Published in final edited form as:
ACM Trans Comput Hum Interact. 2008 ; 2008: 97–104. doi:10.1145/1452392.1452412.

As Go the Feet … : On the Estimation of Attentional Focus from Stance

Francis Quek
Center for Human-Computer Interaction Virginia Tech quek@vt.edu

Roger Ehrich
Center for Human-Computer Interaction Virginia Tech rehrich@vt.edu

Thurmon Lockhart
Center for Human-Computer Interaction Virginia Tech lockhart@vt.edu

Abstract

The estimation of the direction of visual attention is critical to a large number of interactive systems. This paper investigates the cross-modal relation of the position of one's feet (or standing stance) to the focus of gaze. The intuition is that while one CAN have a range of attentional foci from a particular stance, one may be MORE LIKELY to look in specific directions given an approach vector and stance. We posit that the cross-modal relationship is constrained by biomechanics and personal style. We define a *stance vector* that models the approach direction before stopping and the pose of a subject's feet. We present a study where the subjects' feet and approach vector are tracked. The subjects read aloud contents of note cards in 4 locations. The order of `visits' to the cards were randomized. Ten subjects read 40 lines of text each, yielding 400 stance vectors and gaze directions. We divided our data into 4 sets of 300 training and 100 test vectors and trained a neural net to estimate the gaze direction given the stance vector. Our results show that 31% our gaze orientation estimates were within 5°, 51% of our estimates were within 10°, and 60% were within 15°. Given the ability to track foot position, the procedure is minimally invasive.

Keywords

Human-Computer Interaction; Attention Estimation; Stance Model; Foot-Tracking; Multimodal Interfaces

1. INTRODUCTION

The estimation of a subject's zone of attention is important to such domains in human-centered computing as computer-supported collaboration, teaching and learning environments, context-aware interaction, large-scale visualization, smart homes, multimodal interfaces, wearable computing, and analysis of group interaction. Systems that estimate such attention typically involve intrusive sensing technology such as video tracking and wearable technology. In this

Categories and Subject Descriptors H.5 INFORMATION INTERFACES AND PRESENTATION (e.g., HCI) (I.7) I.5 PATTERN RECOGNITION,

General Terms Human Factors

paper, we explore the possibility of estimating one's `zone of attention' by tracking one's footfalls and standing stance. The specific mode of tracking is not the focus of this paper, although one might imagine pressure sensitive carpets with [1,2] thin piezoelectric cables [3], sensorized tiled floors [4–8], and a variety of wearable devices [9–11]. If such zone of attention estimation is possible, a range of multimodal interactive systems will be enabled that present timely information on ambient displays, support human meetings by estimating the zone of attention of one's interlocutor, and create active displays that are attention-sensitive.

By `zone of attention', we mean the sector of visual space of a subject. Our interest here is not in the exact angle of gaze as might be required when using an eye-tracker to support gaze control of a screen cursor. Rather, we want to estimate the general zone of attention (or where the subject is looking) centered on some sector axis as shown in Figure 1. Observe that our estimate does not require that the center of the attention zone be coincident with the `nose-forward' vector.

Our approach is to exploit the anatomical and behavioral constraints of the subject when her feet are set and to employ a biomechanical model from whose parameters we estimate the zone of attention using a classifier or by function estimation. In Section 2, we discuss the rationale for our approach by reviewing the need for gaze/awareness, and reviewing existing attention awareness approaches. We show that there is need for a coarse-scale attention estimation approach that is able to work over a large area. In Section 3, we ground our approach by outlining our biomechanical assumptions and discussing our model. In Section 4, we describe our empirical approach and report our experimental results. We conclude in Section 5.

2. Attention Awareness and Tracking

2.1 Rationale

We make a distinction between gaze tracking and attention awareness. The former typically requires precise detection of the angle or locus of gaze so that, for example, one may control a screen cursor with the output. Attention awareness requires that the zone of gaze be detected to determine the object or area of visual attention. The focus in this paper is the latter. The ability to detect attentional focus is critical to a wide variety of multimodal interaction and interface systems. These include computer-supported group interaction [12–16], wearable computing, augmented reality [17–21], context/attention aware applications [21–26], education and learning [26,27], smart homes [28,29], and meeting analysis [30–36]. In many of these applications, non-intrusiveness in the mode of sensing is more important than the accuracy of gaze angle tracking. We investigate the estimation of the zone of attention from a standing stance unobtrusively using information that may be obtained entirely from a sensing carpet, smart floor, or instrumented shoes. In this section, we review the state of the art in tracking visual attention, discuss the biometric and psycho-social behavioral constraints in direction of visual attention with respect to stance, and advance our model for zone of attention detection and tracking.

2.2 Review of Attention Awareness Approaches

Allocation of attention is a key aspect of collaboration, meeting conduct, and interaction that is a prime determiner of information flow among participants or supporting technologies. Gaze has long been recognized as a primary indicator of zone of attention and as a conversational resource that assists participants in assessing connection, comprehension, reaction, responsiveness, and in interpreting intention [37]. Other researchers have investigated the attentional behavior of subjects who are observing the gaze of others [38,39]. As several researchers point out, however, it is still possible to visually fixate one location while diverting

attention to another [40,41]; even though eye tracking may be highly accurate, gaze direction is not necessarily a highly accurate estimator of attentional focus.

Eye tracking has been by far the technology of choice for gaze estimation. There has been much interest in the use of gaze to control interaction [42,43], to modify information presentation [44–46], and to interact directly with data [47]. Duchowski [40] partitions eye tracking applications into either *diagnostic* or *interactive* categories, depending upon whether the tracker provides objective and quantitative evidence of the user's visual and attentional processes or whether it serves as an interaction device.

Eye tracking has been of intense interest for many years, and many techniques have been proposed [48]. Some, such as electrooculography [49], which involves attaching electrodes near the eyes, and magnetic eye-coil tracking, which involves special contact lenses, are particularly invasive and uncommon. Most current tracking methods are video based and fall into one of two categories, using either infrared illumination or passive tracking. The use of remote fixed cameras is not of great interest except in special studies because of the problems of occlusion, head tracking, and tracking multiple subjects simultaneously. Excellent work has also been done in the case of head mounted trackers to minimize the invasiveness of the camera, mount, and cabling [50], but the apparatus is always present in front of the wearer.

The so-called limbus trackers are usually passive trackers that utilize ambient light to track the limbus, which is the junction between the iris and the white surrounding sclera. These trackers are somewhat simpler since they do not require a special illumination source but suffer from the uncontrolled nature of the ambient light and limitations in vertical tracking due to eyelid movements. On the other hand, the use of an infrared illumination source makes it practical to track the pupil, which is a more sharply defined and less occluded feature than the limbus. However, infrared trackers can suffer in the presence of other infrared sources such as natural sunlight. Other infrared trackers make use of the so-called Purkinje images [51], which are due to reflections from the several optical boundaries within the eye such as the surfaces of the lens and cornea. The measurement of individual features requires that the head position be fixed or tracked; sophisticated trackers can avoid this by tracking features that move differentially when the eye moves as opposed to the head. More important, neither approach is feasible for estimating attention of subjects over a large space.

Some eye trackers separate the process of determining head orientation from the local process of determining eye orientation given the head pose [52–55]. Others have proposed the use of head pose alone as an estimator of gaze direction [41,56,57] in order to eliminate the need for invasive head mounted hardware. Stiefelhagen's results [41,57] provide strong evidence to support the effectiveness of head orientation alone as an estimator of focus of attention.

The prime deterrents to the use of eye trackers for gaze analysis have been the cost of eye tracking, its invasiveness, its lack of robustness, and the difficulty of performing the analysis simultaneously on large groups of meeting or collaboration participants. Yet other problems concern calibration, dynamic range, response time, and angular range. Since gaze direction does not uniquely determine focus of attention, there are many applications in which determining zone of attention to high accuracy (< 1 degree) is unnecessary, applications for which gait, location, identity, and pose are sufficient estimators of the desired information.

The advent of wearable computing has sensitized researchers to the need for deeper context awareness that includes, among other things, the pose and location of the wearer. As always, the dilemma is how to determine this information as noninvasively as possible, simultaneously for multiple individuals, and over a large area where the motion of users are minimally constrained. Our hypothesis is that useful estimates of zone of attention can be obtained from floor stance and approach vector, so that the search for suitable sensor systems can be shifted

to shoe and floor systems. An extensive sensor system for determining floor stance has been proposed by the Responsive Environments Group at the MIT Media Laboratory in which each participant shoe is fitted with a rich complement of wireless sensors [9–11].

In summary, head-mounted and wearable trackers encumber the user; many are still tethered by cabling, which is especially problematic in multi-user environments. Other systems that employ video and electromagnetic technology restrict the movement of the user to the effective tracking volume of the technology.

3. Stance Model

3.1 Biomechanical Orientation Constraint

Floor stance and approach vector during locomotion can provide useful estimates of zone of attention. Figure 3 shows the degrees of freedom available to a human viewer when the feet are set. Human locomotion is guided by *optic flow* and *egocentric direction strategy* utilizing variant degrees of target visual context [58–61]. *Optic flow* describes temporal changes in image structure as a walker moves; and *egocentric direction strategies* describe how one walks in different contexts (e.g. in dimly lit areas one may use egocentric coding that minimizes angular distances to the goal.) The assumption behind both of these locomotion strategies is that the goal is visible, and as such, directly related to the salience of the visual context [62]. The saliency dependence of the visual context suggests that gaze transient (i.e., flow and direction) of a target is an important parameter for goal directed gait. This can be conceptualized as a constraint on the way one approaches a target of focal attention prior to the static (standing) stance or configuration.

Gaze control involves motion coordination of eyes, head and trunk to allow both flexibility of movements and stability of gaze. During straight walking, gaze is maintained in the direction of forward locomotion with small head yaw oscillations in space, despite relatively large oscillations and lateral displacements of the body. A study investigating three-dimensional head, body and eye angles during walking and turning, it was found that the peak body yaw of 3.5° in space was compensated by the relative peak head yaw of 3°, which consequently resulted in a very small head yaw angles (less than 1°) in space. Additionally, the naso-occipital axis of the head was closely aligned with the anterio-posterior direction of locomotion [63]. The head pitch and roll angles peaked at approximately 3° as observed both in over-ground walking [63] and in treadmill walking [64,65]. In terms of gaze behavior, eyes were found to spend the majority of the time (78.8%) fixating the aspects of the environment along the direction of locomotion and a small amount of time (16.3%) searching for possible future routes. What appears to be random point inspecting only took 4.9% of the time during walking [66]. Furthermore, such gazing patterns (fixating along the direction of walking) appeared not to be influenced by individual differences [66].

During turning, gaze is directed in advance of the body heading, and after turning, gaze is returned to align with the direction of motion. During a 90° turn while walking, head yaw was maintained smoothly in space, with a maximum 25° deviation from the heading direction of the body [63]. Eye position, however, was found to shift in saccades in the direction of turn (Figure 2), reaching yaw angles as high as 50° relative to the head. Once the turn was complete, eye position and foot position returned to zero relative to the head [67]. Our goal, then, is to determine the pattern of behavior that relate both the vector of approach and the final pose of the static stance with the likely final focus of attention.

3.2 Modeling Stance and Attention

In addition to biomechanical constraints in the previous section, we add behavioral constraints of instrumental gaze. In our work on meeting analysis [30–32,35,68], we observed that there

is a difference between interactive deployment of gaze and an instrumental one [32,69–71]. Interactive gaze takes place between people, and is influenced by aspects of social behavior such as the avoidance of `nose-to-nose' fixations, and back-channeling behavior. Instrumental gaze involves the deployment of gaze for the purpose of acquiring information (such as reading, or viewing a graphic). Our preliminary analysis of meeting room data suggests that there is greater variation in gaze deflection from the `nose-forward' vector of head orientation for interactive gaze with respect to instrumental gaze. Since our interest is in instrumental use of gaze with technology, our expectation is that eye deflection variability is reduced for such activity. Furthermore, the kind of instrumental gaze necessary to access information requires the deployment of central-foveal vision.

Figure 4 illustrates our base model of stance for the estimation of the zone of visual attention. We call this the base model because we expect that our model will have to evolve as more is known about the relationship between gaze and stance. The reference frame of the model is formed by the connecting line between the centers of mass of the feet, and the normal to that line in the forward direction of the subject (shown as the *x*–*y* reference frame in Figure 4). The orientations of the right and left feet are described by the angles ϕ_r and ϕ_l respectively. The approach angle γ describes the direction of locomotion prior to stopping in the resultant pose. *d* describes the width of the stance. The angle θ_T is the angle of gaze to the target of attention as a deflection off the *stance normal*. By this model, $\mathbf{v}_i = [\phi_r, \phi_l, d, \gamma]$ constitutes an input *stance vector*, and the value θ_T is the output value to be estimated.

4. Experiment

To test the hypothesis that stance may be a predictor of gaze direction, we designed an experiment where subjects are required to read a series of lines of text that are mounted on aluminum posts. The text is small so that the subjects had to move to the target to read the lines. We tracked the feet of the subjects to obtain the stance vector and used a neural net approach to learn the gaze direction. The point of this experiment is not to advance any specific learning approach. It is to ascertain if any patterning exists by which our cross-modal hypothesis may be validated.

4.1 Experiment Design

Figure 5 shows the plan view of our experimental configuration. Since our model describes only horizontal gaze deployment (Figure 4 does not include viewing pitch) the target cards are set at eye height for each subject. Figure 6 shows a picture of our experimental setup in the laboratory. Two gaze targets can be seen. We employ our Vicon near-infrared motion trackers to estimate the parameters in our model to obtain the stance vector, $\mathbf{v}_i = [\phi_r, \phi_l, d, \gamma]$. By tracking the retro-reflector marker configurations on the frame attached to the subject's shoes (see Figure 6 inset), our experiment software produces a time-stamped stream of quaternions from which we derive the basis vectors of the tracked frame for each foot. To simplify the determination of the approach vector, we also track the location of the subject's head (tracked goggles in Figure 6). This also gives us access to the subject's head orientation, although we did not use it for this experiment.

By having the subject place her foot in a box of known coordinates and orientation marked on the floor we obtain the toe-forward vector from the basis frame of each tracked position. Given the unit basis matrix $\boldsymbol{B_C}$ of the calibration box, and the unit basis matrix $\boldsymbol{B_0}$ of the tracked frame attached to a foot, we obtain the tracking transformation $\boldsymbol{M_f} = \boldsymbol{B_C} \times \boldsymbol{B_0}^T$ (for foot *f*, where *f* is *r* or *l* for the right and left foot respectively). Given a subsequent tracked unit basis matrix $\boldsymbol{B_i}$, the toe-forward frame is simply given by $\boldsymbol{M_f} \times \boldsymbol{B_i}$.

The subjects are directed to read lines in 12-point font printed on 3×5 cards placed at in four known coordinates in the laboratory (pictured in Figure 6, and labeled A, B, C, D in Figure 5). Each line of each card contains three columns: a sequential index (A.1, A.2 … for card A, B. 1, B.2 … for card B etc.), a line of text to be read, and the index of the next line to be read. The station for the next line to be read is randomized so that the subject will go from station to station to read the next line. The small fonts ensure that subject must move from one station to the next. The subject reads each line aloud so that we know when her attention is fixed on the target 3×5 cards. With this information, we can extract the parameters described in Figure 4 (Section 3.2). Each time a target is read, we record a *stance vector* $\mathbf{v}_i=[\phi_r, \phi_l, d, \gamma]$ and attention angle θ_T. For each trial, the subject reads 40 lines randomly located at the 4 targets. This trial is repeated 10 times with 10 subjects, yielding a training dataset containing 400 vectors.

4.2 Gaze Estimation from Stance

We employed a standard three-layer backpropagation neural network parameter approach [72] to estimate θ_T from $\mathbf{v}_i$. Kolmogorov [73,74] showed that any continuous function can be represented as a linear additions of multiple continuous functions. In our implementation, the input layer has four neurons for the input of the four parameters ϕ_r, ϕ_l, d, γ. The output layer has one neuron for the parameter θ_T and the hidden layer has 15 neurons (using a rule of thumb of between 4 and 5 times the number of input neurons) [72]. The network was initialized with random weights. After training with samples, the network can learn the relationship $[\phi_r, \phi_l, d, \gamma] \rightarrow \theta_T$. We can apply it to estimate the θ_T for some new $\mathbf{v}_i$. For our study, we divided our dataset into four sets of 300 training vectors and 100 test vectors. We trained our network on the former, and ran the resulting network using the *stance vectors* from the latter group.

4.3 Results and Discussion

Figure 7 is an histogram of the absolute difference $\text{abs}\theta_T'$, where θ_T' is the estimated attention direction and θ_T is the measured direction. For this dataset, 31% of the estimations fell within 5° of the measurements. 51% of estimations were within 10°, and 60% of the estimates were within 15° of error.

Figure 8 shows plots the absolute values measured θ_T against the absolute error $\text{abs}\theta_T'$ for a particular dataset (testing against 100 vectors). This shows that our estimation error increases with the size of deflection. Given the limited size of our dataset (only 400 samples), this might be expected since the data becomes sparser with larger θ_T.

These results show an estimation accuracy far in excess of chance. For example, assuming that the subject is capable of viewing 180° from a particular stance, chance would predict that a 5° estimate of 2.77%, a 10° estimate at 5.56% and a 15° estimate at 8.33%. It should be noted that this experiment did not take individual differences into account, and the training sets are not extensive. Hence, one might expect that the results to improve with more user-specific training. Also, we acknowledge that our *stance vector* is an initial principled guess. One might imagine that extension of the stance vector to include weight distribution, subject parameters (e.g. height), etc. the estimate may be improved. Our purpose here is to advance a proof-of-concept for consideration by the research community.

5. Conclusion and Future Work

We have demonstrated a rather audacious presupposition that one is able to estimate a subject's instrumental gaze direction or attentional focus from her approach vector and standing stance.

We presented our rationale for our research by reviewing the need and the technologies for gaze/attention estimation. We showed that there is need for a non-intrusive coarse scale attention estimation approach that is able to track over a large area.

We ground our proposed stance model and the expectation that we may be able to estimate attention from stance by discussing the biomechanics of approach and gaze fixation. We present our stance model comprising only four parameters.

We present a set of experiments by which we track subjects' feet and approach vector to an attention target using a motion tracking system. Subjects were required to move to one of four stations randomly and read a line of text. We extracted 400 stance vector – attention direction sets, and employed a neural net system to learn the relationship. The results are promising.

While the results are promising, more needs to be done. The approach is to find a mapping between the stance vector and the direction of attention. Our initial stance vector, while arrived at in a principled manner, ignores many other possible vectors that may be deterministic. Examples of these include weight balance (right foot vs left foot, forward lean vs backward lean), dynamics of approach, and duration of gaze.

Also, in our study, our subject approached an initial target and directed visual attention at it. We can think of this as the initial zone of attention from a particular stance. This does not address the retargeting of attentional focus from such a fixed stance after the initial attentional gaze. We conjecture that once a stance is fixed, there is a `zone of comfort' where a subject can redeploy gaze without moving her feet (shifting her stance). This might occur when the subject has selected a stance for a particular initial target and a new target appears in close proximity to the original. Let δx be the distance of some secondary target from the initial target. To characterize the range of δx, a second type of experiment is required that utilizes a large display system such as our tiled wall-sized display (seen in the background in Figure 6). When an initial target is displayed, the subject approaches and reads as before. Secondary targets are displayed at different δx's to determine typical range thresholds that engage adjustment of stance. The range of these `within-stance attention redeployments' may require extension of the stance vector to include balance components, or it may define a zone of uncertainty of secondary gaze targets.

Acknowledgments

This work has been partially supported by NSF grants This research has been supported by the U.S. National Science Foundation NSF ITR program, Grant No. ITR-0219875, “Beyond the Talking Head and Animated Icon: Behaviorally Situated Avatars for Tutoring,” IIS-0624701, CRI-0551610, “Embodied Communication: Vivid Interaction with History and Literature,” and, NSF-IIS- 0451843, “Interacting with the Embodied Mind,” and “Embodiment Awareness, Mathematics Discourse and the Blind.” We also acknowledge Yingen Xiong and Pak-Kiu Chung who conducted the experiments and ran the neural net pattern classification system.

7. REFERENCES

[1]. Bekintex. 2003. [cited 2008 May 20]; http://www.bekintex.com

[2]. Glaser, R.; Lauterbach, C.; Savio, D.; Schnell, M., et al. Smart Carpet: A Textile-based Large-area Sensor Network. 2005. [cited 2008 May 20]; url="http://www.future-shape.de/publications_lauterbach/SmartFloor2005.pdf

[3]. Measurement Specialties. I. Piezo Technical Manual, Piezoelectric Film Properties. 2002. cited; http://www.msiusa.com/PART1-INT.pdf

[4]. Addlesee MD, Jones AH, Livesey F, Samaria FS. ORL Active Floor. IEEE Personal Communications 1997;4(5):35–41.

[5]. Kaddourah, Y.; King, J.; Helal, A. Cost-Precision Tradeoffs in Unencumbered Floor-Based Indoor Location Tracking. third International Conference On Smart homes and health Telematic (ICOST); Sherbrooke, Québec, Canada. 2005.

[6]. Orr, RJ.; Abowd, GD. The smart floor: a mechanism for natural user identification and tracking, in CHI '00 extended abstracts on Human Factors in Computing Systems. ACM; The Hague, The Netherlands: 2000.

[7]. Richardson, B.; Leydon, K.; Fernstrom, M.; Paradiso, JA. Z-Tiles: building blocks for modular, pressure-sensing floorspaces, in CHI '04 extended abstracts on Human Factors in Computing Systems. ACM; Vienna, Austria: 2004.

[8]. Srinivasan, P.; Birchfield, D.; Qian, G.; Kidané, A. A Pressure Sensing Floor for Interactive Media Applications. ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE); 2005.

[9]. Paradiso JA, et al. Sensor Systems for Interactive Surfaces. IBM Systems Journal 2000a;39(Nos.(3 & 4)):892–914.

[10]. Paradiso, JA.; Hsiao, KY.; Benbasat, A. Interfacing to the Foot. ACM Conference on Human Factors in Computing Systems; 2000b.

[11]. Paradiso JJ, et al. Design and Implementation of Expressive Footware. IBM Systems Journal 2000c; 39(Nos. 3 & 4):511–529.

[12]. Hiroshi, I.; Minoru, K. ClearBoard: a seamless medium for shared drawing and conversation with eye contact. CHI '92: Proceedings of the SIGCHI conference on Human Factors in Computing Systems; ACM; 1992.

[13]. Gutwin, C.; Greenberg, S. The Importance of Awareness for Team Cognition in Distributed Collaboration, in Team Cognition: Understanding the Factors that Drive Process and Performance. Salas, E.; Fiore, SM., editors. APA Press; Washington: 2004. p. 177-201.

[14]. Greenberg, S. Real Time Distributed Collaboration, in Encyclopedia of Distributed Computing. Dasgupta, P.; Urban, JE., editors. Kluwer Academic Publishers; 2002.

[15]. Jabarin, B.; Wu, J.; Vertegaal, R.; Grigorov, L. CHI '03: CHI '03 extended abstracts on Human Factors in Computing Systems. ACM; 2003. Establishing remote conversations through eye contact with physical awareness proxies.

[16]. Vertegaal, R. CHI '99: Proceedings of the SIGCHI conference on Human Factors in Computing Systems. ACM; 1999. The GAZE groupware system: mediating joint attention in multiparty communication and collaboration.

[17]. Billinghurst, M.; Kato, H. Collaborative Mixed Reality. First International Symposium on Mixed Reality (ISMR '99). Mixed Reality – Merging Real and Virtual Worlds; Berlin: Springer Verlag; 1999.

[18]. Vertegaal, R.; Dickie, C.; Sohn, C.; Flickner, M. CHI '02: CHI '02 extended abstracts on Human Factors in Computing Systems. ACM; 2002. Designing attentive cell phone using wearable eyecontact sensors.

[19]. Mann S. Wearable computing: toward humanistic intelligence. Intelligent Systems, IEEE [see also IEEE Intelligent Systems and Their Applications] 2001;16(3):10–15.

[20]. Aaltonen, A. A context visualization model for wearable computers. (ISWC 2002). Proceedings. Sixth International Symposium on Wearable Computers; 2002.

[21]. Cheng L-T, Robinson J. Personal contextual awareness through visual focus. Intelligent Systems, IEEE [see also IEEE Intelligent Systems and Their Applications] 2001;16(3):16–20.

[22]. Hyrskykari, A.; Majaranta, P.; Räihä, K-J. Proc. HCII 2005. Las Vegas, NV: 2005. From Gaze Control to Attentive Interfaces.

[23]. Shell J, Selker T, Vertegaal R. Interacting with groups of computers. Comm. ACM 2003;46(3):40–46.

[24]. Crowley, J.; Coutaz, J.l.; Rey, G.; Reignier, P. Perceptual Components for Context Aware Computing. Lecture Notes in Computer Science: UbiComp 2002: Ubiquitous Computing: 4th International Conference; Goteborg, Sweden. September 29 – October 1, 2002; 2002. p. 117Proceedings

[25]. Chou, P.; Gruteser, M.; Lai, J.; Levas, A., et al. IBM Thomas J. Watson Research Center; Yorktown Heights, NY 10598: 2001. BlueSpace: Creating a Personalized and Context-Aware Workspace; p. 20

[26]. Roda, C.; Thomas, J. Computers in Human Behavior Computers in Human Behavior. Elsevier; 2006. Attention Aware Systems: Theory, Application, and Research Agenda.

[27]. Woolfolk A, Galloway C. Nonverbal behavior and the study of teaching. Theory into practice 1985;24:77–84.

[28]. Brumitt B, Cadiz J. Let There Be Light: Comparing Interfaces for Homes of the Future. IEEE personal communications 2000;28:35.

[29]. Feki M, Renouard S, Abdulrazak B, Chollet G, et al. Coupling Context Awareness and Multimodality in Smart Homes Concept. Lecture Notes in Computer Science: Computers Helping People with Special Needs 2004:906–913.

[30]. Xiong, Y.; Quek, F. Head Tracking with 3D Texture Map Model in Planning Meeting Analysis. International Workshop on Multimodal Multiparty Meeting Processing (ICMI-MMMP'05); Trento, Italy. 2005.

[31]. Chen, L.; Rose, T.; Perill, F.; Han, X., et al. VACE Multimodal Meeting Corpus. 2nd Joint Workshop on Multimodal Interaction and Related Machine Learning Algorithms; Edinburgh, UK: Royal College of Physicians; 2005.

[32]. Quek, F.; Rose, RT.; McNeill, D. Multimodal Meeting Analysis, in International Conference on Intelligence Analysis; 2005.

[33]. Chen, L.; Harper, M.; Franklin, A.; Rose, RT., et al. A Multimodal Analysis of Floor Control in Meetings. 3rd Joint Workshop on Multimodal Interaction and Related Machine Learning Algorithms (MLMI06); 2006.

[34]. McCowan, I.; Bengio, S.; Gatica-Perez, D.; Lathoud, G., et al. Modeling Human Interaction in Meetings. ICASSP; Hong Kong. 2003.

[35]. Stiefelhagen R, Zhu J. Head orientation and gaze direction in meetings. Human Factors in Computing Systems (CHI2002). 2002

[36]. Yang J, Zhu X, Gross R, Kominek J, et al. Multimodal People ID for a Multimedia Meeting Browser. ACM multimedia. 1999

[37]. Grayson DM, Monk AF. Are You Looking at Me? Eye Contact and Desktop Video Conferencing. ACM Transactions on Human-Computer Interaction 2003:221–243.

[38]. Langton SRH, et al. Gaze Cues Influence the Allocation of Attention in Natural Scene Viewing. Quarterly Journal of Experimental Psychology 2006;00(No. 0):1–9.

[39]. Vertegaal, R., et al. Eye Gaze Patterns in Conversations: There is More to Conversational Agents than Meets the Eyes. ACM Conference on Human Factors in Computing Systems; 2001.

[40]. Duchowski AT. A Breadth-First Survey of Eye Tracking Applications. Behavior Research Methods, Instruments, and Computers 2002;34(4):455–470.

[41]. Stiefelhagen, R., et al. From Gaze to Focus of Attention. 3rd International Conference on Visual Information and Visual Information Systems; 1999.

[42]. Jacob R. The Use of Eye Movements In Human-Computer Interaction Techniques: What You Look at is What You Get. ACM Transactions on Information Systems 1991;9:152–169.

[43]. Sibert, LE.; Jacob, RJK. Evaluation of Eye Gaze Interaction. ACM Conference on Human Factors in Computing Systems; 2000.

[44]. Nikolov, SG., et al. Gaze-Contingent Display Using Texture Mapping and OpenGL: System and Applications. The Symposium on Eye Tracking Research and Applications; 2004.

[45]. Parkhurst DJ, Niebur E. Variable Resolution Displays: a Theoretical, Practical, and Behavioral Evaluation. Human Factors 2002;44(4):611–629. [PubMed: 12691369]

[46]. Qvarfordt, P.; Zhai, S. Conversing with the User Based on Eye-Gaze Patterns. ACM Conference on Human Factors in Computing Systems; 2005.

[47]. Glenstrup, AJ.; Engell-Nielsen, T. Eye Controlled Media: Present and Future State, in Laboratory of Psychology. University of Copenhagen; 1995.

[48]. Young L, Sheena D. Survey of Eye Movement Recording Methods. Behavior Research Methods and Instrumentation 1975;7:397–429.

[49]. Manabe, H.; Fukumoto, M. Full-time Wearable Headphone-Type Gaze Detector. ACM Conference on Human Factors in Computing Systems; 2006.

[50]. Li D, Babcock J, Parkhurst DJ. OpenEyes: a Low-cost Head-Mounted Eye-Tracking Solution. ACM Symposium on Eye Tracking Research and Applications. 2006

[51]. Cornsweet TN, Crane HD. Accurate Two-Dimensional Eye Tracker Using First and Fourth Purkinje Images. J. Optical Society of America 1973;63(8):921–928.

[52]. Ji Q, Zhu Z. Non-intrusive Eye and Gaze Tracking for Natural Human Computer Interaction. MMI Interaktiv 2003:6.

[53]. Stiefelhagen, R.; Yang, J.; Waibel, A. Workshop on Perceptual User Interfaces; 1997.

[54]. Newman, R., et al. Real-Time Stereo Tracking for Head Pose and Gaze Estimation. 4th IEEE International Conference on Automatic Face and Gesture Recognition; 2000.

[55]. Wang J-G, Sung E. Study on Eye Gaze Estimation. IEEE Trans. SMC 2002;32(No. 3):332–350.

[56]. Gee A, Cipolla R. Determining the Gaze of Faces in Images. Image and Vision Computing 1994:1–20.

[57]. Stiefelhagen, R. Tracking Focus of Attention in Meetings. ICMI; Pittsburg, PA. 2002.

[58]. Harris MG, Carre G. Is optic flow used to guide walking while wearing a displacing prism? Perception 2001;30:811–818. [PubMed: 11515954]

[59]. Rushton SK, Harris JM, Lloyd MR, Wann JP. Guidance of locomotion on foot uses perceived target location rather than optic flow. Current Biology 1998;8:1191–1194. [PubMed: 9799736]

[60]. Warren WH, Kay BA, Zosh WD, Duchon AP, et al. Optic flow is used to control human walking. Nature Neuroscience 2001;4:213–216.

[61]. Wilkie RM, Wann JP. Driving as night falls: The contribution of retinal flow and visual direction to the control of steering. Current Biology 2002;12:2014–2017. [PubMed: 12477389]

[62]. Turano KA, Yu D, Hao L, Hicks JC. Optic-flow and egocentric-direction strategies in walking: Central vs peripheral visual field. Vision Rsearch 2005;45:3117–3132.

[63]. Imai T, Moore ST, Raphan T, Cohen B. Interaction of the body, head, and eyes during walking and turning. Experimental brain research 2001;136(1):18.

[64]. Hirasaki E, Moore ST, Raphan T, Cohen B. Effects of walking velocity on vertical head and body movements during locomotion. Experimental Brain Research 1999;127(2):117–130.

[65]. Moore ST, Hirasaki E, Cohen B, Raphan T. Effect of viewing distance on the generation of vertical eye movements during locomotion. Experimental Brain Research 1999;129(3):347–361.

[66]. Hollands MA, Patla AE, Vickers JN. “Look where you're going!”: gaze behaviour associated with maintaining and changing the direction of locomotion. Experimental Brain Research 2002;143(2): 221–230.

[67]. Land MF. The coordination of rotations of eyes, head and trunk in saccadic turns produced in natural situations. Exp. Brain Res 2004;159:151–160. [PubMed: 15221164]

[68]. Chen, L.; Travis, R.; Parrill, F.; Han, X., et al. VACE Multimodal Meeting Corpus. MLMI 2005 Workshop; Edinburgh. 2005.

[69]. Quek F, McNeill D, Harper M. From Video to Information: Cross-Modal Analysis for Planning Meeting Analysis. 18-Month Presentation. 2005

[70]. McNeill, D.; Duncan, S.; Franklin, A.; Goss, J., et al. MIND-MERGING, in Draft chapter prepared for Robert Krauss Festschrift. Erlbaum Associates; 2006.

[71]. Wathugala, D. A Comparison Between Instrumental and Interactive Gaze: Eye deflection and head orientation behavior, in Computer Science. Virginia Tech; Blacksburg: 2006. expected

[72]. Duda, RO.; Hart, PE.; Stork, DG. Multilayer Neural Networks, in Pattern Classificaton. Duda, RO.; Hart, PE.; Stork, DG., editors. John Wiley and Sons Inc; New York: 2001. p. 282-349.

[73]. Kolmogorov AN. On the representation of continuous functions of several variables by superposition of continuous functions of one variable and addition. Doklady Akademiia Nauk, SSSR 1957;114(5):953–956.

[74]. Kurkova V. Kolmogorov's theorem is relevant. Neural Computation 1991;3(4):617–622.

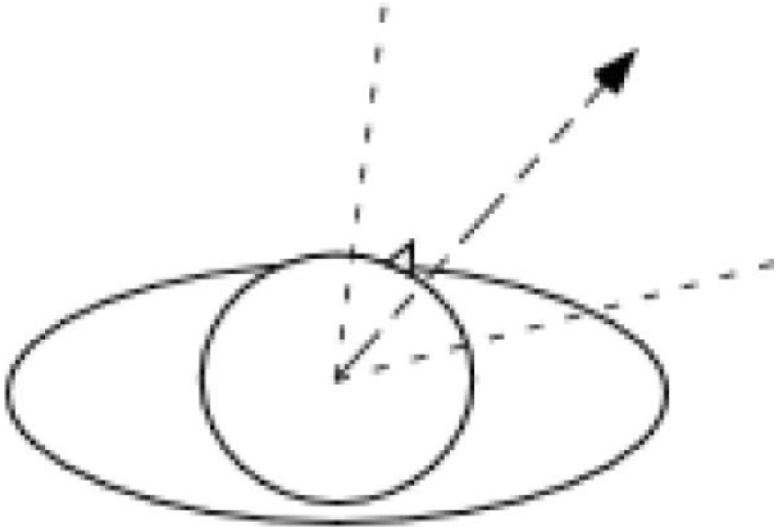

Figure 1.
Top-down view of the one of attention

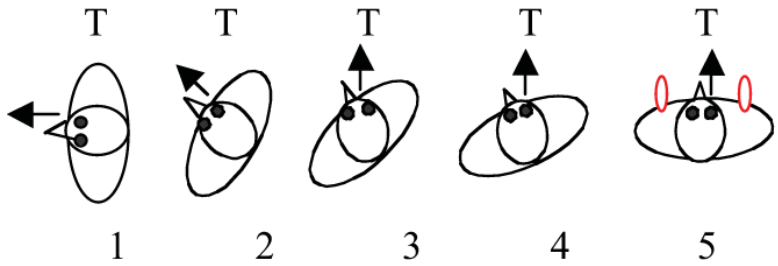

Figure 2.
Coordination of eye, head and trunk rotation during gaze shift. Diagram of a typical 90° turn (five phases) after locomotion approach to a target at T (Land, 2004). Eye movement completed by phase 2, head movement completed by phase 4, and trunk rotates continuously until phase 5 – with feet directed forward

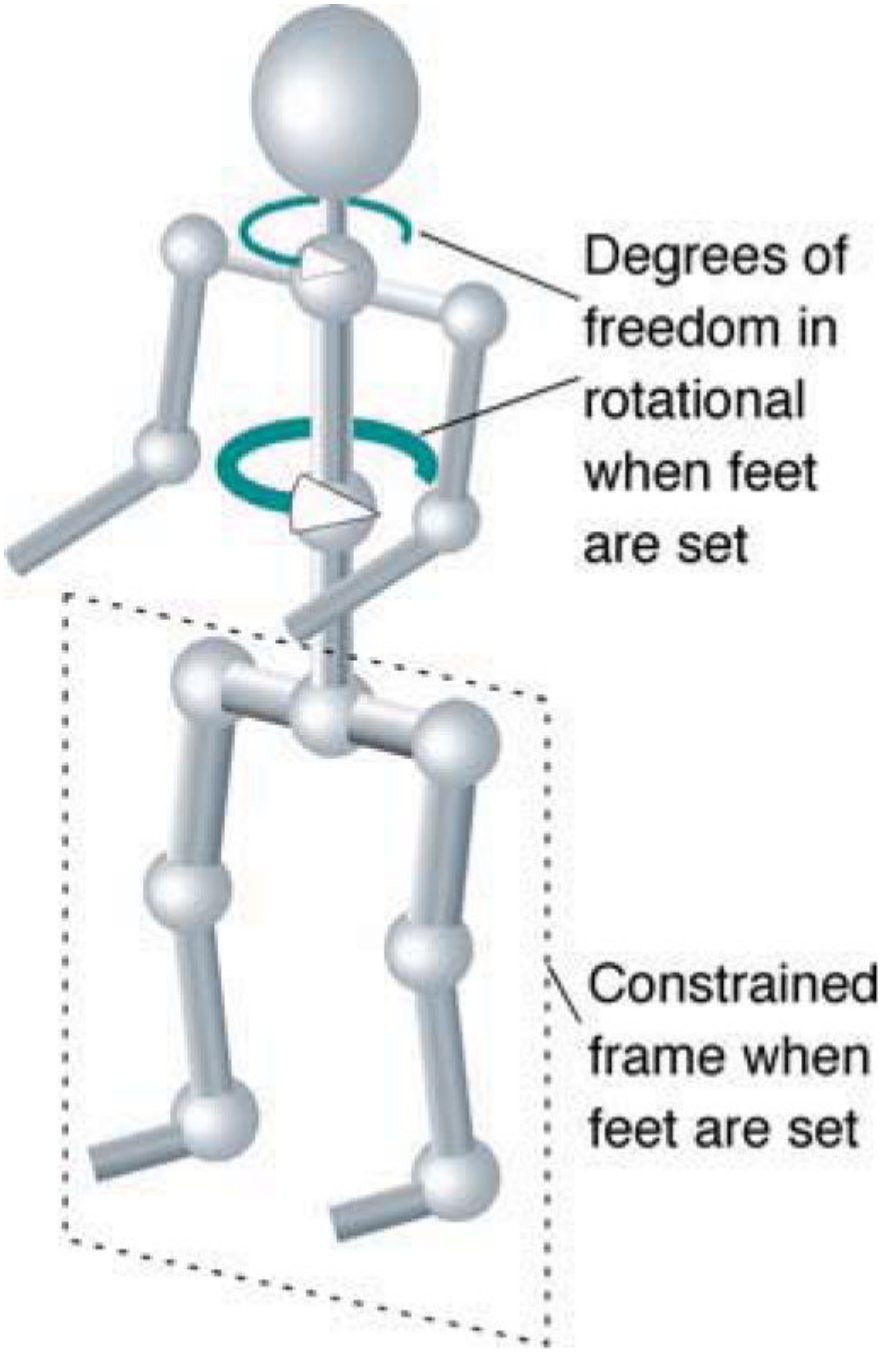

Figure 3.
Degrees of freedom of attention with fixed stance

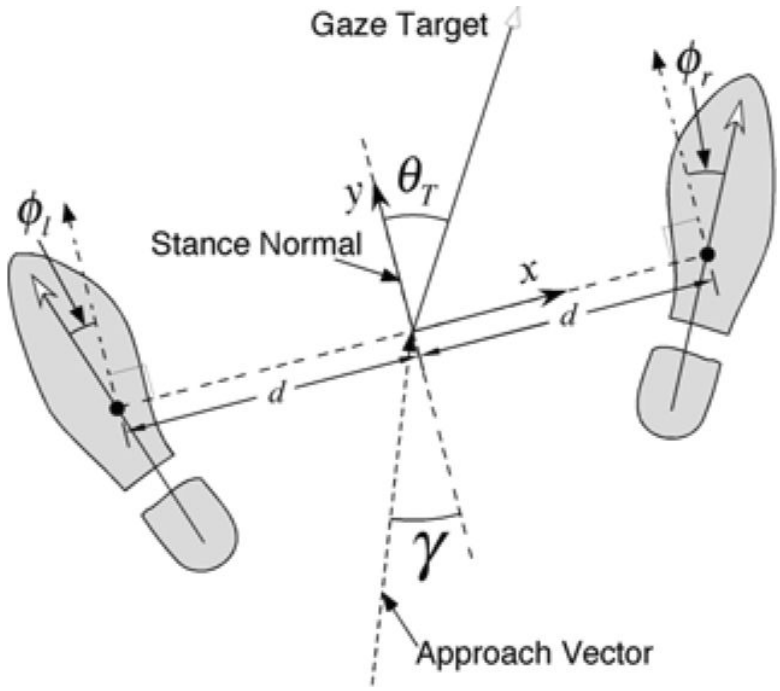

Figure 4.
Subject stance and gaze estimation model

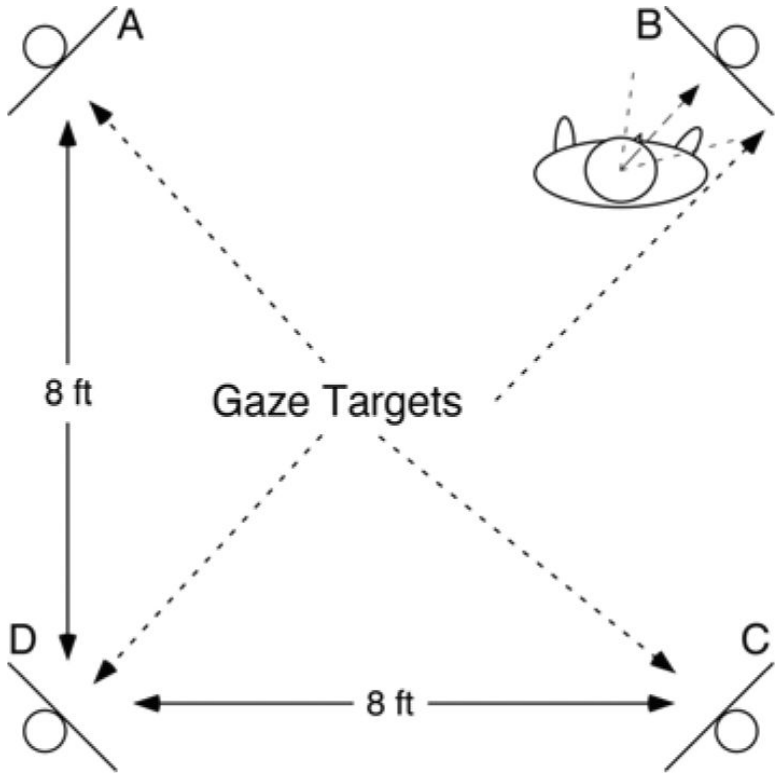

Figure 5.
Plan view of experimental setup

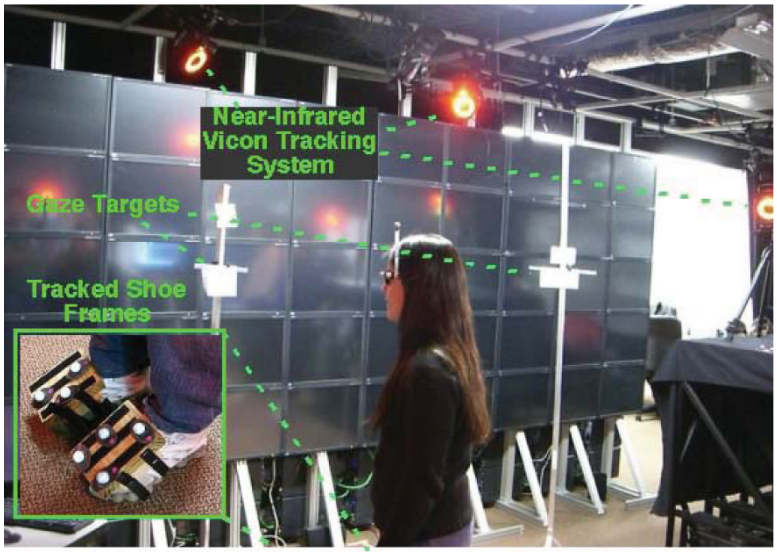

Figure 6.
Experiment setup

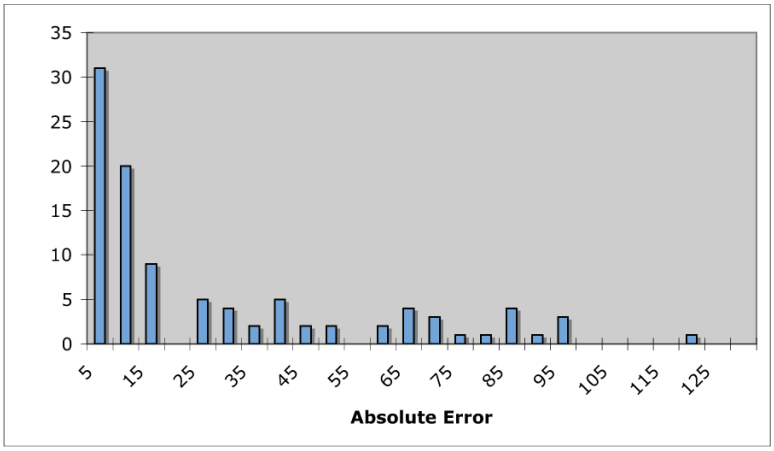

Figure 7.
Histogram of errors of θ_T estimates

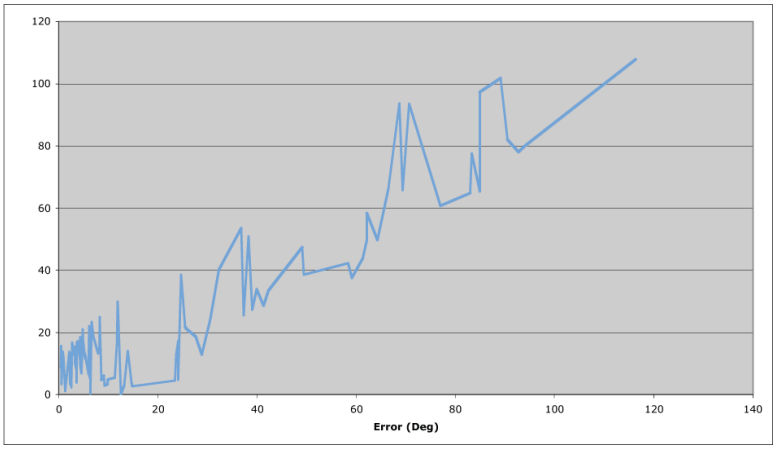

Figure 8.
Error estimates plotted against absolute deflection in test data