To See and to Be Seen in the Virtual Beer Garden - A Gaze Behavior System for Intelligent Virtual Agents in a 3D Environment

Michael Wissner, Nikolaus Bee, Julian Kienberger, and Elisabeth André

Institute of Computer Science, Augsburg Univeristy,
86135 Augsburg, Germany
{wissner,bee,andre}@informatik.uni-augsburg.de

Abstract. Aiming to increase the believability of intelligent virtual agents, this paper describes the implementation of a parameterizable gaze behavior system based on psychological notions of human gaze. The resulting gaze patterns and agent behaviors cover a wide range of possible uses due to this parametrization. We also show how we integrated and used the system within a virtual environment.

1 Introduction

The Virtual Beer Garden is a 3D environment implemented with the Horde3D GameEngine [6]. It serves as a meeting place for virtual agents which can engage in social interaction with each other. Agents can stroll through the beer garden, meet other agents by moving towards them, and interact with them using gestures, facial expression and synthesized speech. Furthermore, agents can interact with the environment itself, such as sitting down on a bench or drinking from a mug. Both kinds of interactions are driven by desires previously assigned to the agents, such as hunger, thirst or sociability.

In previous work, we investigated how to successfully integrate a user's avatar into this environment, letting them partake in the agents' social group dynamics (see [16]). However, the agents were lacking natural gaze behaviors. When a user navigated his avatar through the environment, other agents did not seem to take notice of the avatar and vice versa. We felt that this greatly reduced believability, both of the agents themselves and the virtual environment in general.

As a consequence we present a gaze behavior system that generates natural gaze behaviors for agents in the Virtual Beer Garden. The main challenges in this endeavor are owed to the fact that the Virtual Beer Garden is a detailed and dynamic environment. It contains an abundance of perceivable entities, many of which can also be moving. Also, the perceiving agents might be moving as well.

With this work, we try to achieve two main goals: The first is to counter the above-mentioned lack in believability, while the second is to move towards autonomous behavior generation for user-controlled avatars in virtual environments.

B. Mertsching, M. Hund, and Z. Aziz (Eds.): KI 2009, LNAI 5803, pp. 500–507, 2009.

2 Related Work

There are several psychological notions of human gaze and attention which we should take into account when creating gaze behaviors for virtual characters. One is the control of gaze which can be either top-down or bottom up. Top-down or endogenous control comes from within, is based on an active decision and serves a task- or goal-oriented purpose. Bottom-up or exogenous control is created by an external stimulus, such as a sudden motion (see Posner [15], Wolfe [19], Neisser [9]).

The second is the orientation of attention and how attention can be shifted. Overt attention is the process of actually directing the eyes towards a target, thus shifting the entire field of view. Covert attention, on the other hand, is the mental process of just focusing on a target without moving the eyes (see [15]).

Finally, Wolfe [19] states that a human's visual field contains far more information than the perceptual system can handle. He describes two ways to solve this: One is to limit the amount of information that is sensed, the other is to limit the amount of sensed information that is actually perceived.

Plenty of previous work exists on gaze behavior for virtual characters. Table 1 shows the related work as well as our approach, with the behavior of the respective characters characterized by the following criteria: Embodiment (Talking Head, Torso or Full Body), Number of agents with gaze behavior, Autonomy (completely autonomous agents or semi-autonomous user avatars), Kind of Interaction (towards the user or other characters, with "User (*)" indicating

Table 1. Related Work on Virtual Characters

	Embodiment	Agents	Autonomy	Interaction	Locomotion
Bee et al. [1]	Head	1	Comp.	User	No
Cohen et al. [2]	Head	2 or Many	Semi	User and Character	No
Gillies and Ballin [3]	Full	2	Semi	Character	No
Gillies and Dodgson [4]	Full	1	Comp.	–	Yes
Hill [5]	–	1	Comp.	–	–
Khullar and Badler [7]	Full	1	Comp.	–	Yes
Pedica and Vilhjálmsson [10]	Full	Many	Semi	Character	Yes
Pelachaud and Bilvi [11]	–	2	Comp.	Character	–
Peters and O'Sullivan [12]	Full	1	Comp.	Character	Yes
Peters et al. [13]	–	2	Comp.	Character	–
Poggi et al. [14]	Head	1	Comp.	User (*)	No
Vilhjálmsson and Cassell [17]	Torso	2	Semi	User	Yes
Vinayagamoorthy et al. [18]	Full	2	Semi	User	Yes
Our Work	Full	Many	Comp. or Semi	User or Character	Yes

one-way-communication without input from the user) and Locomotion (whether the characters in question move around). An entry of '–' means that a criterion is not applicable or not explicitly specified. Our system features many fully embodied agents that move around the environment. Their behavior can be either completely autonomous or semi-autonomous. Comparing this to the other entries in the table, we see that our approach covers a wide range of features (such as the possibility to switch the interaction target and autonomy) which makes the resulting agents and their behavior suitable for an equal wide range of applications and scenarios for experiments.

3 Implemented Gaze Behavior System

3.1 Field of View

While a human's visual field (i.e. the area in which the existence of other entities can be perceived) extends to about 95 degrees on either side (see [8]), we chose to limit our agents' field of view to an area in which entities can actually be identified. The corresponding parameter in our model is an angle that extends to either side of a centered and straight "focus line" affixed to the agent's forehead. Thus, as the agent's head turns, its field of view will shift as well.

Besides this horizontal limitation, the agent's field of view is also limited in regard to the distance at which entities can be identified: Each agent can have a certain number of perception areas, each defined by a radius around the agent and an identification probability. The closer an area (and any entities in it) is located to the agent, the higher the probability of successfully identifying an entity. Usually, this probability should be 100 percent for the area closest to the agent, and linearly decreasing in every further area. Outside of the last specified area the probability is zero and thus proper identification is no longer possible.

Figure 1 shows an agent along with its cone-shaped field of view, divided into four perception areas.

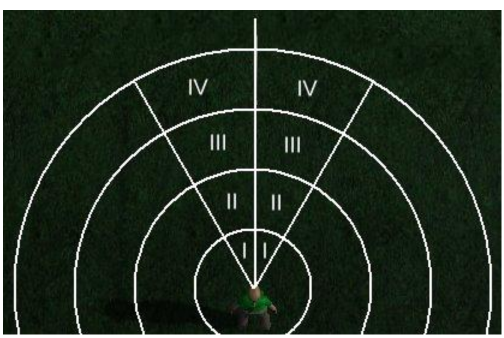

Fig. 1. An Agent's Field of View

3.2 Perception and Identification

In our gaze behavior system, the perception and identification process is distributed across three different modules: The monitoring module surveils an agent's field of view, checking whether other entities enter it, either because they moved on their own or because the agent itself moved or turned its head. The identification module tries to identify such entities, using the identification probability of the respective perception area within the field of view. Finally, the memory module keeps track of perceived entities and the area they were perceived in by putting them in one of two lists, depending on whether the identification was successful. These lists are used to ensure two things: First, there will be no further identification attempts for an already identified entity. Second, the next identification attempt for an unidentified entity will only happen if the entity appears in a nearer perception area. The reason behind this is that it would be unrealistic if the agent, after not being able to identify a target, could suddenly identify it from the same or a greater distance.

3.3 Implemented Gaze Patterns

We implemented four different kinds of gaze patterns for our virtual agents:

- Looking straight ahead: This is an agent's basic gaze behavior, it will just gaze straight ahead without shifting its field of view.
- Idle-viewing: After standing still for a certain, specified amount of time, an agent's gaze will start wandering around, performing what we call idle-viewing. This is done by selecting random points in space and having the agent look at each of them one by one, slowly turning the head in between. These points are not entirely random, however, as they are only generated within the boundaries of the agent's field of view, both on the horizontal and the vertical axis. Other names for this pattern include "undirected attention" [4] or "spontaneous looking" [12].
- Visual search: To improve an agent's chance of spotting entities of interest, the visual search pattern can be employed. Here, the agent will slowly turn its head left and right, as far as the edges of its field of view. Thus, by alternately turning until the "focus line" is aligned with the respective former edge of the field of view, the area covered by the agent's gaze is essentially doubled. As an example, if the agent's field of view was 60 degrees, a search would be started by turning its head 30 degrees to the right. The field of view is still 60 degrees but now it covers the area from zero to 60 degrees, as opposed to the range from -30 to 30 degrees when the agent was looking straight ahead. After that, the agent would turn its head all the way to -30 degrees. Altogether, the agent's gaze now covers the area from -60 to 60 degrees, though not at the same time.
- Fixating: When an agent perceives (but not necessarily identifies) an entity of interest, the agent will stare at it for a short time, fixing its "focus line" on this target. Afterwards, the agent will continue with whichever behavior it was performing before (such as idle-viewing).

The Horde3D GameEngine contains a module for Inverse Kinematics, allowing for agents to point to arbitrary points in space with any of their limbs. This allowed us to very easily implement the above-mentioned patterns by simply turning an agent's head, along with its "focus line", towards a specific point in space.

3.4 Overall Agent Gaze Behaviors

By applying and combining the four gaze patterns mentioned previously, we created three different overall gaze behaviors for our agents. Whenever an agent is added to the virtual environment, it is given one of these behaviors which will govern its actions. However, to create more diversity among the agents (even those employing the same behavior), each behavior can be further customized by adjusting the parameters explained throughout this chapter.

- Stationary: A stationary agent is just perceiving its surroundings, it is standing still and gazes straight ahead, without moving around or turning its head. However, after having identified another entity, it will first stare at them and then react in a previously specified manner appropriate to said entity.
- User-controlled: This agent can be directly controlled by a user and becomes their avatar within the virtual environment. The user can navigate the agent through the environment and the agent will gaze straight ahead while walking. As soon as the user no longer provides any input, the agent will stop and start idle-viewing. Regardless of moving or not, the agent will always react to other entities (as described above for a stationary agent).
- Autonomous: An autonomous agent moves around the environment on its own by following a set of previously defined waypoints, stopping now and then. When doing this, its gaze behavior is similar to that of a user-controlled avatar. However, an autonomous agent can decide to enter a special search mode. While in this mode, the agent will still move between waypoints but it will stop and perform a visual search at some points for a specified amount of time.

Comparing our approach with the psychological notions described in chapter 2, we find that our system contains both endogenous (e.g. visual search) and exogenous (fixating perceived entities) causes of gaze behavior. In case of a conflict between these, exogenous control takes precedence by interrupting other behaviors. Furthermore, our model includes both forms of attention, covert (agents can perceive entities anywhere within their field of view) and overt (agents fix their gaze on entities of interest). Also, we prevent perceptual overload by limiting the amount of sensible entities to a predefined set.

4 Sample Application

As a show case and testing environment for our gaze behavior system, we created a sample application within the Virtual Beer Garden.

This sample application contains eleven stationary agents and one "protagonist" agent that can either be autonomous or controlled by the user. Each of the stationary agents belongs to one of three categories in regard to its relationship to the protagonist: The protagonist can either have positive, negative or neutral feelings towards each of them (similar to [3], where agents' affiliations influence their behavior). Each agent also has a unique name that serves as an identifier. In this scenario there are no other entities of interest besides the other agents.

The application displays additional information for the user, including which behavior the protagonist is currently employing, its current status (moving, idle-viewing, searching), other agents currently within its field of view along with the respective perception areas and the last successfully identified agent. Figure 2 shows this additional information as the protagonist approaches and idly gazes upon a group of other agents. Also note the representation of the "focus line". The user can either start to steer the protagonist into the beer garden and towards the other agents or switch it to autonomous behavior. As soon as the protagonist successfully identifies another agent, it is able to read that agent's name from the environmental annotation. Matching this name against an

Fig. 2. Protagonist perceiving (but not yet identifying) a Group of other Agents

internal representation, the protagonist can then determine its feelings towards that agent and react accordingly. This comprises waving for positive feelings, no action for neutral feelings and suddenly averting gaze for negative feelings.

We used the sample application to conduct an informal study which showed that the integration of our gaze system into the beer garden resulted in a higher believability of the agent behavior: The impression, that agents can really see other agents and can be seen by them while walking around the environment removes of one the main impairments of believability we had with the previous version of the beer garden. Also, although the stationary behavior might not seem very sophisticated, we were satisfied with the results. By having the respective agents play animations that suggested preoccupation (e.g. eating, chatting or standing behind a counter), the fact that they were simply gazing straight ahead did not seem to impair believability. Finally, the visual search seemed a bit too mechanical in its current form with the agent coming to a full stop, then looking around and then moving again.

5 Conclusion and Future Work

We implemented a gaze behavior system for virtual agents in a virtual 3D environment, covering different kinds of gaze behaviors such as idling or those motivated by extrogenous or endogenous factors.

A main feature of our approach is the easy customization for each agent's behavior. This parametrization allows us to experiment with different variations of gaze behaviors that can be used within any virtual environment.

We integrated the implemented gaze behavior system into a test application, the Virtual Beer Garden, and conducted an informal study.

Ideas for future work include an improvement in the visual search pattern, the addition of objects as entities of interest within the virtual environment, testing the gaze behavior system with multiple moving agents, an addition of gaze behavior for communicative purposes as well as a user studies to obtain a formal evaluation of our system and substantiate our findings.

Acknowledgments

This work has been funded in part by the DFG under grant agreement RE 2619/2-1 (CUBE-G).

References

1. Bee, N., André, E., Tober, S.: Breaking the ice in human-agent communication: Eye-gaze based initiation of contact with an embodied conversational agent. In: Intelligent Virtual Agents (IVA) (in press, 2009)
2. Cohen, M., Colburn, A., Drucker, S.: The role of eye gaze in avatar mediated conversational interfaces (2000)

3. Gillies, M., Ballin, D.: Integrating autonomous behavior and user control for believable agents. In: AAMAS 2004: Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems, pp. 336–343 (2004)
4. Gillies, M., Dodgson, N.: Eye movements and attention for behavioural animation. The Journal of Visualization and Computer Animation 13(5), 287–300 (2002)
5. Hill, R.: Perceptual attention in virtual humans: Towards realistic and believable gaze behaviors. In: Proceedings of the AAAI Fall Symposium on Simulating Human Agents (2000)
6. Horde3D GameEngine. University of Augsburg (2008), http://mm-werkstatt.informatik.uni-augsburg.de/projects/GameEngine/
7. Khullar, S.C., Badler, N.: Where to look? automating attending behaviors of virtual human characters. Autonomous Agents and Multi-Agent Systems 4(1-2), 9–23 (2001)
8. Kim, Y., van Velsen, M., Hill, R.: Modeling dynamic perceptual attention in complex virtual environments. In: Panayiotopoulos, T., Gratch, J., Aylett, R.S., Ballin, D., Olivier, P., Rist, T. (eds.) IVA 2005. LNCS (LNAI), vol. 3661, pp. 266–277. Springer, Heidelberg (2005)
9. Neisser, U.: Cognitive psychology. Appleton-Century-Crofts, New York (1967)
10. Pedica, C., Vilhjálmsson, H.: Social perception and steering for online avatars. In: Prendinger, H., Lester, J.C., Ishizuka, M. (eds.) IVA 2008. LNCS (LNAI), vol. 5208, pp. 104–116. Springer, Heidelberg (2008)
11. Pelachaud, C., Bilvi, M.: Modelling gaze behavior for conversational agents. In: Rist, T., Aylett, R.S., Ballin, D., Rickel, J. (eds.) IVA 2003. LNCS (LNAI), vol. 2792, pp. 93–100. Springer, Heidelberg (2003)
12. Peters, C., O'Sullivan, C.: Bottom-up visual attention for virtual human animation. In: Computer Animation for Social Agents, pp. 111–117 (2003)
13. Peters, C., Pelachaud, C., Bevacqua, E., Mancini, M., Poggi, I.: A model of attention and interest using gaze behavior. In: Panayiotopoulos, T., Gratch, J., Aylett, R.S., Ballin, D., Olivier, P., Rist, T. (eds.) IVA 2005. LNCS (LNAI), vol. 3661, pp. 229–240. Springer, Heidelberg (2005)
14. Poggi, I., Pelachaud, C., de Rosis, F.: Eye communication in a conversational 3d synthetic agent. AI Communications 13(3), 169–182 (2000)
15. Posner, M.I.: Orienting of attention. The Quarterly Journal of Experimental Psychology 32(1), 3–25 (1980)
16. Rehm, M., Endrass, B., Wissner, M.: Integrating the user in the social group dynamics of agents. In: Workshop on Social Intelligence Design, SID (2007)
17. Vilhjálmsson, H., Cassell, J.: Bodychat: Autonomous communicative behaviors in avatars. In: AGENTS 1998, pp. 269–276 (1998)
18. Vinayagamoorthy, V., Garau, M., Steed, A., Slater, M.: An eye gaze model for dyadic interaction in an immersive virtual environment: Practice and experience. Computer Graphics Forum 23(1), 1–12 (2004)
19. Wolfe, J.M.: Guided search 2.0: A revised model of visual search. Psychonomic Bulletin and Review 1(2), 202–238 (1994)