Abstract—Interpretation of human implicit intention is crucial in the development of an efficient nonverbal human computer interaction system. According to cognitive visuo-motor theory, the human eye movements and pupillary responses are rich source of information about the human intention and behavior. It has been observed that under conditions of constant illumination and accommodation, pupil size varies systematically in relation to a variety of physiological and psychological factors, such as level of mental effort. It is well known that pupillary responses could be used to measure the differences in cognitive load under various tasks. In this paper, we try to detect the transition between the different human implicit intents based on the pupil state analysis. In real-world environment, the pupillary response can be influenced by various external factors like intensity and size of the image. To overcome the influence of the external factors, we develop a robust baseline model. The proposed approach detects the transition of the human’s implicit intent from navigational intent to informational intent and vice versa during a visual stimulus. The approach also detects the transition among the different states of the informational intent such as informational intent generation, informational intent maintenance and informational intent disappear.

Keywords- human intention; intent generation; disappearance and maintenance; human computer interface & interaction

I. INTRODUCTION

In cognitive psychology an intention refers to the thoughts one has before producing an action [1]. According the theory of mind [2], 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 [3] and internet security [4].

Human intention can be explicit or implicit in nature. Generally, humans express their intention explicitly through facial expressions, speech, and hand gesture. 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 that is vague to facilitate the nonverbal human computer interaction which exists naturally in human beings. Recently, there have been valid attempts to understand a subject’s implicit intention based on electro-encephalogram (EEG) [5, 6], electrooculogram (EOG) [7], and electromyogram (EMG) [7]. In humans, the eye movements and pupillary response are essential motor movements that are controlled by the human cognitive system [8, 9]. The pupil size and its variation have been studied in relation to cognitive processing and visual information [10].

In this work, we propose an approach to detect the transition of the human’s implicit intention in real-world environment based on the pupillary analysis. The influence of the external parameters such as image size and intensity on the pupillary response is minimized by developing a robust base-line model. We try to detect the transition of human’s implicit intent from navigational intent to informational intent and vice versa while visualizing the real-world scenes in both in-door and out-door environments. The transition among the different states of the navigational intent such as informational intent generation, informational intent maintenance and informational intent disappear is also detected using the proposed approach.

This paper is organized as follows. Section 2 presents a brief literature review on the pupillary response and its significance in cognitive psychology. Section 3 describes the robust base-line model and the proposed system to detect the human implicit intent transition based on pupillary response. Section 4 presents the experimental results for the proposed model in indoor and outdoor visual environments. Section 5 concludes the paper.

II. LITERATURE REVIEW - PUPILLARY RESPONSE AND INTERPRETATION

In humans, the primary function of a pupil is to regulate the amount of light entering the eye by dilation or contraction in response to the changes in the incident illumination, which is referred to as pupillary reflex [11, 12]. However, under the conditions of constant illumination, pupil size varies systematically in relation to a variety of physiological and psychological factors, such as fatigue and level of mental effort [13]. Emotional and sensory events also elicit a pupillary dilation [10] and pupil size variation has been studied in
relation to cognitive processing of visual information [14]. A subject involved in visual-motor operations changes its view point and thinks of different goals after gaining the required information regarding the task it is performing. This is referred to as task shift. Recently, it has been shown that the task shift is correlated with a series of abrupt pupil dilations averaged over a period of time [15]. Hence, the magnitude of task-evoked pupil dilations [16] is a function of the cognitive workload and attention required to perform the task [17, 18]. The correlation between a subject’s attention and its pupil size variation in response to a visual detection is studied in [10, 13, 19, 20]. The emotional content embedded in the visual stimulus might also trigger, and be proportional to, the pupil dilation reaction [21]. Therefore, the pupillary response can be considered as reliable indicator for recognizing the human’s implicit intention and behavior.

III. HUMAN IMPLICIT INTENT TRANSITION DETECTION BASED ON PUPIL STATE ANALYSIS

A. Proposed Model

Pupillary response is a good indicator to detect the change in human implicit intent. However, human pupils may dilate for various reasons, including memory load, cognitive difficulty, pain, and so on [16]. Since there are various causes that could trigger pupillary responses, care must be taken to distinguish the exact cause that activated the response, so that the detection can be accurate and reliable. The most important factors that cause of pupillary response are: 1) pupillary light reflex and 2) processing load required to perform the cognitive task.

In the present study, we utilize a robust base-line model to minimize the effect of pupillary light reflex caused by the varying illumination of the visual stimulus images on the measurement of the pupil size variation. With the help of the robust base-line model, we measure the adaptation time taken by the human pupil to adjust to varying illumination of a visual stimulus. It is known the pupil dilation depends on the information content present in a visual image. Therefore, the effect of pupil size variation due to information content can be removed by limiting the size of the visual stimulus images used in the experiment. In the present study, we experimentally determine the size of the images below which the effect of the pupil size variation due to the processing load is minimal.

Fig. 1 shows the block diagram of the proposed model for recognizing the different human implicit intents and the transition between them based on pupil state analysis. We use an eye tracking system to obtain the pupil state raw data during a visual stimulus.

The areas of interest (AOI) in an image presented as a visual stimulus to the subject is preset and the pupillary response data is analyzed for different implicit intents.

B. Eye Tracking System

To obtain the pupillary raw data, we use the Tobii 1750 eye tracker [22], produced by Tobii Technology incorporation. The details of the eye tracker are presented in Fig. 2.

![Tobii 1750 Eye Tracker](image)

During the experiment, the visual stimuli are displayed on the computer monitor located at 40cm from the human head. During the visual stimulus, the eye tracking system tracks the human’s eye and evaluates the features such as fixation length, fixation count, pupil size and visual scan path etc. of the participant.

C. Robust Base-line Model to Measure Pupil Size Variation

The pupil size variation to a particular visual stimulus depends on the average intensity of the stimulus image. Therefore, to accurately measure the pupil size variation due to a particular factor, it is necessary to control luminance during the course of the experiment. However, controlling the luminance may not always be possible. Therefore, researchers have developed several statistical methods to remove the influence due to changes in luminance. Oliveira et al. [23] performed principle component analysis on the pupil diameter
data explicitly to separate the effect of changes in luminance from stimulus relevance in their study.

In the present study, to accurately measure the pupil size variation, we propose a base-line model which is robust to external factors such as intensity of the image. To generate the model, we use visual stimuli images with intensity values in the range 0 to 255 with a difference of 25 between them as shown in Fig. 3. The pupil size variation in 40 subjects during the visualization of each of the image presented in Fig. 3 is noted.

$$\sum_{j=0}^{n} a_j x^{j+k} = \sum_{i=1}^{M} y_i x_i^k, \quad k = 0, 1, \ldots, n$$

The representative curves \( n = 2 \) obtained are shown in Fig. 4 and are generalized using the expressions in Eq. (6)-(7).

$$y = \frac{x^2}{100} - \frac{17x}{10000} + 3.3$$

$$y = \frac{x^2}{100} - \frac{19x}{10000} + 3.4$$

Then the pupil size during a visual stimulus can be obtained as a percentage of the baseline value irrespective of the stimulus intensity.

**D. Adaptation Time Response of the Pupillary Light Reflex to Intensity Variation**

In human vision system, the adaptation to various levels of darkness and light is attained by the variation in the diameter of the pupil in response to the intensity (luminance) of light that falls on the retina of the eye. The adaptation to the reduced light intensity is termed as dark adaptation while the adaptation to an increased light intensity is known as light adaptation. The constriction and dilation of the pupil in response to a light stimulus during the light and the dark adaptations respectively is referred to as the pupillary light reflex. In general, the pupil contracts extensively to more intense stimuli and completes its contraction in about five seconds before beginning to dilate despite the continuous presentation of the stimulus [24].

To study the effect of the intensity variation of the stimuli on the adaptation time response of the pupillary light reflex, we use the robust base-line model proposed, which is adaptive to the external conditions by considering the human’s psychology factor. The sequence of the experimentation performed is shown in Fig. 5. Fig 5(a) and Fig 5(c) are pupil stability images with time duration of 5 seconds each. Fig 5(b) is visual...
stimulus image (intensity value 1) with time duration of 10sec. The intensity values of the visual stimulus image are varied as shown in Fig. 3.

![Image](image1.png)

Figure 5. Experimental sequence to determine the pupil adaptation time.

Fig. 6 shows the representative curves (30 subjects) obtained for different intensity levels using the Eqs. (1) - (5). From the representative curves, it can be observed that as the intensity level of the stimulus image decreases from 255 to 0, the magnitude of the constriction of the pupil is reduced. The red dotted line in Fig. 6 indicates the time step at which the pupil regains its original size indicating the adaptation to the intensity of the stimulus. Therefore the pupil adaptation time can be obtained as the time length between the black and red lines which indicate the onset of the stimulus and the end of the adaptation.

The adaptation times obtained for the stimulus images of different intensities are plotted in Fig. 6. For the intensity values of 19 and 0, the adaptation time obtained is not clear as we cannot identify the point at which the pupil size is regained. By taking the average of the adaptation times (except the last 2), we obtain the average pupil adaptation time as 3.22 second. The red line in Fig. 6 indicated the average adaptation time obtained by experiment.

![Image](image2.png)

Figure 6. Adaptation time of the pupillary response.

**E. Effect of Stimulus Image Size on Pupil Size Variation**

The information content present in the image increases as the size of the image increases. It has been found that the variation in the pupil size depends on the information content present in the visual stimulus. In order to reduce the effect of the stimulus size in the present work, we would like to find the size of the visual stimulus below which the effect is minimal.

Each subject is presented with 6 seconds visual stimulus which consists of 3 seconds bright pixel with intensity 255 at the top left and 3 seconds dark pixel with intensity 38 at the bottom right of the screen respectively. The intensity of the dark pixel (38) is selected based on the experimental study depending on how easily the participants can differentiate between the background (black) and the dark pixel presented to them. The pixel values and the distance between them in the images employed for the experiment are summarized in Table 1. $(D$: distance between the bright and the dark pixel)

The distance $D$ can be approximated using Eq. (8) as:

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

**TABLE 1. DISTANCE BETWEEN PIXEL VALUES FOR DIFFERENT IMAGES**

<table>
<thead>
<tr>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>D4</th>
<th>D5</th>
<th>D6</th>
<th>D7</th>
<th>D8</th>
</tr>
</thead>
<tbody>
<tr>
<td>X pixel</td>
<td>1122</td>
<td>972</td>
<td>832</td>
<td>682</td>
<td>546</td>
<td>400</td>
<td>256</td>
</tr>
<tr>
<td>Y pixel</td>
<td>806</td>
<td>698</td>
<td>600</td>
<td>500</td>
<td>408</td>
<td>296</td>
<td>190</td>
</tr>
<tr>
<td>Distance</td>
<td>1381</td>
<td>1197</td>
<td>1026</td>
<td>846</td>
<td>682</td>
<td>498</td>
<td>319</td>
</tr>
</tbody>
</table>

The variation in the pupil size to the size of the visual stimulus is presented in Fig. 7.

![Image](image3.png)

Figure 7. Experimental sequence to determine the size of the visual image.

From the figure, it can be observed that for an image size greater than 846, the variation in the pupil size is not consistent. For image size less than 846, the pupil size variation decreases with the image size in Fig. 8. So in our experiment we fix the image size to be below the decision level decided by the experiment as 846.

![Image](image4.png)

Figure 8. Variation in pupil size for different sizes of the visual stimulus.
IV. EXPERIMENTS SETUP AND ANALYSIS

Based on the studies conducted in web search [25, 26], we try to identify the human implicit intent 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 view of the image.

- **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. The informational intent is further identified as intent generation, intent maintenance, and intent disappear.

Experiments were designed to examine the pupil responses associated with different intent conditions under different visual stimuli. The size of the visual stimulus images is taken to be less than the threshold value obtained in Section III.E to avoid the unfavorable affects in the pupillary response.

- **Subjects**: A total of 60 subjects without any prior training or knowledge about the type of images and image presentation are investigated.

- **Apparatus**: Tobii 1750 eye-tracking system [22].

- **Stimuli**: The visual stimuli used in the experiment are collected from the real world and are presented in Fig. 9.

- **Procedure**: The experimental sequence is presented in Fig. 10. Each participant is presented with an image as in Fig. 9 and is asked to perform certain tasks given by the operator. When the subject is asked to identify a particular object, the subject has to respond by pressing the keyboard once the task is completed. The experimental steps are:

  - **Step 1** task: search or locate a particular interesting object in the visual stimulus image
  - **Step 2** visual stimulus image
  - **Step 3** task: search or locate the area corresponding to a particular topic in order to obtain information. After finding the object press keyboard button (information intent generation and disappear)
  - **Step 4** visual stimulus image
  - **Step 5** task: search or locate the area corresponding to a particular topic in order to obtain information. After finding the object press the keyboard and imagine something regarding the object (information intent generation and maintain)
  - **Step 6** visual stimulus image

To generalize, the time period between two consecutive tasks of the experiment is sampled into \( n = 50 \) samples as shown in Fig. 11.

Fig. 12 (a) presents the representative curves of the pupil size of different subjects obtained while viewing the stimulus of first three images of Fig. 9 with intensity values 186 (level-4-blue), 166 (level-5-red) and 138 (level-6-green) respectively.
The representative curves for the images with different intensity values are obtained using the Eqs. (6) and (7) of the base-line model proposed in Section III.C. We also delete the time taken the pupil to adapt to the sudden change in the intensity of the different stimulus images. Therefore, by using the baseline model and the deleting the pupil adaptation time during the change of the visual stimulus, the effect of the image intensity on the pupillary response can be minimized. For analysis, the curves in Fig. 12 (b) are obtained by the approximation of every 8 samples of the sampled data. The gradient of pupil size variation (GPSV) at every 8 samples is presented in Fig. 12 (b).

The approximation of average pupil size (PS) is calculated as shown in Fig. 13.

![Figure 12. Approximation of gradient of pupil size variation (GPSV).](image)

### TABLE 2. PS AND GPSV VALUE DURING THE TASK

<table>
<thead>
<tr>
<th></th>
<th>Navigational intent</th>
<th>Informational intent</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Generation</td>
<td>Disappear</td>
<td>Generation</td>
</tr>
<tr>
<td>PS Original</td>
<td>+102.2 ±3.14</td>
<td>+111.3 ±3.91</td>
<td>+109.1 ±3.97</td>
</tr>
<tr>
<td>Normalization</td>
<td>0</td>
<td>34.73</td>
<td>26.34</td>
</tr>
<tr>
<td>PS Original</td>
<td>-0.07 ±0.09</td>
<td>+0.54 ±0.08</td>
<td>-0.38</td>
</tr>
<tr>
<td>Normalization</td>
<td>16.85</td>
<td>50</td>
<td>0</td>
</tr>
</tbody>
</table>

### V. CONCLUSIONS

Pupillary responses have significant relation to a variety of physiological and psychological factors. In this study, we present a computational approach to detect transition between different types of the human implicit intention based on the pupil analysis. We utilize the pupil factors such as the PS and the GPSV to detect the transition of the human implicit intent from navigational intent to informational intent and vice versa. The proposed approach can also detect the transition of intent between the different states of the informational intent such as informational intent generation, informational intent maintenance and informational intent disappear based on the PS and the GPSV. From the experimental results and the statistical analysis, it has been observed that the PS is a good indicator to distinguish the transition between navigational informational intents while the GPSV is good measure to distinguish the transition between the different states of the informational intent.

### ACKNOWLEDGMENT

This research was supported by the Converging Research Center Program funded by the Ministry of Education, Science and Technology (2011K00659) (30%) and also the Original Technology Research Program for Brain Science through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (2011-0018292) (70%).

### REFERENCES


