Evaluating Importance of Information Elements in Graphical Content Using Gaze Points

Ryoji Fukuda
rfukuda@oita-u.ac.jp
Faculty of Engineering,
Oita University
Japan

Junki Iwagami
iwagami@slab.ces.kyutech.ac.jp
Kyushu Institute of Technology
Japan

Takeshi Saitoh
saitoh@ces.kyutech.ac.jp
Kyushu Institute of Technology
Japan

Abstract: *In this study, we evaluate the importance of information elements using gaze points of a person with normal vision. There are many information elements in a mathematical figure. It is necessary to identify important information elements in order to prevent information overload. We focus on elementary geometric figures and express them using directed graphs. non-visual communication is used to analyze the potential of gaze points to express subconscious intentions.*

1. Introduction

There are multiple information elements in most examples of visual content. The human eye-brain system selects some of this information over a short span and obtains the necessary information required in a certain situation. These tasks are often performed subconsciously, and therefore, we take them for granted without understanding their underlying mechanism. However, the visually impaired find it difficult to perform such tasks.

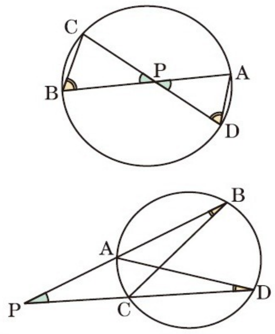
Information is often expressed using directed or undirected graphs. Software tools known as idea processors or outline processors are used to create and arrange ideas or thoughts using graph structures [1], [2]. In these graphs, nodes represent units of thoughts or ideas and edges represent the relations between them. The units of thoughts and their relations act as reminders when we summarize an idea. Consider a figure used in the explanation of a mathematical property. If all the information elements and their relations are expressed in a directed graph, the graph structure will become complicated and difficult to understand. Unnecessary information elements hinder the ability to understand the content. Therefore, a blind student cannot easily understand a tactile graph. Previous studies have attempted to reduce information elements and evaluate the importance of nodes and edges [3]. In this study, we try to evaluate the importance of nodes and edges to reduce the number of information elements by using a gaze point (a point on which a person's eyes are focused) of a person with normal vision. Technologies to obtain gaze points (eye-tracking technologies) are used in various fields, and some of them have succeeded in obtaining human intentions. These are the starting points of our research.

Graphical content is a collection of various information elements. We use a gaze point to select the relevant content. When a person with normal vision reads mathematical content containing figures, he/she obtains information from the figures subconsciously. This make it difficult for him/her to identify an important information element. In this study, we determine the subconscioous intention of a person from his/her gaze points.

We use two cameras to detect a gaze point: an eye camera, which captures the user's eye, and a scene camera, which captures the user's visual field. The captured image are stored and used to estimate the user's gaze points. Then we obtain time-series data of the gaze points. Using the data on positions and directions, we evaluate the importance of the information elements and their relations.

2. Expression Rules

We focus on elementary geometric figures. Information elements are represented by visual shapes that are connected to each other. In non-visual communication, we often use expressions, a tactile output system, a tactile display, a Braille printer, and other devices. However, such methods may not reproduce the figures effectively, and hence, they cannot convey the relevant information. Therefore, we consider verbal expression methods for such content.



2.1 Nodes of Information Graphs

In this study we consider the node categories listed in Table 2.1. We can construct most figures using these concepts; however, they are not sufficient for expressing mathematical properties. For example, a triangle consists of three line segments; however, the concepts of a triangle and a set of three line segments are very different. The concept of area calculation is associated with the triangle, whereas the ratio of edge lengths is associated with a family of edges.

Node category	Sufficient condition for drawing
Point	*x, y* coordinates
Angle	1 point, start and end radian.
Line Segment	2 points.
Circle	center point and radius
Arc	center point, radius, and angle
Triangle	3 points
Quadrangle	4 points

Table 2.1 Categories

To describe various properties, we define different types of elements for each node category, as listed in the table below.

Node category	Type of element
Point	Origin(base point)
Angle	Right angle, Flat angle
Line Segment	Horizontal line, Vertical line
Triangle	Regular, Isosceles, Right angled, Obtuse, Acute
Quadrangle	Regular, Rectangle, Lozenge, Parallelogram, Trapezium

2.2 Edges of Information Graphs

Some concepts expressed in a figure are interlinked owing to the relations between two elements.

We consider the following relation types:

1. Typical parts of figure:
   The center of a circle, vertices of a triangle or quadrangle, vertex point of an angle, start and end point of a line segment, and other similar parts.
2. Included in other, picked up parts:
   A point on the circumference of a circle, a mid-point of an interval, an angle of a triangle or a quadrangle, the base of a perpendicular line, and other similar part.
3. Simple positional relation:
   Inscription and circumscription of a circle and a triangle, intersection of two line segments, a line segment that is a chord of a circle, or a line segment that is tangential to a circle.
4. Delicate positional relation:
   Two orthogonal or parallel line segments, two congruent triangles, the mid point of a line segment, and so on.

Although there are many concepts, we have listed only a few commonly used above.

3. Extraction of Gaze Point

A gaze point is a point on which a person's eyes are focused. Gaze points are used to analyze human intentions. We try to evaluate the importance of information elements using gaze points and the position of a pupil. In this section, we outline the gaze point extraction method. Our system does not support real-time analysis. The experiment starts with a system that stores movie files, and all results are obtained using these data.

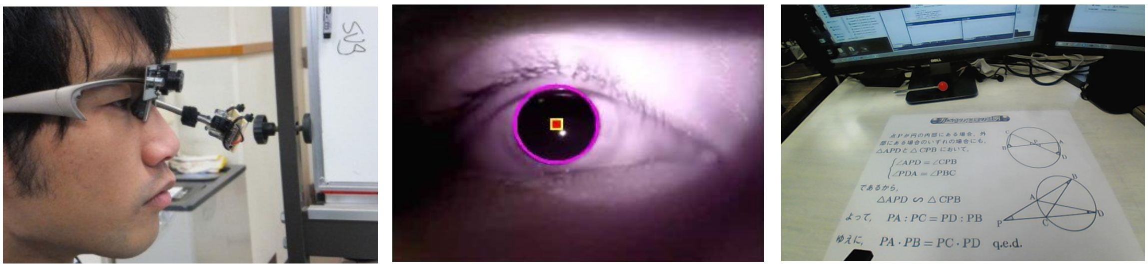


Figure 3.1 Inside-out Camera, Image of Eye Camera, Image of Scene Camera

3.1 Inside-out Camera

The eye-camera and the screen camera are used to obtain images of the eyeball and the target figure, respectively. Together, these two cameras are called an “inside-out camera” (Figure 3.1).

The eye camera captures images of the eyeball under infrared light. Infrared light has very low intensity and is not visible to the naked eye. Hence, it does not affect the user's vision. We can obtain a clear pupil image using infrared light and a camera.

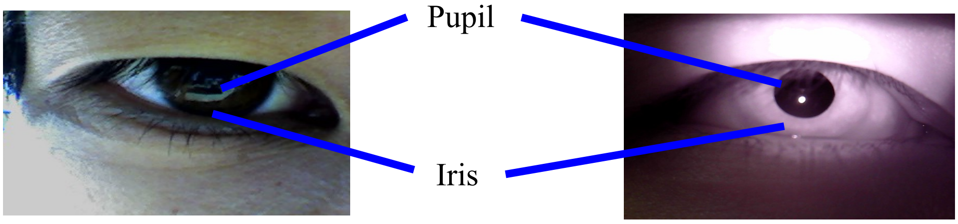


Figure 3.2 Eyeball Images without and with infrared

3.3 Extraction of Pupil

We use the center point of a pupil to estimate a gaze point. For the first step of the estimation, we extract the pupil in an image obtained by the eye-camera by the following procedure.

1. Capture image.
2. Obtain the temporal eye position using the combined separability filter.
3. Obtain edge lines using simple separability filter.
4. Using ellipse estimation, localize the pupil in the image.

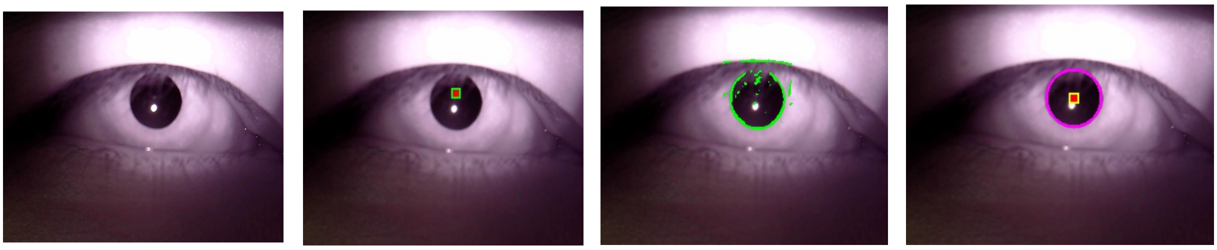


Figure 3.3 Eyeball images (original, temporal position, ellipse estimation, and pupil detection).

3.4 Calibration

We have to adjust the position of the center of the pupil in the image captured by the eye-camera with the gaze point captured by the scene-camera. In our system, the user adjusts these positions using his fingertip in the image captured by the scene-camera. First, the system obtains the two images from two cameras. Next, we carry out pupil and fingertip detection using these images. For the detection of a fingertip, we use the shape and color features until the pupil and fingertip are successfully detected and a pupil and gaze point pair is obtained. After obtaining several pairs, we estimate the relation using a linear trans form using a standard linear regression by the least squares method.

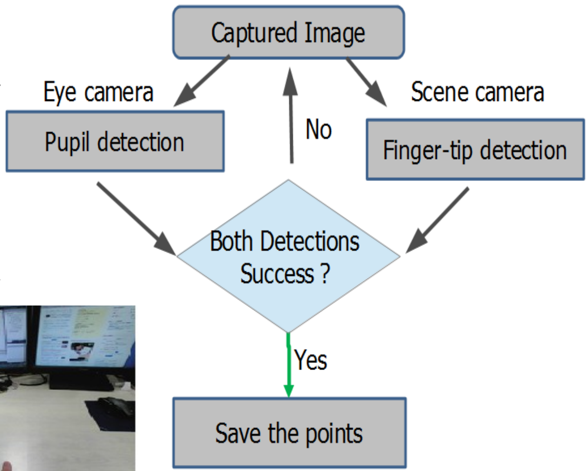


Figure 3.4 Procedure of calibration.

3.5 Extraction Results

To obtain the position of a gaze point, we need two phases: input and analysis. First, we obtain two moving images using the inside-out camera. After that the system estimates the positions of the input gaze points using a linear transform with parameters obtained in the above calibration, In an actual procedure, we need some adjustments that consider several distortions of both picture images. The extraction results are expressed as a sequence of two dimensional points. A target image is given by the scene camera and fixes one image file After extraction, each point is translated to image coordinates. Thus we obtain a time series of the gaze points.

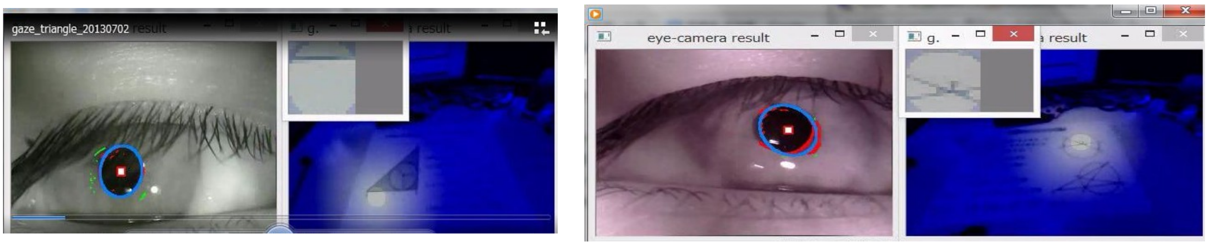


Figure 3.5 Extraction results.

4. Experimental Results

Using the extracted gaze points, we estimate the user's intention. The target figure includes some graphical information elements. We evaluate their importance using the positional relation and movement of the gaze point.

4.1 Description of Graphical Elements

Our target graphical elements are expressed by ellipses or sequences of line segments a in target image. Assume that the photo images are linear transformations of original images, then we can express all of them by an ellipse or a sequence of line segments.

An ellipse is expressed with four arcs. First we select a center point $\vec{c}$, a major axis, and minor axis. Fix two base vectors $\vec{v}_1$, $\vec{v}_2$, and consider an arc expressed by

$$\left\{\vec{c}+\cos(\theta)\vec{v}_1+\sin(\theta)\vec{v}_2 : \theta\in[0,\pi/2]\right\} .$$

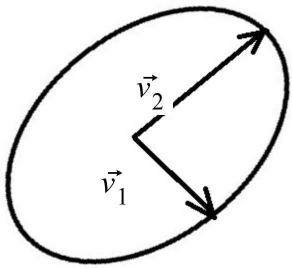


Figure 4.1 Ellipse

Using a center point and 4 points on the ellipse, we obtain 4 arcs, and an ellipse is expressed by these 4 arcs. Thus, we describe an ellipse with 5 points, and a polyline (a piecewise linear line) is also described with a sequence of two dimensional vectors. In the case where the polyline is a closed curve, the start point is same with the end point.

4.2 Evaluation Function for Importance

Let $\{g_j\}_{j\leq N}$ be a sequence of gaze points, p_j be the nearest point to the target element T, from gaze point g_j,

$$d_j=(g_j-g_{j-1})/|g_j-g_{j-1}| ,$$
$$d'_j=(p_j-g_{j-1})/|p_j-g_{j-1}|$$

for each j, where $|v|$ denotes the Euclidean norm of a two dimensional vector v. Then we define an evaluation value as follows.

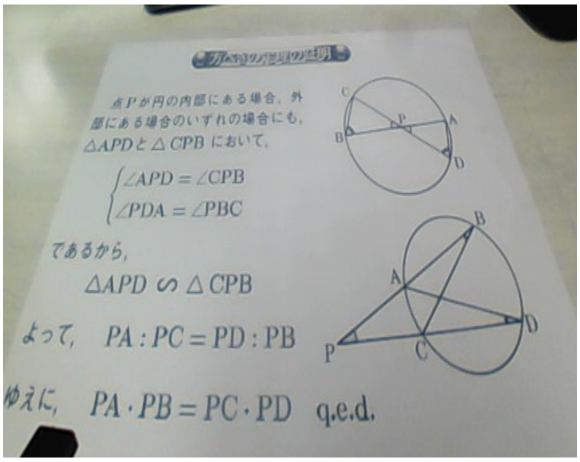
$$ev_j=\sum_{k=j-a}^{j+a}\frac{c_1}{|g_j-p_j|+c_2}+\frac{c_3}{|d_j-d'_j|+c_4}$$

,

where a, c_1, c_2, c_3, and c_4 are experimentally defined constants ($a=20$, $c_1=300$, $c_2=1$, $c_3=1$, $c_4=0.01$).

4.3 Target Figure

Figure 4.2 is our target figure. In this version of the system, gaze point extraction is adjusted by one person. The test subject looked at this figure, under two conditions. For case 1, he compared two triangles. First, he looked at the triangle PAD in the area. After that, he looked at triangle PAD in the area below. For case 2, he checked the condition that a triangle internally touches a circle. In the above area, the triangle ABD internally touches circle $DACB$. Actually, triangle ABD is not complete. Then, the test subject imagines the shape of the triangle.



Case 1: Compare △PAD (above) and △PAD (bellow).

Case 2: Check that △ABD (above) internally touches a circle.

Figure 4.2 Target Figure

4.4 Evaluation Results

Figure 4.3 shows a graph of these evaluations**.** The left-hand graph is for case 1, and we can recognize a change in concentration for two elements in this graph. First the subject concentrated on the upper triangle and after that the target changed to the lower triangle. The right-hand graph is for case2. We can also recognize the change in concentration. Between two peaks of concentration for the ellipse, there is a peak for the triangle. In this graph, the evaluated values for a triangle are smaller than those of an ellipse. We cannot find a clear answer for this phenomenon. These values may be very delicate, and come under the influence of various elements of graphical information.

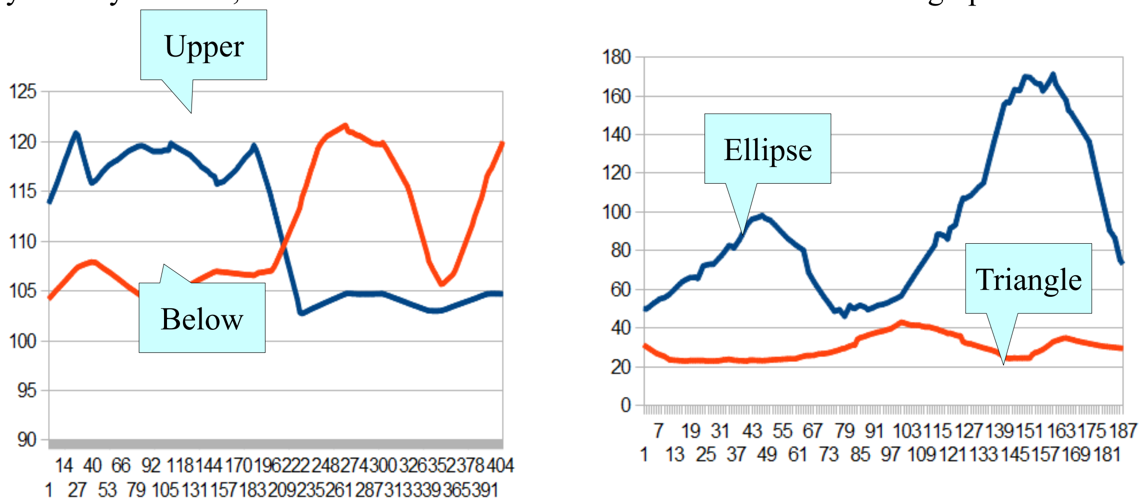


Figure 4.2 Evaluation Results

5. Conclusions

We developed a system to extract gaze points in a figure of elementary geometry and create an evaluation function for the importance of information elements. We will optimize the information structures of graphical contents using this method in the near future, however, this is a first step to complete this task. For the next step we will try to solve the following problems.

1. Improve the gaze contraction system to be used by unspecified persons.
2. Real time treatment of the gaze points.
3. Analyze the total structure of information for a graphical content.
4. Grasp various aspects in the movement of gaze points corresponding to the concentrations for information elements or their relations.
5. Analyze relations with other information, explanations, knowledge database, and other sources.

References

[1] Chik V., Plimmer B., Hosking J., Intelligent mind-mapping Proceedings of the 19th Australasian conference on Computer-Human Interaction: Entertaining User Interfaces ACM New York, NY, USA e2007

[2] Midmap editor: web page: http://freemind.sourceforge.net/wiki/index.php/Main_ Page

[3] Fukuda, R., Expression Rules of Directed Graphs for Non-visual Communication, Lecture Notes in Computer Science, Computers Helping People with Special Needs, 13th International Conference, ICCHP 2012,pp. 182-185