Research Article

Eye Gaze Assistance for a Game-Like Interactive Task

Tamás (Tom) D. Gedeon, Dingyun Zhu, and B. Sumudu U. Mendis

Department of Computer Science, Australian National University, Canberra ACT 0200, Australia

Correspondence should be addressed to Tamás (Tom) D. Gedeon, tom@cs.anu.edu.au

Received 27 February 2008; Accepted 14 May 2008

Recommended by Kok Wai Wong

Human beings communicate in abbreviated ways dependent on prior interactions and shared knowledge. Furthermore, humans share information about intentions and future actions using eye gaze. Among primates, humans are unique in the whiteness of the sclera and amount of sclera shown, essential for communication via interpretation of eye gaze. This paper extends our previous work in a game-like interactive task by the use of computerised recognition of eye gaze and fuzzy signature-based interpretation of possible intentions. This extends our notion of robot instinctive behaviour to intentional behaviour. We show a good improvement of speed of response in a simple use of eye gaze information. We also show a significant and more sophisticated use of the eye gaze information, which eliminates the need for control actions on the user’s part. We also make a suggestion as to returning visibility of control to the user in these cases.

Copyright © 2008 Tamás (Tom) D. Gedeon et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

We consider an interactive task, which involves a human-controlled “robot” (called a “foreman,” effectively an onscreen agent) and a pair of assistant robots. We introduce the technology of fuzzy logic which is used by the robots to infer human intentions. We briefly reprise the fundamentals of fuzzy logic in a nonmathematical fashion, followed by definition of fuzzy signatures. The application of our fuzzy signatures to modelling eye gaze paths is discussed. Eye gaze is important in human communication [1] and is likely to become important in communication with computers, especially games. Finally, we discuss our previous results and the improvements which derive from the addition of eye gaze to allow the assistant robots to better predict human intentions. In the conclusion, we summarise the contributions of this work and suggest how a perception of extra control by the user may be a useful consideration for future systems of this nature.

2. BACKGROUND—EYE GAZE IN COMPUTER GAMES

The increasing availability of relatively inexpensive and reliable eye-gaze detectors has sparked much interest in their potential use in games. These uses have been in 4 main areas as follows.

1. Use of eye gaze to improve rendering, where the region of interest or attention is determined from the eye gaze and more detail is shown in these regions [2, 3]. This has been demonstrated for still images and proposed for games.

2. Avatar realism for enhanced interaction with humans, where an avatar will use some plausible eye gaze motions [3], or to actually track the user’s eye gaze [4]. The merely “plausible” eye-gaze motions suffer from a lack of real interaction. This can be compensated for by reducing the visual realism of the avatar to reduce the behavioural realism expected with regards eye gaze, but this is opposed to current game trends towards increased realism. Actual tracking has obvious advantages over generation of “plausible” eye-gaze motions. It also appears that “humans do not conform to social norms of politeness when addressing an agent” [5], which may further limit the degree to which such agents can enhance games.

3. Eye gaze can be used in therapeutic games, for example, in training of autism spectrum disorder children [6, 7].

4. Use the user’s eyes as active game controls via their eye gaze has generated significant interest. Use of eye
gaze for pointing [7] is faster than pointing using hand gestures, but is at the cost of spatial memory. There is also the use of eye gaze in first-person shooter games. This use can end up reactive, as one’s eyes tend to track objects rather than aim to where they will be soon [8], which is what is required to shoot something.

The first 3 of these areas should be seen as incremental enhancements of current games technologies or an enhanced application. The last of these is a novel development. This notion has its potential disadvantages, of user fatigue, as fine eye control can be tiring [9], and eye gaze-based control is mostly suitable [10] for “applications in which a high-control resolution is not a requirement.”

None of these uses of eye-gaze information is similar to our notion of unobtrusive inference from the user’s eye gaze. To enhance the immersive experience of a good computer game, we believe our unobtrusive technique of inference with fuzzy signatures and possibility will be of significant benefit, as it does not depend on a user using a normal human action out of its usual context.

3. GAME-LIKE INTERACTIVE TASK

Scenario. The scenario of cooperating intelligent robots [11], which we style as context-dependent reconstructive communication [12, 13]: there are a set of identical oblong objects in a space. Various configurations can be built from them, such as a large U-shape, a large T-shape, and a very large oblong, rows of objects.

A group of autonomous intelligent robots is supposed to build the actual configuration according to the exact instructions given to the “robot foreman” (R0). The other robots have no direct communication links with R0, but they are able to observe the behaviour of R0 and all others, and they all posses the same codebook containing all possible object configurations. This scenario could correspond to a real-world firefighting scenario, where interrobot communication could be hampered by the adverse environment or a multiplicity of reasons why this was not allowed in a game environment.

The individual objects can be shifted or rotated, but two robots are always needed to actually move an object, as they are heavy. If two robots are pushing in parallel, the object will be shifted according to the joint forces of the robots (see Figure 2).

If the two robots are pushing in the opposite directions positioned at the diagonally opposite ends, the object will turn around the centre of gravity. If two robots are pushing in parallel, and one is pushing in the opposite direction, the object will not move.

Under these conditions, the task can be solved, if all robots are provided by suitable algorithms that enable “intention guessing” from the actual movements and positions, even though they might not be unambiguous.

Figure 3 illustrates the first configuration. For the other target configurations, we used the same initial configuration of objects and robots for simplicity. This configuration was chosen before any experiments were run, and we are not aware that it has any special properties.

The tasks/experiments performed using our simulator ranged from the following:

(i) the control case of two humans operating their own robots,
(ii) to a human operated foreman and assistant robot throughout,
(iii) to a human operated foreman and assistant robot to start the task which is then completed by the two assistant robots.

Most of the experiment subjects/users we had commented that this task could be easily generalised to a game, particu-
larly in the last case as they could see the benefit of assistants who could complete tasks once the decision was made and (implicitly) communicated by the human operator of the foreman.

It seems to us that the key benefit comes from the ability to implicitly communicate the task. There seems to be a benefit from starting and then having some automatic process finish for you. The ability to “show what needs to be done” allows us to eliminate any need for some kind of language (which the user would need to learn) to specify what is to be done. It is appropriate to note at this point that if the assistant robots make a mistake and are clearly trying to complete the wrong task, the foreman can intercede by blocking an action which leads to reevaluation of the whole guessing-of-intentions by the assistant robots.

To make the implicit communication effective, we have implemented a form of “instinctive behaviour” for the robots. These behaviours are selected by a robot, when it is unable to work out a more intentional action based on the current situation. At this time, these behaviours only include following the foreman around, and moving to the nearest object. Thus, the assistant robots are more likely to be “nearby” when they can infer the human’s task. Hence, their assistance arrives sooner.

4. FUZZY SOLUTION FOR TASK INFERENCE

We first reprise fuzzy logic/signatures.

4.1. What is fuzzy logic?

Fuzzy logic is an extension to traditional binary sets. A fuzzy set can have any value from 0 to 1. So, a set of tall people would include someone only very slightly taller than average with a value of 0.51 and someone slightly shorter who would have a value of 0.49. If we were to use their height to predict their weight, we would get much better results using 0.51 and 0.49, rather than if we used the rounded up values of 1, 0 (i.e., 0.51 is quite similar to 0.49, while 1 is not very close to 0).

4.2. What are fuzzy signatures?

A hierarchical structure to organise data, which is similar to the way human beings structure their thoughts [14]. Fuzzy signatures are vector-valued fuzzy sets, where each vector component can be a further vector-valued fuzzy set, and use aggregations to propagate the fuzzy values from low levels to high levels in the structures, resulting in effective and efficient fuzzy inference. Such aggregations encompass the simple classic fuzzy conjunction and union operations, but still maintain transparency of fuzzy reasoning.

Each signature corresponds to a nested vector structure/tree graph [15, 16]. The internal structure indicates the semantic and logical connection of state variables corresponding to leaves of the signature graph. For example, contains 9 variables \( x_1, \ldots, x_9 \) which describe the problem.

\[
\begin{bmatrix}
    a_{11} \\
    a_{12} \\
    a_2 \\
    a_{31} \\
    a_{32} \\
    a_{33} \\
    a_{411} \\
    a_{412} \\
    a_{42}
\end{bmatrix}
\begin{bmatrix}
    x_1 \\
    x_2 \\
    x_3 \\
    x_4 \\
    x_5 \\
    x_6 \\
    x_7 \\
    x_8 \\
    x_9
\end{bmatrix}
\]

The variables have tightly connected subgroups \( \{x_7, x_8\}; \{x_1, x_2\}; \{x_4, x_5, x_6\}; \{x_7, x_8, x_9\} \), the subscripts of \( x \) are the sequential positions in the signature. In many cases, not all data has full signatures, as shown in (2).

\[
\begin{bmatrix}
    a_1 \\
    a_2 \\
    a_{31} \\
    a_{32} \\
    a_{33} \\
    a_{41} \\
    a_{42}
\end{bmatrix}
\begin{bmatrix}
    x_{1,2} \\
    x_3 \\
    x_4 \\
    x_5 \\
    x_6 \\
    x_{7,8} \\
    x_9
\end{bmatrix}
\]

This is not missing data, instead it means only an aggregated value is available for \( \{x_1, x_2\} \) and for \( \{x_7, x_8\} \).

The key components are combined in a structured way at the top, and each component can have some substructure, and each of those can have their own substructure and so on. We construct these fuzzy signatures directly from data [17]. Fuzzy Signature has been successfully applied to different applications such as cooperative robot communication, medical diagnosis, and personnel selection models [13, 18, 19].

In order to construct the fuzzy signatures for inferring the foreman’s following action, we need to figure out which “attributes” will be essentially related to foreman’s intentional action based on the current situation. Therefore, by measuring these “essential attributes,” the other robots might be able to know what type of actions the foreman
is going to carry out, so that they can go and work with the foreman cooperatively. Since the current situation is that there are a set of objects, if the foreman is intending to do something, he can go and touch a particular object first or get closer at least. So, the first “essential attribute” should be the “distance” between the foreman and each object in the environment. Figure 5 illustrates the membership function of “distance.”

Proximity is used to infer intention, however there exists a possible situation that cannot be handled by “distance” only; if the foreman moves towards an object then touches it, but after that he moves away or switches to another object immediately, the other robots still cannot infer what the foreman is going to do. In order to solve this problem, we need to add another “essential attribute” called “waiting time” (the membership is similar in shape to Figure 5 and is not shown) which is used to measure how long a robot stops at a particular spot. The reason why we need to measure the stopping time is that it is too difficult for a robot to perceive the meaning of the situation using instantaneous information (a snapshot) only [20].

By combining the “waiting time” with the previous item “distance,” the final fuzzy signatures for intention inference can be extended. Under this circumstance, other robots will be able to infer the foreman’s next action according to his current behaviour. For instance, if the “distance” between the foreman (R0) and a object (Ti) is touched, meanwhile foreman’s “waiting time” at that spot is long, then it implies the foreman is “waiting for help” which means another robot (Ri) should go to Ti and help the foreman. Otherwise, if neither of the conditions is satisfied, which means other robots will not assume the foreman is going to carry out any intentional action because they cannot figure it out by observation of the foreman’s current behaviour.

**Pattern matching with possibility calculation**

So far, we have discussed the problem of inferring the foreman’s intentional action by constructing the fuzzy signatures based on the foreman’s current behaviour. In some sense, it means other robots still have to count on the foreman completely. Probably now, these robots might be able to work with the foreman cooperatively, but it actually does not show that these robots are intelligent enough that can help the foreman to finish the final task effectively and efficiently as well as to truly reduce the cost of their communication.

In order to improve our modeling technique, it is important for us to consider the current situation after each movement of an object, which means other robots should be able to guess which object shape is supposed to be the most possible one according to the foreman’s previous actions and the current configuration of objects.

The solution here is that we need to measure how close the current object shape matches each of the possible shapes after the foreman’s intentional actions.

Therefore, apart from the previous fuzzy signatures, we construct another data structure to model robots’ further decision making (see Figure 6).

---

**Figure 5: Membership function of distance.**

---

**Figure 6: Structure of pattern matching with possibility calculation.**

---

Figure 6 shows a tree structure with all the leaves representing each possible object shape as well as its possibility value, respectively. The following strategies show how this structure works: we have a set of objects: T1, T2, . . . , Tn so that the total number of objects is n:

1. if foreman and a robot push an object to a place which matches one of the possible object shapes: Si, then increase the possibility value of Si: PVSi = 1/n;
2. if foreman and a robot push an object to a place which does not match any of the possible object shapes, then none of the possibility values will change;
3. if foreman and a robot push an object which matched Si to a place which does not match any of the possible shapes, then decrease the possibility value of Si: PVSi = 1/n;
4. if foreman and a robot push an object which matched Si to a place which matches another possible shape: Sj, then decrease the possibility value of Si: PVSi = 1/n, and increase the possibility value of Sj: PVsj = 1/n;
5. if two robots (where neither is the foreman) push an object to a place which matches one of the possible object shapes: Si, then the possibility value of Si, that is, PVSi, will not change.

The above points recognise that only the user can substantially alter possibility values (the representation of user intentions) for task goals.
The inherently hierarchical nature of fuzzy signatures and the ability to deal with complex and variable substructure make them ideally suited to the abstract representation (in the form of a polymorphic version of fuzzy signatures) or storage of eye-gaze path data (simple fuzzy signatures).

A possible mapping of eye-gaze path data for a set of artificial eye-gaze paths (see Figure 7) using simple fuzzy signatures to create “eye gaze signatures” is shown below.

**Partial raw eye-gaze path:** the numbers are screen coordinates for the left image (Figure 7), for the path between the first 2 fixations (repeated values indicating fixation duration not shown); see Table 1.

**Fixation sequence signature**, with the first two fixations shown in bold, is given in the equation below:

\[
\begin{align*}
  x: & \quad [155, 298, 251, 285, 214, 159, 48] \\
  y: & \quad [311, 309, 426, 478, 549, 509, 466] \\
  p: & \quad [28, 32, 12, 29, 28, 27, 13] \\
  d: & \quad [28, 13, 24, 23, 45, 7, 19]
\end{align*}
\]

**Region of interest encoded signature**, with the hierarchical structure shown by bracketing, is given below:

\[
\begin{align*}
  r: & \quad [\begin{bmatrix} 155 \\ 311 \\ 28 \end{bmatrix}, \begin{bmatrix} 161 \\ 321 \\ 38 \end{bmatrix}, \begin{bmatrix} 298 \\ 309 \\ 13 \end{bmatrix}] \\
  i: & \quad [\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}, \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}, \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}] \\
  x: & \quad [\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}, \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}, \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}] \\
  y: & \quad [\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}, \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}, \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}] \\
  p: & \quad [\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}, \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}, \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}] \\
  d: & \quad [\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}, \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}, \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}] \\
  t: & \quad [\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}, \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}, \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}]
\end{align*}
\]

**6. EYE-GAZE PATH**

The raw eye-gaze path is quite complex and difficult to interpret.

The scenario we use is quite simple, and the recording of eye gaze was from presentation of the scenario to the user-identified decision point. The instructions were to indicate when they would otherwise start moving the foreman. Figure 9 shows the raw eye-gaze path for this user.

We can readily identify the most important part of this path. We have found from their eye gaze (in our restricted domain) that users seem to automatically recheck their decision just prior to implementation.

This is consistent with work on decision making in soccer [22] with distinct pre and postdecisional phases of user action: once a target player has been chosen, the visual field narrows around this target for postdecisional action planning.

In Figure 10, we show one user checking the object they intend to move and the destination.
In the next section, we describe in detail how we perform task inference using fixation information from the eye-gaze path.

7. FUZZY INFERENCE ON EYE GAZE

The duration of fixations is represented by the size of black circle overlaid on the image. We show in Figure 11 just part of the window for a horizontal rows task overlaid with the eye-gaze task. Note that the initial configuration of robots and objects is the same for all tasks.

The size of fixation circles on the scene representing fixation duration was kept small to allow the scene to be visible, and so do not represent the probable area of interest shown by the user. In Figure 12, we show detail of the fuzzy inference for eye-gaze fixations.

We will consider partial inference via the columns in our scene.

Fixation 1 projects onto a trapezoidal fuzzy membership function which represents the degree of possibility from 0 to 1 of the user’s eye gaze indicating user interest in that region. As mentioned earlier, the size of black circles was kept small to not obscure the scene. So, the size of the circles does not represent the size of the likely region of interest. Thus, the core of the fuzzy membership function (horizontal top part with value 1) is wider than the size of the fixation circle.

We perform the same projection for fixation 2.

We then sum the two fuzzy membership functions, using a union operator. This can result in a concave or convex result. In this case, as the cores overlap, the result is convex, a wider trapezoidal fuzzy membership function. This is projected on the scene. As our scene is discretised, we can represent the result visually as areas truncated from discrete rectangles.

An identical process is followed with regards inference along the rows. Clearly, the next step is combining the
horizontal and vertical results of the inference. We show the results in Figure 13.

We combined the results of the vertical and horizontal fuzzy inference using intersection, and show this with intersection of the diagram representations of the results we used before. The intersections are shown shaded in Figure 13, with the potentially significant squares outlined. Note that one of these outlined squares has the highest possibility value, with all the other high values (≥ 0.5) being adjacent to this square.

We have made a number of simplifying assumptions above. For example, the graphical treatment used required rectilinear regions of interest. In Figure 14, we demonstrate some more plausible regions of interest.

On the left in Figure 14, we show circular regions of interest for both fixations. On the right in Figure 14, we show an ovoid region of interest for fixation 1 only, displaced towards the location of the next fixation. This is consonant with [23], as attention, is displaced ahead of the eye. Much of this work is from reading [24] with smaller regular saccades than the long saccades we find between object and target, so the second fixation region of interest may be a different shape. We can also see from the diagram that for our simple discrete domain, changing the shape of the region of interest in inference would have had no significant effect.

We list the remaining simplifying assumptions here: use of union and intersection only, while many other aggregation operators are available in this range [21]; we used a single level fuzzy signature to combine variable numbers of fixations which were preidentified, in practice we would use the full fixation sequence which would require a more complex fuzzy signature structure; and we omitted any discussion of propagation of possibilities/iterative thinning in cases where there are multiple potential significant areas for the user’s eyes to chose.

8. RESULTS

We reprise our previous results for comparison purposes. We performed 3 initial experiments.

Although we allowed players to have verbal communications in experiment 1 (see Table 2), the human-controlled robots still took the most steps on average to finish each of the test tasks. The reason for this phenomenon is that players might make different decisions in dynamic situations. Therefore, it is possible for them to decide to move different objects at the same time rather than aiming at the same target or placing the same object with different route plans, which will cost them extra steps to reach the common target or correct previous incorrect actions. That is, even with the explicit communication (talking) possible, it may be that it is only after incompatible moves that humans notice that they are following different plans.

The result in experiment 2 is quite good compared with the other two experiments (see Table 3). Since the robot with the codebook could infer the human-controlled foreman robot’s action by observation and cooperate with it, it is not necessary for the player to communicate with the other robot directly, which is different from the situation in experiment 1.

So the players can make their own decision without any other disturbance, which may be what leads to an improvement in all the costs, including robots steps, object movements, and time.

Apart from the second test case (vertical rows), the robots in experiment 3 made the most object movements in the rest of the test cases (see Table 4). The main reason here would
### Table 2: Two human-controlled robots (control case).

| Experimetal 1 | Horizontal rows | Vertical rows | T-shape | U-shape |
|---------------|-----------------|---------------|---------|---------|
| Robot A (human) | 163.0           | 136.8         | 149.2   | 127.4   |
| Robot B (human) | 141.6           | 159.0         | 151.2   | 143.4   |
| Total steps    | 304.6           | 295.8         | 300.4   | 270.8   |
| Shifting movements | 40.0        | 43.0          | 42.8    | 38.6    |
| Rotating movements | 7.2           | 6.8           | 7.2     | 5.6     |
| Total movements | 47.2           | 49.8          | 50.0    | 44.2    |
| Time (s)       | 74.6’           | 75.0’         | 77.6’   | 62.2’   |

### Table 3: One assistant robot helps human throughout.

| Experimental 2 | Horizontal rows | Vertical rows | T-shape | U-shape |
|---------------|-----------------|---------------|---------|---------|
| Foreman (human) | 112.4           | 110.6         | 113.6   | 108.4   |
| Robot A       | 153.6           | 141.4         | 156.4   | 143.2   |
| Total steps    | 266.0           | 252.0         | 270.0   | 251.6   |
| Shifting movements | 39.2        | 40.6          | 41.0    | 36.8    |
| Rotating movements | 6.8           | 4.8           | 4.8     | 4.8     |
| Total movements | 46.0           | 45.4          | 45.8    | 41.6    |
| Time (s)       | 66.0’           | 56.0’         | 61.8’   | 55.8’   |

### Table 4: Human and assistant start task, robots finish.

| Experimental 3 | Horizontal rows | Vertical rows | T-shape | U-shape |
|---------------|-----------------|---------------|---------|---------|
| Foreman (human) | 28.6            | 26.8          | 29.0    | 24.4    |
| Robot A       | 115.6           | 103.8         | 118.2   | 106.8   |
| Robot B       | 143.4           | 142.8         | 150.0   | 132.0   |
| Total steps    | 287.6           | 273.4         | 297.2   | 263.2   |
| Shifting movements | 42.0        | 40.2          | 43.6    | 41.8    |
| Rotating movements | 7.4           | 4.8           | 7.2     | 6.0     |
| Total movements | 49.4           | 45.0          | 50.8    | 47.8    |
| Time (s)       | 69.0’           | 65.0’         | 71.4’   | 64.0’   |

### Table 5: Eye-gaze indication of object and target.

| Experimental 4 | Horizontal rows | Vertical rows | T-Shape | U-Shape |
|---------------|-----------------|---------------|---------|---------|
| Foreman (human) | 29.6            | 26.8          | 25.2    | 28.0    |
| Robot A       | 89.6            | 89.6          | 106.4   | 92.6    |
| Robot B       | 131.8           | 118.0         | 132.4   | 125.4   |
| Total steps    | 251.0           | 234.4         | 264.0   | 246.0   |
| Shifting movements | 41.2        | 40.4          | 43.6    | 41.4    |
| Rotating movements | 6.4           | 4.0           | 6.0     | 6.0     |
| Total movements | 47.6           | 44.4          | 49.6    | 47.4    |
| Time (s)       | 60.7’           | 55.8’         | 62.6’   | 56.9’   |

### Table 6: Eye gaze used to initialise task goal.

| Experimental 5 | Horizontal rows | Vertical rows | T-Shape | U-Shape |
|---------------|-----------------|---------------|---------|---------|
| Robot A       | 132.2           | 115.4         | 131.4   | 124.0   |
| Robot B       | 143.2           | 123.6         | 135.0   | 130.2   |
| Total steps    | 275.4           | 239.0         | 266.4   | 254.2   |
| Shifting movements | 43.0        | 41.0          | 44.0    | 42.0    |
| Rotating movements | 8.0           | 4.0           | 6.0     | 6.0     |
| Total movements | 51.0           | 45.0          | 50.0    | 48.0    |
| Time (s)       | 65.1’           | 56.6’         | 64.2’   | 62.4’   |
be suboptimal strategies of route planning and obstacle avoidance.

In most of the test cases, the total steps made in experiment 3 are more than experiment 2 but still better than robots totally-controlled by humans. This is of course the key benefit of our work, to be able to complete the task, and faster than 2 humans is an excellent result.

We now continue and report the results of our experiments using eye gaze, with two further experiments. We immediately report the results below.

The results in experiment 4 show that on average, the tasks were completed some 14% faster than in experiment 3 (see Table 5). This is a significant time saving.

This reduction is due to a helper robot starting to move to the correct object at the same time as the foreman and due to time saved during rotation.

Since the robot can only infer human actions, once a sequence of shifts is complete and a rotation is needed, the human-controlled foreman robot stops pushing the object and just waits.

Then, the assistant robot normally takes some time to move around the object once it infers the shifting task is complete. While some of our users noticed some boredom in some tasks, our results from experiment 2 show that the overall task was still completed quickly.

Here, the robot “knew” (can infer this from identification of the object and target, and constantly updates its path planning) the target so could immediately move to the correct spot on the object to make the required rotation.

In experiment 4, after the first object was in place, the assistant robots completed the task, achieving results comparable to those in experiment 2.

In this experiment, the eye gaze was used to initialise the initial possibility values of the targets, and the 2 assistant robots started and completed the task, without the user moving the foreman (see Table 6).

The results are 5% worse than experiment 4, but with no guidance via the foreman robot, all inference was via the eye gaze.

9. CONCLUSIONS

Our experiments have demonstrated that with suitable AI techniques (our fuzzy signatures), it is possible for robots (agents) to correctly infer human actions in a game-like interactive and cooperative task.

We have further shown that we can extend the notion of inference from the actions of the human to inference from their eye gaze. Such work is the first steps towards computing devices which understand what we want in ways similar to other human beings.

In our experiments, we had a clear indication of the decision point due to the way we structured the recording of eye-gaze information. In practice, while extracting the decision point may be possible from the eye gaze, we propose an alternative, a control device which the user could use to see a display of their inferred intentions, as well as a “do it” button. Of course, it is always possible that some will choose the “do it” immediately.

There seem to be two benefits of such a generic “show me and do it” control device. Firstly, to return the visibility of control to the user. That is, while the user is always in control, we want that to be explicitly visible to the user. This seems to us particularly significant for immersive computer games.

Secondly, human beings are often multitasking. In a general, setting beyond our experiments it is likely that users would receive phone calls and make eye gaze gestures with their eyes while talking—and would not like such actions to control their game or editing task. Hence, the need for a control which says “take note of my recent eye gaze behaviour and act appropriately.”

So, how useful will eye gaze be for computer games? In the areas, we identified avatar realism has some advantages, but the difference between how humans respond to other human versus avatars may limit the degree to which such agents can enhance games. In the area of active game controls, there is some scope; however the need to remain close to “normal” uses of people’s eye gaze again limits their use (except perhaps in rehabilitation and therapeutic games). In many first-person shooter games, the action is always at the centre of the screen, so there seems little eye gaze can add, except perhaps in our fashion to infer user intentions and to rotate the world for the user automatically.

Finally, we conclude that our techniques have some benefit in an assistive fashion, and that eye-gaze technology can be used to further enhance the immersive quality of games, but that eye gaze is unlikely to lead to a qualitative change in the nature of computer games.

ACKNOWLEDGMENTS

James Sheridan was invaluable in the collection of eye-gaze data for this experiment. Thank you very much James. The seeing machines eye-gaze hardware and faceLab software were purchased using an ANU MEC Grant.

REFERENCES

[1] H. Kobayashi and S. Kohshima, “Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye,” Journal of Human Evolution, vol. 40, no. 5, pp. 419–435, 2001.
[2] A. Nguyen, V. Chandran, and S. Sridharan, “Gaze tracking for region of interest coding in JPEG 2000,” Signal Processing: Image Communication, vol. 21, no. 5, pp. 359–377, 2006.
[3] V. Vinayagamoorthy, M. Garau, A. Steed, and M. Slater, “An eye gaze model for dyadic interaction in an immersive virtual environment: practice and experience,” Computer Graphics Forum, vol. 23, no. 1, pp. 1–11, 2004.
[4] S. Raidt, G. Bailly, and F. Elisei, “Analyzing and modeling gaze during face-to-face interaction,” in Proceedings of the International Conference on Auditory-Visual Speech Processing (AVSP ’07), Hilvarenbeek, The Netherlands, August-September 2007, paper P23.
[5] M. Rehm and E. André, “Where do they look? Gaze behaviors of multiple users interacting with an embodied conversational agent,” in Proceedings of the 5th International Working Conference on Intelligent Virtual Agents (IVA ’05), pp. 241–252, Kos Island, Greece, September 2005.
[6] M. Hopkins and F. Biasini, “Computer game helps autistic children recognize emotions,” 2007, ScienceDaily, http://tinyurl.com/3q4q3x.

[7] V. Tanriverdi and R. J. K. Jacob, “Interacting with eye movements in virtual environments,” in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ‘00), pp. 265–272, ACM, The Hague, The Netherlands, April 2000.

[8] E. Jönsson, If looks could kill—an evaluation of eye tracking in computer games, M.S. thesis, Department of Numerical Analysis and Computer Science, Royal Institute of Technology, Stockholm, Sweden, 2005.

[9] A. Kenny, H. Koesling, D. Delaney, S. McLoone, and T. Ward, “A preliminary investigation into eye gaze data in a first person shooter game,” in Proceedings of the 19th European Conference on Modelling and Simulation (ECMS ‘05), Riga, Latvia, June 2005.

[10] C. L. Calhoun, “Gaze-based control,” in International Encyclopedia of Ergonomics and Human Factors, W. Karwowski, Ed., pp. 234–236, CRC Press, Boca Raton, Fla, USA, 2001.

[11] T. Terano, "Research Project at LIFE," Laboratory for International Fuzzy Engineering Research, Yokohama, Japan, 1993.

[12] D. Zhu and T. D. Gedeon, “Fuzzy logic for cooperative robot communication,” in Proceedings of the 2nd KES International Symposium on Agent and Multi-Agent Systems: Technologies and Applications (KES-AMSTA ‘08), vol. 4953 of Lecture Notes in Computer Science, pp. 401–410, Incheon, Korea, March 2008.

[13] T. D. Gedeon and D. Zhu, “Automating game activities: a proposal based on experiments,” in Proceedings of the Annual International Conference & Symposium on Computer Games; Animation, Multimedia, IPTV, Entertainment and Security (CGAT ’08), Singapore, April 2008.

[14] L. T. Kóczy, T. Vámos, and G. Biro, “Fuzzy signatures,” in Proceedings of the 4th Meeting of the Euro Working Group on Fuzzy Sets and the 2nd International Conference on Soft and Intelligent Computing (EUROPUS-SIC ’99), pp. 210–217, Budapest, Hungary, May 1999.

[15] T. D. Gedeon, L. T. Kóczy, K. W. Wong, and P. Liu, “Effective fuzzy systems for complex structured data,” in Proceedings of the LASTED International Conference on Control and Applications (CA ’01), pp. 184–187, Banff, Canada, June 2001.

[16] T. Vámos, L. T. Kóczy, and G. Bíró, “Fuzzy signatures in data mining,” in Proceedings of Joint 9th IFSA World Congress and 20th NAFIPS International Conference, vol. 5, pp. 2842–2846, Vancouver, Canada, July 2001.

[17] A. Chong, T. D. Gedeon, and L. T. Kóczy, “Projection based method for sparse fuzzy system generation,” in Proceedings of the 2nd WSEAS International Conference on Scientific Computation and Soft Computing, pp. 321–325, Creta Island, Greece, July 2002.

[18] K. W. Wong, T. D. Gedeon, and L. T. Kóczy, “Construction of fuzzy signature from data: an example of SARS pre-clinical diagnosis system,” in Proceedings of IEEE International Conference on Fuzzy Systems, vol. 3, pp. 1649–1654, Budapest, Hungary, July 2004.

[19] B. S. U. Mendis and T. D. Gedeon, “A comparison: fuzzy signatures and choquet integral,” in Proceedings of IEEE World Congress on Computational Intelligence (WCCI ’08), Hong Kong, June 2008.

[20] Y. Inagaki, H. Sugie, H. Aisu, and T. Unemi, “A study of a method for intention inference from human’s behavior,” in Proceedings of the 2nd IEEE International Workshop on Robot and Human Communication, pp. 142–145, Tokyo, Japan, November 1993.

[21] B. S. U. Mendis, T. D. Gedeon, J. Botzheim, and L. T. Kóczy, “Generalised weighted relevance aggregation operators for hierarchical fuzzy signatures,” in Proceedings of the International Conference on Computational Intelligence for Modelling, Control and Automation (CIMCA ’06), Jointly with the International Conference on Intelligent Agents, Web Technologies and International Commerce (IAWTIC ’06), p. 198, Sydney, Australia, November-December 2006.

[22] H. Kösling and O. Höner, “Bend it like Beckham—mindsets and visual attention in decision-making in soccer,” in Proceedings of the 12th European Conference on Eye Movements (ECEM ’03), Dundee, UK, August 2003.

[23] J. E. Hoffman, “Visual attention and eye movements,” in Attention, H. Pashler, Ed., chapter 3, pp. 119–154, Psychology Press, Hove, UK, 1998.

[24] K. Rayner, “Eye movements in reading and information processing: 20 years of research,” Psychological Bulletin, vol. 124, no. 3, pp. 372–422, 1998.
Submit your manuscripts at http://www.hindawi.com