Neurocomputing 128 (2014) 421–432

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Human intention recognition based on eyeball movement pattern and pupil size variation

Young-Min Jang [a], Rammohan Mallipeddi [a], Sangil Lee [b], Ho-Wan Kwak [b], Minho Lee [a,*]

[a] School of Electrical Engineering, Kyungpook National University, 1370 Sankyuk-Dong, Puk-Gu, Taegu 702-701, Republic of Korea
[b] Department of Psychology Kyungpook National University, 1370 Sankyuk-Dong, Puk-Gu, Taegu 702-701, Republic of Korea

ARTICLE INFO

Article history:
Received 6 October 2011
Received in revised form
30 April 2013
Accepted 12 August 2013
Communicated by: Bijaya Ketan Panigrahi
Available online 8 November 2013

Keywords:
Eyeball movement
Human implicit intention
Navigational intent
Informational intent
Intention recognition
Human computer interface and interaction
Pupil size variation

ABSTRACT

To develop an efficient nonverbal human computer interaction system it is important to interpret the user's implicit intention, which is vague. According to cognitive visuo-motor theory, the human eye movements are a rich source of information about the human intention and behavior. According to Beatty's study, a task-evoked pupillary response is a consistent index of the human cognitive load and attention. In this paper, we propose a novel approach for a human's implicit intention recognition based on the eyeball movement pattern and pupil size variation. Based on the Bernard's research, we classify the human's implicit intention during a visual stimulus as informational and navigational intent. In the present study, the navigational intent refers to the human's idea to find some interesting objects in a visual input without a particular goal while the informational intent refers to the human's aspiration to find a particular object of interest. The proposed model utilizes the salient features of the eye such as fixation length, fixation count and pupil size variation as the inputs to classify the human's implicit intention. The experimental results show that the proposed model can achieve plausible recognition performance.

1. Introduction

Human intention recognition is crucial for an efficient human computer interaction. Recently, intention modeling and recognition are being perceived in psychology and cognitive science to create a new paradigm of human computer interface (HCI) and human robot interaction (HRI) [1,2]. Human intention can be explicit or implicit in nature. Generally, humans express their intention explicitly through facial expressions, speech, and hand gesture. In HCI and HRI, the user intention such as “copy this file” or “close the door” can be explicitly conveyed through a keyboard and a computer mouse [3], which can be easily interpreted. However, the human's implicit intention is vague and is difficult to understand. Interpreting the user's implicit intention, which contains valuable information in addition to the explicit intention, is vital in developing an efficient human computer interaction system. Recently, there have been valid attempts to understand a subject's implicit intention based on Electro-Encephalogram (EEG) [4,5], Electrooculogram (EOG) [6], and Electromyogram (EMG) [6].

In humans, the eye movements are the essential motor movements that are controlled by the human cognitive system [7,8]. In other words, the eye movements and its position are not random but directly related with the visual information present in the scene and provide a rich window into the human's sensory processing, intentions and thoughts. Therefore, being a “window to the mind,” the eye and its movements are tightly coupled with human cognitive processes. Therefore, in humans, when viewing a visual scene, different implicit intentions result in different eye movement patterns. Hence, the eyeball movement patterns can be considered as possible factors for recognizing the human's implicit intention. The notion to take advantage of the information present in the eye-gaze leads to the development of efficient eye tracking equipment [9] which attracted many researchers in human-computer interaction [10–13]. In [14], the authors present an eye-typing interface based on eye fixation tracing, an automated method to map the eye movements into a process model prediction using hidden Markov model. In [15], the authors develop an eye based activity recognition (EAR) system using a support vector machine (SVM) and eye movements such as fixations, saccades and eye blinks. In addition to the eyeball movements, the pupil size variation has been studied in relation to cognitive processing and visual information [16]. The pupil size can be used as a measure of the human attention [17].

In this work, we develop a methodology to classify the human's implicit intention in real-world environment based on the eyeball movement pattern and pupil size variation. Based on Bernard's research [18], we categorize human's implicit intention as navigational

* Corresponding author. Tel.: +82 53 950 6436; fax: +82 53 950 5505.
E-mail addresses: ymjang@ee.knu.ac.kr (Y.-M. Jang), mallipeddi.ram@gmail.com (R. Mallipeddi), tarsys@nate.com (S. Lee), kwak@knu.ac.kr (H.-W. Kwak), mholee@gmail.com, mholee@knu.ac.kr (M. Lee).

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.08.008

and informational intent while visualizing the real-world scenes in both in-door and out-door environments by the proposed method. To analyze the human eyeball movement patterns and the pupil size variation, we use the Tobii 1750 eye-tracking system [9]. Fixation length, fixation count and pupil size variation are used as reliable indicators to monitor the human's implicit intention. To classify the user's intention we used two different types of classifiers: (1) nearest neighborhood (NN) and (2) SVM. The proposed system with each of the classifiers shows a reliable performance in classifying the user's implicit intention in both in-door and out-door environment with average recognition accuracy over 85%.

This paper is organized as follows. Section 2 presents a brief literature review on the intention recognition and eye tracking. Section 3 describes the proposed human implicit intention recognition system based on eyeball movement pattern and pupil size variation using a nearest neighborhood classifier. Section 4 presents the experimental results for the proposed intention recognition model on visual images obtained from indoor and outdoor environments. Section 5 concludes the paper.

2. Intention recognition and eye tracking – a review

2.1. Definition of user intention

In cognitive psychology an intention refers to the thoughts one has before producing an action [19]. According the theory of mind [20], human beings have a natural way to represent, predict and interpret the intention expressed explicitly or implicitly by the others. For efficient human computer interaction system it is necessary to understand the human's intention. Intention recognition is a relatively new field that is being widely used in web applications [21] and internet security [22].

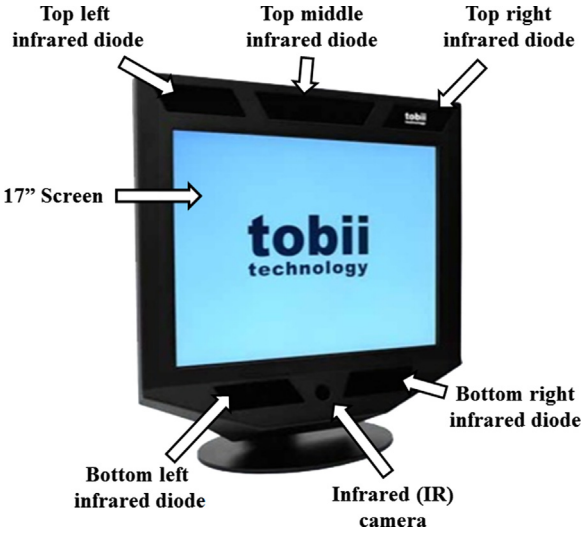

Fig. 1. Tobii 1750 eye tracking system.

During an interaction human beings tend to express their intention explicitly through speech, gesture and facial expression. A significant work is done to understand the human explicit intention [23,24]. However, the explicit expressions alone may not be enough to understand the human intention accurately. Therefore, it is critical to understand the implicit human intention which is vague. Understanding the implicit human intention facilitates the nonverbal human computer interaction which exists naturally in human beings.

The humans' implicit intention can be identified through various biomedical signals which can be acquired from various tissues, organs, or cell systems like the nervous system. Examples include Electro-Encephalogram (EEG) [4,5], Electrooculogram (EOG) [6], and Electromyogram (EMG) [6].

Several researchers [18,25–27] have examined the elements of user intention in web searching using a variety of controlled studies, surveys, and direct observation. In [25], the authors classified the human intention in web searching into (1) *search-oriented browsing* which is the process finding information relevant to a fixed task; (2) *review browsing* which is the process of scanning to find interesting information, and (3) *scan browsing* which is the process of scanning to find information with no reviewing or integration involved. Marchionini [26] articulated similar browsing patterns as directed browsing, semi-directed browsing, and undirected browsing. The authors in [18,27], classify the user intention in web search as (1) informational, (2) navigational and (3) transactional or resource. In the present work, we redefine and present a comprehensive classification of human intention as

- Navigational intent: refers to the human's idea to find some interesting objects in a visual input without a particular goal. In other words, the subject glances over the input scene to get an overall picture.
- Informational intent: refers to the human's aspiration to find a particular object of interest. In other words, the subject searches the input scene for a particular object assuming that it exists.

2.2. Eyeball movement pattern and pupil size variation – correlation to human implicit intention

To visualize an object clearly, it is necessary to move the eyeball to such a position that the object appears on the fovea, a small area

Table 1
Technical specifications of Tobii 1750.

Characteristic	*Technical specifications*
Physical design	• 17 in. TFT monitor, • Max. resolution 1280 × 1024 pixels
Freedom of head-movement ($W \times H \times D$) Field of view of the camera Accuracy/Spatial resolution	• $30 \times 16 \times 20$ cm^3 at 63 cm from tracker. • $21 \times 16 \times 20$ cm^3 at 60 cm from tracker. • 0.5°/0.25°
Data output	• Time stamp • Gaze position relative to stimuli for each eye (X and Y) • Position in camera field of view of each eye (X and Y) • Distance from camera of each eye • Pupil size of each eye • Validity code of each eye