Head Orientation and Gaze Direction in Meetings

Rainer Stiefelhagen
Interactive Systems Laboratories
University of Karlsruhe
Germany
stiefel@ira.uka.de

Jie Zhu
Interactive Systems Laboratories
Carnegie Mellon University
Pittsburgh, PA
jzhu@cs.cmu.edu

ABSTRACT

Detecting who is looking at whom during multiparty interaction is useful for various tasks such as meeting analysis. There are two contributing factors in the formation of where a person is looking at : head orientation and eye orientation. In this poster, we present an experiment aimed at evaluating the potential of head orientation estimation in detecting who is looking at whom, because head orientation can be estimated accurately and robustly with non-intrusive methods while eye orientation can not. Experimental results show that head orientation contributes 68.9% on average to the overall gaze direction, and focus of attention estimation based on head orientation alone can get an average accuracy of 88.7% in a meeting application scenario with four participants. We conclude that head orientation is a good indicator of focus of attention in human computer interaction applications.

KEYWORDS

Attention-based Interfaces, Head Pose, Gaze Tracking.

INTRODUCTION

Detecting were a person is looking at could be useful for various human-computer interaction applications, such as multimodal interfaces, intelligent meeting rooms or shared collaborative workspaces. Many researchers have pointed out the close relation of gaze and a person direction of attention. The role of gaze as a signal of attention during multiparty social interaction had been emphasized already in the 70s in the work by Argyle [1] and has also been confirmed in more recent work [2].

In our work we want to track at whom participants in a meeting are looking at. This information could be used to analyze meetings, to identify the addressees of speech acts and potentially to monitor activity in meetings [3,4].

When trying to find out where a person is looking at, a first solution would be to measure the person's gaze and then determine the target based on the measured gaze. Gaze is defined as the direction where the eyes are pointing in space. It is the sum of head orientation and eye orientation. This means we must know both two orientations to figure out the exact gaze direction. The problem of estimating unconstraint head orientation has already been solved by several vision-based approaches. In [4], for example, a neural network approach is applied to simultaneously track four meeting participants' head orientation robustly with poor-quality input images captured by a standalone omni-directional camera. However, unfortunately, there is still no non-intrusive vision-based method to estimate eye orientations of several participants in a meeting. Current eye orientation estimation systems either cumber users with head-mounted equipments, including cameras and special light sources, or set heavy restrictions on users' behavior. However, our goal is to find out the focus of attention target, not the exact gaze point. Given this situation, naturally we want to find out the possibility of using head orientation estimation only, instead of the overall gaze direction to estimate a users' focus of attention.

In this study, we conducted experiments to answer the following questions:

1. How much does head orientation contribute to the overall gaze direction in meetings?
2. How good can we predict users' visual focus of attention based on head orientation?

DATA COLLECTION

The scenario in our experiments is a round-table meeting. There are four participants in the meeting, and we collected a session of data for about ten minutes with each participant. In every session, as the subject, one of the participants wears a head-mounted ISCAN system. The system uses a magnetic pose and position tracking subsystem to track the subject's head position and orientation. Another subsystem uses a head-mounted camera to capture images of the subject's eye. Software provided with this system can estimate and record the following data with a frame rate of 60 Hz: the subject's head position, head orientation, eye orientation, eye blink, and the overall gaze (line of sight) direction. All these estimations have precision of less than one degree, which is far beyond the capability of any current non-intrusive tracking methods. Figure 1 shows an image taken during the data collection. Note that the second right person, who wears the head-mounted equipment, is the subject at this time.

Figure 1. Data collection with eye and head tracking system during a meeting.

CONTRIBUTION OF HEAD ORIENTATION TO GAZE

First, we analyzed the contribution of head orientation and eye orientation to the overall gaze direction along the horizontal axis. On the data from the four participants we found that in 87% of the frames head orientation and eye gaze pointed in the same direction (left or right). In the remaining 13% of the frames, the head orientation is opposite to eye orientation. For the frames in which head orientation and eye gaze point to the same direction, we calculated the contribution of head orientation to the overall line of sight orientation. Since the horizontal component of the line of sight los_x is the sum of horizontal head orientation ho_x and horizontal eye orientation eo_x, the percentage of head orientation to the horizontal direction of gaze is computed as:

$$Head\ contribution = \frac{ho_x}{los_x} \qquad (1)$$

Table 1 summarizes the results of four experiment sessions. From the results, we can see several interesting points:

1. Most of the time, the subjects rotate their head and eye in the same direction to look at their focus of attention.
2. The subjects vary much in their usage of head orientation to change gaze direction: from Subject 2's 53% to Subject 4's 96%, with an average of 68.9%.
3. Even for Subject 2, whose head contribution is the least among the four participants, head orientation still contributes more than half of the overall gaze direction.
4. Eye-blinks (or eye-tracking failures) take about 20% of the frames, which means even for commercial equipments as accurate as the ISCAN system we used, eye orientation, and thus the overall gaze direction cannot be obtained in about a fifth of the time.

Table 1. Eye blinks and contribution of head orientation to the overall gaze direction.

Subject	#Frames	Eye blinks	Same direction	Head contribution
1	36003	25.4%	83.0%	62.0%
2	35994	22.6%	80.2%	53.0%
3	38071	19.2%	91.9%	63.9%
4	35991	19.5%	92.9%	96.7%
Average		21.7%	87.0%	68.9%

From these points we can conclude that head orientation is the most important and sometimes the only measure in gaze direction estimation. A plot of one subject's change of horizontal head orientation, eye orientation, and overall gaze direction over time is shown in Figure 2.

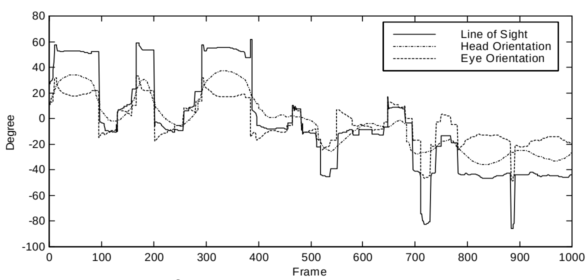

Figure 2. Change of one subject's head, eye, and overall gaze orientations.

ESTIMATING THE GAZE TARGET BASED ON HEAD ORIENTATION ONLY

We approached the second question we proposed before in this particular meeting application: we analyzed how often the real target person (the one the subject is looking at) can be detected correctly based on only the head orientation data.

Labeling Based on Gaze Direction

We labeled each frame as to which target person the subject was looking (focus of attention) at by using the recorded line of sight data.

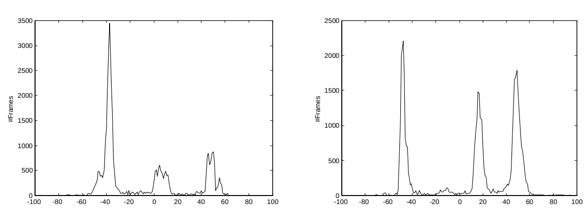

Figure 3. Histogram of horizontal gaze direction of two subjects.

Figure 3 shows the histograms the horizontal gaze direction of two subjects. In the histograms, it can be seen that there are three peaks. These correspond to the directions where the other three participants were seated. We determined these directions using k-means algorithm. Then, for each frame, its focus of attention label was assigned based on the least distance of the actual horizontal line of sight to the three target directions.

Estimation of Focus of Attention from Head Orientation Alone

We used a mix of Gaussian modeling approach to estimate the focus of attention target from only the head orientation data [4], and compared the results with the labels that we obtained from the analysis of the overall gaze direction data. The accuracy of the head orientation only result is summarized in Table 2.

Table 2. Accuracy of focus of attention estimation based on head orientation data alone.

Subject	Accuracy
1	85.7%
2	82.6%
3	93.2%
4	93.2%
Average	88.7%

The accuracy result shows that the focus of attention target can be correctly estimated with only head orientation data in more 82.6% (Subject 2) to 93.2% (Subject 3 and 4) of the frames, with an average of 88.7%. This can be seen as the upper limit of accuracy that we can get in head orientation based focus of attention estimation. We find this percentage very impressive, because as we have seen in the previous analysis, even a state of the art intrusive gaze tracking system will fail to give the correct overall gaze direction due to eye blinks, whereas most head orientation estimation methods, such as the magnetic sensor method in ISCAN and the neural network method in [4], don't have this restriction.

CONCLUSION

We analyzed the head orientation's contribution in overall gaze direction and application potential in focus of attention estimation. Experimental results show that head orientation contributes 68.9% to the overall gaze direction on average, and focus of attention estimation based on head orientation alone can get an average accuracy of 88.7% in a meeting application scenario. Considering the robustness of head orientation estimations, these results show that head orientation estimation can be of great help in focus of attention detections.

REFERENCES

1. Michael Argyle and Mark Cook, Gaze and Mutual Gaze, Cambridge University Press,1976.
2. Vertegaal, R., Slagter, R., van der Veer, G., and Nijholt, A. Eye Gaze Patterns in Conversations: There is More to Conversational Agents than Meets the Eyes, *Proceedins of CHI 2001* (Seattle, WA, April 2001)
3. Stiefelhagen, R., Yang, J., and Waibel, A. Estimating Focus of Attention Based on Gaze and Sound, *Proceedings of PUI 2001* (Orlando, FL, Nov. 2001).
4. Stiefelhagen, R., Yang, J., and Waibel, A., Estimating Focus of Attention for Meeting Indexing, *Proceedings of ACM Multimedia '99, pp. 3-10, 1999.*,