Fuzzy Control for Gaze-Guided Personal Assistance Robots: Simulation and Experimental Application

Carl A. Nelson
Department of Mechanical and Materials Engineering
University of Nebraska-Lincoln
Lincoln, NE, USA
email: cnelson5@unl.edu

Xiaoli Zhang, Jeremy Webb, and Songpo Li
Department of Mechanical Engineering
Colorado School of Mines
Golden, CO, USA
email: xzhang@mines.edu, jwebb@mymail.mines.edu, soli@mines.edu

Abstract—As longer lifespans become the norm and modern healthcare allows individuals to live more functional lives despite physical disabilities, there is an increasing need for personal assistance robots. One of the barriers to this shift in healthcare technology is the ability of the human operator to communicate his/her intent to the robot. In this paper, a method of interpreting eye gaze data using fuzzy logic for robot control is presented. Simulation results indicate that the fuzzy logic controller can successfully infer operator intent, modulate speed and direction accordingly, and avoid obstacles in a target following task relevant to personal assistance robots. The fuzzy logic approach is then validated through navigation experiments using a small humanoid robot.

Keywords—gaze tracking; fuzzy logic; autonomous robot; obstacle avoidance; personal assistance robot

I. INTRODUCTION

With lifespans increasing worldwide due to advancements in healthcare and related technologies, the importance of care for the elderly and disabled is increasing. In particular, there is a shifting emphasis in technology development towards improving quality of life in the face of diminishing physical capabilities. One of the burgeoning areas of this trend is personal assistance robotics. In a typical scenario, a robot assistant may be present in the home to help with basic day-to-day tasks (e.g., object retrieval), especially those tasks requiring navigation throughout the home, since age- or disability-related mobility limitations may keep an individual from performing all these tasks personally. In extreme circumstances, it can even be challenging to give instructions to the robotic assistant, as in the case where the individual is not physically able to type, speak, or otherwise provide clear inputs to the human-robot interface. Here, we build on work presented in [1] and present progress towards a robotic assistance system which relies on gaze tracking, including eye blinking patterns, to infer a person’s intent and thereby create instructions for the robot. In this paper, we specifically focus on the intelligent inference of intent based on gaze and blinking input.

This problem is an extension of the task of robotic target following and path planning. Significant work has been done in this area of service robotics, where a robot is to follow a moving target. For instance, some have used computer vision, using optical flow algorithms to track the target [2][3]. Other computer vision-based approaches have used Kalman filters for improving the accuracy of tracking [4]. Other tracking methods include the use of depth images with verification via a state vector machine [5], or following acoustic stimuli [6]. Control approaches in these target-following scenarios include potential field mapping [7] and a variety of other techniques. Of particular interest are fuzzy logic controllers [6][8][9], which tend to be used primarily for steering, but can easily be adapted to handle various types of linear and nonlinear systems [10]. Here, we will describe a fuzzy logic controller which not only determines the robot’s heading based on the location of the target, but also avoids obstacles and modulates speed based on the perception of intent from the combined gaze direction and blink frequency inputs. The authors believe this perception of intent combined with heading, speed, and obstacle avoidance to be unique with respect to the state of the art in robot guidance. This is conceptually based in part on recent work demonstrating how such a combined input using operator gaze could be used for automatic control of endoscope positioning in surgical tasks [11] using a commercially available eye tracking system, which is also similar to the work described in [12]. Existing examples of robot control using gaze input are relatively scarce. In [13][14][15], specially identified eye movements, such as looking up, down, left and right, were mapped to wheelchair steering commands to drive it forward, backward, left and right. In [16], on-screen buttons were created to activate joint rotation of an articulated robot arm such that a user could steer the arm by gazing at the buttons. However, steering a robot arm in this manner is very inefficient and can be exhausting for the user, as he/she has to explicitly control every movement of the arm. In [17], the user’s gaze vector was estimated and served as a pointing line along which the robot could search to retrieve the first found object. However, as the exact location of the object was not calculated, extensive searching had to be carried out along the gaze vector to locate the object. The gaze-based robot control approach proposed in this article extends beyond the most typical uses of eye gaze, which tend to be for two-dimensional human-computer interfaces [18], to interacting in the three-dimensional context using gaze tracking for activities of daily living, i.e., using personal assistance robots. Although the robotic assistance scenario clearly would involve more subtasks (such as object manipulation), we limit our treatment in this paper to development of controllers which modulate robot heading and speed while avoiding obstacles, for navigating in a potentially cluttered environment using gaze as the input data stream.
The remainder of this paper is organized as follows. In Section II, the eye gaze data and the fuzzy logic controller are described. In Section III, simulation results are presented, followed by experimental validation using a small humanoid robot platform. Section IV includes conclusions and recommendations for future work.

II. Methods

A. Test Dataset and Simulation

A gaze dataset was artificially generated to have spatiotemporal characteristics similar to those described in [11], in a planar workspace. The data were arbitrarily assumed to be sampled at 10 Hz and included a logical blink data channel in addition to the $x$ and $y$ gaze target channels on the interval $[-0.5, 0.5]$, providing a total of over 23 seconds of simulated robot tracking. Due to the noisy nature of gaze data, the target $X$ was determined by a linear weighted average of the previous $n$ data points $P$, with $n = 20$:

$$X_k = \frac{2}{n} \sum_{i=k-n}^{k} \left(1 - \frac{K-n}{n}\right) P_i.$$  

(1)

In this particular dataset, there are five intended target locations, characterized by dwelling gaze and higher blink frequency, and it is assumed that a supplementary action such as object placement or retrieval would follow target acquisition (although this supplementary action is beyond the scope of this preliminary study). Within the workspace, three round obstacles were defined to test the ability of the simulated robot to avoid obstacles while seeking a target. The data were imported into MATLAB (The MathWorks, Natick, MA) for simulation of gaze-based robotic target tracking.

B. Fuzzy Logic Controller

A Mamdani-type fuzzy logic controller [19] with five inputs and three outputs was created using the Fuzzy Logic Toolbox in MATLAB; the Mamdani-type model handles multi-input, multi-output (MIMO) problems better than the Sugeno-type alternative. The inputs, shown in Table I, were intended to take into account the control objectives: to track a target at an appropriate speed based on uncertain data while avoiding obstacles. Distance to the target is captured by target variability and target variability is quantified by the obstacle distance. The blink frequency is used to capture operator intent and desired speed. The outputs, also shown in Table I, were used to control the speed and heading of the robot, including steering adjustments for obstacle avoidance. All of the membership functions were triangular, as shown in Fig. 1, and their parameters were tuned by hand using a minimal amount of trial and error.

The target was determined using a weighted average of the gaze data as in (1), the target and obstacle distance variables were then calculated using the Pythagorean theorem, and target variability was represented by the standard deviation of the gaze input data over the averaging window. (It is noteworthy that target variability is likely to be the input parameter most sensitive to individual characteristics, and therefore would need to be tuned for each individual’s gaze “signature.”) In this case, it was tuned to accommodate the characteristics of the dataset described in Section II.A.)

![Membership functions for target variability (zero, low, and high).](image)

**Figure 1.** Membership functions for target variability (zero, low, and high).

Blink frequency was normalized to the interval $[0, 1]$ by assuming that four blink events within the 20-sample averaging window was high (achieving a value of 1), and lower blinking rates in the same window of time receive a proportionally smaller membership value. If no obstacles were detected in the direct path between the robot and target, the obstacle distance was set to its maximum value of 1. The other distance-based variables did not need to be explicitly normalized since the workspace was set up as a unit square. It should also be noted (referring to Table I) that in certain cases (e.g., the “zero” membership functions for target variability, obstacle distance, and blink frequency), negative values (which do not have physical meaning) were used in order to create membership

<table>
<thead>
<tr>
<th>Input/Output</th>
<th>Variables</th>
<th>Membership Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Target $\Delta x$</td>
<td>distance</td>
</tr>
<tr>
<td>I</td>
<td>Target $\Delta y$</td>
<td>distance</td>
</tr>
<tr>
<td>I</td>
<td>Target variability</td>
<td>distance</td>
</tr>
<tr>
<td>I</td>
<td>Blink frequency (normalized)</td>
<td>-</td>
</tr>
<tr>
<td>I</td>
<td>Obstacle distance</td>
<td>distance</td>
</tr>
<tr>
<td>O</td>
<td>Speed</td>
<td>distance/time</td>
</tr>
<tr>
<td>O</td>
<td>Heading</td>
<td>rad</td>
</tr>
<tr>
<td>O</td>
<td>Heading adjustment</td>
<td>rad</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Units</th>
<th>Membership Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target $\Delta x$</td>
<td>distance</td>
<td>negative [-1, -0.5, 0]</td>
</tr>
<tr>
<td>Target $\Delta y$</td>
<td>distance</td>
<td>zero [-0.1, 0, 0.1]</td>
</tr>
<tr>
<td>Target variability</td>
<td>distance</td>
<td>positive [0, 0.5, 1]</td>
</tr>
<tr>
<td>Blink frequency (normalized)</td>
<td>-</td>
<td>zero [-0.4, 0, 0.4]</td>
</tr>
<tr>
<td>Obstacle distance</td>
<td>distance</td>
<td>low [0, 0.3, 0.6]</td>
</tr>
<tr>
<td>Speed</td>
<td>distance/time</td>
<td>zero [-0.4, 0, 0.4]</td>
</tr>
<tr>
<td>Heading</td>
<td>rad</td>
<td>up [0.125, 0.25, 0.375]</td>
</tr>
<tr>
<td>Heading adjustment</td>
<td>rad</td>
<td>zero [-0.4, 0, 0.4]</td>
</tr>
</tbody>
</table>

International Journal on Advances in Intelligent Systems, vol 8 no 1 & 2, year 2015, http://www.iariajournals.org/intelligent_systems/ © Copyright by authors, Published under agreement with IARIA - www.iaria.org
functions for which the lower bound is also the mode, without causing errors in the software.

Concerning the output variables, the maximum speed was constrained to a value of 0.25 (covering one-fourth the workspace in one second at maximum speed), and the maximum heading adjustment for obstacle avoidance was set at ±100°. The heading variable was scaled to allow the robot to steer within the full 360° range.

Fifteen rules were defined to characterize the influence of the five input variables on the three outputs. In particular, four rules capture the influence of the inputs on the output variable speed, eight rules accommodate the division of heading into eight regions in polar coordinates, and the remaining three rules govern obstacle avoidance. The rules defining the fuzzy logic controller are as follows:

1. IF blink IS high THEN speed IS high
2. IF target Δx IS positive OR target Δx IS negative OR target Δy IS positive OR target Δy IS negative THEN speed IS high
3. IF target Δx IS zero AND target Δy IS zero THEN speed IS zero
4. IF target variability IS high OR blink IS low THEN speed IS low
5. IF target Δx IS positive AND target Δy IS zero THEN heading IS right
6. IF target Δx IS positive AND target Δy IS positive THEN heading IS up/right
7. IF target Δx IS positive AND target Δy IS negative THEN heading IS down/right
8. IF target Δx IS negative AND target Δy IS zero THEN heading IS left
9. IF target Δx IS negative AND target Δy IS positive THEN heading IS up/left
10. IF target Δx IS negative AND target Δy IS negative THEN heading IS down/left
11. IF target Δx IS zero AND target Δy IS positive THEN heading IS up
12. IF target Δx IS zero AND target Δy IS negative THEN heading IS down
13. IF obstacle distance IS zero THEN heading adjustment IS high
14. IF obstacle distance IS low THEN heading adjustment IS low
15. IF obstacle distance IS high THEN heading adjustment IS zero

The first four rules govern the robot’s speed. Higher blink rates imply a more focused operator intent and cause increased speed (rule 1). Conversely, high gaze variability or low blink rate imply a less sure target and lead to lower speed (rule 4). The higher the distance to the target, the higher the necessary speed to reach it in a timely manner, and speed should drop to zero as the target is reached (rules 2-3). It should be noted that lower speeds are sometimes desirable to conserve energy either when the goal is unclear or has been reached.

Rules 5-12 pertain to heading. These are relatively straightforward and use the four cardinal directions and the four semi-cardinal directions to navigate in the planar map based on the relative target distance in the x and y directions. This can be thought of as a fuzzy calculation of inverse tangent for the heading angle using target Δx and target Δy as inputs.

The remaining three rules (rules 13-15) constitute the robot’s obstacle avoidance behavior. The closer the obstacle, the larger the heading adjustment applied to go around it. Whether this adjustment is added or subtracted from the heading variable is determined by whether the obstacle centroid is to the right or the left of the straight line along the robot’s heading.

C. Experiments

In addition to simulation, the fuzzy logic controller was implemented on a commercially available small humanoid robot (NAO, Aldebaran Robotics). This platform was chosen, in part, because it has computer vision and sonar sensors suitable for obstacle detection as a built-in feature, so it is expected to scale well towards more advanced demonstrations in the future. The overall architecture, illuminating the concept and the correlations of components in the system, is shown in Fig. 3. The system contains four parts: an eye tracking system which can track where the user is looking on a monitor, a camera which provides video feedback to the user and functions as a global tracking system, the host system which is responsible for collecting and interpreting the gaze data into robot motion commands and sending these motion commands wirelessly to the robot for navigation, and the NAO robot. A rectangular workspace area (3m x 4m) was defined, with two round obstacles marked, similar to the setup of the simulation experiments.

Figure 2. Interaction in the gaze-based system for robot navigation.

In the presented system, the user, sitting in a chair, watches the live video fed from the camera (shown in Fig. 3). The user’s gaze on the video is sensed, from which the user’s visual attention is detected. The detected visual attention represents
the target position that the user wants the robot to approach. Then the visual attention position as well as the online collected gaze and blink data are interpreted by the fuzzy controller into a series of control commands to generate the optimal trajectory for the robot to approach the target position. The response of the NAO robot was recorded.

A GP3 Eye Tracker (GazePoint) is used to track where the user is looking on a monitor. GP3 is a video-based remote eye tracking system which allows head movement of a user in a volume of 25×11×30cm³, without significant degradation of the tracking accuracy. It can report the gaze data at 60 Hz with an accuracy of 0.5° - 1° and drift of less than 0.3°. Calibration is required before it can provide accurate eye tracking. (Although the long-term objective is to track 3D gaze rather than planar gaze on a screen, and hardware is under development to achieve such 3D gaze tracking, the work presented here is focused on validation of the control concept, and as such, planar gaze tracking consistent with the simulation approach presented in Section II.A-B is appropriate for these experiments.)

Human eyes are capable of making many different movements. Some are involuntary, such as rolling, nystagmus, drift or microsaccades; these involuntary movements are superimposed on the voluntary eye movements such as gaze fixation. The gaze data estimated from human eye movements will thus include noise from the superimposed involuntary eye movements, and must therefore be filtered to extract the gaze data that is related to the attentional processes of the viewer. An adaptive sliding window filter is employed to eliminate this noise, and a dwell time method is utilized to extract visual attention. Due to the complexity of human eye movement, this filter is slightly more complex than the one used in simulation (Eq. (1)).

The adaptive sliding window filter is illustrated in Eqs. (2) and (3). \( N \) is the size of the sliding window, \( P_i \) and \( \tilde{P}_i \) are the \( i^{th} \) gaze point before and after filtering, respectively. \( E_i \) is an influence coefficient, calculated using Eq. (3), which indicates the degree of influence a newly received gaze point has on the attention extraction. The influence coefficient of a gaze point is determined by the relative distance from that point to the mean of all the gaze points in the current sliding window. The output of the filter is the mean of all the weighted gaze points. This filter is intended to remove the effects of blinking, attention shifting, and tracking failure. At the same time it can smooth the gaze points by eliminating effects of involuntary eye movements such as rolling, nystagmus, drift and microsaccades.

\[
P_i = \frac{1}{\sum_{k=1}^{N} E_{i-k+1}} \left( \sum_{j=1}^{N-1} \tilde{P}_{i-j} * E_{i-j} + P_i * E_i \right)
\]

(2)

\[
E_i = \begin{cases} 
1, & \left| P_i - \frac{1}{\sum_{k=1}^{N} E_{i-k}} \sum_{j=1}^{N} \tilde{P}_{i-j} * E_{i-j} \right| \leq \text{threshold} \\
0, & \left| P_i - \frac{1}{\sum_{k=1}^{N} E_{i-k}} \sum_{j=1}^{N} \tilde{P}_{i-j} * E_{i-j} \right| > \text{threshold}
\end{cases}
\]

(3)

Using gaze as an input signal for human-robot interaction requires the differentiation of normal behavioral eye movements and intentional eye “commands,” which is known as the Midas touch problem [20]. The “select” or “click” command is usually derived from either blink or gaze dwell time, which is used as a confirmation of a specific command from the eyes. The dwell time method is derived from the fact that a person’s eyes stay focused on a target when he/she concentrates on a visual target. In this paper, a dwell time of 2 seconds was used. Once the user stares at an object for more than 2 seconds, the system considers that object as the visual attention point of the user, and this triggers a series of motion commands of the robot.

Due to the challenges in implementing the proposed control method on a real robot, a few changes were made to the simulated control flow. In particular, since the NAO robot travels relatively slowly and it would be cumbersome for a user to have to continuously focus on a target location until the robot reached it, the target location was acquired at the beginning of each movement by fixating on a single location for two seconds. Once the target was known, NAO would first turn towards the target and then begin moving. At this point, the fuzzy rules took over, controlling the heading based on the current location of the robot, and the speed based on the distance from the target, the gaze variability, and the blink rate. Note that the user is not required to focus on the target the whole time so if he/she is looking at many different locations while the robot is moving, the gaze variability will be high and rule 4 above will be invoked. On the other hand, if the user is engaged with the task at hand and focused on the robot or the target, the gaze variability will be low. Additional differences are in the speeds used. The output from the fuzzy logic rules was calculated and sent to the robot at 8 Hz while the gaze data were collected at 60 Hz with an averaging window size of 70 data points.

III. RESULTS

The simulations described in Sections II.A-B and the experiments described in Section II.C generally produced similar outcomes validating the approach. These outcomes are described as follows.
A. Simulation Outcomes

Simulation in MATLAB revealed the ability of the fuzzy logic controller to simultaneously determine human intent from the combined gaze location and blink data, use this intent to modulate robot speed, follow a moving target, and avoid obstacles. In Fig. 4, it can be observed that the robot (whose position is indicated by red diamond markers) can start at a location somewhat removed from the initial target, quickly acquire the target, and then follow it consistently without colliding with obstacles in the workspace. It can be noted that filtering the raw gaze data (black dot markers) smoothes but does not significantly alter the target path (blue circle markers), and that the robot follows the target reasonably closely when it is not engaged in obstacle avoidance.

![Figure 4](https://example.com/figure4.png)

Figure 4. Target following behavior: robot (red diamond markers) follows target (blue circles) while avoiding fixed environmental obstacles. Green circles indicate target location with a blink event. Targets of definite interest (based on dwell duration and blink frequency) are at approximately (-0.3, 0.3), (0.1, 0.3), (0.3, 0), (-0.1, -0.3), and (0, 0). Raw gaze data are shown as black dots.

The more interesting outcomes of the simulation are highlighted in Figs. 5-8, in which the input/output model parameters from the simulation of Fig. 4 are plotted separately to elucidate the effects of the fuzzy rule set. In Fig. 5, one can see that robot speed tends to increase with blink frequency, as intended (rules 1 and 4). High speed at low blink value can be attributed to the effects of target distance (particularly at the beginning of the simulation, rule 2). Note that the results in Fig. 5 are striated at discrete levels, since blinking is a discrete, logical event; this could be smoothed by applying an averaging method similar to that used in target determination.

![Figure 5](https://example.com/figure5.png)

Figure 5. Output speed as a function of blink membership function value: a positive correlation is noted.

![Figure 6](https://example.com/figure6.png)

Figure 6. Output speed as a function of target distance: a positive correlation is noted.

![Figure 7](https://example.com/figure7.png)

Figure 7. Output speed as a function of distance to current gaze location: correlation is much less pronounced.
Target distance also has an important effect on speed (rule 2), as shown in Fig. 6. In contrast, Fig. 7 illustrates that the relationship between robot speed and distance from the robot to the actual gaze point is less pronounced, since the target is based on a weighted average of the gaze point and is thus a less noisy signal. The interdependence of speed on multiple input parameters is evident in Figs. 5 and 6. The effectiveness of the obstacle avoidance behavior (rules 13-15) is shown in Fig. 8 by the clean heading adjustment curve. These results illustrate the suitability of the fuzzy controller for satisfying the control objectives noted in Section I.

B. Experimental Outcomes

The experimental setup is shown in Fig. 9. It includes the NAO robot with a unique marker, and two obstacles with a different unique marker, in a rectangular area. The target gaze position is indicated by a small blue square on the figure. Feedback to the control computer (not shown in the figure) is done through overhead camera capture of the scene. In this experiment shown, there are two target locations at which the gaze lingers (indicated in the latter two parts of Fig. 9).

The data recorded in the experiment of Fig. 9 are shown in Fig. 10 as a time lapse, similar to the simulation results of Fig. 4. It is clear that the fuzzy logic controller allows the robot to effectively seek targets while avoiding obstacles, just as in the simulation. Additional evidence of this is illustrated in Figs. 11-13. In Fig. 11, one can observe that at low blink rates, other rules pertaining to target distance (rules 2-3) dominate the determination of robot speed, but at higher blink rates, the interpretation of user intent becomes more influential to increase speed (rule 1). In Fig. 12, the target distance is seen to have a positive correlation with speed, and speed increases up to the hardware-limited threshold. The adjustment of heading with obstacle proximity is shown in Fig. 13, where avoidance behavior is more extreme for closer obstacles (rules 13-15). All these behaviors are as intended and are consistent with the results of simulation.
IV. CONCLUSIONS

In this paper, a technique for gaze-based guidance of personal assistance robots has been illustrated via simulation and experiments. Fuzzy logic allows the robot to simultaneously manage multiple behaviors, practicing energy conservation when appropriate but pursuing the target when human intent to do so is clear. Combined use of the eye gaze point and blinking data is a pivotal feature of the fuzzy logic controller. Basic obstacle avoidance is demonstrated as an integrated behavior within this controller. Additionally, the fuzzy controller was successfully used to control a real-time robot using actual gaze data acquired from human users using an eye tracking system.

The results presented in this paper suggest promise for additional future work, which could focus on incorporating the controller into a system that actively detects the robot’s location and obstacles without the need for special markers. The controller should also be tuned for improved performance, and some of its more basic rules may be replaced by a more sophisticated steering and obstacle avoidance rule set based on recent research in inference modeling [21][22]. Performance comparison with other MIMO control approaches will then be appropriate. More advanced work will focus on detailed implementation for a broader variety of personal assistance tasks (e.g., object pick-and-place, operating on a static object) in a true 3D environment.

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under Grants No. 1264504 and 1414299.

REFERENCES


Figure 10. Robot trajectory from the experiment of Fig. 9.

Figure 11. Modulation of speed based on blink rate demonstrates effective interpretation of user intent in the experiment of Fig. 9.

Figure 12. Modulation of speed based on target distance demonstrates effective target seeking in the experiment of Fig. 9.

Figure 13. Modulation of heading adjustment based on obstacle distance demonstrates effective obstacle avoidance in the experiment of Fig. 9.


