ESANN'2007 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 25-27 April 2007, d-side publi., ISBN 2-930307-07-2.

An Estimation of Response Certainty using Features of Eye-movements

Minoru Nakayama and Yosiyuki Takahasi

CRADLE, Tokyo Institute of Technology
Ookayama, Meguro-ku, Tokyo 152-8552 Japan

Abstract.
To examine the feasibility of estimating the degree of "strength of belief (SOB)" of viewer's responses using support vector machines (SVM) trained with features of gazes, the gazing features were analyzed while subjects reviewed their own responses to multiple choice tasks. Subjects freely reported the certainty of their chosen answers, and these responses were then classified as high and low SOBs. All gazing points of eye-movements were classified into visual areas, or cells, which corresponded with the positions of answers so that training data, consisting of the features and SOB, was produced. A discrimination model for SOB was trained with several combinations of features to see whether performance of a significant level could be obtained. As a result, a trained model with 3 features, which consists of interval time, vertical difference and length between gazes, can provide significant discrimination performance for SOB.

1 Introduction

Common belief is a key concept for communication between humans and with machines and devices; we feel certainness or uncertainness when the degree of shared common belief is strong or weak during communication [1]. An estimation of the certainness may help us to understand human behavior, such as searching for information while browsing the Internet, measuring confidence in online testing processes, or other activities without the need for further surveys or interviews. Using the phenomenon that unstable eye-movement is associated with uncertainness [2], low "strength of belief (SOB)" answers for multiple choice tasks were extracted by analyzing scan paths of eye-movements [3]. In this approach, all scan paths and transitions are analyzed to extract low SOBs, because viewer's uncertainness affects scan-path patterns. Poulamäki et al. have applied the Hidden Markov Model (HMM) to detecting relevant text using eye-movement data during document retrieval experiments [4]. These extractions were based on some of the characteristics of eye-movement while viewers were looking at a text [5]. As the HMM is based on the probabilistic transition function with a finite set of states, it is not easy to create a discrimination model using appropriate transition states, however. Shimodaira et al. have suggested the possibility that discrimination performance using support vector machines (SVM) is comparable to the performance using HMM [6]. The building of a discrimination model using SVM is easier than using HMM when the training data, consisting of examples of features and desired target categories, is provided. In

ESANN'2007 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 25-27 April 2007, d-side publi., ISBN 2-930307-07-2.

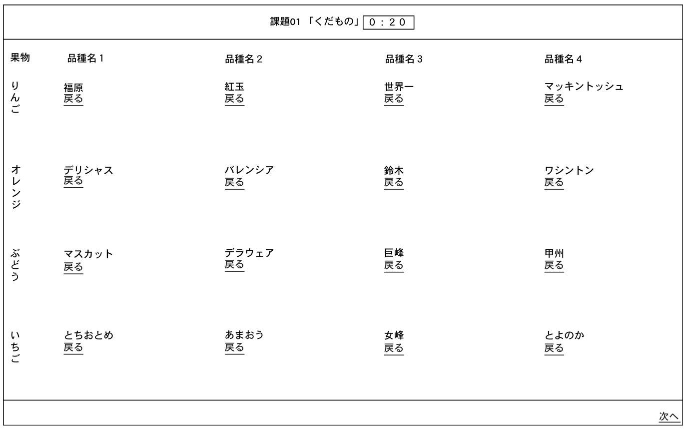

Fig. 1: Answer chosen display in Japanese for the reviewing session.

this paper we examine the feasibility of whether an index of the "strength of belief (SOB)" for answers to multiple choice tasks [3] can be estimated using SVM discrimination of features of eye-movements for gazing points where subjects were looking at answers they had chosen.

2 Method

2.1 Reviewing experiment

A set of multiple choice tasks, arrayed in a 4 by 4 matrix of 4 questions each with 4 answer choices, was prepared as a full screen web page. This task requires the subject to select an appropriate response from the 4 answer choices using a computer mouse. Each column on the web page contains an item, such as a type of fruit, for example. In the 4 rows of multiple choice answers are varieties of each kind of fruit (strawberries, grapes, apples, and oranges), such as Aroma, Delaware, Delicious, and Valencia. Subjects are required to click on a variety if such kind exists for that fruit, such as choosing the variety Valencia for oranges.

The contents of the questions were selected in accordance with the results of a preliminary experiment. The task consisted of an answering session and then a reviewing session. A screen shot of the display for the reviewing session is shown in Figure 1 [1]. Five subjects participated in both sessions. For the first minute of the experiment, answers were selected, followed by the reviewing session in the second minute, where subjects reviewed their own answers without making corrections. A practice session was prepared for all participants to throughly

[1]The figure shows the varieties of fruits chosen by a subject. The rows and the columns correspond to the category of the fruits and questions respectively. In each question, the subject had to classify 4 options for the varieties of the fruits into the correct categories of the fruits. The leftmost column shows the names of the categories.

understand the complete experimental procedure. After the reviewing session, subjects freely noted their own subjective certainty, or SOB, for each answer on a scale between 0 and 100. Score distributions of SOBs depend on subjects, therefore.

In this experiment, we examined the relationship between SOBs and eye movement behaviors in the reviewing session. Each subject reviewed three sets of tasks so that in total 48 SOBs were reported.

2.2 Eye-movement measuring

During the experiment, subject's eye-movements were observed using a video-based eye tracker (nac:EMR-8NL). The task was displayed on a 17 inch LCD monitor positioned 65 cm from the subject. The subject rested his or her head on a chin rest and a small infra-red camera was positioned between the subject and the monitor, 40 cm from the subject. The subject's hands are always free, so that he or she is not restricted during the task. The tracker was calibrated at the beginning of every session. Eye-movement was tracked on a 800 by 600 pixel screen at 60 Hz. Eye-movement data was recorded on a PC as time course data. The tracking data was converted into visual angles according to the distance between the viewer and the display.

Eye-movements were divided into saccades and gazes. A gaze is defined as eye-movement staying within a 0.3 degree visual angle and at a velocity of 3 degrees/sec. or less, using an analyzing program. However it is not easy to measure accurate saccade speeds because the eye tracking rate of this equipment is too low for ballistic eye movement. Therefore, mainly gazing points and gazing time are analyzed in this paper.

To specify the gaze points for question items or chosen answers, all gaze points were classified into 4 cells for question items and 16 cells for chosen responses to the questions, as shown in the grid in Figure 2. The cells were set up in accordance with the positions of the question items and the chosen answers.

3 Features of gazing and training procedure

3.1 Certainness evaluation

According to the relationship between subjective reports of certainty and answer correctness for questions, subject's reports (as SOBs) show the correctness of their answers. Since the answer correctness can be classified as a hit or a miss for this question type, subject's reports of SOB were also divided into two levels using a threshold as the overlap point of two normal distributions. These distributions were estimated from mean and standard deviations (SD) for hit and miss responses by each subject (overall mean(SD), hit:75.0%(7.1); miss:38.4%(8.7)). In comparing average percentages of correct answers, the average for high SOB (64.6%) is significantly higher than for low SOB (35.4%).

These two classes of high and low SOBs are the targeted class for the estimation of subject's responses. Here, each cell for an answer choice is defined

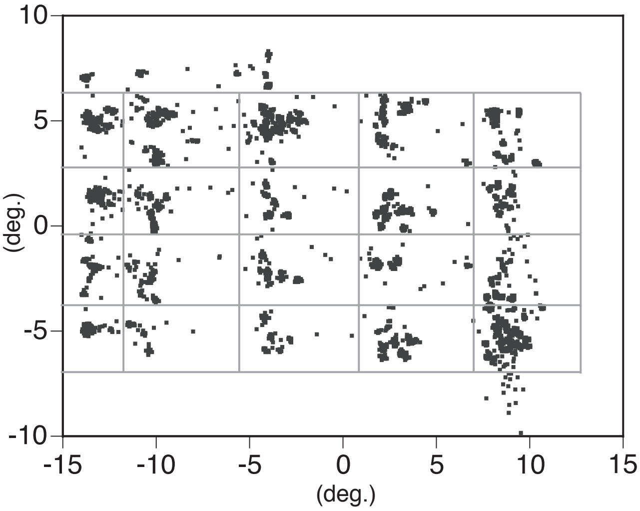

Fig. 2: Gazing point classification into cells of question items and chosen answers.

as $Q_{ij}, (i, j = 1, \ldots, 4, j = 0 : \textit{question item})$, and subjective SOB is defined as $z_{ij} = \{\textit{high SOB}, \textit{low SOB}\}$.

3.2 Features of gazing as eye-movement

Gazing points (p_k) were extracted as a time series using the number k for the above algorithm of eye-movements. Each gazing point has several attributes: horizontal and vertical position in visual angle, and gazing time. Gazing positions may not be affected by SOB because they depend on the task style and layout. To consider the interval characteristics of gazes, we looked at the differences between two successive gazing points. These were the horizontal and vertical differences, and the length of the two points. The results of eye-movement analyses suggest that a category of saccade for eye-movement includes saccadic movements and smooth pursuit movements, because some interval times between the two gazing points are too long for ballistic movements [7]. Therefore, the interval times may show behavior for acquiring further information, and they are affected by the gazing point of the eye when reviewing the preceding answer chosen.

As a result, feature vectors (V_k) for gazing points (p_k) were produced using the following 5 components:

$$V_k(gt(\textit{gazing time}), it(\textit{interval time}), dx, dy, length) \tag{1}$$

The SOB of a cell Q_{ij} which includes a gazing point (p_k), can be noted as $t_k = \pm 1(+1 : \textit{high SOB}, -1 : \textit{low SOB})$ for $p_k \in Q_{ij}$. Therefore, the problem is defined as the prediction of SOB value t_k from the feature vector V_k of the

Table 1: Discrimination performance.

Subject	Scan path	3-dim$_{(\mathrm{it,dx,dy})}$	2-dim$_{(\mathrm{it,dx})}$	2-dim$_{(\mathrm{it,dy})}$	3-dim$_{(\mathrm{it,dy,length})}$
A	**64.6**	50.0	47.9	56.3	58.3
B	**66.7**	47.9	52.1	**77.1**	54.2
C	58.3	**64.6**	60.4	60.4	**64.6**
D	**79.2**	**68.8**	52.1	52.1	**68.8**
E	**75.0**	**77.1**	54.2	47.9	**75.0**
Total	**68.8**	61.7	53.3	61.3	**64.2**

Bold figures show significantly correct percentages ($p < 0.05$).

gazing point in p_{ij}. This classification problem was solved using LIBSVM [8], which was based on training data (V_k, t_k). In the training procedure, the penalty parameter of the error term C for the soft margin, and the γ parameter as a standard deviation of the Gaussian kernel, were optimized. The data were classified completely into the correct SOB classes by a model trained with the data of all subjects at once. Therefore, the performance evaluation was conducted as a Leave-one-out method. To examine estimation performance for a set of answers of a subject, a model was trained with all other data, and estimated SOBs for a set of answers were predicted using the trained model.

After the training, estimated SOB $\hat{t}_k$ for a feature vector V_k was given using a trained SVM model G as $\hat{t}_k = G(V_k)$. The predicted SOB $\hat{z}_{ij}$ for each cell Q_{ij} is derived as a logical sum (OR operation) for all $\hat{t}_k$ for $p_k \in Q_{ij}$. If there was no gazing point in a cell, the SOB for the cell is defined as high SOB. This means the SOB for a chosen answer is high because the subject does not mind the chosen answer.

4 Results and Discussion

According to the scan-path analysis for eye-movements, the total rate of correct responses was 68.8%, and was significantly higher than chance ($p < 0.05$) [3].

After examining the SVM's discrimination performance using appropriate combinations of feature components for the prepared training data, total performances are summarized in Table 1. This table compares performance using scan-path analysis. As a result, only the SVM performance with 3 features (interval time, vertical difference (dy) and length of successive gazing points in reviewing) provide a significant total rate of correct responses (64.2%), and are comparable with the correct percentages using scan path analysis. This result suggests that "strength of belief (SOB)" for a chosen answer can be estimated from features of gazes in eye movement data.

In comparing the performance amongst individual subjects, the performance of each subject depends on the features of their gazes. All performances using the original method, which discriminates low SOB answers using scan-path analysis, are significant, except for one subject. The SVM performance with 3 features

(interval time, horizontal difference (dx), and vertical difference (dy)) are not completely significant, but the performance of three subjects are significant. Detailed analysis suggests that the horizontal difference varies significantly between high and low SOBs. These results demonstrate the hypothesis that SOBs affect eye-movements, but this tendency is not significant for a subject.

Therefore, in looking at horizontal differences (dx), the performance with 2 features (interval time and dx) was examined. All performances were worse than for those with 3 features, but the significance of the vertical differences (dy) is worth examining. The performance with 2 features (interval time and dy) is not completely significant, but performance is significant for one subjects, namely B. This suggests that features of eye-movement are significantly different amongst subjects, and therefore the directional difference of eye movement may not be appropriate as a feature for discrimination of SOB.

According to the results, eye-movement length between successive gazing points was introduced. The performance with 3 features (interval time, dy, and length) is completely significant as well as the performance using scan-path analysis, and most performances for subjects are significant. It is interesting that vertical differences (dy) are significant features for discrimination, although the task may mainly affect horizontal differences.

The SVM discrimination performances of subject A are not significant across combinations of features of eye-movements, however. Features of uncertainness surely affect eye-movement behavior, but this representation may be different between individuals.

The development of a more robust discrimination method with appropriate features may be the subject of our further study.

References

[1] N. Iwahashi, A method for the coupling of belief systems through human-robot language interaction, *Proc. of IEEE ROMAN 2003*, pages 385-390, 2003.

[2] G. Underwood, Eye fixations on pictures of natural scenes: Getting the gist and identifying the components, In Underwood Ed. *Cognitive processes in Eye Guidance*, pages 163-187, Oxford University Press, 2005.

[3] M. Nakayama, Y. Takahasi, An Estimation of Certainty for Multiple Choice Responses using Eye-movement, *Proc. of COGAIN 2006 'Gazing into the Future'*, pages 67-72, 2006.

[4] K. Puolamäki, J. Salojärvi, E. Savia, J. Simola, S. Kaski, Combining Eye Movements and Collaborative Filtering for Proactive Information Retrieval, *Proc. of ACM-SIGIR 2005*, pages 145-153, 2005.

[5] J. Salojärvi, K. Puolamäki, S. Kaski, Relevance Feedback from Eye Movements for Proactve Information Retrieval, *Workshop on Processing Sensory Information for Proactive System*, 2004.

[6] H. Shimodaira, K. Noma, M. Nakai, S. Sagayama, Dynamic Time-Alignment Kernel in Support Vector Machine, *Advances in Neural Information Processing Systems 14*, Vol.2, pages 921-928, 2001.

[7] S.E. Palmer, *Vision Science*, The MIT Press, 2000.

[8] C.C. Chang, C.J. Lin, LIBSVM: A Library for Support Vector Machines, 2001, http://www.csie.ntu.edu.tw/˜cjlin/libsvm