Document Similarity Judgment for Interactive Document Clustering

[1]*Yasufumi TAKAMA,* [2]*Minghuang CHEN,* [3]*Seiji YAMADA*

[1,2]Graduate School of System Design, Tokyo Metropolitan University, Japan
[3]National Institute of Informatics, Japan
[1,2]6-6 Asahigaoka, Hino, Tokyo 191-005, Japan
[3]2-1-2 Hitotsubashi, Chiyoda, Tokyo 101-8430, Japan

Email: [1]ytakama@sd.tmu.ac.jp, [2]chenminghuang@krectmt3.sd.tmu.ac.jp, [3]seiji@nii.ac.jp

***Abstract*— This paper investigates the task of document similarity judgment for interactive document clustering. We suppose one of the promising approaches for developing next generation of web search engines is to incorporate user feedback mechanism into constrained clustering. As a basis for designing such search engines, it is important to study the interface design that can reduce user' burden of giving feedback to a system. This paper focuses on the task of judging the similarity of two documents as the primitive task for user feedback, and compares 3 types of information to be presented to users: snippet, topic terms, and original text. In particular, snippets suitable for document similarity judgment are proposed, which consist of two kinds of snippets: common snippets showing the common part of documents, and specific snippets showing the difference between documents. An experiment is conducted with 21 test participants, who were asked to judge the similarity of document pairs based on the 3 conditions. Those conditions are compared in terms of judgment time and accuracy with ANOVA and chi-square analysis. The typical judging behavior of the participants is also investigated by an eye-tracking system.**

***Keywords*—snippets, document clustering, eye-tracking.**

I. INTRODUCTION

This paper proposes a snippet generation method that is suitable for document similarity judgment. Recent growth of the Web has brought us huge volume of information, which has already exceeded human capacity of information processing. In order to make use of available information on the Web, performance of web interfaces should be improved. As typical interfaces for accessing information on the Web are currently search engines, next generation of web search engines has been studied.

One approach for improving the performance of search engines is to redesign search engine result pages (SERPs), and document clustering is one of promising approaches for improving the design of SERPs [6]. A clustering method is used to divide the vast number of retrieved pates into groups based on similarity. By displaying the retrieved result based on the clustering result, users can easily access a set of similar web pages by selecting a cluster of interest. The yippy[1] is one of existing search engines that employ clustering approach.

Another approach is to introduce the mechanism of obtaining user feedback. When we are not satisfied with the retrieved result by using existing search engines, we have to modify queries manually. Although some search engines support users by suggesting candidate keywords, it is not an active approach in the sense a user has to select appropriate keywords from the suggestion. As for more active approaches, relevance feedback has been studied in the field of information retrieval [14]. Relevance feedback obtains the result of user's relevance judgment of documents as the feedback for improving the retrieval performance.

In order to realize the next generation of web search engines, this paper focuses on the introduction of feedback mechanism into clustering-based search engines. That is, feedback from a user is used to guide the clustering algorithm so that generated clusters can correspond to what s/he expects.

Introducing feedback information from a user to clustering process can be done by constrained clustering [1]. It uses two kinds of constraints; must-link that indicates two documents connected with the link must be in the same cluster, and cannot-link that indicates two documents having the link must not be in the same cluster. By converting the user feedback into those links, the feedback mechanism can be incorporated into clustering-based search engines.

When obtaining feedback from users, the workload of users providing feedback should be considered. Although much feedback information improves the performance of systems (i.e. search engines), it forces heavy burden on users. In order to solve this tradeoff, the concept of Minimal User Feedback (MUF) [11] has been proposed, which aims at decreasing the cost of a user providing feedback information. The MUF employs two approaches; minimizing the quantity of feedback information and minimizing the cost of generating each of feedback information (i.e. relevance judgment for a single document). A related work with the former approach is active learning [12, 13], in which a learner actively gathers training data effective for learning. The MUF concept is also important

[1] http://clusty.com/

for incorporating feedback mechanism into clustering-based search engines.

This paper addresses the latter approach, i.e., minimizing the cost of generating each of feedback information for constrained clustering. In order to provide must-links and cannot-links for constrained clustering, a user has to judge the similarity between target objects. Therefore, decreasing the cost of judgment is important. The user's task of relevance judgment that is essential for document retrieval including web search has been studied as introduced in Section II-C. However, to our best knowledge, the task of document similarity judgment that is essential for constrained document clustering has not been investigated.

A snippet, which is a fragment of original documents that contains query keywords, has been widely used by most of commercially available search engines. This paper also employs this approach as one of clues for document similarity judgment. We suppose that information identifying the difference and commonality of documents is effective. Therefore, we propose two kinds of snippets: common snippets showing the common part of documents, and specific snippets showing the difference between documents. The effectiveness of original text and terms are also investigated in the paper.

The effectiveness of the above-mentioned clues is evaluated with test participants. They are asked to judge the similarity of two documents. Given a pair of news articles, a participant judges whether those articles relate with the same topic or not. As the clue for judging similarity, one of 3 types of information: original text, the proposed snippets, and terms, is mutually provided.

The judgment accuracy and judgment time are compared between these 3 conditions (text, snippet, term). The result shows judgment time with using the proposed snippets is less than reading original text, and the judgment accuracy with the snippet can be improved with experience. The typical judging behavior of the participants is also investigated by an eye-tracking system.

II. RELATED WORKS

A. Constraint Clustering

Compared with supervised learning such as pattern learning and classification, clustering is called unsupervised learning in the sense that no training data is given. Usually, a clustering problem has no unique result, i.e., there is several possibilities of clustering data. It is also usual that users have background knowledge on the data space. For example, s/he might want some instances to be in the same cluster. Alternatively, s/he might know 2 instances belonging to different clusters. If such information can be considered in clustering process, the result would be improved.

The basic idea of constrained clustering [1, 18, 17] is to incorporate such users' background knowledge or preference on data space into clustering process as constraints. Typical constrained clustering uses two kinds of constraints: must-link that indicates two documents connected with the link must be in the same cluster and cannot-link that indicates two documents having the link must not be in the same cluster.

Various constrained clustering method has been proposed based on several clustering methods, such as COBWEB [17] and K-means [18]. Okabe has proposed a constrained clustering method based on similarity learning [10].

Constrained clustering can be viewed as the method for introducing feedback information from a user to clustering process. By converting the user feedback into those links, the feedback mechanism can be incorporated into constrained clustering. Interactive visual clustering has been proposed based on constrained clustering [5], in which constraints are generated from the results of user's moving instance. However, it does not mention how to support users' task of judging similarity/dissimilarity of instances.

B. Snippet Generation

There are a number of search engines on the Web. Most of them have the similar design of SERPs, which shows a title, URL, and a snippet for each retrieved page. A snippet is an extracted fragment of a retrieved page, which includes the keywords inputted as a query. In a snippet, query keywords are usually highlighted. All of the presented information helps users to make decisions on whether or not to visit the corresponding page. Among them, a snippet is important for users to guess the content of a page.

In order to generate a snippet that is useful for users, it is important that how to pick out the fragments of a web page that reflects the major topics of the page. It is also important to reflect the contents of a page that relate with a user's query. There are various methods for generating snippets. Typical methods calculate the score of sentences in a page according to the query, and the sentences with the highest score are chosen as a snippet. These methods are further divided into various methods based on algorithms and attributes used for score calculation. One of the simple methods employing such a score calculation is based on the terms in a given query. As a query consists of terms that reflect a user's search purpose, a sentence that contains the terms is expected to refer to the topic of interest. Takami [16] has classified snippets into 4 types, for each of which appropriate word weighting scheme, such as TF, TF-DF, and TF-iDF, is proposed.

Li [7] has proposed a different algorithm from the above-mentioned algorithm based on score calculation. It employs the language model to determine the segment of a document that is suitable as a snippet. Xue [19] has proposed image snippets, which extracts representative image from a page based on visual cue-based page segmentation and text-based similarity between query and the segments.

C. Investigating Effect of Snippets

It is expected that snippets are useful for users to find relevant pages without viewing the page itself. The effect of snippets in document retrieval including web search is usually evaluated in terms of judgment time and accuracy [2]. However, when the evaluation should be done in practical situation, such as actual SERPs, the effect of snippets should be discriminated from that of other constituents such as titles and URLs. Recent trend for performing such detailed analysis is to use eye-tracking systems [3, 4, 8, 9, 15].

An eye-tracking system can record user's eye movement on a computer screen. It can record eyes movement and focusing area (area of interests, AOI) on the screen. The focusing time is also recorded. The recorded data are usually visualized as a gaze plot and a heat map. A gaze plot shows the trajectory of focusing point. In the gaze plot, the size of node represents the focusing time. A heat map visualizes the accumulated focusing time for areas on a screen. The heat map can be used for displaying areas of a Web page that are frequently scanned by readers. The most frequently focused areas are highlighted with red.

Such eye-tracking systems have been widely used for studying users' behavior in viewing SERPs. Cutrell [3, 4] has investigated the effect of task type, snippet length, and the position of the best result in the SERP of MSN search engine on users' viewing behavior. Interesting results were obtained, such as that snippet length has different effect between navigational and informational tasks. That is, for informational task, longer snippet improved click accuracy while reducing the task time, but opposite effect was observed for navigational task. This result was explained based on the analysis of eye-tracking record that user performing informational task tended to rely on snippets.

Rodden [15] has explored the relationship between mouse movements and eye movements when performing a search task with using Google. Various interesting patterns are observed, such as keeping the mouse still while reading and using the mouse as a reading aid.

Lorigo et al. [9] has investigated users' search and evaluation behavior based on the analysis of scan path recorded by using eye-tracking system. The results have shown that users tended to make decision on performing new search before viewing entire page of retrieved result, and that they tended to reexamine top 1 and 2 results frequently.

III. Snippet Generation Methods

This section describes the proposed snippet generation method. The point of the method is to generate two types of snippets, common snippets and specific snippets. Common snippets are the fragment of a document that shows common parts of two documents. Specific snippets are the fragment of a document, from which difference between two documents can be confirmed. The proposed algorithm receives a pair of documents as an input, and outputs common and specific snippets for both documents.

The proposed method is divided into 2 processes: topic term extraction and sentence extraction, which are described in the following subsection.

A. Topic Term Extraction

Terms that represent the topic of a document are supposed to work as a clue for judging similarity of documents. In particular, the terms indicating the difference and commonality between documents should be presented to a user. Based on this consideration, we classify the topic terms into common and specific terms, which are extracted with the following two steps:

Step A-1: Extraction of topic terms from a document
Step A-2: Extraction of common and specific terms

In step A-1, given a set of documents D, terms that have high TF-IDF values are extracted as topic terms. Among various definitions of TF and IDF, we employed the following equations.

$$TF_{ij} = \frac{n_{ij}}{\sum_{t_k \in d_j} n_{kj}}, \tag{1}$$

$$IDF_i = \log \frac{|D|}{|\{d \in D \mid t_i \in d\}|}. \tag{2}$$

where n_{ij} is the number of occurrences of the considered term (t_i) in document d_j, and the denominator of (1) is the sum of the number of occurrences of all terms in document d_j. The denominator of (2) shows the number of documents in which the term t_i appears. It should be noted that we calculate TF-IDF score for only the terms appeared at least once in D. To be more exact, all of the terms in D are extracted and the TF-IDF values are calculated except the terms that are contained in a stop word list. In the experiment, we employed the stop word list available from Wikipedia[2].

In step A-2, for a pair of documents that are to be compared, the topic terms that occur in both of the documents are selected as common terms, whereas the terms exclusively occur in either of the documents are selected as specific terms.

B. Sentence Extraction for snippet generation

We suppose that the snippet generated based on the topic can help users to make a judgment easily on whether to read the corresponding documents or not. This merit is supposed to be valid for similarity judgment. Based on this consideration, the proposed method extracts two types of sentences, each of which correspond to common or specific snippets. Given the set of common and specific terms as the result of Sec. III-A, the snippets are generated based on the following steps.

Step B-1: Score calculation for each sentence.
Step B-2: Extraction of a set of snippets.

In step B-1, the score of a sentence in the document is calculated based on the TF-IDF values of specific/common topic terms that are contained in the sentence. Two kinds of score, specific and common scores, are calculated for each sentence.

In step B-2, 2 sentences with the highest common / specific score is selected as a common snippet and a specific snippet, respectively. A sentence with the highest common (specific) score is called common (specific) snippet. For each input document, both of specific and common snippets are generated.

IV. EXPERIMENTAL RESULTS

A. Outline of Experiments

An experiment is conducted for evaluating the effect of the proposed snippets, topic terms, and original text on document similarity judgment. The evaluation is performed based on the comparison with those 3 conditions, in terms of judgment accuracy and time by test subjects. Furthermore, the viewing behaviors of test subjects are examined by using an eye-tracking system. This paper uses T60 from Tobii technology to investigate the user viewing behavior in document similarity judgment.

The task of test participants is to judge the similarity of two documents. Given a pair of documents, they are asked to judge whether those documents relate with the same topic or not.

For the experiment, we implement an experiment system that is written in VB language as ASP pages. Figure 1 shows the screenshot of the experiment system, which can be accessed with ordinary web browsers. On the screen, information about 2 documents is arranged side by side. Figure 1 shows when the proposed snippets are displayed. Common snippets are displayed at upper part of the screen, whereas specific snippets at lower part. In the snippets, topic terms (common/specific terms) are highlighted with red. There are 3 buttons at the lower part of the screen: same topic, unclear, and different topic. Each test participant judges whether the topic of the 2 document is the same or not, and selects the corresponding button. In the topic term conditions, terms are displayed in the same layout as Fig. 1. In the case of original text condition, original text of a document is displayed at the left side of the screen, and another document is displayed at the right side of the screen.

The documents and topics are selected from Reuter Test Collection[3]. It includes 21578 documents (articles) with 135 topics. In the experiment, we prepared the document set by selecting a few topics and randomly picking up the corresponding documents. If the documents of different topics are obviously different, test participants can judge the similarity of documents without carefully reading displayed information. Therefore, topics that are to be used in the experiment should relate with each other. Based on this consideration, we selected the following 3 topics, which are overlapping each other, i.e., several documents belong to two of those topics: Coffee, Cocoa, and Corn.

Fourteen documents that belong to only one of those topics are collected from each topic, and total 42 documents are used in the experiment. In the experiment, the document pair to be displayed is extracted randomly from the 42 documents.

[2] http://en.wikipedia.org/wiki/Stop_words
[3] http://www.daviddlewis.com/resources/testcollections/

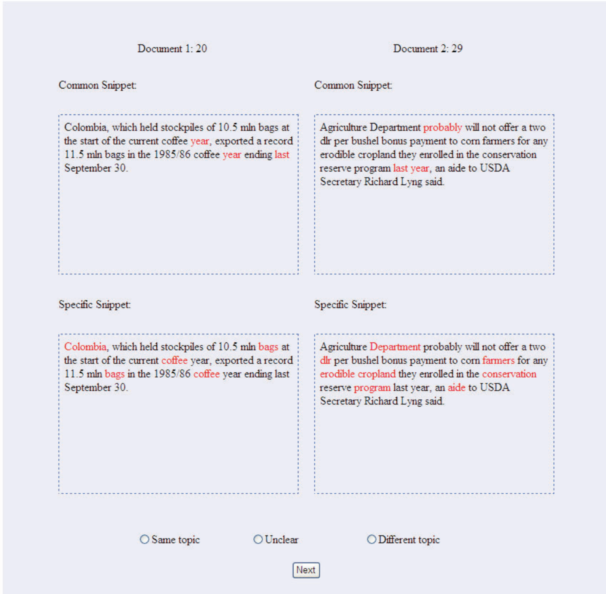


Fig. 1. A screenshot of the experiment system

B. Results: Performance of similarity judgment

In the experiment, 21 participants including graduate / undergraduate students and researchers in engineering field took part in the experiment. A participant is asked to judge 3 document pairs for each condition (snippet, original text, term). A pair of documents is generated randomly from the document set containing 42 documents as noted in Sec. 4-A.

The experimental results are analyzed by two approaches. First, the performance of participants' similarity judgment is compared among 3 conditions in terms of judgment accuracy and judgment time. After that, the behavior of the participants in performing the experiment is analyzed based on the record of the eye-tracking system.

When a user performs similarity judgment in an actual application, the judgment has to be repeated several times. Therefore, user's adaptability is one of important factors for evaluating the effect of the proposed snippets. In the experiment, a participant judged the similarity of documents 3 times for each condition. In order to consider the participants' adaptability, we separately analyzed the results in the 1st and the 3rd trials.

The difference of snippet, text (original text), and term in judgment time of the 1st trial is analyzed using one-factor repeated measures analysis of variance (ANOVA). As a result, we found statistically significant differences in the mean judgment time among snippet, text, and term ($F(2,40)=16.52$, $P=5.9E-06$). In the case of the 3rd trial, the assumption of equality of variance was rejected. Therefore, we conducted nonparametric test (Kruskal Wallis Test) and confirmed the difference is statistically significant ($\chi^2=7.023$, $P=0.030$).

In order to examine the effectiveness of each condition, multiple comparison tests are conducted. In the case of the 1st trial, assumption of equality of variance could not be rejected. Therefore, Tukey's test and Fisher's LSD is used. In the paper,

* and ** indicate the significant level of 5% and 1%, respectively. Table I summarizes the result. The result shows that the participants could judge the similarity of document using snippet and terms more quickly than reading original text.

TABLE I. Multiple comparison test

	1st trial		3rd trial	
	p-value (Tukey)	p-value (LSD)	p-value (Scheffe)	p-value (Steel-Dwass)
Snippet vs. Text	0.0042**	0.0015**	0.5393	0.4264
Snippet vs. Term	0.9287	0.7152	0.3112	0.2206
Text vs. Term	0.0014**	0.0005**	0.0307*	0.0341*

As already noted, the assumption of equality of variance was rejected in the case of the 3rd trial. Therefore, we conducted nonparametric tests: Scheffe test and Stell-Dwass test, of which the results are shown in Table I. In this case, only the difference between text and term is statistically significant. From both the results of the 1st and the 3rd trials, it is shown that providing terms is more effective in terms of the time cost of similarity judgment.

A chi-square analysis on the number of correct answers and mistakes is performed in order to investigate the effect of type of information on the accuracy of similarity judgment. Although we found no significant difference among 3 types of information in both of the 1st (χ^2=0.382, P=0.826) and the 3rd trials (χ^2=4.672, P=0.097), we can see the tendency that the difference in the 3rd trial tends to be larger than the 1st trial. In particular, it was observed that the judgment accuracy in the 3rd trial when snippet is provided gets improved from the 1st trial. That is, in the snippet condition, 11 among 21 participants judged correctly in the 1st trial, and 16 participants in the 3rd trial. However, the judgment accuracy when terms are provided is low both in the 1st and the 3rd trials, i.e., the number of participants who judged correctly was 10 in the 1st trial and 11 in the 3rd trial, respectively.

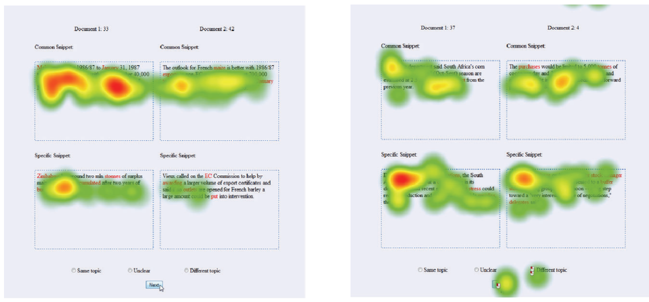
We suppose this result indicates that snippets and original text are easier for the participants to adapt than terms. In order to consider it in more detail, typical judging behavior of the participants is investigated in the following subsection.

C. *Analysis of viewing behavior in similarity judgment*

In order to confirm the results obtained in Sec. IV-B, the viewing behaviors of the test participants are analyzed. By analyzing the eye-tracking data, we found that the position and topic of documents affected participants' viewing behavior. We also found the viewing behavior that is specific to the term condition. It is noted that all of the 3 trials for each condition per participant are analyzed in this subsection. Among the 21 participants, 6 participants were excluded from the analysis, because of stability problems with the eye tracking, leaving us with 15 participants.

Table II shows the number of participants who focused on the corresponding (common or specific) area more frequently than another areas. It is counted in two categories: when documents of the same topic are presented and those of different topic are presented. Figure 2 shows the typical heat maps of viewing snippets, when 2 documents belong to same topic (Fig. 2(a)) and different topic (Fig. 2(b)), respectively.

Table II shows that in the same topic condition, participants more frequently looked at common snippet / terms than specific ones. On the contrary, in the different topic condition, specific snippet / terms were more frequently focused by participants than common ones. From the result, it is supposed that participants first assume whether given documents relate with same topic or not, then examine the assumption by reading the corresponding information. That is, participants would carefully read specific terms / snippet when they assume the documents relate with different topic from each other. This result suggests participants need different kind of information according to assigned task. In that sense, separately providing common/specific information (snippets and terms) as proposed in the paper is effective for supporting the task of similarity judgment.



(a) (b)

Fig. 2. Heat map of viewing snippets when (a) 2 documents belong to the same topic, (b) different topic

TABLE II. Effect of topic on viewing behavior

Topic	Snippet		Term	
	Common	Specific	Common	Specific
Same topic [persons]	22	5	19	12
Different topic [persons]	7	11	4	10

(a) (b)

Fig. 3. Comparison of gaze plot between (a) term condition, and (b) snippet condition

As noted in IV-B, the test participant could get used to the proposed snippet through experience. On the other hand, there was no significant difference of judgment accuracy between

1st and 3rd trials in the case of term condition. In order to examine the reason, viewing behaviors between term and snippet condition are compared. Figure 3 (a) shows a gaze plot of viewing behavior that is specific to the term condition. The figure shows that the participant frequently switches AOIs between left-hand and right-hand document areas. This behavior was frequently observed when terms were presented. On the contrary, as shown in Fig. 3 (b), such a switching behavior was less observed when text or snippet was presented.

In order to investigate this tendency in more detail, we calculated the number of "switches"; a switch occurs when participants changed AOI from left-hand to right-hand document areas or vice versa. Table III shows the average switching frequency that is counted per condition. It shows that participants most frequently switched when terms are presented.

TABLE III. Average frequency of switching AOIs

	Text	Snippet	Term
Average frequency	4.49	7.06	12.34
STDEV	3.28	3.54	6.82

It is supposed that each term can be examined independently, whereas participants have to read a sequence of terms when snippets or original text is displayed. We think such a difference between terms and other conditions is one of the reasons that participants reading terms can judge the similarity of document more quickly than other condition as shown in Sec. IV-B. On the other hand, the improvement of judgment accuracy with experience was not observed when terms are presented. This implies that the context formed by a sequence of terms is important to grasp the contents of a document.

V. CONCLUSION

This paper investigates the task of document similarity judgment for interactive document clustering. Three types of clues, snippet, topic terms, and original text are compared regarding judgment time and accuracy. The result shows that the test participants could judge the similarity of documents with using the snippets and terms faster than reading original text. It was also observed that judgment accuracy was improved with experience when the snippets were presented.

Viewing behaviors of test participants are also analyzed using eye tracking system, and the result shows that viewing behavior when reading terms is different from other two clues, which corresponds to the result obtained from the analysis of judgment accuracy. The advantage of providing two types of snippets, common / specific snippets, is also confirmed by the analysis of the behaviors. The obtained result will contribute to the design of interface that can minimize the user's feedback cost for constrained clustering.

ACKNOWLEDGMENT

This work was partially supported by a grant from National Institute of Informatics (NII).

REFERENCES

[1] S. Basu, I. Davidson, and K. Wagstaff eds., Constrained Clustering: Advances in Algorithms, Theory, and Applications, Chapman & Hall, 2008.

[2] M. Chen, S. Yamada, and Y. Takama, "An Investigation of Snippet Generation for Minimum User Feedback," 23rd Annual Conference of the Japanese Society for Artificial Intelligence (JSAI2009), 2B2-1, 2009.

[3] E. Cutrell and Z. Guan, "Eye tracking in MSN Search: Investigating snippet length, target position and task types", Microsoft Technical Report, MSR-TR-2007-01, 2007.

[4] E. Cutrell, Z. Guan, "An eye-tracking study of information usage in Web search: Variations in target position and contextual snippet length," CHI'07, pp. 407-416, 2007.

[5] M. desJardins, J. MacGlashan, J. Ferraioli, "14: Interactive Visual Clustering for Relational Data," Constrained Clustering: Advances in Algorithms, Theory, and Applications, Chapman & Hall, pp. 329-356, 2008.

[6] E. D. Giacomo, W. Didimo, L. Grilli, G. Liotta, "Graph Visualization Techniques for Web Clustering Engines," IEEE Trans. Visualization and Computer Graphics, Vol 13, No. 2, pp 294-304, 2007.

[7] Q. Li, K. S. Candan, Y. Qi, "Extracting Relevant Snippets from Web Documents through Language Model based Text Segmentation," WI-IAT2007, pp. 287-290, 2007.

[8] Q. Li, L. Sun, and J. Duan, "Web Page Viewing Behavior of Users: An Eye-Tracking Study," Int'l Conf. on Services Systems and Services Management (ICSSSM'05), pp. 244-249, 2005.

[9] L. Lorigo, B. Pan, H. Hembrooke, T. Joachims, L. Granka, and G. Gay, "The influence of task and gender on search and evaluation behavior using Google," Information Processing and Management Vol. 42, No. 4, pp. 1123-1131, 2006.

[10] M. Okabe, S. Yamada, "Clustering with Constrained Similarity Learning," IWI2009 (WI-IAT Workshop), pp. 30-33, 2009.

[11] M. Okabe and S. Yamada, "Semi-supervised Query Expansion with Minimal Feedback," IEEE Trans. Knowledge and Data Engineering, Vol.19, No.11, pp.1585-1589, 2007.

[12] M. Okabe and S. Yamada, "Interactive Spam Filtering with Active Learning and Feature Selection," IWI2008 (WI-IAT Workshop), pp. 165-168, 2008.

[13] T. Onoda, H. Murata, and S. Yamada, "SVM-based Interactive Document Retrieval with Active Learning," New Generation Computing, Vol. 26, No. 1, pp. 49-61, 2007.

[14] J. J. Rocchio, "Relevance feedback in information retrieval," In The SMART Retrieval System: Experiments in Automatic Document Processing, 1971.

[15] K. Rodden and X. Fu, "Exploring How Mouse Movements Relate to Eye Movements on Web Search Results Pages," SIGIR 2007 Workshop on Web Information Seeking and Interaction (WISI), pp.29-32, 2007.

[16] S. Takami and K. Tanaka, "Customizing Web Search Results by Dynamic Re-Generation of Web-Snippet based on Search Purpose," DBSJ Letters, Vol. 6, No. 2, pp. 33-36, 2007.

[17] K. Wagstaff, C. Cardie, "Clustering with instance-level constraints," Proc. 7th Int'l Conf. on Machine Learning, pp. 1103-1110, 2000.

[18] K. Wagstaff, C. Cardie, S. Rogers, S. Schroedl, "Constrained K-means Clustering with Background Knowledge," Proc. 18th Int'l Conf. on Machine Learning, pp. 577-584, 2001.

[19] X. Xue, Z. Zhou, Z. Zhang: Improve Web Search Using Image Snippets, Proc. 21st National Conference on Artificial Intelligence, pp. 1431-1436, 2006.