6th International Conference on Multimodal Interfaces, 2004.

EyePrint: Support of Document Browsing with Eye Gaze Trace

Takehiko Ohno
NTT Communication Science Laboratories, NTT Corporation
3-1 Morinosato Wakamiya, Atsugi, Kanagawa, 243-0198, Japan
takehiko@brl.ntt.co.jp

ABSTRACT

Current digital documents provide few traces to help user browsing. This makes document browsing difficult, and we sometimes feel it is hard to keep track of all of the information. To overcome this problem, this paper proposes a method of creating traces on digital documents. The method, called EyePrint, generates a trace from the user's eye gaze in order to support the browsing of digital document. Traces are presented as highlighted areas on a document, which become visual cues for accessing previously visited documents. Traces also become document attributes that can be used to access and search the document. A prototype system that works with a gaze tracking system is developed. The result of a user study confirms the usefulness of the traces in digital document browsing.

Categories and Subject Descriptors: H.5.2 [**Information Systems and Presentation**]:[User Interfaces] Graphical user interfaces (GUI)

General Terms: Design

Keywords: Document browsing, EyePrint, reusability problem, gaze-based interaction, information retrieval, readwear

1. INTRODUCTION

The amount of digital documents available to us is huge and growing. A personal computer allows us to access many types of documents including online manuals, articles, dictionaries, and of course the ubiquitous e-mail. The Internet allows us to access a virtually unlimited amount of information. Printed documents are being replaced more and more by digital documents. Compared to physical documents, digital documents have many advantages in terms of accessibility. Full-text search enables us to mine valuable information from large databases. Hypertext enables us to jump from one place to another with little effort. After using digital documents, it is sometimes really hard to go back to the world of printed documents. Unfortunately, we still find that managing digital document is too difficult.

In this paper, we focus on *the reusability problem* of digital documents. This means that, unlike physical documents, repeated use of a digital document does not increase the accessibility of the document. For example, even if we know that the required information is contained in a previously visited document, we sometimes can not find it. To overcome this problem, technologies that support simple but effective document reuse are required.

ICMI'04, October 13–15, 2004, State College, Pennsylvania, USA.

1.1 Design Motivation

We have the ability to retrieve information even if we saw it a long time ago. Digital documents, however, provide poor reusability. For example, as mentioned before, we can not remember which document contains the previously acquired information. There are three reasons for this.

- Physical insignificance of digital documents. Unlike printed documents, digital documents have few clues that support information retrieval. For example, creases and stains on the pages help the user access the previously visited information. On the other hand, a digital document shows no wear marks even if it is used frequently.
- Human recognition vs. recall. In general, the recognition of information (information recognition) is much easier than retrieving the exact information (information recall). Because digital documents have little physical substances, limited to the text color, size and figures. Therefore, recall must be used instead of recognition to access the information in general.
- Poor ability to judge absolute time. We often find it hard to remember when we saw some piece of information. This complicates information retrieval because absolute time is a strong clue for retrieving previously acquired information, but the user can rarely use it.

One approach to overcome the reusability problem is to add *traces* to the digital documents. Each type of trace is a clue that helps users navigate through and keep track of the document. By adding a trace explicitly by hand, it is easy to access the document later. This paper calls the trace manually created by the user the *active trace*.

A typical example of an active trace is the bookmark function of the WWW browser, which is commonly used to retrieve previously visited pages. Annotation is another example. We often annotate printed documents for later reference [8]. In digital documents, a number of systems have been developed to support this function. Some of them provide a Post-It like function and others allow users to write in free form. There are additional functions that are superior to physical annotation. For example, annotated words can be used for searches, and hyperlinks to matched results can also be created [11]. The problem with these methods are that one has to judge whether the information is valuable or not just after seeing it. This is not easy, and the criteria for making the judgment may

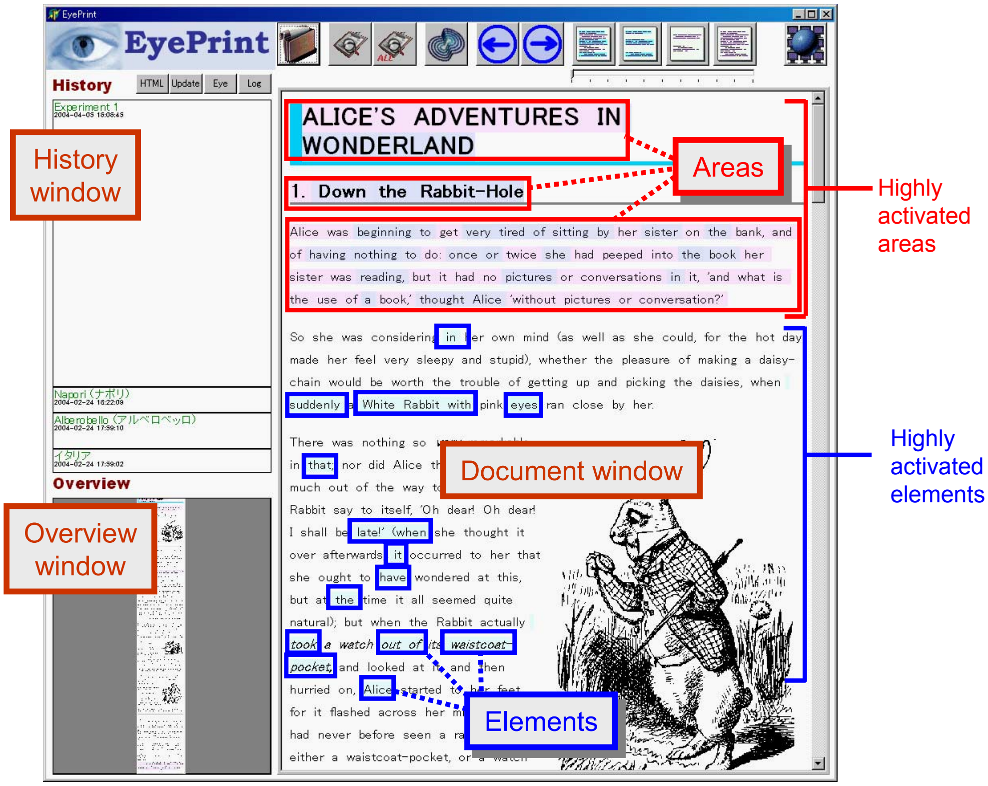


Figure 1: Screen shot of the EyePrint browser. The digital document is displayed in *the document window***; the highly activated areas are highlighted with pink background, and the highly activated elements are highlighted with blue background (the boxes around the highlighted words were added by the author). In the upper-left** *the history window***, the visited documents are displayed in order of the access data. The lower-left** *overview window* **shows a scaled image of the entire document.**

change over time. We do not want to waste effort in setting up the information, since it may or may not necessary in the future.

The passive method is an alternative way of overcoming the reusability problem. In the passive method, traces are automatically generated on the basis of the user's behavior. Therefore, the user does not need to manually add any trace. An example is the history data of a WWW browser. This data is created automatically and helps the user return to a previous page. In some systems, the history helps the user to understand the structure of hyperlinks [3, 5]. Wexelbrat et al. proposed six properties of interaction history [12] and argued the importance of passive history collection. In their system, *Footprints*, the trace emphasizes the information hidden under the link and supports the sharing of tacit knowledge of sites that appears during browsing. The concept of visualization of user activity comes from Hill et al.'s approaches called "Edit Wear" and "Read Wear" [6]. They proposed the basic idea of recording user activity and offering otherwise unavailable information to guide the work by visualizing it. Those passive traces are created from the user's activities. The problem of these approaches is the trace granularity. The minimum unit size of the trace is a page or screen, which is too rough to retrieve the exact information.

This paper proposes a method of adding traces to digital documents, called *EyePrint*. A trace, named *eyeprint*, is generated from the user's eye gaze, and it indicates the manner in which the user viewed or read the document. The name comes from the analogy to fingerprints, which remain on everything one touches. Like fingerprints, eyeprints appear everywhere the user looks, and it helps the user's document browsing task. EyePrint is categorized as a passive method. The most obvious difference between EyePrint and existing methods is that EyePrint uses the user's eye gaze, which indirectly represents the user's reading behavior. We do not normally pay attention to which part of a document we read, so it is hard for us to remember if we have read it before or not. Sometimes we are unaware of skipping a number of paragraphs. EyePrint helps the user find unread areas and supports selective browsing.

This paper presents the architecture of EyePrint, discusses its design requirements, and introduces its functions that support digital document browsing. We also report a user study conducted to observe user behavior with a prototype system. Finally, conclusions and future work are mentioned.

2. EYEPRINT

An eyeprint is a trace generated from the user's eye gaze, which indirectly represents the user's reading behavior. To represent the reading behavior, an eyeprint is defined as the activation of the document. *An element*, which is equal to the word in the document, is the minimum unit of the trace. For each element, *element activation* is calculated from the user's eye gaze data. The wider area unit, which is called *an area*, is equal to the block structure comprising paragraphs and sections. An area consists of several elements. *The document activation* is also defined as the total amount of the area activation contained in the document

These three eyeprints, generated from eye gaze data, are stored in the computer, and are retrieved to support document browsing when the user accesses the document again. Here, we first present two user scenarios to explain EyePrint. We then classify the user's reading behavior, and then describe the eyeprint calculation algorithm and browsing support methods that are based on eyeprint.

2.1 User Scenario

While there are many situations where the reusability problem becomes conspicuous, we prepared two user scenarios as the typical of browsing digital documents. These scenarios were used to design the functions of EyePrint.

Sequential Reading. The first scenario considered to develop the EyePrint browser is to read an online document sequentially with the browser. Reading an online document is a typical browsing task. For example, we read articles, essays, and novels that are provided online.

When a user reads an article, she generally reads it from beginning to end. Sometimes she skips some part of the document unconsciously. In such a case, we are unable to easily confirm what parts of the document we have actually read. If some easy way of showing skipped areas was available, our reading style could be improved. With this function, it is also possible to find the reading termination point without consciously inserting a bookmark.

Collecting Scattering Information. Our second scenario focuses on the information collecting task in which the user refers to information on different pages. Typical examples are to browse an online manual to find the required functions for programming, to browse a travel guide for creating a travel plan, and navigating online catalogs to select some gift.

In those cases, the user visits many documents to gather scattered information. The user sometimes browses many similar documents, and it becomes difficult to relate the document title to the information contained in the document. Sometimes, a search function is provided, but the large number of candidates makes information retrieval difficult.

2.2 Relationship between Eyeprint and Reading Behavior

We use element and area activation to record the user's reading behavior. Here, we mention the relationship between activation and the user's document reading behavior. The user's document reading behavior is categorized into three processes: *scanning*, *skimming*, and *reading* (Fig. 2). Scanning is the fastest process to acquire information; for example, it is used when the user tries to gain an overview of the document, or to locate some specific terms. In the scanning process, the eye moves down the page, so the gaze data appears discontinuously on a few elements.

Skimming is a quick information access process and yields the point of the paragraph. In this process, the user gazes at just part of each paragraph. It is known that user often looks at just the top and bottom of the paragraph.

Reading is the most detailed and slow process. In the reading process, the user reads each sentence word by word, so gaze data appears at every word in every sentence.

The read and skimmed areas provide the vast majority of the information acquired from a digital document; especially the read areas. Scanned areas provide some fragments of information. No information can be acquired from areas other than one of these three. Therefore, we can estimate the amount of information the user acquired from her reading behavior.

These three processes can be distinguished from the distribution of element and area activation. In the read areas, area activation

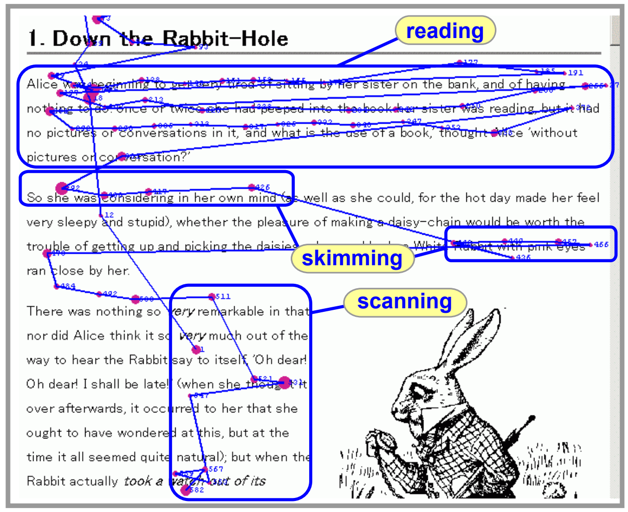


Figure 2: Three types of reading behavior: scanning, skimming, and reading. Circles on the text are fixation points derived from the gaze data. Circle radius is proportional to fixation duration.

is high as is element activation. Skimmed areas have lower area activation and a few elements have high activation. Scanned areas have low area activation but element activation is high. This makes it possible to distinguish the scanned areas from the unviewed areas.

2.3 Document Browsing with EyePrint

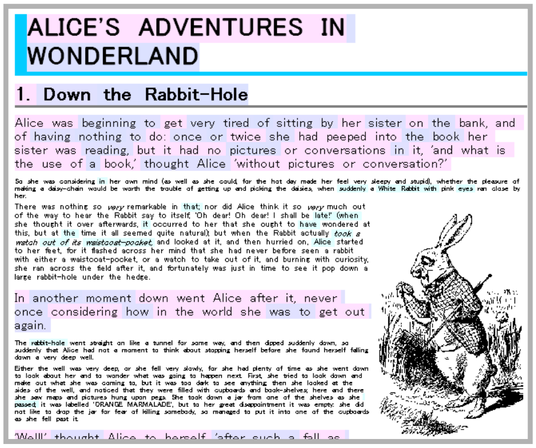
The three types of eyeprints generated from the user's eye gaze are used to support user's document browsing activities. Here, we describe the browsing support methods that use eyeprints.

Visualization of Highly Activated Areas and Elements. The most direct method of using eyeprints is to visualize the highly activated elements and areas by highlighting them. An example is shown in Fig. 1, where the title, the first chapter, and the first paragraph are highlighted with the light pink background. Pink areas represent the areas with high area activation. The intensity of the color depends on the degree of area activation; higher activation areas are darker.

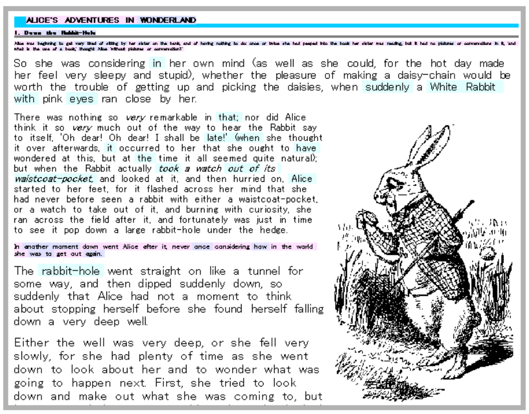
Words highlighted with blue background represent elements with high element activation. Like the areas, the darkness of the color represents the degree of element activation. Some elements in the highly activated areas have a purple background; this means that the element has high element activation as well as high area activation.

Highlighted areas and elements help the user quickly understand which part of the document she has already read. For example, highly activated areas can be quickly found even if the user scrolls down the document rapidly. Another role of visualization is to help the user understand her reading history intuitively. For example, in the first scenario, the user can find the last read text, and can continue reading from that position. The scanned areas, which may contain important information, are also easily found. A paragraph with scattered blue elements should be reread more carefully.

The background color of areas and elements are refreshed only when the document is displayed. Even though area and element activation changes in real time, background color is held unchanged until the user jumps to the next document. It seems to be intuitive to update the background color as the user looks the document; our



(a). High activation areas are displayed in normal size, and low activation areas are displayed in small size.



(b). Low activation areas are displayed in normal size, and high activation areas are displayed in small size.

Figure 3: Document zooming via area activation. Both the text and the figure size are small.

initial system used this design approach. However, we found that updating the background color sometimes attracts the user's eye too much and disturbs the reading process. Offline updating does not disturb the user's reading process.

The entire document is also displayed, small size, in the overview window (Fig. 1). In the window, the highly activated areas and elements are highlighted; the user is able to understand which parts of the document have already been browsed without scrolling the page.

Zooming. Zooming allows a lot of texts to be displayed in the limited screen space or show only the most important part of the document [7]. In EyePrint, the user is able to use two zooming modes in the main window. One mode shrinks low activation areas (Fig. 3(a)), and the other mode shrinks the high activation areas (Fig. 3(b)). The distortion factor is changed by user control. The zoomed document is restored to normal when the user clicks an icon. The zooming function helps the user in both scenarios. In reading an article, already read areas are not necessary, and so can be shrunk. In correcting information, only highly activated areas should be displayed. Moreover, the increment of the available information in the limited screen size will be helpful to locate the required information.

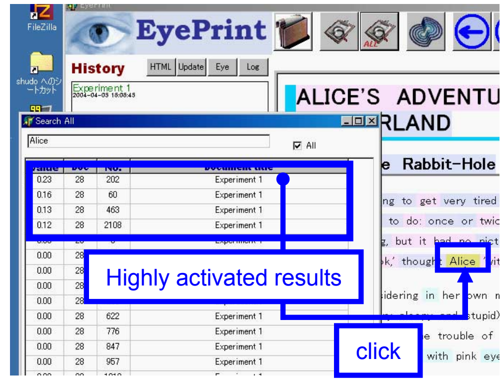


Figure 4: Restricted keyword search. Search results are sorted by element activation order. The corresponding keyword is highlighted when the user clicks a search result.

Restricted Keyword Search. Information in online documents can be located by the keyword search function. Unfortunately, all instances are assigned equal importance so it is difficult to find the desired information. Eyeprints make it possible to restrict the search target to highly activated elements. This function, called *restricted keyword search*, is very convenient if the user knows that he has already read the desired information. Fig. 4 shows an example of restricted keyword search. In this example, there are twenty seven search results, but only four results have high element activation. Therefore, to find the previously read documents, the user just selects from among the four candidates.

History. The history function is a common way to keep track of previously visited documents; the list of documents is usually sorted in order of access date. This function is used by many WWW browsers. In EyePrint, that function is enhanced by the addition of document activation. In the history window (Fig. 1), each box represents a document, and visited documents are ordered by date. The height of the box is proportional to document activation. Documents that fall under a document activation threshold are not displayed. When the user clicks a box, the document is retrieved and displayed.

2.4 Calculation of Eyeprint

We describe the algorithms for calculation of element, area, and document activation.

Element Activation. Element activation represents how strongly the user looked at a word. In general, the document reading process consists of saccadic eye movement and fixation. Most information acquired during reading comes from the fixation points and their surrounding areas. Little information is acquired during saccadic eye movement. Therefore, element activation is defined for the words that have fixation points.

The element activation of element i at time t is defined as $e_i(t)$. When the user looks at element i for Δt, the element activation of element i at time $t+\Delta t$ is defined as

$$e_i(t+\Delta t) = \max\left((e_i(t),\ G\cdot\log(1+B\cdot\Delta t)\right), \quad (1)$$

where G and B are constant values, and Δt is the fixation duration.

Longer fixation durations increase the element activation value. This is because the fixation time is related to the user's cognitive process. There is a tendency to acquire more information if the fixation time is long. The increase in element activation is not linear to Δt. Sometimes fixation duration is much longer than normal. In general, fixation duration in reading lies between 100ms to 300ms, but sometimes it is longer than 500ms, and in rare case, it is longer than 1000ms.

The gazed element is determined from the coordinates of the fixation point. The activation values of its upper and lower elements also increase with $\Delta t/2$. This is because the horizontal length of the boundary area of the element exceeds the vertical length, which may cause error in detecting the gazed element in the vertical direction.

Area Activation. Area activation is derived from the ratio of gazed elements in the area. If the user looks at N_{gaze_i} elements in area i which consists of N_{total_i} elements, the area activation value $a_i(t+\Delta t)$ is given by

$$a_i(t+\Delta t) = \max\left(a_i(t), C \cdot \frac{N_{gaze_i}}{N_{total_i}}\right), \quad (2)$$

where C is a constant, and Δt is the time user spent looking at area i. The area activation is updated if and only if the ratio of gazed elements exceeds the previous ratio.

The number of gazed element N_{gaze_i} is counted by the following algorithm.

$N_{gaze_i} := 0$;
for $i := 1$ **to** $total_i$ **do**
begin
  if $e_i > th$ then $N_{gaze_i} := N_{gaze_i} + 1$;
end;

th is a threshold that eliminates low-activated elements.

Our first version used eye movement direction to calculate the area activation because horizontal eye movement appears during the reading process. However, we found that this choice was not adequate in certain conditions. For example, when an user reads a narrow area repeatedly, area activation became too high. There is no such problem in the current version.

Document Activation

Document activation indirectly represents the amount of information the user acquired. It is calculated from the area activation. If the document activation of the document i in time t is $d_i(t)$, $d_i(t)$ is defined as

$$d_i(t) = \sum_{j=1}^{n} a_j(t), \quad (3)$$

where n is the number of areas that appear in document i.

3. IMPLEMENTATION

To validate the effectiveness of EyePrint, we implemented a prototype system. The goal was to confirm the validity of the activation calculation and to observe the user's document browsing behavior with this system.

The EyePrint browser consists of the three components shown in Fig. 5: a document browser, which displays the digital document on the screen; a document database, which records both digital documents and personal information; and a gaze tracking system, which records the user's eye movement.

3.1 Document Browser

The document browser has a function to display the digital document and allows the user to browse the document via a graphical user interface. The document browser, implemented in Borland Delphi, uses ActiveX components of Microsoft Internet Explorer to display digital documents. It has the ability to display HTML files. The current version does not have a function to download HTML files from the WWW server so all HTML files were initially saved in the host machine. Supporting HTTP protocol transfer is an on-going task.

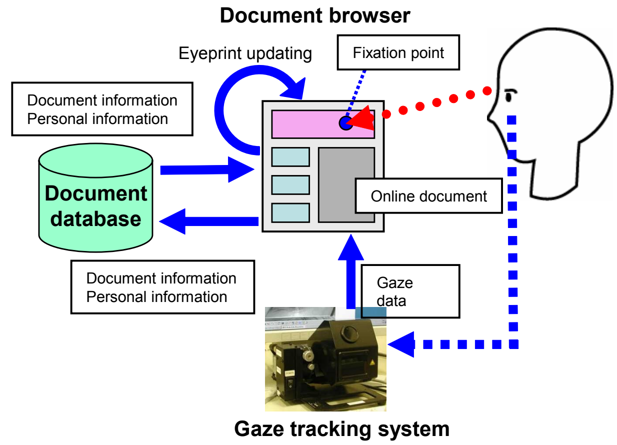


Figure 5: Three components of the EyePrint browser: document browser, document database, and gaze tracking system.

```
<body>
<h1>ALICE'S ADVENTURES IN WONDERLAND</h1>
<h2>1. Down the Rabbit-Hole</h2>
```

(a). Original HTML file (partial).

```
<BODY vLink=blue aLink=blue link=blue>
<H1>
<SPAN id=id0 style="BACKGROUND-COLOR: #ffe1ff" areaid="0" wordid="2667"> ALICE'S </SPAN>
<SPAN id=id1 style="BACKGROUND-COLOR: #ffe1ff" areaid="0" wordid="2668"> ADVENTURES </SPAN>
<SPAN id=id2 style="BACKGROUND-COLOR: #ffe1ff" areaid="0" wordid="2669"> IN </SPAN>
<SPAN id=id3 style="BACKGROUND-COLOR: #dce1ff" areaid="0" wordid="2670"> WONDERLAND </SPAN></H1>
<H2>
<SPAN id=id4 style="BACKGROUND-COLOR: #ffe1ff" areaid="1" wordid="0"> 1. </SPAN>
<SPAN id=id5 style="BACKGROUND-COLOR: #e1e1ff" areaid="1" wordid="2671"> Down </SPAN>
<SPAN id=id6 style="BACKGROUND-COLOR: #dce1ff" areaid="1" wordid="2672"> the </SPAN>
<SPAN id=id7 style="BACKGROUND-COLOR: #dbe1ff" areaid="1" wordid="2673"> Rabbit-Hole </SPAN></H2>
```

(b). Modified HTML file (format is modified manually to improve readability).

Figure 6: HTML modification by the document browser. <SPAN> tags are inserted to segment the paragraph into the elements.

Another role of this component is to determine which element of the document the user is looking at. The Document Object Model (DOM) is used to implement this function. When a HTML file is read from storage by the component, the DOM tree generated from the HTML file is analyzed. For each node of the DOM tree, the lexical analyzer tokenizes the text into a set of elements. This browser is designed for both English and Japanese. When the text is written in English, the element is equal to the word. In case of Japanese, the Japanese Mrphological analyzer 'ChaSen' [1] tokenizes a sentence into a set of phrase, which is defined as the element.

Areas are also defined by the DOM tree structure. Each node of the tree, which is equal to regions separated by HTML tags, is taken as one area. A unique ID is then assigned to each element and area. At this moment, the word information of each element is queried to the database, and a unique word ID is assigned to each unique word.

Next, a new HTML tag that contains information about the element ID, area ID, and word ID is generated. Each element of the HTML is tagged, and the DOM tree is updated using the tags.

The third role of this component is to highlight the highly activated elements and areas, and to change the font size in cooperation with the zooming function. This process is realized by modification of the HTML tags in the DOM tree.

Fig. 6 shows an example of HTML modification. As shown in Fig. 6(b), each word is identified by `<SPAN>` tag, which contains the element identifier, the area identifier, background color and the word identifier. Word identifier, where a unique number is assigned to each different word, is used for restricted text search. When the font size of the element is changed, the character size attribute is also inserted into the tag.

3.2 Document Database

The role of the document database is to record the relation between the digital document and the personal information. When new document information is sent from the document browser, a unique ID is assigned to the document. A unique ID is then assigned for each area and element of the document generated by the document browser. The relation between an area and its elements is also recorded, and is used to retrieve the document from the element ID.

A relational database PostgreSQL 7.3.2 [2] was used to manage both documents and personal data. The database, running on a Linux PC, is connected with the document browser via a TCP/IP socket. When a new document is displayed by the document browser, information about the document including title, filename, and all elements contained in the document are sent to the database and recorded. Personal data of the HTML file including the element, area and document activation is updated when the user jumps to a different HTML file.

3.3 Gaze Tracking System

The role of the gaze tracking system is to determine which element the user looks at. A gaze tracking system developed by the author's group [10] was used to detect the user's eye gaze. The gaze tracking system, which uses the corneal reflection method for gaze detection, is placed under the display. The sampling rate of the system is 30 Hz., and the accuracy is about 0.8 degrees in view angle. The control software and the document browser were run on the same Windows PC. The gaze tracking system limits the user's eye position to a 4cm by 4cm area, so the user has to fix her head while the gaze is being detected [1].

When the user looks at the document, gaze position on the display is calculated. Fixation point is also calculated when the user looks at the same position continuously, and this information is sent to the document browser. The document browser determines the element equivalent to the user's fixation position, and the element and the area activation data are updated.

3.4 PC and Display

A 18.1 inch LCD display EIZO FlexScan L675 was used to display the EyePrint browser. Screen size is 358mm wide by 288mm high, and its resolution is 1280×1024. A PC (AMD Athlon XP 2100+ CPU, 512MB of memory) was used to control the EyePrint browser and the gaze tracking system.

4. USER STUDY

Two user studies were conducted to evaluate the validity of eyeprint calculation and to observe the reading behavior while browsing the digital documents with the EyePrint browser.

[1]This limitation will be solved in our next version.

$2.0 < \alpha$
$1.0 < \alpha \leq 2.0$
$0.0 < \alpha \leq 1.0$
$\alpha = 0.0$

Area No.	Subject A	B	C	D	E	F	G	H
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.22	0.33	0.11	0.00	0.00	0.00	0.22
3	0.00	0.20	0.33	0.13	0.07	0.07	0.00	0.00
4	0.07	0.21	0.23	0.05	0.17	0.13	0.03	0.23
5	0.25	0.20	0.22	0.20	0.24	0.17	0.05	0.22
6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	0.19	0.24	0.27	0.25	0.27	0.18	0.08	0.28
8	0.14	0.27	0.25	0.27	0.29	0.21	0.00	0.17
9	0.20	0.21	0.22	0.23	0.27	0.22	0.00	0.24
10	0.21	0.23	0.26	0.28	0.25	0.24	0.14	0.26
11	0.19	0.23	0.24	0.28	0.30	0.17	0.12	0.29
12	0.20	0.24	0.26	0.22	0.23	0.21	0.16	0.26
13	0.20	0.25	0.24	0.27	0.02	0.13	0.15	0.00
14	0.18	0.26	0.01	0.00	0.00	0.11	0.18	0.00
15	0.17	0.00	0.00	0.00	0.00	0.14	0.11	0.00
16	0.15	0.00	0.00	0.00	0.00	0.14	0.14	0.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	0.19	0.00	0.00	0.00	0.00	0.16	0.20	0.00
19	0.13	0.00	0.00	0.00	0.00	0.25	0.18	0.00
20	0.21	0.00	0.00	0.00	0.00	0.12	0.17	0.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	0.18	0.00	0.00	0.00	0.00	0.10	0.21	0.00
23	0.17	0.00	0.00	0.00	0.00	0.05	0.09	0.00
24	0.00	0.00	0.00	0.00	0.00	0.14	0.19	0.00
25	0.13	0.00	0.00	0.00	0.00	0.05	0.15	0.00
26	0.12	0.00	0.00	0.00	0.00	0.19	0.21	0.00
27	0.00	0.00	0.00	0.00	0.00	0.12	0.03	0.00
28	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

(a). Area activation computed by the EyePrint browser.

$\alpha = 1.0$
$0.5 < \alpha < 1.0$
$0.0 < \alpha \leq 0.5$
$\alpha = 0.0$

Area No.	Subject A	B	C	D	E	F	G	H
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	1.00	1.00	1.00	1.00	1.00	0.75	0.00	1.00
5	1.00	1.00	1.00	1.00	1.00	1.00	0.25	0.75
6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	1.00	1.00	1.00	0.70	1.00	0.35	0.45	0.80
8	1.00	1.00	1.00	1.00	1.00	0.50	1.00	0.50
9	1.00	1.00	1.00	1.00	1.00	0.33	1.00	1.00
10	1.00	1.00	1.00	0.80	1.00	0.90	0.80	1.00
11	1.00	1.00	1.00	1.00	1.00	1.00	0.50	0.75
12	1.00	1.00	0.40	0.10	1.00	0.50	0.20	0.40
13	1.00	1.00	1.00	0.50	1.00	0.70	0.30	0.00
14	1.00	1.00	0.71	0.00	1.00	0.64	0.71	0.00
15	1.00	0.14	0.00	0.00	0.29	0.57	0.14	0.00
16	1.00	0.00	0.00	0.00	0.00	0.75	0.00	0.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	1.00	0.00	0.00	0.00	0.00	1.00	0.57	0.00
19	1.00	0.00	0.00	0.00	0.00	1.00	0.33	0.00
20	1.00	0.00	0.00	0.00	0.00	0.71	0.14	0.00
21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	1.00	0.00	0.00	0.00	0.00	0.65	0.76	0.00
23	1.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
24	1.00	0.00	0.00	0.00	0.00	1.00	0.50	0.00
25	0.13	0.00	0.00	0.00	0.00	0.13	1.00	0.00
26	0.00	0.00	0.00	0.00	0.00	1.00	0.71	0.00
27	0.00	0.00	0.00	0.00	0.00	1.00	0.44	0.00
28	0.00	0.00	0.00	0.00	0.00	0.60	0.00	0.00
29	0.00	0.00	0.00	0.00	0.00	0.75	0.00	0.00
30	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00

(b). Area coverage ratio deraived by the self evaluation of the read area.

Figure 7: Result of the experiment 1. The area activation as calculated by EyePrint browser, and the result of the self evaluation of read area for each subject. Area No. 6, 17 and 21 represent illustrations.

4.1 Experiment 1: Detection of Highly Activated Areas

The motivation of this experiment is to confirm the validity of eyeprint calculation. In the experiment, subjects were asked to read a document, and were then asked to identify that part of the document that described the story. The result was compared with the area activation computed by the EyePrint browser.

Subjects. Eight subjects participated in the experiment (four males and four females); all had normal vision. All subjects used Japanese for communication. Seven subjects used WWW browsers more than once a week, while one subject had little experience in using computers.

Material. One chapter of a novel written in Japanese, 30 areas (24 paragraphs, 3 illustrations, 2 titles and 1 chapter title), was used in the experiment.

Tasks. At the beginning, the gaze tracking system was calibrated. The material was then displayed on the EyePrint browser, and the subjects were asked to read the document for five minutes. During the task, the subjects' head position was fixed using a chin stand. After browsing the document, subjects were asked to write an outline of the story on an answer sheet, (this task was mentioned at the beginning of the experiment). Finally, each subject was given a printed version of the material, and was asked to mark those areas that they remembered as having been read.

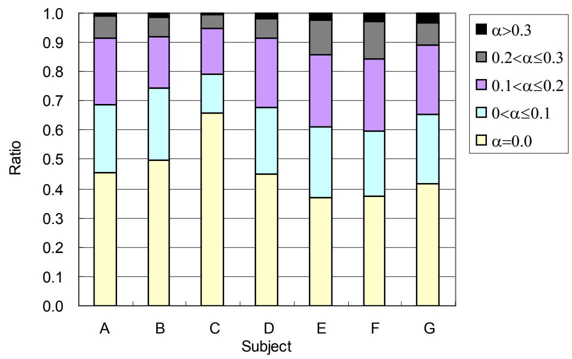


Figure 8: Distribution of area activation for each subject in Experiment 2.

Results. The area activation computed by the EyePrint browser for the each subject is shown in Fig. 7(a). In the thirty areas, area numbers 6, 17, and 21 represent illustrations; , in the current version the area activation is zero even if the user looks at the area. The number of activated areas is different among the subjects because the reading speeds were different. Subject H, who had no activated areas after area No. 12, might be the slowest reader in the experiment. Subjects D and E also had no activated areas after area No. 13. On the other hand, in case of subjects A, F, and G, most areas were activated, but the activation level was lower than those of the other subjects. This result implies that fast readers use the skimming process instead of the reading process.

Fig. 7(b) shows the result of the self highlighting of the document. Area coverage ratio indicates that how many lines the subject marked for each area. If the mark appears for all lines of the area, the coverage ratio is 1.0. Even if the highlight covers only part of the line, it is counted as one line. Therefore, the same coverage ratio does not mean that two subjects marked the same elements. However, it is valuable to understand the tendency of the user's document recognition behavior.

The question we should investigate in this experiment is whether eyeprint represents the user's reading process or not. The correlation between the area activation and the area coverage ratio is 0.78. Between the three fast reading subjects (A, F, and G) and the remaining subjects (B, C, D, E, and H), there is a significant difference in average area activation ($p < .01$), where no activated areas are eliminated. This results implies that it is reasonable to estimate how carefully the user read from area activation. We also found that there was very little conflict, 6.3%, between no activation areas and no recall areas. Those results suggest that recallable areas can be estimated from area activation.

4.2 Experiment 2: Document Navigation Task

The second experiment is a document navigation task, which requires the user to scan many documents for collecting information. The aim of this experiment is to observe browsing behavior in the navigation task with the EyePrint browser.

Subjects. Seven subjects participated in the experiment (four females and three males). They also participated in Experiment 1. They had normal vision. All subjects used WWW browser more than once a week. They were paid for participating in the experiments.

Table 1: Correct answer ratio for the questionnaires in Exp. 2.

	1st stage	2nd stage (group A)	2nd stage (group B)
Mean	.81	.94	.94
S.D.	.22	.07	.12

Materials. We created a homepage of 'Travel information of southern Europe' to use in this experiment. The homepage consisted of 25 pages (1 top page, 4 country information pages, and 20 town information pages). In the 24 pages where the top page is exclusive, there were 12.9 areas in average (9 in minimum, 18 in maximum) and 284.5 elements on average (156 in minimum, 463 in maximum). The pages had, in total, 337 areas and 6971 elements. All pages were written in Japanese.

Tasks. This experiment consisted of two stages. At the beginning, the gaze tracking system was calibrated. Then, in the first stage, they were asked to browse the traveling homepage with the EyePrint browser to find five favorite cities that they would like to visit in their spring vacation. At the beginning of the experiment, they were asked to write down the city name on the answer sheet. They were also asked to fill in a questionnaire that had ten questions about the cities that appeared on the homepage. All answers were also presented on the homepage. For each question, the country name of the city was printed, which was prepared to make the question not too difficult. Typical questions are as follows (original questions used in the experiment were written in Japanese):

- (Portugal) What is the name of the westernmost point of the Eurasian Continents?
- (France) In what century was the beautiful castle standing at the beautiful village in Luberon built?

The browsing time was thirty minutes. The subject had to terminate the task when the time was over, however, if the subject had not finished selecting five cities, an additional five minutes was given to select them. After the first stage, they took a short break.

Next, in the second stage, they were asked to answer a different questionnaire by browsing the traveling homepage with the EyePrint browser. This questionnaire consisted of twenty questions about the cities; ten of them were also presented in the first stage (*group A*), the others were new (*group B*). In this stage, eyeprint was not updated, so that the gaze tracking system was not used.

In the second stage, they used the EyePrint browser with eyeprint visualization. They took thirty minutes to answer the questionnaires.

In both stages, no zoom function or search function were permitted to use for browsing.

We found that one question used in both stages had no correct answer. It was eliminated in the following analysis.

Results. The distribution of area activation for each subject is shown in Fig. 8. On the average, 45.9% areas had no area activation, which implies that they did not see those areas.

The correct answer ratio of the first stage was 0.81, and that of the second stage was 0.94. There was no significant difference between the two groups (Tab. 1). In the second stage, there was no significant difference between two conditions. This result shows that they could find the correct answer even if the number of the question was doubled. It is supposed that, in the second stage, they had enough time to find the answer because only the scanning task was needed to find the answer. On the other hand, in the first stage, they had to select the favorite cities, which would require reading

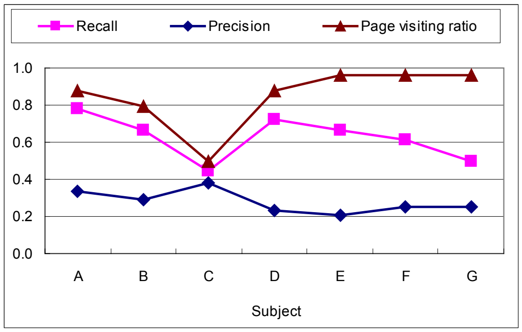


Figure 9: Recall and precision of restricted keyword search. Page visiting ratio (fraction of visiting pages out of 24 pages) is also plotted.

and skimming. For example, in case of subject C, she visited only one third of the areas in the first stage.

Performance of Restricted Keyword Search.

Here, to confirm if eyeprint really helps the user improve search performance, we compared full text search to restricted keyword search. For each question, one to three keywords that were relevant to find the answer, were selected. The total number of keywords was 18.

The result of the restricted keyword search is presented in Fig. 9. The average recall fraction was 0.62 (62 out of 100 truly relevant pages were found), and the average precision was 0.28 (28 out of 100 pages were relevant). In the case of full text search, the recall was 1.0 and the precision was 0.15. Therefore, restricted keyword search improved precision by an average of 79%.

The recall seems to be affected by the page visiting ratio. There is a weak correlation between the precision and the page visiting ratio ($r = 0.48$). It is obvious that if the user has not seen an area, keywords in the area never appear in the search result.

5. DISCUSSION

5.1 Comparison with Other Passive Methods

The difference of EyePrint from other passive methods is that, EyePrint uses the user's eye gaze to observe the document reading behavior. Therefore, the granularity of the trace is much smaller than the other methods. For example, if we use the page visiting information as the trace, the improvement in precision with restricted keyword search is just 18% [2].

Another method to be considered is to display time and appearance area on the screen as the passive trace instead of EyePrint. However, they are too rough to represent the information acquisition process. This is because we often do different activities while displaying the document, so such passive traces provide inaccurate clues for document browsing.

To detect the user's eye gaze, we use an eye gaze tracking system, which is known to be expensive. However, recent advances in computer power and vision research will soon make gaze tracking systems more reasonable. Improvements in technology will also reduce the complexity and difficulty of gaze detection. For example, one system allows the user to move his head freely[4, 9], and another reduces the complicated set up procedure normally needed at the beginning of gaze detection[10]. These improvements will make the gaze tracking system easy-to-use and practical.

5.2 Integration with Other Browsing Methods

The use of eyeprints is not limited to functions implemented in the EyePrint browser. For example, they will be useful in combination with active traces like bookmarks and annotations. This is because eyeprints can act as attributes of the active traces. A bookmark list with document activation will help the user access frequently visited documents. While browsing a document with annotation and area activation, it is possible to understand the degree of past information acquisition, that will help the user judge the importance of the annotation.

6. CONCLUSION AND FUTURE WORK

A document browsing method EyePrint, which generates the trace of document browsing from the user's eye gaze, is proposed. The trace, called eyeprint, is the reflection of the user's reading behavior. It helps the user to find previously visited documents, and also to overview the documents. It also improves the precision of full text keyword search.

An ongoing project is to implement the HTTP protocol for the EyePrint browser to access the WWW, which will enable the browser to generate eyeprints for all documents the user sees through the WWW in everyday life. This extension has the potential to change our document browsing behavior. Observing the impact of EyePrint on everyday browsing remains as a future task.

7. REFERENCES

[1] Japanese morphological analysis system chasen. http://chasen.aist-nara.ac.jp/.
[2] Postgresql. http://www.postgresql.org/.
[3] E. Z. Ayers and J. T. Stasko. Using graphic history in browsing the world wide web. In *Fourth International World Wide Web Conference*, 1995.
[4] D. Beymer and M. Flickner. Eye gaze tracking using an active stereo head. In *Proceedings of Computer Vision and Pattern Recognition*, volume 2, pages 451–458, 2003.
[5] R. R. Hightower, L. T. Ring, J. I. Helfman, B. B. Bederson, and J. D. Hollan. Graphical multiscale web histories: A study of padprints. In *Proceedings of the 9th ACM Conference on Hypertext and Hypermedia: links, objects, time and space - structure in hypermedia systems*, pages 58–65, 1998.
[6] W. C. Hill, J. D. Hollan, D. Wroblewski, and T. McCandless. Edit wear and read wear. In *Proceedings of CHI'92*, pages 3–9, 1992.
[7] K. Hornbæk and E. Frøkjær. Reading patterns and usability in visualizations of electronic documents. *ACM Transactions on Computer-Human Interaction*, 10(2):119–149, 2003.
[8] K. O'Hara and A. Sellen. A comparison of reading paper and on-line documents. In *Proceedings of CHI'97*, pages 335–342, 1997.
[9] T. Ohno and N. Mukawa. A free-head, simple calibration, gaze tracking system that enables gaze-based interaction. In *Proceedings of Eye Tracking Research & Application (ETRA2004)*, pages 115–122, 2004.
[10] T. Ohno, N. Mukawa, and A. Yoshikawa. Freegaze: A gaze tracking system for everyday gaze interaction. In *Proceedings of Eye Tracking Research & Application (ETRA2002)*, pages 125–132, 2002.
[11] B. N. Schilit, G. Golovchinsky, and M. N. Price. Beyond paper: Supporting active reading with free form digital ink annotations. In *Proceedings of CHI'98*, pages 249–256, 1998.
[12] A. Wexelblat and P. Maes. Footprints: History-rich tools for information foraging. In *Proceedings of CHI'99*, pages 270–277, 1999.

[2] In the first stage of experiment 2, subjects visited 85% pages on average. Therefore, the number of search results was reduced by approximately 15%, which contributes to improve the precision by 18% ($1/0.85 = 1.18$).