MEASUREMENT AND QUANTITATIVE ANALYSIS OF HUMAN VISUAL INTERPOLATION ABILITY FOR PARTIALLY ERASED OBJECTS

YINLAI JIANG AND SHUOYU WANG

Department of Intelligent Mechanical Systems Engineering
Kochi University of Technology
185 Miyanokuchi, Tosayamada-cho, Kami, Kochi 782-8502, Japan
096404j@gs.kochi-tech.ac.jp; wang.shuoyu@kochi-tech.ac.jp

Received September 2007; accepted December 2007

ABSTRACT. The ability of human visual system to efficiently and accurately identify an object based on object parts is defined as the human visual interpolation ability. A letter recognition experiment was designed to evaluate the human visual interpolation ability for partially erased objects in the present study. The experiment results indicated that the features and correlations between different features significantly influence letter identification. An algorithm is proposed to quantitatively evaluate the importance of letter features and the human visual interpolation ability for partially erased letters based on the experiment.

Keywords: Visual interpolation, Letter recognition, Quantitative analysis

1. Introduction. Brain, which interferes with nearly all the function of human body, is considered as the most complex and important part of the human body. However, brain health is not paid as much attention to as health of the other body parts. It is suggested that one possible reason for ignorance of brain health is that brain health is difficult to be checked and monitored. The existing brain imaging technologies, such as EEG, CT and MRI, require high-level professional knowledge and are not feasible to utilize in daily life. Neuropsychological examinations, such as MMSE (Mini Mental State Examination) and FAB (a Frontal Assessment Battery at bedside) [1,2], is impossible to carry out without experienced doctors. Therefore, it is necessary to develop a method, which is easy and convenient for common people to use, to check the brain functions in daily life.

In this study, a method to check the brain functions based on the human visual interpolation ability is being investigated. Visual interpolation ability is defined as the ability that humans can recognize an object based on object parts [3]. Vision provides more than 80% of sensory information that we get from surroundings to the brain [4]. Many parts of the brain, such as visual cortex and frontal association area, are engaged in processing visual information and respond. Therefore, it is considered possible to measure the brain’s visual information process ability and check the health of brain parts associated with vision through proper visual stimuli, regarding the visual system as a “black box”. In this paper, a partially erased letter recognition experiment was presented in order to evaluate the human visual recognition ability for partially erased objects. Based on the results of the experiment, we proposed an algorithm to quantitatively evaluate the importance of letter parts and the human visual interpolation ability with information entropy.

2. Letter Recognition Experiment.
2.1. **Apparatus and materials.** A Hewlett-Packard Compaq nx9000 notebook computer was used for letter recognition experiments. LCD resolution was $1024 \times 768$ pixels, and color was 32 bits. Letters were extracted from the Microsoft Paint program installed in Windows 2000 environment. The font was MSP Gothic, and font size was 72. Letter color was black and the background was white. Letters were presented in the middle of 26 bitmap images of 128-pixels length and width, as shown in Figure 1.

![Figure 1. Letter image](image)

2.2. **Subjects.** Six students at Kochi University of Technology served as volunteer subjects. Subject age was 20-25 years. Vision was normal or corrected to normal.

![Figure 2. Partially erased letter R (L: length, W: width)](image)

2.3. **Stimuli.** A program was developed to partially erase the letters, in which rectangles were used to partially erase the letters simulating the procedure of erasing a letter with a rubber. First, rectangle length, width, position over the letters, and gradient were randomly determined. Black letter pixels covered by the rectangles were then erased. This procedure was continued till the ratio of the number of erased pixels to the number of the black pixels in the original image reached a value of 0.7, 0.8, 0.86, 0.9 or 0.92. Erasure was categorized into three levels according to rectangle size: In the first level, the rectangle size was 1 pixel $\times$ 1 pixel. Thus, letters were erased one pixel by one pixel. In the second level the rectangle length was 4 $\sim$ 8 pixels and width was 2 $\sim$ 4 pixels. In the
third level the rectangle length was 8 ∼ 16 pixels and width was 4 ∼ 8 pixels. The erased letters are illustrated in Figure 2.

2.4. Procedure. In the experiment, firstly, a letter was selected. Secondly, it was erased and displayed in the center of the screen for a short period. And then a dialog was presented for the subject to select an answer and next letter would be erased and displayed. Experiments were completed in 3 days, one erasure level per day. Five levels of erasure ratios (0.7, 0.8, 0.86, 0.9 or 0.92) were set at each erasure level, and the display duration was set to 300 ms, 200 ms and 100 ms at each erasure ratio. Thus the experiment was divided into 45 sections according to three parameters: rectangle size, erasure ratio and display duration. In each section, 26 letters were recognized in a random order.

3. Results and Discussion. Six subjects’ average rates of correct letter identification in the experiment are plotted as a function of the experimental parameters in Figure 3. The three curves show that the rate of correct letter identification decreased as the erasure ratio increased. And the smaller the rectangle, the higher is the correct rate, which illustrates that the correct rate decreased with the increase of rectangle size. The correct rate also decreased with the shortening of display duration.

Humans recognize and remember objects according to their features [5]. When a partially erased letter is recognized, the features of the image are sampled and compared with the memorized features, and a decision is made based on similarity [6]. Therefore, as the erasure ratio increases, fewer and fewer features are left, leading to a decrease in correct rate. Features of an object are generally correlated and the correlation between features plays an important role in recognition [7]. In the present study, the bigger the rectangle, the farther the remaining parts of the erased letter become as shown in Figure 2, which reduced the correlations between the remaining features. Therefore, the correct rate decreased with the increase of rectangle size. It takes time to process visual information across the visual system [8]. Therefore, when the display duration was prolonged, the features that can be sampled during the duration increased and the correct rate was enhanced.

Letters with similar features were frequently mistaken for each other in the experiment. Results for C, G, O, Q, and A are shown in Table 1. Each letter was recognized 270 times, in which C was recognized as C 155 times, as G 17 times, as O 40 times, as Q 12 times,
and as other letters 46 times. A and other letters have few features in common, so A was easy to recognize and was seldom mistaken for other letters.

Table 1. Characters frequently identified incorrectly

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>G</th>
<th>O</th>
<th>Q</th>
<th>A</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>155</td>
<td>17</td>
<td>40</td>
<td>12</td>
<td>0</td>
<td>46</td>
</tr>
<tr>
<td>G</td>
<td>4</td>
<td>208</td>
<td>3</td>
<td>17</td>
<td>0</td>
<td>38</td>
</tr>
<tr>
<td>O</td>
<td>21</td>
<td>14</td>
<td>174</td>
<td>10</td>
<td>0</td>
<td>51</td>
</tr>
<tr>
<td>Q</td>
<td>3</td>
<td>34</td>
<td>8</td>
<td>187</td>
<td>0</td>
<td>38</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>261</td>
<td>9</td>
</tr>
</tbody>
</table>

4. Evaluation of Importance of Letter Features and Human Visual Interpolation Ability. Features and correlations between features play important roles in letter recognition. Features differentiate objects. Therefore, the fewer objects with which a feature is associated, the more important the feature becomes. Pixels are the elementary features used in computers to represent objects. Letters are comprised of black pixels. Taking black pixels as the basic features of a letter, an algorithm based on the idea of information entropy is proposed to calculate the amount of information associated with a single black pixel. Then, based on the information associated with the black pixels, the human visual interpolation ability is quantitatively defined.

Firstly, whether one pixel in a letter image is black is calculated according to the following equation:

$$\phi_\eta(x, y) = \begin{cases} 1, & (x, y) \in \eta \\ 0, & (x, y) \notin \eta \end{cases}, \quad \eta \in A, \ldots, Z$$  \hspace{1cm} (1)

where $0 \leq x \leq 127$, and $0 \leq y \leq 127$. $(x, y)$ is the position of a pixel in the letter image, and $\eta$ is an English letter. As shown in Figure 1, letter images are bitmaps of 128-pixel width and length. If point $(x, y)$ is black in $\eta$’s letter image, $\phi_\eta(x, y)$ is 1. Otherwise, it is 0.

Secondly, the probability that one pixel in the letter image is black is calculated by:

$$p(x, y) = \frac{1}{26} \sum_{\eta=A}^{Z} \phi_\eta(x, y)$$  \hspace{1cm} (2)

Because the stimulus letter image is randomly selected from the 26 English letters with the same probability, the probability $p(x, y)$ equals the product of $1/26$, which is the probability that one letter image is selected, and the number of letters in whose image point $(x, y)$ is black.

Third, the information associated with one black pixel in the letter image is given by:

$$s(x, y) = -\log_2 p(x, y)$$  \hspace{1cm} (3)

Eq.(3) is the equation used in information entropy to measure the amount of self-information contained in a probabilistic event. The smaller this probability is, the more self-information is associated with receiving information that the event indeed occurred. Therefore pixels that belong to fewer letters have a lower probability of being black, and carry more information.

And then, the sum of the information carried by the black pixels in letter $\eta$ is given by:

$$s(\eta) = \sum_{x=0}^{127} \sum_{y=0}^{127} s(x, y) \phi_\eta(x, y)$$  \hspace{1cm} (4)
The sums of information carried by the 26 English letter images used in our experiments are listed in Table 2. Avg denotes the average information associated with one pixel in a letter image (bits/pixel). For example, $s(A) = 2326.6$ bits, and $s(O) = 1404.9$ bits, whereas the number of the black pixels in the $A$ letter image and that of the $O$ letter image are 995 and 1068, and their Avg values are 2.3 bits/pixel and 1.3 bits/pixel, respectively. The $O$ letter image contains more black pixels but less information. This is because several letters, such as $C$, $G$, $Q$ and $D$, are similar to $O$, increasing the probability that the positions of black pixels in $C$ letter image are black and reducing the amount of information associated with the black pixels in the $O$ letter image according to Eq.(2) and Eq.(3). If a letter image has a higher Avg value, it tends to be recognized more accurately.

Then, supposing that $\eta$ is a partially erased letter, the information carried by the black pixels shared by $\theta$ and $\eta$ is given by:

$$s(\eta \cap \theta) = \sum_{x=0}^{127} \sum_{y=0}^{127} s(x, y) \phi_{\eta}(x, y) \phi_{\theta}(x, y)$$  \hspace{1cm} (5)

where $\theta \in A, \ldots, Z$.

If $\theta$ is recognized as $\eta$, the correctness of this recognition is defined as $c(\theta \Rightarrow \eta)$:

$$c(\theta \Rightarrow \eta) = \frac{s(\eta \cap \theta)}{s(\theta)} \quad c \in [0, 1]$$  \hspace{1cm} (6)

Eq.(6) judges whether the black pixels in the partially erased letter image $\theta$ are all included by $\eta$’s letter image. If they are, then $c(\theta \Rightarrow \eta) = 1$, and the recognition is considered correct. Otherwise $c(\theta \Rightarrow \eta) < 1$, and the recognition is considered wrong. If $c(\theta \Rightarrow \eta) = 1$, namely, the recognition is correct, the interpolation ability indicated by this recognition is defined as $h(\theta \Rightarrow \eta)$.

$$h(\theta \Rightarrow \eta) = 1 - \frac{s(\theta)}{s(\eta)}$$  \hspace{1cm} (7)

Otherwise, $h(\theta \Rightarrow \eta) = 0$. The interpolation ability is defined as the percentage of $\eta$’s information which is interpolated by the subject. Examples are shown in Figure 4. For the right image, 90% of its black pixels are erased. If it is recognized as $L$, $c(\theta \Rightarrow L) = 1$ and $h(\theta \Rightarrow L) = 0.85$, which means 85% of $L$’s information is interpolated. The interpolation ability of this recognition is lower than the erasure ratio. This is because black pixels
with relatively less information are erased. Generally, the calculated interpolation ability is around erasure ratio because all the pixels have the same probability to be erased in the erasure algorithm.

Finally, for a section in which there are \( n \) times recognition, the interpolation ability indicated by this section is defined as the average:

\[
h = \frac{1}{n} \sum_{i=1}^{n} h(\theta_i \Rightarrow \eta_i)
\]

Six subjects’ average visual interpolation ability of each section is plotted in Figure 5. As shown in Figure 3, when the rectangle size was 1 pixel \( \times \) 1 pixel, the average correct rates were nearly 100%. So the interpolation abilities of those sections are almost the same to the erasure ratio according to Eq.(7) and Eq.(8). With the increase of rectangle size the correct rate decreased and accordingly the calculated interpolation ability decreases.

This algorithm uses pixels as the elementary features of a letter. Actually, pixels are not the features used by humans to recognize and remember letters. The algorithm will be more accurate if it is based on the basic features of letters such as point, line, angle, shape, and stroke. However, features are composed of features, and the level of elementary features used by humans to recognize object remains to be determined.

5. **Conclusion.** In this study, the human visual interpolation ability for partially erased objects was investigated with an erased letter recognition task. The results showed that
as features and correlations between features in letter images were decreased by erasing, the rate of correct letter identification decreased, which indicated the importance of features and correlations between features in object recognition. An algorithm based on the idea of information entropy was proposed to quantitatively evaluate the importance of letter features and the human visual interpolation ability for partially erased letters. The algorithm is now being used in measurement of the correlation between age and the human visual interpolation ability [9].

REFERENCES


