumping through time intervals. Further, it should be presented as a standalone, clean representation of the important things going in the video that helps researchers focus on what they are looking for, or an overlaying layer on top of the actual video to give additional proxemic information.

R2. Visualize individual’s proxemic information

Proxemic dimensions were considered highly relevant information for analyzing activities in a space [31]. Therefore, it was necessary to visualize 5 proxemic dimensions which were distance, orientation, movement, identity, and location. For distance information, it should be able to visualize both continuous distance and the movement when entering and leaving discrete zones [1]. The tool should also show the activities a person does. Ideally, it should visualize a wide range of activities such as standing, talking, sitting, using a phone, using a tablet, or interacting with other objects. However, in this research, we did not focus on the algorithm to identify different activities. We only visualized the activities provided by data capturing tool.

R3. Visualize fixed and semi-fixed interactive and non-interactive objects

From the design concept used to design proxemic interactions in the past research, Ballendat el al. suggested that designers considered the places and environment where the technology would be used [1]. These included fixed
interactive objects such as displays, computers, digital surfaces, or fixed non-interactive objects such as walls, doors, windows, pavement, or semi-fixed objects such as chairs, table, books where its location might change over time. There should be an option for researchers to choose to add each object to the visualization if they decided to include that object. Whether a study was conducted in a laboratory or public settings, this information could potentially provide more context to the actions a person is doing such as a person entering a room from a door.

**R4. Identify and visualize people’s attention on another people or interactive devices**

From another design concept in Ballendat et al.’s work [1]. Knowing whether a person was giving an attention to another person or devices could be used to initiate implicit actions. For instance, a video player might automatically pause the video when a person’s attention deviated from the screen, either to look at the phone or to have a conversation with another person. Our tool should identify the current points of attention and visualize them. In certain scenarios, multiple people might share the same point of attention and form a group. Some researchers might be interested in the holistic interaction as a group, rather than the individual [44]. It was important that people who shared the same point of attention were visualized as the same group.

**R5. Create summary images of proxemic data over a time period**

This requirement was inspired by position heat map in GIAnT [48] and finger movement trace in VisTACO [44]. A single image could be simple yet a powerful tool to quickly give a rough idea of what happened in a 15-min period within a glance. While the overview image might not completely neglect the need to watch video playback, it could serve as an outline that leads researchers directly to the key activities they were interested. The overview of time period also added the time dimension to the data, which could give new insights. Hence, it was important that our tool included the summary visualization of key proxemic data over a time period.

**R6. Aim to be easily understood by expert users given an application training**

This tool aimed to be used by researchers who were considered experts. A training, explaining how to use each feature, was expected to be given to users. After the introduction and training phase, expert users should be able to understand how to use the interfaces as well as the meaning of each visualization.

### 3.2 Design principles

In this section, we chose four design principles that were considered when making design decisions. These principles were used in the design of user interfaces, visualizations, and the workflow of EagleView.

**Flexible.** As different research might have different goals when analyzing proxemic interactions, our tool should be able to adapt to those aims. The flexibility was important as the lack of it meant that users could not get the information they wanted. For example, the visualization components should be given options to adjust for size, color, or other relevant parameters. Each component should also be able to show or hide based on user’s need.

**Representative.** Our tools focused on the visualizations of proxemic information. It used shapes, lines, and colors to represent the data. It also added a time dimension to the data to create new useful information. An example could be seen in GIAnT’s timeline where it combined 2D position and time dimension in an enhanced line graph [48].

**Instructive.** Visualizations should be straightforward and intuitive, but some design could be inevitably complex. To aid the complex elements, an introduction on how to use each visualization component was provided in the application training. Moreover, the mechanism of how each visualization was explained, which helped users to understand the system and allowed them to be more creative with the tool [41].

**Responsive.** Our tool was designed to help users complete the tasks quickly. It aimed to be fluid, and not having lagging user interfaces. The tool minimized wait time and allowed users to add or remove visualizations without having to pause or restart the video. The tool was able to achieve real-time performance that matched the visualization with the video footage.

### 3.3 Design Methodology

In this research, we followed user-centered design (UCD) as described by Rubin and Chisnell [39]. There were 3 basic principles: the early involvement of users in the design process, the evaluation of the prototypes by users, and the subsequent enhancement through multiple iterations of redesign. A conventional UCD iteration consisted of 4 phases: study, design, build and evaluate (Figure 6). We did not use extended UCD as described by Harper because of the lack of access to users in the earlier part of the design [16].

![Figure 6. Iterative design process involving multiple idea generation and selections](image-url)
4. DESIGN & PROTOTYPE
The design went through an iterative process where ideas were created, refined, prototyped, and tested with expert users. This chapter discusses the detail of study, design, and build stages.

There were 4 phases in the design process (Figure 7). First, in the study phase, we explored opportunities for possible research area, then we narrowed it down by making design requirements. Second, we generated ideas through sketches which turned concepts into rough design. Third, visualization and interaction details were clarified and built into the prototype. Lastly, the prototype was tried out by researchers who had experienced in video analysis. Table 1 described different activities and their objectives in each phase.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Activities</th>
<th>Objective</th>
<th>Granularity of detail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study</td>
<td>Background research</td>
<td>Address concepts</td>
<td>Concept</td>
</tr>
<tr>
<td></td>
<td>Design requirements</td>
<td>Select concepts</td>
<td>Concept</td>
</tr>
<tr>
<td>Design</td>
<td>Ideation &amp; sketches</td>
<td>Generate ideas</td>
<td>Coarse</td>
</tr>
<tr>
<td></td>
<td>Initial selection</td>
<td>Select ideas</td>
<td>Coarse</td>
</tr>
<tr>
<td>Build</td>
<td>Low-fidelity prototype</td>
<td>Clarify details</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>Hi-fidelity prototype</td>
<td>Clarify interactions</td>
<td>Fine</td>
</tr>
<tr>
<td>Evaluate</td>
<td>Expert evaluation</td>
<td>Identify problems</td>
<td>Fine</td>
</tr>
</tbody>
</table>

Table 1. Activities and objectives in the design process

4.1 Background research
At the early time of the research, we did not yet have an access to expert users. Alternatively, the direction of the research was set by reviewing past literature. We looked at design implications and guidelines from past studies [1,14,27,28,45], and established the design requirements as discussed in section 3.1. After that, the features associated with each requirement were identified (Table 2).

4.2 Ideation & sketches
After the features needed in the application were listed, we created sketches to visualize each feature. In this stage, the focus was on creating as many ideas as possible. Then, early decisions were made to select the design to be used for prototyping in the next step.

Table 2. Features in response to design requirements.

<table>
<thead>
<tr>
<th>Design Requirements</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1. Playback visualizations in real-time</td>
<td>Real-time visualizations</td>
</tr>
<tr>
<td>R2. Visualize individual's proxemic information</td>
<td>Distance, orientation, identity, movement, location, attention</td>
</tr>
<tr>
<td>R3. Visualize fixed and semi-fixed interactive and non-interactive objects</td>
<td>Interactive and non-interactive objects</td>
</tr>
<tr>
<td>R4. Identify and visualize people’s attention on another people or interactive devices</td>
<td>Face-to-face and side-by-side grouping, phone and tablet indicators</td>
</tr>
<tr>
<td>R5. Create summary images of proxemic data over a time period</td>
<td>Overview visualizations</td>
</tr>
<tr>
<td>R6. Aim to be easily understood by expert users given an application training</td>
<td>Application training</td>
</tr>
</tbody>
</table>

Distance. There were different ways to visualize a distance. Our ideas included direct line between two people distance label above the line (Figure 8a), outline circle around a person with and without gradient colors (Figure 8b), adapting distance line colors to different distances (Figure 8c), and overlaying grid (Figure 8d). Based on our design principles, we chose three design. First, the distance line with labels was a simple and straightforward representation of distance. Second, we selected adaptive colors, although made it simpler to only two colors: red or black. Third, the circle outline was useful for constantly monitor whether anything got closer than the specified distance, but the gradient was not used because simple circle outline was enough to represent the distance. The grid distance was not selected because it would make the interface clutter.
Figure 8. Sketch of distance visualization

Orientation. This visualization needed to convey the direction a person was facing. In order to show the direction, we created variations of an arrow (Figure 9a). Another design was more realistic, representing head and body of a human with added lines of sight to indicate the direction (Figure 9b). In the selection process, we found that simple arrow shapes were too small and not suitable to represent how human body occupied space. As a result, we chose the body outline visualization which occupied larger space and was suitable for overlaying on top people in the video.

Figure 9. Sketch of orientation visualization

Identity. There were mainly 2 types of visualization for identity. The first was to use different colors (Figure 10a) and the second was to use name label (Figure 10b). The use of different colors was easier for users to recognize than texts [46]. However, we had associated many colors in other visualization, which might conflict each other and make users confused. Eventually, we decided to use text label, either with a name, or a number.

Figure 10. Sketch of identity visualization

Movement. To visualize movement, we used tracing lines, which could be represented by different shapes such as simple lines (Figure 11a), shadow (Figure 11b), footprints (Figure 11c), and dots (Figure 11d). Colors could be added to differentiate between different people (Figure 11e). Other than showing movement in the past, we also thought of showing movement in the future, which might help aid researcher to know where a person is moving toward. We selected a line visualization for movement because it showed all information continuously, unlike other design where information was shown discretely, resulting in the loss in minor movements.

Figure 11. Sketch of movement visualization

Location and objects. Location included interactive objects such as displays, phones, tablets, and laptops. We explored different shape for displays with the key idea to make it distinguishable between the front and the back (Figure 12a). We chose a rectangle design with one bold side and a name label which rotated together with the shape to indicate the direction it was facing (Figure 12a\b). Mobile devices were represented by ‘p’, ‘t’, or ‘l’ letter for phone, tablet, and laptop respectively (Figure 12c). Later,
we did not choose to go with letter label design, but with a more realistic icon of each device instead.

When a person uses a mobile device, there would be icons appearing to indicate that (Figure 14c). An alternative using text was considered but was not used because when rotating the text would be difficult to recognize Figure 14d). However, there was no color highlighting because the icon would always attach to one person.

![Figure 12. Sketch of interactive and non-interactive objects](image)

Zones were another aspect of location, which could be used to highlight areas of their interests. These zones could be picked in different shapes: horizontal, vertical, circle, ellipse, or half-ellipse (Figure 13a and b). Since our design principles aimed to be flexible, all shapes would be supported for different analysis needs. Labels could be added to represent the name of each zone.

![Figure 13. Sketch of designated zones](image)

Attention grouping. A person’s attention could be on objects, mobile devices, or other people. If two or more people shared the same attention, they would be a visualization to indicate the group relationship. Figure 14a shows grouping using color highlight. Figure 14b shows grouping using black and white. The later design was not suitable because it only showed one group at a time.

![Figure 14. Sketch of attention grouping](image)

Overview visualization. Overview was a single image which could give an information of data over a time period. Figure 15 shows the ideas of different overview data.

a) Heat map shows the position where a person stays the longest in comparison other positions.
b) Movement trace shows either past or future movement of a person using lines.
c) Zone overview shows which zone a person is staying with respect to time.
d) Distance overview shows the how far a person is to an object during a time interval.
e) Attention overview shows how much time a person spent on another person or objects.

Some overview visualizations could be placed directly over real-time visualization, while others might require a dedicated space. In the next step, prototyping, we chose to implement only heat map and movement trace due to time limitation.

In summary, in this ideation stage, we created sketches of different design for each proxemic feature. Then, we selected some ideas by ourselves based on the reasons described in each section. In the next stage, we combined all the features together.

4.3 Low-fidelity prototype
The prototypes in this stage had an increase in details which helped further shape the design of each component. Adobe Illustrator was used to create and combine shapes. The tool allowed for flexibility for the drawing visualizations which we found suitable for creating a low-fidelity prototype.
Figure 15. Sketch of overview visualizations

Figure 16 shows a typical setting of two people sitting on a sofa in front of a television with some furniture in a room. A person’s body was represented by a white ellipse, and the surrounding grey area was personal space belongs to that person. The size of body shape was adjusted to match the realistic size of people in relative to a sofa. The lines of sight worked well to give the orientation information. For interactive objects, a name label was used for its identity. For non-interactive objects, different patterns were used on each object. However, we found the pattern to be confusing because it did not directly tell what an object was. Later on, we changed the pattern to name label.

Figure 16. First low-fidelity prototype

Due to time limitation, three most common shapes were picked for zone prototypes: horizontal, vertical, and ellipse (Figure 17). Its implementation using different color seemed to differentiate each zone clearly. Though, we found that a name label for each zone should be added to make the recognition for each zone easier.

Figure 17. Zone low-fidelity prototypes

In the movement prototype (Figure 18), it might be difficult to tell when a person was staying at a location. To aid this problem, we added transparency to the line. The lighter parts indicated a movement earlier in time and the darker parts indicated a movement later in time. Both past and future movement were shown in the prototype using different colors (red and yellow for person 1, pink and purple for person 2). Another problem was the lines might obscure the name label of a person. As a result, we decided to move the body in front of the lines in the next prototype.

Figure 18. Past and future movement low-fidelity prototype

The visualization used grey color as the default color. This allowed the use of color for grouping as in Figure 19. If a person was determined as in the same group as another object, the color of shapes and labels turned into a color of that group. Other people remained at default color. The position of mobile devices was placed in front of users to mimic natural usage position.
After the details were clarified in low-fidelity prototype, we had to determine whether we would conduct an evaluation of the prototype, or continue with a high-fidelity prototype. Because the tool we were developing was meant to be highly interactive, a static image would not allow us to test the interactions. We needed an interactive prototype in order to understand how users use each function of the tool.

The need for interactive prototype led us to create high-fidelity prototype. Initially, a mock prototype was considered, where visualizations were to be created manually frame by frame to match a recorded video. This option was exhaustive to produce, so we dropped the idea. Alternatively, we chose to create a functioning prototype, where visualizations were automatically generated by computers according to the input data. The details of the high-fidelity prototype are presented in the next chapter.

5. EAGLEVIEW VISUALIZATION TOOL

This chapter discussed the technical aspects and user interface of EagleView. EagleView was a visualization tool for supporting video analysis of interaction in a space. It aimed to quicken the analysis process by visualizing key components that would highlight fundamental data in the video, and using visualizations to make patterns and behavior easier to recognize by users.

5.1 Implementation

**System platform and frameworks.** In the first step, we selected JavaScript language and web-application platform for building a prototype. To make the application real-time, all visualizations were calculated and drawn on client-side. Visualization libraries were considered to provide fundamental capabilities for drawing. We selected to use EaselJs as this framework provide a low-level library of lines and geometric shapes which were necessary for the implementation of our design. Other frameworks such as D3.js only provided high-level manipulations, therefore, were not suitable for the work.

**Input data.** EagleSense, a platform that made top-down tracking more readily available, was used to capture input data [47]. EagleSense connected with a Kinect v2 camera to monitor top-view depth and infrared information within a desired area. It simultaneously tracked the location and head direction of multiple people. Further, it recognized activities and posture people were performing. The activities included standing, sitting, pointing, using a phone, using a tablet, and reading a paper.

The area captured by Kinect and EagleSense was approximately 2.0 x 2.0 meters. This area was quite small and limiting for activities. During data capturing, we needed to be extra careful that a person was staying in the tracked area. After we recorded the data, the output file was exported from EagleSense. The data comprised of person’s location, orientation, identity, and activity (Figure 20). After an initial inspection of the whole set of data, we found that some data did not have accurate values. Examples of the errors found were:

- Occasionally, a person’s orientation was 180 degrees off.
- When two people moved close to each other, the data identified them as one person.
- The data detected wrong activity in some cases.

To correct the errors, we performed data cleaning. However, it was labor-intensive to correct all raw data manually. We decided to reduce the number of frames from 30 to 4 frames per second to make the manual operation possible.

In addition to the data, videos were also recorded. Kinect captured a top-view video, while phone cameras captured videos from side views. These videos were used for analyzing fine-grained interactions.
Real-time visualization.

Real-time visualization was located in the visualizing pane (Figure 21a). Data of each proxemic dimension was visualized in this area.

**Distance** – A line between person and person, or person and object could be added from the preferences pane (Figure 22). The distance was calculated from the head of a person. A calibration of the size of the tracked area was needed to calculate the distance. On top of the line was a label identifying Hall’s proxemic zone according to the current distance shown (Figure 2). Another visualization of distance, a distance circle around the body was represented by an outline. The radius of the distance circle could be adjusted on preferences pane.

**Orientation** – Two ovals represented the head and body of a person (Figure 22). Two lines in front of the head projected the sight a person is viewing. There was a limitation of the implementation. The body orientation actually displays the same value as head orientation due to the lack of a separate body orientation value from the source data.

**Identity** – People and devices could be distinguished by name labels on top of the visualization (Figure 22).

**Movement** – Future and past movement were shown by the lines in different colors in Figure 23. Users had an option to set the interval of the movement, for example, 15 seconds to the past, and 20 seconds to the future. Different colors could be assigned to past movement line and future movement line.

Figure 21. EagleView user interface: (a) visualizing pane showing real-time visualizations and playback control, (b) video pane showing videos from different angles, and preferences pane showing options and overview visualizations.
Past and future movement, zones, and attention grouping

Location and objects – EagleView supported rectangle and half-ellipse zones (Figure 24). When a person enters the zone, that zone color’s become darker and a bold outline is added around the border. We did not implement the option to add new zones due to its complexity. Only the options to hide or show the zones were provided. Fixed interactive objects and non-interactive objects were supported, although there were no options to add or edit objects through the user interface.

Attention grouping - There were three conditions which were used to determine an attention grouping.

- A person is facing directly to fixed interactive objects
- A person is facing directly to another person.
- Two people are standing next to each other and they are facing at the same object.

People or objects who belong in a group would have the same colors. Multiple groups were supported in EagleView.

Overview visualization. In this prototype, we chose to implement only 2 visualizations, heat map and movement trace, due to time limitation. Previews of the heat map and movement trace were presented in smaller boxes (Figure 27b and Figure 27c) where users could choose to show one or multiple selections on the main visualization pane (Figure 26). Heat map was useful for finding an area where users were the most active, while movement trace made the pattern of movement easier to identify. A time interval slider was introduced to limit the start and the end data for overview visualizations (Figure 27a). This range can be useful when analyzing a lengthy video. It would allow researchers to focus on interactions in a period of interest.

Overview visualization overlaying main visualization pane a) Position heat map b) Movement trace
Preferences pane. This interface was where users could make adjustments to variables and selected options to show or hide visualizations. There were 4 groups for preferences: overlay, distance, movement, and groups (Figure 28).

Overlay group contained elements to be displayed in the visualization pane (Figure 28a). Users could choose to show a background video and adjust its opacity. The background video is in black and white (Figure 22). This was to avoid having too many colors in the visualization pane, which could create confusion. If users wanted to see the color variation of the video, they could look at the video pane (Figure 21b). Another option in this group was to toggle showing zones.

In the distance group, users could adjust person’s sight angle (Figure 28b). This property would help simulate the vision of a person. The size of the circle around the body could be used for visually determining whether someone was within a specified distance. Also, users could choose to show multiple distance lines between people and objects.

Movement group allowed configuration of the number of seconds to show the past or future movement, along with the color customization for each person (Figure 28c).

Attention grouping allowed the change of color for each group (Figure 28d). If more than one group was detected, each group would be automatically assigned a different color.

In this section, we described the user interface of EagleView and its functions. The design was then evaluated by expert users.

6. EVALUATION
This chapter discusses the methodology of a qualitative study conducted to evaluate EagleView prototype. The objectives of the study were to:

1. Conduct user study on how researchers use video analysis tool to aid an analysis of study in space.
2. Evaluate usability of the prototype given a face-to-face training
3. Evaluate usefulness of the prototype
4. Explore possible improvements and future directions

Before going into the detail of the evaluation, we will first present the findings from the pilot studies.
6.1 Pilot studies
We conducted 2 pilot studies with a PhD student to solve technical issues and to refine the methodology used in the evaluation. There were 2 main findings:

Resolve major usability issues. The earlier version of EagleView prototype lacked a separate timeline. Participant was forced to use a timeline on a small video in the video pane. This made navigation of the video difficult and imprecise.

Another major issue found was that four videos were not synchronized together. Instead, each video started playback as soon as it finished buffering enough frames. This led to a mismatch between visualization and the actual video. Participant was confused to which video should be taken as the correct one when observing movements in the video.

Later on, we introduced a synchronous video playback by making sure every video was ready to play, then, play all video at the same time.

Change to evaluation task. At first, the study was designed to give a training to participants and allow them to explore the tool freely. We observed the lack of focus and meaningful goal. As a result, we created a scenario and questions to challenge participants to use each feature of the tool.

6.2 Participants
Seven participants (5 males, and 2 females) were recruited from universities in UK and Canada (Table 3). Their age ranged between 24 and 34 (μ = 27.9, σ =3.8). One participant was a master student. Four participants were PhD candidates in Human Computer Interaction (HCI). The last two participants were post-doctoral researchers in HCI. All participants had past experiences conducting research in HCI that involved video analysis except for one participant who was brought in based on the experience in EagleSense system. Their experiences were necessary to collect expert opinions and suggestions that were most relevant to the end user. Participants who participated in face-to-face meeting received £10 GBP, while participants who participated remotely received £15 GBP.

6.3 Design
Participants in face-to-face and remote conditions received the same set of questions for questionnaire and interview, also the same instructions for training and user evaluation task. In the remote setting, participants shared their working screen to enable observation of activities during the application usage.

6.4 Materials
EagleView prototype was accessed on a Chrome browser, which was the same browser used to test the application during the development process. The evaluation was conducted either locally, or remotely through Skype. For the local study, a MacBook with 13-inch display was used with a mouse provided for tracking. An iPad was used to capture video and audio of the participant and the interviewer. Additionally, the interaction on-screen was recorded through QuickTime player. For remote studies, participants used their own computers. Skype was used for video call, and screen sharing. We recorded the audio and video of the screen through an iPad pointing toward interviewer’s laptop.

Two set of videos and interaction data were captured through top-view Kinect and EagleSense API. Two additional side-angle videos were recorded in the first set while only one was recorded in the second set. The data from EagleSense was manually adjusted to correct the position, orientation, identity, and activity. The first set was used for training, while the second set was used for evaluation task.

6.5 Procedure
The study took about 60 minutes for the face-to-face meeting, and 90 minutes for Skype meeting. In the beginning, participants were given an information sheet for details of the study. After a consent to participate was given voluntarily, we started video recording on iPad, and audio recording on iPhone.

Pre-task questionnaire. Participants were asked to complete a questionnaire about their demographic information and past experiences. It was given on paper (local study), and via Google Docs (remote study). The questions focused on how they conducted past studies, how video analysis software were used, and the critical reflections of the tools. We expected this opportunity to help participants to recall the details of relevant experiments before talking about them in the interview.

Pre-task semi-structured interview. Upon the completion of the questionnaire, we followed up the answers with interview questions. We asked participants to elaborate on each answer and to give details about one past experiment

<table>
<thead>
<tr>
<th>No.</th>
<th>Occupation</th>
<th>Years experience in HCI</th>
<th>No. of video analysis conducted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Post-doctoral researcher</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Post-doctoral researcher</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>PhD candidate</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>Master student</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>English teaching assistant (also PhD candidate)</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>PhD candidate</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>PhD candidate</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3. Participants demographic information
they conducted. When an interviewer came across interesting topics, additional questions were asked.

**Scenario-based application training.** At the beginning, participants were told that they were free to explore the tool and to ask questions while the interviewer explained functions of the application. We started from explaining the overall layout, then, continued to describe the meanings of basic components in the visualization pane (Figure 21a). After that, a description of a scenario in the video was given along with 7 use cases to demonstrate the functionality of the tool.

**Scenario-based user evaluation.** After application training, participants were given a task sheet (local study), or Google Docs (remote study) with 8 questions to be completed on their own without any help. We did not use the same scenario which was in the training, instead, a new scenario and a new set of video were introduced. Participants were asked to think aloud while completing the task. We did not limit the time to be used by participants. During the task, we made an observation whether each question was easy or difficult to answer for later follow up questions.

**Post-task questionnaire.** Participants were asked to complete a 2-page questionnaire on paper (local study), or Google Docs (remote study). The first part of the questionnaire was about the usability of the tool, while the second part was about the usefulness of each visualization component. Both parts used a 5-point Likert scale for the answer.

**Post-task semi-structured interview.** The questions in the interview were divided into four parts. The first was the question to follow up difficulties during the evaluation task, and other usability issues. The second was the question about usefulness with references from their answers in the post-task questionnaire. Third, we asked about how they would incorporate the tool into their current workflow. Lastly, participants were asked to suggest possible features or changes.

**6.6 Data analysis**

Data from the questionnaire were analysed quantitatively. Due to the small sample size, we did not aim to claim any significant results. Rather, we aimed to use the quantitative data to aid the qualitative explanation from task observation and interview. Thematic analysis was used. The audio and video recording were transcribed, then coded with labels. Lastly, we grouped them into a theme.

7. RESULT

In this chapter, the findings from the evaluation were reported in 3 separate parts: researchers in proxemic interactions, application review, and future improvements. Participants are referred to using an abbreviation (P1 to P7).

7.1 Researchers in proxemic interactions

From pre-task questionnaire and interview, we asked participants to point out specific goals they were looking for in the video analysis, and the different strategies adopted by researchers.

**What researchers looked for in video analysis.** A video could capture a lot of information that it could become overwhelming for the analysis. Multiple things could happen simultaneously and the video length could be from minutes to hours. Consequently, researchers needed to set a specific focus. P3 mentioned specific hand gestures that were the focus of observation during the video analysis: “...grabbing the cube from partner, handing it over, setting it on the surface.” P2 looked for the change of attention: “...when does a person change their focus. If someone stares for 2 minutes into their devices on the table, and then look up and look to another person.” P2 further explained another type of search goal that was more abstracted, described as a collection of consecutive actions rather than one single action: “One focus was how a visitor progress from watching to become an active player.”

**Video analysis activities.** Despite the data in video analysis could be vastly different, we found that most researchers followed similar steps for video analysis. Almost all of the participants described the first step in video analysis was to run a single playback to get the overall knowledge of what happened: “I just want to play a video and get the sense of overall interaction” (P1). If they see interesting events related to their search objectives, they would mark the time and add a short description for a later thorough watch: “At first, I looked through the video and make comments on paper and pen to get an idea of what’s happening at each point of the task” (P3). P2 had a video recording of over 25 hours, therefore he sped up the preview using faster playback: “I do a preview. I might get away with watching the video at 16 times the speed.”

The second step was to visually search for the pre-defined objectives. Some participants defined a tag and applied it by setting the start and end time when an instance of such activity was found: “For thematic analysis, you can have your tags, and you can select part of the timeline and drag and drop onto them and then it’s colored so you can see which tag is applied to which time in the video” (P1). Another participant took a qualitative approach by describing the events with text: “I see 2 people entering the camera view, then I would describe how these people interact as long as they are in the focus until they leave” (P2). A portion of the video could contain multiple search objectives, requiring researchers to repeat the playback several time until they complete going through all details in the video: “...I could have multiple focuses happen at the same time that I would observe and describe differently. So, I watched certain pieces with a lot of people several times and transcribed several instances out of the same video material” (P2).
After the marking and annotation of the instances, quantitative researchers might analyse the data with statistical tools: “I exported the file into .csv and tried to look at which interaction occurred most frequently in term of 2 to 3-minute segment” (P3). They might compare the number of occurrences of each interaction, or tried to find a pattern out of the data by using visualizations. P7 described how he looked at the event markers on the timeline to reveal a usage pattern: “Frequency, duration, and different mobile devices usage pattern...”

To sum up, researchers first set up specific instances which they looked for in the video. Then, they made the first play through and made a note of interesting parts. After that, they watched each part in detail and marked or tagged the occurrences of interactions. Finally, the data was imported into statistical tools for further analysis.

7.2 Application review
The application was evaluated in two dimensions, usability and usefulness, with the use of both quantitative and qualitative data.

7.2.1 Usability
Overall, EagleView prototype received positive feedbacks for being easy to use given the training participants received during the evaluation (Figure 29). P7 described the visualizations as “...simple, not overly complicated”. This result fulfilled the usability requirement for expert users (R6). However, there was one participant who faced difficulties trying to understand the system even after the training: “I didn’t remember a lot of things that you told me...” (P1). Nonetheless, this could be a result of the time limit during the training. Many functions were introduced that it was difficult to remember all of them at once.

Grouping visualization. When asked in more detail if there was any visualization that its meaning seemed unclear, many participants pointed out the need for separation of each condition used in grouping visualization. As to remind, the three conditions used to determine grouping could be seen in Figure 25:

- A person is facing directly to fixed interactive objects
- A person is facing directly to another person.
- Two people are standing next to each other and they are facing at the same object.

Three participants argued that the mix of conditions made them unable to determine if a person was grouped because one condition was met, or two conditions were met: “If both are looking at the screen, it was really difficult to see if they also talk to each other” (P2). P5 asked in more detail about how the grouping was decided by the application: “I did not understand what those group meant and why they were selected that way. Why does it think that’s a group?”

One participant described in detail two different perspectives of grouping, one from the perspective of the display, another from the perspective of a person. The following quote described an event where two people who were not together, each was looking at the same display: “From the point of view of a person, they are only interested in the display, and not in the other person ... From the point of view of a display, you just create a triangle. I want to have a way of knowing that the groups actually, whose point of view is from.” (P5). To make the visualization clearer, P5 further suggested that each individual person should have its own color, and tagged that color to the display if they were using it. If two people looked at the display, there would be 2 colors on the display: “If two people are both looking at the display then one person is blue, one person is red, then the display will show red and blue, so we know that both are looking at the display. (Otherwise) It’s hard to tell them apart. Especially with 2 or 3 people.”

Reliability of visualization. When asked if participants would rely on the plain visualization without other videos, most participants explained that they preferred to have accompanying videos for other information which the visualizations did not provide: “I don’t think I would totally rely on the clean visualization just because for something like talking, you actually need the video” (P3). She further explained that cross-checking was still required to make sure the application visualized data correctly: “...as a researcher, I don’t 100 percent trust. There is kind of a pre-analysis have been done. It’s nice having the actual video next to it just so I know...

![Figure 29. Result of post-study questionnaire on EagleView’s usability using 5-point Likert scale](image-url)
that this has been done correctly.” – P3

P5 noticed that the visualizations relied on top-view camera, which had a vertical distortion. There would be an offset between the position of the head and the feet: “It’s interesting because you would think that the person is closer to the display all the time than they actually are, but I wonder if that’s because of the distortion because if you look at the images, the camera was exactly right above them.” The magnitude and the direction of the offset for each person varied depending on the relative position to the Kinect camera: “...if you look at the actual distortion of the image, they are kind of like V-shape” (P5).

Figure 30. The offset of positions between the head and the feet created by vertical distortion

There was also another type of offset distance. The distance as measured by the system was from one head to another head, but arguably, the distance could also be measured from other parts of the body: “This person’s finger is very close to this person’s tablet, so arguably they are within 50 cm range, and that’s not something that I will get from the system because the system is just not showing me that.” (P5). Nevertheless, it was worth noting that only one participant was able to notice the offset of distance, while other participants did not question the distance information given by the system.

Integration into video analysis workflow. Participants identified two methods for using EagleView with other video analysis tools. For participants who used a simple note application for annotation, EagleView could replace video playback application: “If this would work like a player, where I would load my video. It would be not a big gap.” (P2). Participants who made annotations on the video analysis application would not use EagleView as a main video analysis application because it lacked the annotation features. Instead, P1 imagined importing a visualization video into other tools which supported video annotations: “May be I could save the video after applying all these things and then import into Nvivo and somehow combine these features with another thematic framework.”

7.2.2 Usefulness

Post-task questionnaire result showed that all visualization components were perceived as useful for video analysis (Figure 31). P7 mentioned the activities he searched for in his previous research that with the visualizations, it would be easier to identify these activities: “It’s good to know when are people together, when are they looking at the screen.” He also mentioned that other tools would not provide the information visualized by EagleView: “You get to see when people use devices, when people are talking to each other, when people meet each other, which I wouldn’t able to get from the other tools” Unanimously, all participants agreed upon the usefulness of head, distance line, sight and heat map visualization. This result supported the need for each visualization components as in design requirement (R2-R5).

P2 emphasized on the overview visualizations that they gave data which could not be retrieved by watching video playback: “...if I’m interested in heat map. It’s very difficult to get out of a transcription because how would you do that?” He also commented on the use of movement trace: “with the traces, you can do prediction that you couldn’t do.”

Figure 31. Result of post-study questionnaire using 5-point Likert scale to the question - “I perceive the following components to be beneficial to video analysis.”
Some participants thought that the real-time visualizations could be used as a simplified version of the raw video. This idea could be seen during the evaluation task, where some participants did not turn on video background, instead they chose to look at just the visualizations because it was cleaner: “The way it is gives the most objective view of the relationship that are happening in theory.” (P7).

7.3 Future improvements
Additional features were suggested for EagleView. We put the ideas into 4 categories (Table 4). These features could be implemented in future works.

<table>
<thead>
<tr>
<th>Category</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>New visualization</td>
<td>Overview visualization on timeline</td>
</tr>
<tr>
<td>Enhancement for existing visualization</td>
<td>F-formation</td>
</tr>
<tr>
<td></td>
<td>Vision visualization</td>
</tr>
<tr>
<td></td>
<td>Public heat map</td>
</tr>
<tr>
<td>Improvement for usability</td>
<td>On-visualization navigation</td>
</tr>
<tr>
<td></td>
<td>Export video and data</td>
</tr>
<tr>
<td>Fine-grained tracking</td>
<td>Vertical head tracking</td>
</tr>
<tr>
<td></td>
<td>Fine-grained body movement</td>
</tr>
<tr>
<td></td>
<td>Object tracking</td>
</tr>
</tbody>
</table>

Table 4. Features for future improvement

7.3.1 Additional visualizations
Overview visualization on the timeline. From the observation of how participants completed each question in the evaluation task, we found that participants relied mainly on visual search. For example, the question asked to identify the period where a person in the video started to use a phone. Participants started playback from the beginning until the end to look for a phone icon. This strategy was time-consuming, especially if a video was lengthy (a few hours), visual search might even be considered impractical. When asked how we could improve the visualization, P7 thought that the application should provide that information directly without having to manually search for it: “Have they been using phones, or looking at the screen? I want to see that on the timeline without having to go through the video.”

Other participants had similar ideas that expanded to other data. P6 talked about the timeline visualization of attention and interaction of people: “What I want to see in the summary view is the summary of interactions that happen, like interacting with devices, two people talking, one person using phone, then turn back, then do something together side-by-side. It’s like high-level description of what happened” (P6). Similarly, P5 considered the visualization of f-formation on the timeline: “...the system automatically detects if there is f-formation happening, then it will show you somewhere in the timeline that took place.” (P5)

Two participants further emphasized the benefits of using timeline for visualizations. P3 favoured that she could know exactly when events start and end without repeating the playback: “When you annotate it, you remember exactly when it occurred in relation to the time of the video.” P7 used timeline visualization to help emerging pattern of interactions: “I used a timeline. Just by looking at the timeline, you can see different patterns.” Overall, overview visualization could offer the benefits of reducing the needs for playback, show a summary of interactions, and reveal patterns from data.

7.3.2 Enhancement for existing visualization
F-formation. We reported the usability problem of grouping visualization in section 7.2.1. P6 further suggested the use of a combined circle to show the co-interaction between two people: “When two people are standing close together, instead of individual proxemic (circle around the body), show something that group them together like circle the group.” The combined circle is in another word, f-formation (see Figure 4).

Vision visualization. As a refined version of sight visualization, P5 suggested the use of flashlight concept to imitate the actual sight of a person to show what was in the view, and what was out of the view: “May be in term visual, I would have done an actual field triangle with very low opacity that goes all the way to see everything that they see. Make it fill instead of stroke like a flashlight. If it’s occluded, then that’s not the line of sight anymore.”

Public heat map. P7 wanted to know the position which all people in the area tended to stay. Instead of showing multiple individual heat maps, only one heat map could be plotted from the combined data of all people: “Combine heat map to reveal where multiple people stand ... If you notice that people are often stand in the same place, there might be a reason for that. May be people always stand on the right-hand side. That allows the interface to take that into consideration.”

7.3.3 Improvement for usability
On-visualization navigation. Instead of turning visualizations on and off by a separate menu, P4 suggested that each visualization could be directly clicked to reveal configurable options. It would offer the option menu at the position where it was the most related to the context: “If there was a way to interact on the visualization pane itself ... rather than having options on the sidebar, I could have two or three options on the visualization pane itself.”

Export video and data. In section 7.2.1, we gave the details how EagleView could be used with other video analysis tool. P7 suggested the application should be able to export the main visualization video with different options: “The main visualization. You should be able to export with different layers.”, and the statistical data for further
analysis: “Being able to export the data, so people can import to any package they want to measure how long people are doing stuff. Export the data at least in .csv format.”

7.3.4 Fine-grained tracking
EagleView aimed at providing information related to proxemic interaction. However, the study conducted by participants also involved other data, which were necessary for their interaction analysis. The following features require additional data that EagleSense did not provide [47]. Future works were required for implementation of these features.

Vertical head tracking. One of the key activities to look for in the study conducted by P2 was whether a user looked downward to use a device, or looked upward and changed focus to other things. EagleView’s prototype only showed a direction a person looked, but not the vertical position of the head: “How would it detect where he is actually looking? ... I could look like this (head up), but not look down, or look at the display” (P2). A front-view Kinect could be added for tracking vertical head position and other body movements, both of which are important for an interaction analysis.

Fine-grained body movement. P1 looked for body movement while standing or sitting, for example, leaning forward or backward, sitting straight or with leg crossed: “...if it could capture more fine grained movement, so they are going to be sitting around, may be people get out of their chairs, may be they lean forward.” P3 conducted a study that involved hand and arm movement such as stretching or retracting the arm, grabbing or releasing an object: “...grabbing the cube from partner, handing it over, setting it on the surface”.

Object tracking. Another interest in the video analysis was the object. P1 proposed this feature when asked about a dream tool for video analysis: “select a region and say “this is somebody’s shoe” and I want to track that over time.” This could be done by image recognition, or by giving some identities that can be tracked when capturing a video.

8. DISCUSSION
Overall, the evaluation found that researchers had positive thoughts toward the usability and usefulness of real-time and overview visualizations. We used the new understandings of how researchers conducted video analysis to make recommendations for improvement of the visualizations.

8.1 Real-time visualization
The implementation of real-time visualizations was mostly perceived as responsive and clear. We achieved the level of responsiveness of the visualization that was fast enough for video playback and video seeking as desired by the first design requirement (R1). No participants reported the lag or frustration from the performance issue. For the requirement to visualize proxemic dimensions and objects (R2, R3), people and object visualizations were designed to be straightforward. Participants were able to understand the meaning of these visualizations given a 10-minute application training.

The visualization of attention between groups of people and devices (R4) was more complex to users. One participant explained the issue that there were two points of view when determining which person was in a group. From a device perspective, if two people were looking at a display, they both should be highlighted as using the display (Figure 32a). However, this visualization created a false information that two people were also interacting with each other. Instead, the system should take a perspective of a person, if one person was not interacting with another person, then they should not be placed in the same group. An accurate representation of this scenario should have 2 separate groups (Figure 32b).

Figure 32. a) grouping highlighting people who are watching a display b) grouping highlighting the absence of interaction between 2 people

The design of real-time visualization was praised for being a clean and objective representation of data. It removed all distractions from the view while highlighting subjects of interest. In the evaluation task, most participants preferred to use real-time visualizations without any background. However, they still needed to watch the actual video when the information provided by visualization was insufficient, or to look for fine details such as hand and body movement. The visualization would not replace the need to watch a raw video, but rather to accompany the video as a high-level representation of data for coarse-grained search. After finding an interesting part of the video, researchers could continue watching the actual video for fine-grained details.

Figure 33. An example of perspective distortion correction [37]
The distance visualization provided a distance measurement that could not be obtained by watching a regular video. However, the distortion of a top-view camera made the measurement of distance inaccurate. Looking from the camera view, objects were pushed away from the centre of the camera (Figure 30). The more they were to the edge of the camera, the larger the offset distance. Nevertheless, perspective distortion could be corrected by applying orthographic projections, which were supported by video editors such as Adobe Premier Pro. Figure 33 shows an example of perspective distortion correction of a building taken from front view. A similar technique could be applied to Kinect camera for accurate distance tracking.

8.2 Overview visualization
The expectations for overview visualization to be “information at a glimpse”, and “reveal new insights” were met (R5). Participants commented that the visualizations introduced a new way for video analysis that had not been considered before. For example, it required an intensive observation to gain the knowledge of the position a person stayed the longest, whereas a heat map tool simply gave that information right away. Similarly, a movement trace tool could be used for movement pattern analysis just by looking at a single image. Participant provided positive feedbacks that both overview visualizations were beneficial for researchers.

The potential of overview visualization did not stop at the two tools we implemented, rather participants requested for more overview visualizations. The ideas brought by researcher’s feedback used timeline as a medium. Timeline was a feature that most participants found crucial to video analysis. This was because of its ability to accurately convey the events which happened in a video, to identify the start and the finish of an event, to recognize patterns from a combination of multiple events, and finally to export the events data for statistical analysis. All these benefits made timeline a robust visualization design suitable for video analysis. What we could do in the future works was to add information of various events happen onto timelines.

Figure 34 shows a possible design of overview visualization on the timeline using event brackets and graph. This design would make answering the questions in the evaluation task less reliant on watching playback. For example, one of the questions was “What is the distance when 2 people are talking together. Using the timeline visualization, one could seek to the start of the event where P1 and P2 were in the same group and look for more details in the video.

The concept of filtering a portion of a video using multiple conditions was similar to that of EXCITE [30]. However, there was a major difference in how the two tools could be used. The direct query in EXCITE would match a top-down analysis methodology where researchers have already known what to look for, and want to extract those events from the video. EagleView’s timeline visualization can be an alternative for researchers who use a bottom-up analysis method similar to Grounded theory. The visualization does not require users to specify a condition at the start of the analysis. Instead, it allows researchers to explore the visualization as a summary of events which take place in the video. When researchers have identified interesting events or patterns, they can use a query tool to further find those events. From the evaluation interview, we found that both strategies were used by participants in the past research. P1 and P2 knew specific things to look for, while P3 and P7 watched video playback to uncover behavioural pattern. Overall, the overview visualization on the timeline can be the next implementation which aligns with the goal to provide researchers with more insights.

8.3 Integration into video analysis workflow
There were two functions which participants requested to enable using EagleView with other tools for video analysis:
exporting visualization video and event data. The reason why participants wanted to export a video was to allow them to import the video into other tools for making annotations. We thought this process would introduce some drawbacks. An exported video would lose the flexibility of applying or removing each visualization components when not required. This behaviour was observed in the evaluation, where participants would show and hide background videos and distance line as needed. To maintain the flexibility, an annotation feature should be implemented directly in the visualization tool. The second feature was to export data for statistical analysis. This feature was a logical next step to complete the analysis workflow.

9. CONCLUSION
This research implemented new visualization design that helped researchers quicken the analysis of video recordings in a study of interaction in a space. Visualization components were designed to represent the proxemic dimensions and people-device relationship that allowed researchers to focus on fundamental information when watching video playback. Time dimension was added to the distance data to create overview visualization (heat map and movement trace) which gave additional information that was useful for the analysis. The evaluation of EagleView suggested that the representation of grouping visualization and the reliability of distance visualization could be improved. Other visualizations were easy to understand by expert users. The research ended with proposing new features from the findings. Researchers wanted to see more of overview visualization, particularly the overview visualization of events on the timeline, which could be highly time-saving from reducing the needs to watch through the whole video playback.

ACKNOWLEDGMENTS
Thank you to the supervisors, Nicolai Marquardt and Frederik Brudy for constantly providing assistance and feedbacks required to successfully conduct this research. I was sincerely grateful for the help from PhD and post-doctoral participants to provide genuinely helpful feedbacks during the evaluation. Lastly, thanks to HCI 2016-2017 classmates that kept encouraging the spirit during the hard time.

REFERENCES


Additional image references from thenounproject.com
1. Arrow by arjuazka from the Noun Project
2. Arrow by Landan Lloyd from the Noun Project
3. User by Mello from the Noun Project
4. Idea by Numero Uno from the Noun Project
5. Wireframe by mikicon from the Noun Project
6. Clipboard by Creative Stall from the Noun Project
APPENDIX
A1. Pre-task questionnaire

1. What is your gender?
   - Female
   - Male
   - Prefer not to say
   - Other (please specify)

2. What is your age?

3. What is your current occupation?

4. How many years of experience do you have in HCI field?

5. How many interaction design research have you conducted that involves an analysis through videos?

6. How many video analysis have you conducted?

7. What methodology did you use for video analysis?

8. What were the key activities during video analysis?
9. Which software or tools did you use for video analysis?

10. What did you find useful about these tools?

11. Did you experience any difficulties while using these tools? Briefly list, if any.

12. Were there any missing features which you wish they had? Briefly list, if any.
A2. Evaluation script

Evaluation Script

Introduction
Hi, thank you for participating in the user evaluation session. This will take about 60 minutes. If you feel uncomfortable during the period, you can choose to leave at any time without giving an explanation. We will use the iPad right there for video recording, and iPhone for voice recording, but don’t worry, we haven’t started recording yet.

Let me introduce you to my research. I am working on a video analysis tool to be used by researchers. It will help them to analyze proxemic interactions, or activities which happen in a space. For example, a display which reacts according to how close you are to the screen. The tool focuses on visualization to give information and insights that a raw video does not provide. We will try out this tool later.

The session will start from having you answer some questions on the questionnaire. Then, we will have an interview about your past research experiences. Next, I will introduce the prototype as well as how to use it. After that, you will try out the tool in a mock scenario. Finally, we will have a post-task interview.

During the task, a video will be recorded. Any data or images used in the research report will be anonymized. You will get 10 Pounds after the session finish. Do you have any questions?

Here is the consent form. Please read and sign if you agree.

(Consent form collected)

Thank you, I will now start video recording.

(Start video recording)

(Start audio recording)

Here is some questions to ask about your past experiences.

(Hand in questionnaire and give a pen)

(Questionnaire completed)

Thanks for your information. Next, we will ask you questions to follow up your given answers.
Pre-task interview (Time ______________)

Objectives and context of video analysis

1. Can you tell me about one past experiment and describe the goal of the experiment?

2. Why did you need to use video analysis?

3. What were you looking for in the video analysis?

4. [From number 7 to 12 in the questionnaire] Can you explain your answers?

Existing method and tool for video analysis

a. What methodology did you use for video analysis?

b. What were the key activities during video analysis?

c. Which software or tools did you use for video analysis?

Advantages and disadvantages of existing tool

d. What did you find useful about these tools?

e. Did you experience any difficulties while using these tools?

f. Were there any missing features which you wish they had?

That’s all the questions I have. Next, I will introduce you to the prototype.
Application introduction (Time ______________) 

(Open EagleView)  
(Start screen recording with pointing enable) 

During the introduction, you are free to try out each function as I explain.

This tool is called EagleView. As you can see, it is designed to support a video analysis. It takes data from Kinect as an input and turn those data into visualizations. On the left is the visualization windows, in the center is the video from different angles, and to the right is the configurations. The videos can be controlled using the control bar on the left. You can try it out. The oval represent body and head of a person. Each person has a unique number label on it. The circle represents social distance of 1 meter radius. If a person is using a phone or tablet, there will be icons to indicate that. Background video can be added by clicking on the show background button in overlay menu. You can try to add background video and adjust its opacity.

For now, these are the basic information you need to know. Next, we will walk you through the research scenario in this video. A public display is situated in public area where pedestrians are walking by. The display tries to attract people’s attention through on-screen adaptive elements such as user shadow and floating pictures. There are 3 zones which triggers different on-screen animation as people move across zones: less than 1.2m, 1.2-1.5m, and above 1.5m. You can show zones on the screen by clicking on the right pane.  

I have recorded a mock video of this situation and loaded it to the prototype. Please note that there are a few limitations of this prototype. First, due to the room we used was quite small, people are forced to stand closer than they normally are. Second, each video might not perfectly synchronize. There might be about 1 second overlap. 

If you have any questions during the walkthrough of the application, please feel free to ask. Please also speak aloud what you are thinking while using the tool.

Let’s start with case 1 – Time (____________)

In this experiment, we want to know the distance where a pedestrian start to turn to the display. You can choose to display the distance line between a person and the display in the right bar, under the Distance group. Click on Person 1 and TV. Now restart the video and observe the distance which a person starts to notice the display. You should be able to measure the distance at about 1.3 meters.

Case 2

Next, we want to know the zone where the first person stands to interact with the display. You can toggle showing zones by clicking on the right bar. Continue to observe the video from 15th second onward. (0.5-0.7 meters, zone <1.2m)

Case 3

We want to know whether the observer’s attention is on the display. We can determine this by two methods. First, a line of sight provides the direction in which a person is facing. You can try to adjust the sight angle in the right pane.

Second, if a person is watching the screen, the person color will change to that of the screen. Let’s watch person 3 from 1-minute mark onward.
Case 4
We want to search for a time when the third person is interacting with the display. You can set future movement to 15 seconds and drag the video seek bar to search. The color is changeable through the right pane. (1:05)

Case 5
We want to know an area where people spent the most time standing during an interaction with the display. You can use the heat map tool which is located in the summary view tab. Let’s take a look at person 1’s heat map. You can also choose to show one or multiple heat map at a time.

Case 6
We want to know the walking path of the person who interact with the display. You can use show future and past movement in the preferences pane. If you are more interested in an overview, you can use movement trace tool in the summary view just below the heat map. Let’s activate person 1’s trace. The lighter color indicates the position earlier in time, while the darker color indicates the later positions.

Case 7
We want to know the duration which person 3 stays in the area. You can adjust the beginning time interval until the lighter tracing line starts to disappear and adjust the ending time interval until the darker tracing line starts to disappear. (0:55 – 1:19)
Evaluation task – Time (_______________)

Now that you have learned about the tool. We will now give you another scenario and tasks and leave you to complete it without my help. Please think-aloud while you are completing the tasks.

Scenario. A person is interacting with a display. The second person shows up to have a conversation. Then both of them starts to interact with the display.

Task

1. What is the distance when the two persons are talking to each other? (1.40m)

   Easy           Difficult ________________________________

2. Search for the time where the first person stops interacting with the display and turn back. (0:34)

   Easy           Difficult ________________________________

3. Search for the time where 2 persons start to talk face to face. (0:33)

   Easy           Difficult ________________________________

4. Search for the time where 2 persons talk start to talk side by side. (1:04)

   Easy           Difficult ________________________________

5. From what time does person 2 starts to use a phone? (0:44)

   Easy           Difficult ________________________________

6. Who walks around more, person 1 or person 2? (Slightly, person 1)

   Easy           Difficult ________________________________

7. Have 2 persons ever gotten closer than 50 cm?

   Easy           Difficult ________________________________
8. Which areas do person 1 stand the longest?

Easy  Difficult ________________________________

Post-task interview – Time (______________)

Well done for completing all the tasks, please rate the usability of the tool using this sheet.

(hand in post-task questionnaire)
(questionnaire completed)

There are a few questions I want to ask to get your feedback. Please don’t hesitate to be critical. Any positive or negative opinions is welcome.

Usability questions

1. Did you face any difficulty during the tasks? Please describe.

2. Is there any visualization that its meaning seems unclear to you?

3. Would you prefer clean visualization or with the background video?

Usefulness questions

4. Did you find the following visualizations to improve productivity for video analysis task?
   a. Head

     Useful  Not useful ________________________________

   b. Body

     Useful  Not useful ________________________________
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>c. Sight</strong></td>
<td></td>
</tr>
<tr>
<td><strong>d. Proxemic distance circle</strong></td>
<td></td>
</tr>
<tr>
<td><strong>e. Phone / tablet icons</strong></td>
<td></td>
</tr>
<tr>
<td><strong>f. Distance line between persons or objects</strong></td>
<td></td>
</tr>
<tr>
<td><strong>g. Zones</strong></td>
<td></td>
</tr>
<tr>
<td><strong>h. Future and past movement</strong></td>
<td></td>
</tr>
<tr>
<td><strong>i. Color grouping when directly facing a person or object</strong></td>
<td></td>
</tr>
<tr>
<td><strong>j. Color grouping when 2 persons are close together</strong></td>
<td></td>
</tr>
<tr>
<td><strong>k. Heat map</strong></td>
<td></td>
</tr>
</tbody>
</table>
1. Movement trace

Incorporating tool into workflow

5. If you are to incorporate this tool into your video analysis, what would you use it for?

6. How would this tool replace or work together with the tool you are currently using?

Future directions

7. What other information in the left visualization window do you want to see?

8. What other information in the summary view do you want to see?

9. Can you suggest any functions to add to the tool?

10. Let’s forget about the tool we just used. If you can create your dream tool for video analysis, how would it look like?

11. Do you have any other suggestions?
This is the end of the evaluation. Thank you for your time and feedback. Please write your name and sign on this paper for money collection.

(sign receipt form)
(hand off payment)
(stop screen recording)
(stop video recording)
A3. Task sheet for participants

Evaluation task

Scenario. A person is interacting with a display. The second person shows up to have a conversation. Then both of them starts to interact with the display.

Task

1. What is the distance when the two persons are talking to each other?

2. Search for the time where the first person stops interacting with the display and turn back.

3. Search for the time where 2 persons talk face to face.

4. Search for the time where 2 persons talk side by side.

5. From what time does person 2 starts to use a phone?

6. Who walks around more, person 1 or person 2?

7. Have 2 persons ever gotten closer than 50 cm?

8. Which areas do person 1 stand the longest?
A4. Post-task questionnaire

**1. Please rate the usability of the tool.**

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I think that I would like to use this system frequently.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I found the system unnecessarily complex.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I thought the system was easy to use.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I would imagine that most people would learn to use this system very quickly.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I found the system very cumbersome to use.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel comfortable using this system.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I needed to learn a lot of things before I could get going with this system.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2. I perceive the following visualization components to be beneficial to video analysis.

<table>
<thead>
<tr>
<th>Component</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Body</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Sight</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Proxemic distance circle</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Phone / tablet icons</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Distance line between persons or objects</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Zones</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Future and past movement</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Color grouping when directly facing a person or object</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Color grouping when 2 persons are close together</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Heat map</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Movement trace</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>