Design concepts and algorithms were developed to address the eye tracking analysis issues that arise when (1) participants interrogate dynamic multielement objects that can overlap on the display and (2) visual angle error of the eye trackers is incapable of providing exact eye fixation coordinates. These issues were addressed by (1) developing dynamic areas of interests (AOIs) in the form of either convex or rectangular shapes to represent the moving and shape-changing multielement objects, (2) introducing the concept of AOI gap tolerance (AGT) that controls the size of the AOIs to address the overlapping and visual angle error issues, and (3) finding a near optimal AGT value. The approach was tested in the context of air traffic control (ATC) operations where air traffic controller specialists (ATCSs) interrogated multiple moving aircraft on a radar display to detect and control the aircraft for the purpose of maintaining safe and expeditious air transportation. In addition, we show how eye tracking analysis results can differ based on how we define dynamic AOIs to determine eye fixations on moving objects. The results serve as a framework to more accurately analyze eye tracking data and to better support the analysis of human performance.

1. Introduction

Eye tracking research is useful for evaluating usability or analyzing human performance and more importantly understanding underlying cognitive processes based on the eye-mind hypothesis [1]. This hypothesis asserts that what we observe when performing a task is highly correlated with our cognitive processes. Thus, eye tracking research has been conducted in diverse fields to investigate how objects or spatially fixed areas are interrogated [2–7]. For example, an air traffic control specialist (ATCS) must timely detect and control multiple aircraft on a radar display in order to maintain a safe and expeditious flow of air traffic. Through eye tracking data, we can identify which aircraft the ATCS interrogates and what visual search pattern the ATCS applies.

However, the analysis of eye tracking data for a task that requires interrogating moving objects (e.g., an ATCS controlling multiple moving aircraft on a radar display or a weather forecaster determining whether to issue a warning by observing the weather features on a radar display) can be difficult due to the different characteristics of the moving objects and the limited capabilities of the eye tracking system. Furthermore, eye tracking analysis becomes more difficult if the object’s overall shape can change due to the shape change of the object’s elements or the physical relocation of its elements (e.g., an aircraft on a radar screen is composed of elements such as a vector line and a data block, and the length of the vector line can change due to the aircraft speed change, or the data block can be repositioned by the ATCS). The details of the issues are as follows.

In order to map and analyze the eye tracking data for such a task described above, different characteristics of those moving objects need to be identified (Figure 1). Objects can have irregular shapes and sizes and different movement
In this paper, we present designs and algorithms to address the issues raised to facilitate the analysis of the eye fixations at times $t$ (eye fixation 1) and $t + 1$ (eye fixations 2 to 5).

Figure 1: Characteristics of multiple moving objects: each object is in motion except for object “B.” “$A_t$” indicates a circular object at time $t$ and “$A_{t+1}$” indicates the change of its location at time $t + 1$. “$D$” is an object rotating clockwise, and “$E$” is an object changing its shape. The red dots on and around “$C$” indicate the order of eye fixations of interest. Specifically, we need to consider the following aspects. First, the eye fixation should be affected by the experimental conditions. Therefore, observing the eye fixation may have occurred on object “$C$.” In addition to the inherent error of eye tracking systems, accuracy error can also be affected by experimental conditions.

For example, in the actual air traffic control rooms, ATCSs sit close to a large monitor (i.e., 19.83 × 19.83 inches) in order to better detect and control multiple (i.e., sometimes up to 50 or more) aircraft within their sector. For such an environment, the accuracy of the eye tracker can drastically decrease. These issues occur when measuring eye tracking data not only in air traffic control tasks, but also in other various tasks such as during driving or during a virtual simulation of offshore oil and gas operations. Therefore, the visual angle accuracy is not fixed at 0.5° and can vary based on the environmental conditions when we pursue high face validity.

In addition, the mapping of eye tracking data to moving objects can be difficult if there are multiple small objects moving on the display and each object is composed of several elements (e.g., the aircraft position symbol (or target), vector line, and data block). To accommodate the complex shapes of objects as well as the visual angle accuracy, the concept of an area of interest (AOI) can be applied. An AOI is a convex shape that can approximate and represent the complex object shape and can be simple shapes such as circles and rectangles. For example, the AOI can be fixed rectangular areas [5, 6, 13] or moving rectangular areas [9, 14] on a display based on the task types. Note that the size of an AOI should be slightly enlarged to consider the visual angle accuracy [9, 14].

To determine whether an eye fixation occurred on an object, we need to consider two aspects. First, the eye fixation should have occurred within the visual angle error range (e.g., 0.5° from all edge points of an object). Second, there should be no other object or background image to which the eye fixation occurred. In other words, if two objects are in close proximity, it can be difficult to determine which object the participant was interrogating. Even if the objects arrived from different locations, they can come into close proximity and even overlap as time progresses (Figure 3). Although considerable research was conducted to investigate the eye movements of air traffic control operations [15–18], it was limited to creating spatially fixed AOIs or did not elaborate on how overlapping issues were addressed.

Additionally, the mapping issue becomes more complex if the shapes of the multielement objects change. For example, if two aircraft are approaching close proximity, the aircraft position symbols (or targets) as well as the data blocks can overlap, and then an ATCS can reposition the data block (Figure 4). The data block can be repositioned in eight directions relative to the aircraft position symbol (e.g., from the bottom of the target itself to the top or right upper corner of the target) as well as increased in distance (e.g., from 0.5 cm away to 5 cm away).

In this paper, we present designs and algorithms to address the issues raised to facilitate the analysis of the eye fixations.
tracking data for tasks that involve interrogating multielement moving objects that can change their overall shape and overlap with one another by considering different shapes and sizes of the AOIs that are fitted to represent the objects.

2. Conceptual Designs and Algorithms

The main features of our approach are to (1) develop dynamic AOIs that continuously fit the multielement objects into convex or rectangular shapes whenever the objects’ overall shapes or locations change, (2) modify the size of the AOIs (through the concept of AOI gap tolerance) to consider the visual angle error, (3) map the pixel-coordinate based eye fixations with the AOIs, and (4) define eye fixations on overlapping AOIs. Specific to air traffic control operations, the designs and algorithms create AOIs based on matching the pixel-coordinates from the flight data block, target, and vector lines with the pixel-coordinates of the eye fixations. Figure 5 represents the data processing flowchart of the overall methodology. The flowchart consists of seven major steps, which are discussed in detail in the subsequent sections. Note that the introduced algorithm is based on discretized movements of the moving objects, and the background (scene) is fixed.

**Step 1.** Collect and preprocess simulation and eye tracking data.

**Step 1.1 (collect and preprocess simulation data).** Assume the simulation scenario is of $m$ duration in minutes, given an update rate (UR) in seconds (e.g., 1 second), defined as the refresh rate of the objects’ locations and shapes on a display; the total duration of a scenario can be divided into $UR \times m \times 60$ time frames in seconds. Thus, if we want to represent the $m$ minutes scenario into discrete time frames we can represent it as

$$ T = \{UR, UR \times 2, UR \times 3, \ldots, UR \times m \times 60\},$$

where $T$ represents the time frame counter in seconds.
(1) Collect simulation and preprocess simulation and eye fixation data
(2) Develop different types of AOIs
(3) Map eye fixation data with AOIs
(4) Visualize plot eye fixation data on the simulated scenarios
(5) Calculate different metrics of interest for the initial (or minimal) AOI gap tolerance value
(6) Change (increase) AOI gap tolerance value

All AOI gap tolerance values covered?
Yes

(7) Find optimal AOI gap tolerance value

End

Start

Start time

Frame “1”
(Total elapsed time 1 sec)

Frame “2”
(Total elapsed time 2 secs)

Frame “3”
(Total elapsed time 3 secs)

End time

Frame “n”
(Total elapsed time “n” secs)

Figure 5: Data processing flowchart.

Figure 6: Discretization of the simulation video into time frames.

Figure 6 represents an example of the discretization process of the simulation output for a 20-minute duration. Note an observable (or systematic) discrete movement of the object (e.g., aircraft or radar display). In other words, no change in position occurs within a time frame; for example, suppose the simulation starts at 0 seconds, the next change in position of the aircraft will occur at the end of the first second, and the next change will be at the end of two seconds
Table 1: Example of eye fixation data.

<table>
<thead>
<tr>
<th>X pos (pixels)</th>
<th>Y pos (pixels)</th>
<th>Start time (secs)</th>
<th>Stop time (secs)</th>
<th>Duration (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1047</td>
<td>1668</td>
<td>0.35</td>
<td>0.466</td>
<td>0.116</td>
</tr>
<tr>
<td>628</td>
<td>1255</td>
<td>0.816</td>
<td>1.15</td>
<td>0.334</td>
</tr>
<tr>
<td>852</td>
<td>1174</td>
<td>1.5</td>
<td>2.233</td>
<td>0.733</td>
</tr>
<tr>
<td>1150</td>
<td>1690</td>
<td>2.666</td>
<td>2.916</td>
<td>0.25</td>
</tr>
<tr>
<td>1162</td>
<td>1721</td>
<td>2.933</td>
<td>3.416</td>
<td>0.483</td>
</tr>
</tbody>
</table>

and so on. After discretizing the time frames as part of the simulation data preprocessing step, the corresponding multielement object data are identified for each time frame. Let \( P \) be the set that contains all the information of the multielements for the total time duration. Then \( P \) can be represented as

\[
P_{N_{ef},T} = \{ p_{n_{UR},UR}, p_{n_{UR},UR+2}, \ldots, p_{n_{UR}} \text{and } m_{UR} \text{and } m_{UR+60} \},
\]

where \( P_{N_{ef},T} \) is the set of multielement objects present for each time frame.

**Step 1.2** (collect and preprocess eye fixation data). The eye fixation data needs to be processed according to the time discretization strategy used for processing the simulation data. Table 1 represents a small sample of eye fixation data. The first and second columns represent the horizontal and vertical pixel-coordinates of the eye fixations, respectively. The third and fourth columns show the start and stop time of an eye fixation. The fifth column represents the time duration of an eye fixation. The start and stop time values can be used to determine the time frame in which the eye fixations occurred.

The eye fixations during a time frame can be described as

\[
E_{M_{ef},T} = \{ e_{m_{1},UR,UR+2}, e_{m_{2},UR+2}, \ldots, e_{m_{UR+60}} \},
\]

where \( E_{M_{ef},T} \) is the set of eye fixations that occurred for each time frame.

Figure 7 shows an example of eye fixation durations that occurred over the time frames. The time frames are based on the object movement update rate (i.e., objects would make discrete short burst of movements), and eye fixation durations can either fall within a time frame or stretch over more than one time frame.

**Step 2** (develop different types of AOIs). Based on the preprocessed data from Step 1, different types of AOIs were developed. Two types of dynamic AOIs are considered: convex AOI and rectangular AOI. The rectangular AOI is an adaptation from [9], and in this research the shape and size of the rectangular AOI change based on each time frame. The convex AOI was developed by calculating the convex hull of the set of coordinate points used to represent each multielement object. The convex AOIs change their shapes and sizes based on each time frame as well. Figure 8 represents the two different types of AOIs (convex and rectangular) for a multielement object. Thus, if an eye fixation occurs within a dynamic AOI, then we conclude that an eye fixation occurred on the multielement moving object.

To define a parameter that governs the size of the buffer, we define the buffer as the “AOI gap tolerance (AGT).” Since any given AOI corresponds to only one multielement object, \( P_{N_{ef},T} \) can be substituted by \( A OI_{N_{ef},T} \), the set of AOIs during a time frame, as

\[
AOI_{N_{ef},T} = \{ aoi_{n_{UR},UR}, aoi_{n_{UR+2},UR+2}, \ldots, aoi_{n_{UR+60}} \}.
\]

Step 3 (map eye fixation data with AOIs). The “AOI mapping (AM)” performs a match between the eye fixation set and the AOI set during the same time frame. AOI mapping identifies whether the eye fixations fell within the boundaries of the AOIs by comparing the coordinates. The AM can be expressed as

\[
AM : E_{M_{ef},T} \rightarrow AOI_{N_{ef},T}.
\]

The functional mapping described in (5) is called a many-to-many mapping. Many-to-many mapping refers to the fact that eye fixations can be mapped to more than one AOI index and similarly AOIs can also be mapped to more than one eye fixation during a time frame. For example, in a single time frame, two or more eye fixations (that have different pixel-coordinates) can occur within a single AOI, or two or more AOIs can share a single eye fixation (when overlapping). The resulting mapped AOIs during the time frame \( t \) can be expressed as \( AM(e_{m_{j},t}) = aoi_{n_{j},t} \). The collection of all mapped AOIs can be defined as a “mapped AOI set (MA)” and be written as

\[
MA_{I,T} = \{ ma_{j} \mid AM(e_{m_{j},t}) = aoi_{n_{j},t}, ma_{j} \in aoi_{n_{j},t} \},
\]
where $\text{MA}_{TF}$ is the set of mapped AOIs during a time frame and $I$ is index.

Figure 9 represents a mapping example where the rectangular and convex AOIs are shown in green. The red “+” symbol represents the eye fixation point that falls within the AOI boundary. There may be situations when an eye fixation falls inside the boundary of more than one AOI simultaneously. In other words, the eye fixation falls into a region that is in the intersection of several AOI boundaries, thus giving rise to the concept of “overlapped AOI mappings.” Thus, in this example, the mapped AOI set for this eye fixation will include three elements, which can be shown as $\text{MA}_{TF} = \{a_{12}, a_{22}, a_{32}\}$.

Another important concept, which will be useful in the analysis, is the cardinality of the MA set, where cardinality is the number of elements present in that set. This can be expressed as follows:

$$|\text{ma}_j| = c, \quad c \in \{0, 1, 2, \ldots, n_t\},$$

where $|\cdot|$ is the cardinality function and $n_t$ is the number of multielement objects present at time $t$.

Thus, if “$c$” is the cardinality of the $\text{ma}_j$ set we can say that the corresponding eye fixation index has been mapped to “$c$” number of AOIs simultaneously. The larger the cardinality of the $\text{ma}_j$ set, the greater the difficulty in analyzing those eye fixations.
fixations. Therefore, an important consideration in the data
analysis is the frequency distribution of different cardinal
values of the \( m_{A_{ij}} \) set.

**Step 4** (visualize plotted eye fixation data on the simulated
scenarios). After the mapping process, the eye fixation data
is overlaid on the simulated display as a function of time
using the update rate. This process requires subsequently
plotting both the eye fixations and AOI data pertaining to the
same time frames and covering the time frames sequentially.
Example cases are shown in Figure 12.

**Step 5** (investigate the mapping effects for different AOI gap
tolerance (AGT) values). Some of the metrics that are of
particular interest for this study are (1) the “percentage of
the number of eye fixations falling inside AOIs (PNFIA)” and
(2) the “percentage of the duration of the eye fixations falling
inside AOIs (PDFIA).”

PNFIA is defined as

\[
\text{PNFIA} = \frac{\text{Number of eye fixations falling inside AOIs}}{\text{Total number of eye fixations}}, \quad \text{(8)}
\]

where the number of eye fixations falling inside the AOIs (in
(8)) is

\[
\text{Number of eye fixations falling inside AOIs} = \sum_{t=1}^{\text{max}(T)} \sum_{i=1}^{m_{t}} A_{i,t} \quad \text{(9)}
\]

where \( \text{max}(T) \) is maximum value of the time frame count and
\( m_{i} \) is the number of eye fixations during time frame \( t \):

\[
A_{i,t} = \begin{cases} 1 & \text{if } |m_{A_{ij}}| \neq 0 \\ 0 & \text{otherwise} \end{cases}, \quad \text{(10)}
\]

where the cardinality function is expressed as \(| \cdot |\) (e.g., \( |m_{A_{ij}}| \)).

\( A_{i,t} \) is the indicator function that becomes 1 if the
cardinality of the corresponding set \( m_{A_{ij}} \) is nonzero; in other
words this function takes the value of 1 if the associated eye
fixation falls at least within one AOI boundary. Therefore,
using (9) and (10) we get

\[
\text{PNFIA} = \frac{\sum_{t=1}^{\text{max}(T)} \sum_{i=1}^{m_{t}} A_{i,t}}{N} \quad \text{(11)}
\]

where \( N \) is the total number of eye fixations.

PDFIA is defined as

\[
\text{PDFIA} = \frac{\text{Time duration of eye fixations falling within AOIs}}{\text{Total time duration of eye fixations}}, \quad \text{(12)}
\]

The time duration of eye fixations falling within AOIs is
calculated as

\[
D = \sum_{t=1}^{\text{max}(T)} \sum_{i=1}^{m_{t}} d_{i,t} \quad \text{(13)}
\]

where \( d_{i,t} \) is time duration of eye fixation index \( i \) during time
frame \( t \) and \( m_{t} \) is the number of eye fixations that occurred
during time frame \( t \).

For the purpose of calculating the time duration of eye
fixations falling within an AOI, we need to consider only
those eye fixation indexes for which the cardinality of their
corresponding AOI mapped set is nonzero. Therefore we
can use the indicator function described in (10) to take into
account only those specific eye fixation indexes that fall at
least within one AOI boundary. Thus we get the following:

\[
D' = \sum_{t=1}^{\text{max}(T)} \sum_{i=1}^{m_{t}} A_{i,t} \times d_{i,t} \quad \text{(14)}
\]

Using (13) and (14) we get that the percent time duration of
eye fixations falling within an AOI to be

\[
\text{PDFIA} = \frac{D'}{D} \quad \text{(15)}
\]

The next metric of interest is the frequency distribution
of \( m_{A_{ij}} \) of various cardinalities. In other words, it is the
frequency distribution of various possible “c” values, where \( c \)
is as described in (7). This can be found by counting the
number of occurrences of various possible values of “c.” This
frequency distribution is an important metric because it is a
qualitative measure of the difficulty associated with the
analysis of the eye fixation sequence.

**Step 6** (change AOI gap tolerance (AGT) values). Due to the
visual angle error, the choice of the AGT value depends on the
discretion of the analyst. In absence of any established rela-
tionship between the AGT values and the relevant eye fixation
parameters discussed in an earlier section, the optimal range
of the AGT value becomes very much context dependent. As
a result, it becomes important to study this relationship for
the present context. Thus, the next step involves varying the
AGT value to investigate its impact on the relevant metrics of
interest. The equation governing the change in AGT can be
written as

\[
\text{AGT}_{r+1} = \text{AGT}_{r} + \delta \quad \text{(16)}
\]

where \( \text{AGT}_{r} \) is AOI gap tolerance value for the iteration
value \( r \) and \( \delta \) represents increments of AGT values (e.g., \( \delta = 5 \)

pixels).

Table 2 shows the various values of the iteration counter
\( r \) and the associated AGT values. All the above-mentioned
steps need to be performed from Steps 2–5 for each \( r \) value.

**Step 7** (find optimal AOI gap tolerance value). Assuming that
a participant or a group of participants interrogate one object
at a time, one method to find the optimal AGT value is to
select the AGT value that provides the highest frequency of
the mapped AOI set of cardinality 1, or in other words we can
identify the optimal AGT value for which the number of eye
fixations on single AOIs is maximum.
### Table 2: AGT values defined for each iteration (r).

<table>
<thead>
<tr>
<th>r</th>
<th>AGT (pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>7</td>
<td>35</td>
</tr>
<tr>
<td>8</td>
<td>40</td>
</tr>
<tr>
<td>9</td>
<td>45</td>
</tr>
<tr>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>11</td>
<td>55</td>
</tr>
<tr>
<td>12</td>
<td>60</td>
</tr>
<tr>
<td>13</td>
<td>65</td>
</tr>
<tr>
<td>14</td>
<td>70</td>
</tr>
<tr>
<td>15</td>
<td>75</td>
</tr>
<tr>
<td>16</td>
<td>80</td>
</tr>
<tr>
<td>17</td>
<td>85</td>
</tr>
<tr>
<td>18</td>
<td>90</td>
</tr>
<tr>
<td>19</td>
<td>95</td>
</tr>
<tr>
<td>20</td>
<td>100</td>
</tr>
</tbody>
</table>

The equation to find the optimal AGT value ($AGT_{optimal}$) is as follows:

$$AGT_{optimal} = \arg \max_{AGT \in \{5,10,\ldots,100\}} \left[ freq(c) : c = 1 \right],$$

(17)

where $c$ is cardinality of the mapped AOI set and $freq(c)$ is frequency of set with cardinality value $c$.

Note that we can also obtain an overall single near optimal AGT value recommended for an experiment if we used the aggregated eye tracking data obtained from multiple participants.

Pseudocode 1 shows the simplified pseudocode based on the algorithmic flowchart shown in Figure 5.

### 3. Implementation

The developed approach was benchmarked through retired professional air traffic control specialists (ATCSs) who primarily work as instructors for the Federal Aviation Administration (FAA). The experiment was held at the FAA Civil Aerospace Medical Institute (CAMI), located in Oklahoma City, OK.

#### 3.1. Participants

Ten certified ATCSs with over 32 years of experience participated in the experiment. In addition, three FAA employees participated as pseudo pilots who maneuvered the aircraft based on the controllers’ clearances. Eye tracking data were collected from the certified controllers.

Due to the unforeseen technical issues when using the eye tracking system and the air traffic control simulator, the data obtained from the first five participants were discarded, and only the data obtained from the subsequent five participants were used.

#### 3.2. Apparatus

The experiment environment closely resembled the actual environment in the field (Air Route Traffic Control Center) in order to obtain high face validity. The simulated air traffic scenarios were displayed using a 19.83 × 19.83-inch monitor (2048 × 2048-pixel active display area). The size and resolution were equivalent to the actual display size used in the field. An additional monitor was placed to the right of the simulation monitor to display the En Route Automation Modernization (ERAM) tool, a decision support tool that provides text data with respect to aircraft data, trajectory, and possible conflicts. A keyboard was placed beneath the simulation monitor for an ATCS to input commands.

The eye tracking data were only collected from the simulation monitor to test our designs and algorithms. FaceLab 5 eye tracker system [11] was used to collect the eye tracking data with a sampling rate of 60 Hz. The threshold for defining a fixation was set at 100 ms. The accuracy of the eye tracker was in the range of 0.5°–1° of visual angle error. Each participant’s eyes were approximately in the range of 55–70 cm from the simulated display. Kongsberg-Gallium 1-Sim software, internally outsourced and used by the FAA, was used for generating three different air traffic scenarios. The refresh rate of the simulated radar display was 1 second. Obtained raw eye tracking data was exported through the Eyeworks software [21], and the data output was similar to that shown in Table 1.

The structure of the air traffic simulation file is provided in Table 3 (sample data). The output file contains the details of the aircraft movements, their coordinates, and other relevant details of the aircraft representation used for the simulation. The data update rate (UR) was 1 second. In Table 3, the first and second columns show the elapsed time from the start of the experiment and the actual time of day, respectively. The third column named “aircraft code” shows the code name of the aircraft under consideration. The fourth column is the “target” column which shows the horizontal (X pos) and vertical (Y pos) coordinates of the targets (aircraft) in pixels. The fifth column is the “data block” column which has three subparts: (1) top left corner coordinates of the data block, (2) bottom right coordinates of the data block, and (3) direction column that represents the relative location of the aircraft with respect to the target position (N (north), NE (northeast), E (east), SE (southeast), S (south), SW (southwest), W (west), and NW (northwest)). The last column provides the position coordinates in pixels of the vector line’s end point.

#### 3.3. Task and Scenarios

The task was a high fidelity representation of air traffic control as performed in the U.S. National Airspace System’s Air Route Traffic Control Centers. Controlling simulated traffic such as this requires an experienced ATCS to observe the radar screen and give
for \( r = 1 \) till max(\( r \)) (loop to cover all iterations)
   for \( t = 1 \) till max(\( T \)) (loop to cover all time frames)
      for \( j = 1 \) till \( n \) (loop to cover all multi-element objects for the current time frame \( t \))
         Plot \( j \)th plane elements for time frame \( t \)
         Plot \( j \)th AOI boundary for time frame \( t \)
      end for
   end for
   for \( i = 1 \) till \( m \) (loop to cover all eye fixation \( e_{ij} \) for the current time frame \( t \))
      Plot \( i \)th eye fixation \( e_{ij} \) for time frame \( t \)
      for \( j = 1 \) till \( n \) (loop to check whether the current eye fixation falls within the AOI list of the current time frame \( t \))
         find whether \( e_{ij} \) falls inside \( AOI_{ij} \)
         store the result: store 1 for inside, 0 for outside AOI
         store the time duration of the eye fixation
      end for
   end for
   calculate percent number of eye fixations within AOI
   calculate percent time duration of eye fixations within AOI
   calculate the frequency distribution of mapped AOI sets of various cardinalities
   calculate the optimal AGT value

Pseudocode 1: Pseudocode used for the overall process.

### Table 3: Air traffic simulation sample output data.

<table>
<thead>
<tr>
<th>Scenario time</th>
<th>Time of the day</th>
<th>Aircraft code</th>
<th>Target X pos</th>
<th>Target Y pos</th>
<th>Top left X pos</th>
<th>Top left Y pos</th>
<th>Bottom right X pos</th>
<th>Bottom right Y pos</th>
<th>Direction</th>
<th>Vector line end point X pos</th>
<th>Vector line end point Y pos</th>
</tr>
</thead>
<tbody>
<tr>
<td>00:00:14</td>
<td>11:55:19</td>
<td>DAL1268</td>
<td>988</td>
<td>939</td>
<td>1050</td>
<td>922</td>
<td>1141</td>
<td>994</td>
<td>E</td>
<td>957</td>
<td>892</td>
</tr>
<tr>
<td>00:00:14</td>
<td>11:55:19</td>
<td>EGF1819</td>
<td>599</td>
<td>1248</td>
<td>608</td>
<td>1282</td>
<td>699</td>
<td>1354</td>
<td>S</td>
<td>624</td>
<td>1237</td>
</tr>
<tr>
<td>00:00:14</td>
<td>11:55:19</td>
<td>N21LD</td>
<td>732</td>
<td>1300</td>
<td>792</td>
<td>1334</td>
<td>883</td>
<td>1406</td>
<td>SE</td>
<td>742</td>
<td>1277</td>
</tr>
<tr>
<td>00:00:15</td>
<td>11:55:20</td>
<td>DAL1268</td>
<td>987</td>
<td>938</td>
<td>1050</td>
<td>922</td>
<td>1141</td>
<td>994</td>
<td>E</td>
<td>957</td>
<td>892</td>
</tr>
<tr>
<td>00:00:15</td>
<td>11:55:20</td>
<td>EGF1819</td>
<td>599</td>
<td>1248</td>
<td>446</td>
<td>1231</td>
<td>537</td>
<td>1303</td>
<td>W</td>
<td>624</td>
<td>1237</td>
</tr>
<tr>
<td>00:00:15</td>
<td>11:55:20</td>
<td>N21LD</td>
<td>732</td>
<td>1299</td>
<td>792</td>
<td>1334</td>
<td>883</td>
<td>1406</td>
<td>SE</td>
<td>742</td>
<td>1277</td>
</tr>
</tbody>
</table>

### Table 4: Characteristics of different simulation scenarios.

<table>
<thead>
<tr>
<th>Scenario name</th>
<th>Average unique aircraft per frame</th>
<th>Min unique aircraft per frame</th>
<th>Max unique aircraft per frame</th>
<th>Std dev unique aircraft per frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderate traffic</td>
<td>20</td>
<td>7</td>
<td>30</td>
<td>7</td>
</tr>
<tr>
<td>Moderate traffic + weather feature</td>
<td>20</td>
<td>6</td>
<td>29</td>
<td>6</td>
</tr>
<tr>
<td>Busy traffic</td>
<td>24</td>
<td>8</td>
<td>37</td>
<td>7</td>
</tr>
</tbody>
</table>

clearances to aircraft adjusting their altitudes, headings, or speeds so as to maintain aircraft-to-aircraft separation and route aircraft through the sector or to their destination airport within the sector. The ATCSs gave voice commands, via the communication system, to pseudo pilots who were situated in a remote room. The pseudo pilots followed the clearances and provided read-back to the ATCSs. Three scenarios were used (moderate traffic, moderate traffic with convective weather, and busy traffic). The duration of each scenario was 20 minutes. Table 4 and Figure 10 show the details of the scenarios. In Figure 10(b), the blue patch represents the weather feature.

3.4. Data Analysis. The analysis of convex and rectangular AOIs was automated as follows: Based on the provided simulation output and the eye tracking output, both data sets were synchronized (step (1) in Figure 5). After the preprocessing steps, the two types of AOIs (convex and
rectangular AOI) were created using the aircraft coordinates at every second (step 2). Then, mapping was performed using the eye tracking data and the simulation data (step 3). The mapped data was visualized (step 4), and relevant metrics including the PNFIA and PDFIA were calculated by varying the AGT values (steps 5 and 6). Finally, the optimal AGT value was obtained by identifying the highest percentage of the eye fixations on single AOIs (step 7).

The complexity of the data processing time was $O(n_1 n_2 n_3 n_4 n_5 n_6)$, where $n_1$ is the number of participants, $n_2$ is the number of scenarios, $n_3$ is the number of AOI types, $n_4$ is the number of AGT values, $n_5$ is the number of AOIs per time frame, and $n_6$ is the number of eye fixations per time frame. Each eye fixation was compared with each AOI per time frame.

In the Results, the total eye fixation numbers and durations on the display (without using AOIs) were plotted in order to investigate the oculomotor trends. Then, aggregated PNFIA and PDFIA values for all participants were plotted based on the AGT values. Then, the number of eye fixations
that occurred on single and multiple overlapping AOIs was plotted based on the AGT values. The optimal AGT value was computed, and examples of different scanpath sequences (resulting from either different AOI types or AGT values) were identified.

4. Results

The oculomotor trends are shown in Figure 11. Figure 11(a) shows the total number of eye fixations and Figure 11(b) shows the total duration of eye fixations with respect to scenario difficulties: moderate traffic (Mod), moderate traffic with weather feature (Mod+W), and busy traffic (Busy). The legends in Figure 11 showing 1, 2, 3, 4, and 5 represent the participant numbers.

Figure 12 displays example snapshots of the visualization process (see Step (4) in Figure 5) for both AOI types. The example snapshots show the dynamic AOIs with the AGT value set to 40 pixels. In Figure 12, the AOIs are highlighted in green and the order of eye fixations along with the associated saccades (connections between eye fixations when moving from one to the next) are highlighted in red. Note that the automated illustrations of the ordered eye fixations (shown in numbers) and the saccades linking the eye fixations are accumulated, meaning that the illustrations show all eye fixations from the scenario start time (time frame 1) until the indicated time frame such as time frame 120 or 1200.

Figure 13 depicts the effect of changing the AGT values on (1) the percentages of the numbers of eye fixations that fall within AOIs (PNFIA) shown in grey and (2) the percentages of the durations of the eye fixations that fall within AOIs (PDFIA) shown in black. The plots show the mean and standard error associated with every AGT value. In addition, the fitted polynomial equations and the $R^2$ values are provided.

Figure 14 depicts the change in the frequency of mapped AOI sets, of various cardinalities, with respect to the change in AGT values for convex and rectangular AOIs, respectively. The plots show the mean and the standard error associated with the coverage percent values. The maximum possible observed cardinality of the mapped AOI set is 8. A general trend among the various plots is that the frequency count of the $m_{c}$ set having cardinality 1 (or in other words $c = 1$ (shown in red)) increased and then decreased. As the AGT values increased, the number of overlapping AOIs also increased, and the eye fixations on a single AOI subsequently decreased.

The near optimal (or recommended) AGT values (by considering all participants and scenarios) are provided in Table 5. The AGT value of 40 pixels captures approximately 70–80% of the total eye fixations that fall within the AOIs. Note that the participants can freely observe other areas that are not defined as AOIs within the display.

Figure 15 depicts the change in the frequency of mapped AOI sets, of various cardinalities, with respect to the change

Table 5: Mean and standard error for the optimal AGT values for different AOI types.

<table>
<thead>
<tr>
<th>Optimal AGT</th>
<th>Convex AOI</th>
<th>Rectangular AOI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (pixels)</td>
<td>40</td>
<td>38.3</td>
</tr>
<tr>
<td>Standard error (pixels)</td>
<td>1.8</td>
<td>1.2</td>
</tr>
</tbody>
</table>
in AGT values for convex and rectangular AOIs, respectively. The plots show the mean and standard error associated with the frequency values for every AGT value. The maximum possible observed cardinality of the mapped AOI set is 8. In many cases the frequency of cardinality values higher than five was zero. Thus the curves for these cardinalities might not be exclusively visible on the plots as they are overlapping each other. A general trend among the various plots is that the frequency count of the \( m_{\alpha} \) set having cardinality 1 (or in other words \( c = 1 \) (shown in red)) increased and then decreased. As the AGT values increased, the number of overlapping AOIs also increased, and as a result, the eye fixations on a single AOI subsequently decreased.

Figure 16 shows examples of how different AGT values can affect the resulting AOI-based scanpath sequences. More relevant eye fixations were captured when using the optimal AGT value of 40 (obtained from our experiment) than the AGT value of 5. As shown in Figure 16, the identified scanpath sequence “FFCC(A,B)E” (Figure 16(b)) shows much more pertinent mappings compared to the scanpath sequence “CCA” (Figure 16(a)). Again, note that the scanpath sequences can be further collapsed into “FC(A,B)E” and “CA,” respectively.

5. Discussion

An approach was developed that automatically (1) created rectangular and convex AOIs around multielement objects, (2) mapped eye fixations with different types of AOIs, (3) systematically evaluated the mapping characteristics by increasing the size of the AOIs to consider the fidelity of the eye trackers, and (4) investigated how the increase of the AOI sizes affects the overlapping of multiple AOIs. This approach was applied to the collection of visual scanning data from a...
Figure 13: Plots of coverage percentages of the numbers and durations of the eye fixations that occurred within the AOI versus AGT values: the figures on the left column are the results for the convex type, and the figures on the right column are the results for the rectangular type.
Figure 14: Distribution of the number of eye fixations on single or overlapped AOIs based on AGT values: the top red line shows the change of the number of eye fixations for a single AOI. The subsequent lines show the change of the number of eye fixations on overlapping AOIs (increasing from 2 to 8).
high fidelity simulation of an air traffic control task. The task required ATCSs to interrogate multielement moving objects (that can change their overall shapes) on a radar display. The approach was applied to eye tracking data collected from the ATCSs as they performed the conflict detection and control task through interrogating multiple moving aircraft within their sector.

The oculomotor statistics on different types of scenarios show that the overall eye fixation numbers and durations on the display (without considering AOIs) did not significantly differ among the scenarios. The results differ from previous aircraft conflict detection research [22, 23]. In [22], eye fixation numbers and durations increased as the difficulty level increased (easy: many aircraft had different altitudes; moderate: many aircraft had similar altitudes; difficult: many aircraft changed altitudes), while setting the number of aircraft on the display at twelve for all scenarios. In [23], eye fixation numbers and durations increased as the number of aircraft on the display was increased from twelve to twenty. A major difference in the scenario settings was that there was no time limit on detecting possible collisions for [22, 23], whereas the experiment in this research had a time limit of twenty minutes.

Regarding the ATCSs’ cognitive processes, one reason that similar oculomotor trends could be found is that the ATCSs were constantly vigilant on interrogating and controlling the aircraft throughout the experiment. In addition, the reason for a marginal decreasing trend on eye fixations and durations may be due to the order effect of the scenarios being performed in a sequence of moderate traffic, moderate traffic with convective weather, and busy traffic. The participants could have become more comfortable with the situation as they continued to control the multiple aircraft. Another possibility is that the ATCSs may have spent more time on looking at the ERAM display as well as the keyboard. Unfortunately, the exported eye tracking data only provides pixel-based eye fixations that occurred within the defined display; therefore, it is difficult to know where the eye fixations occurred outside the display.

The convex and rectangular AOI types did not generally affect the amount of mapped eye fixations among the participants and the scenarios due to the relatively small size of AOI-based scanpath sequence: FFCCAE
(a) AGT = 40 (pixels)

AOI-based scanpath sequence: FFCC(A,B)E
(b) AGT = 40 (pixels)

Figure 15: Examples illustrating how AOI types can affect the AOI-based scanpath sequences: the red “+” shows the location of the eye fixations, and the numbers are the corresponding eye fixation orders. For (a), eye fixation 5 only falls inside AOI “B,” whereas for (b) eye fixation 5 falls inside both AOIs “A” and “B.”

AOI-based scanpath sequence: CCA
(a) AGT = 5 pixels

AOI-based scanpath sequence: FFCC(A,B)E
(b) AGT = 40 pixels

Figure 16: Examples illustrating how AGT values can affect the AOI-based scanpath sequences: the red “+” shows the location of the eye fixations, and the numbers are the corresponding eye fixation orders. For (a), eye fixations 1, 2, 6, and 7 fall outside the AOIs, whereas for (b) only eye fixation 7 falls outside the AOIs.
the objects as well as the accuracy of the eye tracking system for a high face validity experiment. However, we were able to identify specific examples of different AOI types affecting the resulting scanpath sequence (Figure 15). The analysis of human performance using the scanpath sequences may have substantially differed for the same experiment if the analysts applied different AOI types. The effect may have been overall significant if the size of the multielement objects was bigger due to the increased unnecessary area (Figure 8) created by the rectangular AOI type. The unnecessary areas would also result in creating more overlapping AOI areas.

The AGT values substantially impacted the amount of covered eye fixations and durations on both AOI types and the trends fitted to polynomial equations. Up to a certain point, the increase of the AGT value was able to accommodate many eye fixations that occurred around the objects; then the increase rate (of the amount of included eye fixations) began to reduce since lower amount of eye fixations occurred further away from the objects. The eye fixation numbers and durations were highly correlated for our experiment. Note that the AGT values also affected the resulting scanpath sequences (Figure 16). The use of too tightly fitted AOIs resulted in missing many eye fixations that occurred around the object. Note that if we used AOIs that were too large, then the cardinality of the mapped AOI set would increase, leading to either inaccurate mapping or an increase in the complexity of the scanpath sequences by having more overlapping AOIs.

Thus, the selection of the AGT value gives rise to a trade-off between the coverage (amount of eye fixations) versus complexity (overlapping AOIs) of the algorithm because the more we increase the coverage, the more we increase the complexity. As the AGT value increases, the coverage of the overlapping AOIs increases accordingly, but the coverage of the single AOIs starts to decrease (Figure 14). The reason is that overlapping AOIs begin to take away the amount of eye fixations that occurred within single AOIs. Therefore, we were able to determine the near optimal AGT value by identifying the coverage peak of single AOIs. Having an adequate AOI size to map an eye fixation to a single AOI is more preferred to having larger AOIs that would result in creating unnecessary overlapping areas. In other words, the more we increase the coverage, the more we increase the complexity for multielement moving objects that can overlap.

In addition, the near optimal AGT values were obtained from aggregated data across the whole experiment and among the participants. The limitation to this approach is that we apply a constant AGT value for the whole duration. The optimal AGT value might not be a constant for all the time frames, and further detailed analysis might help to segregate time segments from the whole experimental duration (i.e., identify the amount of variations for different segregated time segments). Note that we would not be able to obtain a trend to identify the optimal value if the time length was too short (e.g., for a 1-second time frame, we would only obtain 1 or 2 eye fixations). To investigate how it would vary, we would first need to define the time segments that we should apply.

Another limitation is that we assumed that the multielement objects make discretized movements and that the scene (background) is fixed. If the background is moving or the objects make rapid movements (e.g., from one end of the screen to another end of the screen in a very short time), then our approach would not work. These issues are difficult to solve and should be addressed in our subsequent research.

The overarching goal of our research is to obtain more accurate mappings between the eye movements and the moving objects in order to better support the analysis of human performance. This research concentrated on prototyping, implementing, and evaluating new conceptual designs and algorithms to obtain more accurate mappings. Based on the obtained results in this research, we are currently analyzing the human performance based on the obtained AOI-based scanpath sequences through the Directed Weighted Networks [24, 25].

Furthermore, the results can be a basis to develop better scanpath analysis methods that build upon existing methods [22, 26–31], mimic human performance [32], and develop data visualization methods for active learning using the experts’ visual scanning patterns [33]. In addition, the visual scanning data could be combined with EEG analysis [34] to better understand how the different types of tasks or incidents affect brain response and visual scanning and how the brain response data is correlated with visual scanning data.

7. Conclusion

To address the issue of mapping eye fixations with multielement objects (that move, can change their shape, and overlap over time), we proposed and implemented dynamic AOIs that represent the multielement objects. During the process, we showed a way to map eye fixations to overlapping AOIs. In addition, the concept of AGT was applied in order to address the issue of the fidelity of the eye trackers. Our approach was automated and applied to data collection from a high-fidelity simulation of an air traffic control task. The benchmark showed that eye tracking data analyses can substantially differ based on how the AOIs are defined and how we can obtain near optimal values to better define the AOIs.

Competing Interests

There are no competing interests to declare.
Acknowledgments

This research was funded based on a cooperative agreement with the FAA NextGen Organization’s Human Factors Division, ANG-CI (Award no. 15-G-006) and conducted through collaboration with researchers at the Civil Aerospace Medical Institute’s Aerospace Human Factors Division. The authors deeply appreciate the support from Dr. Carol Manning.

References


