Transdisciplinary Engineering: A Paradigm Shift
C. Chen et al. (Eds.)

doi:10.3233/978-1-61499-779-5-639

Development of Presentation Slide Retrieval System Based on Visual Information

Yoshiaki OIDA[a,d,1], Kazuo HIEKATA[a], Taiga MITSUYUKI[b], Hiroki KAMBA[c] and Isaac OKADA[a,d]
[a] *Graduate School of Frontier Sciences, The University of Tokyo*
[b] *Graduate School of Engineering, The University of Tokyo*
[c] *School of Engineering, The University of Tokyo*
[d] *Service Technology Unit, Fujitsu Limited.*

Abstract. Sharing and reusing knowledge in presentation form is very important because more and more presentations are made and accumulated in many companies. However, most of existing presentation slide retrieval systems are document file unit searches, which deal with only text information. Therefore, it is difficult to search enough reusable slides comprehensively when users don't memorize clear text information about necessary slides. The objective of this research is to develop a slide retrieval system which can comprehensively search slides necessary for presentation creation in a short time. First, Presentations are decomposed into slides. A Uniform Resource Identifier (URI) is assigned for each slide and managed in metadata repository in order to search across databases. Second, relevant slides can be retrieved based on the similarity of text or visual image such as colors and shapes so as to search enough slide candidates for a reusable slide. Third, a retrieval function based on handwritten graphic information is developed in order to search based only on visual memory. In case study, the effectiveness of each function is illustrated. In addition, it is illustrated that this system reduces the time to make a presentation by 19.7%.

Keywords. Presentation Knowledge Sharing, Slide Retrieval, Bag of Visual Words, Bag of Words, Handwritten Graphic Information

Introduction

Sharing and reusing knowledge in presentation form is very important because the more and more proposals and reports in presentation form are accumulated in companies and it takes much more time to make a presentation slide, which includes not only text information but also visual information such as colorful diagrams, graphs and charts in various layouts in order to express the message clearly.

However, there are three problems. First in companies, many presentation slides are accumulated across databases in a presentation file unit. Therefore, unnecessary work occurs such as switching of the database to be searched, opening and closing the presentation file, and browsing extraneous slides in the file. Second, there are many similar slides edited according to the various purposes based on a slide. It is difficult to

[1] Corresponding Author, Mail: yoida@s.h.k.u-tokyo.ac.jp

comprehensively check all of them when users search the reusable slide candidates. Third, most of the existing presentation slide retrieval systems deal with only text information. Users cannot obtain appropriate search results adequately when they don't memorize the text information with sufficient amount and accuracy to identify the required slide. Therefore, since it is inevitable to decide which slide to reuse from the limited search results, despite the fact that better slids are accumulated, the slide which is not the best is reused.

The objective of this research is to develop a slide retrieval system which can comprehensively search slides necessary for presentation creation in a short time. Specifically, the slide retrieval system using visual information which can search the slide most similar to the necessary slide based on fragmentary and ambiguous information of the necessary slide imaged in the user's head. First, in order to shorten the opening and closing time of the presentation file, the slide unit search is realized by decomposing the presentation file into slide units and managing them by Uniform Resource Identifier (URI) and metadata. Second, in order to comprehensively and efficiently search reusable slide candidates, the similar slide search function is developed which can search similar slides by using the specified slide itself as a search query. Third, in order to realize slide retrieval based only on visual memory when a user cannot remember the character string as the search query, a retrieval function based on handwritten graphic information is developed.

In case study, the effectiveness of each function is illustrated by comparing each developed function with the full text search function. In addition, it is shown that this system reduces the time to make a presentation by 19.7%.

1. Related work

In this chapter, existing presentation slide retrieval systems are outlined. UPRISE [1] and DocMIR [2] were developed as content retrieval systems for lectures with a slide-based search function. However, these systems only have a full-text search function. In order to improve the accuracy of full-text search, Kushki et al. [3] proposed an indexing method for weighting slides based on structured information expressed in OpenXML and it is demonstrated that the search accuracy and the coverage of the search result are higher than UPRISE.

Liew et al. [4] considered a slide as an image and proposed a slide retrieval system SLIDIR applying image processing technology. SLIDIR gives text labels by machine learning to the area excluding the background from the slide image. A slide search can be performed using text as a search query. However, it is a problem that a large amount of correct training data is required for learning.

Tanaka et al. [5] proposed a slide retrieval method focusing on the shape and arrangement of figures in slide based on analyzing OpenXML included in presentation file and. It is not practical because it is necessary to create a slide as a search query. Sakuragi et al. [6] also analyzed OpenXML and estimated the meaning and the hierarchical structure of the group of figures by analyzing the arrangement and the relationship between the figures, and added metadata for search to each slide. Since this method can be applied only to the OpenXML format, the information such as pictures and images contained in the slide is not utilized for a slide retrieval system. It is difficult for users to convert the characteristics of the necessary slide into the metadata such as the number of elements of a figure group or the shape of elements of a figure group and so on.

In addition, Sakuragi et al. proposed the similar slide re-search function. However, the similarity calculation method has only one criterion, which is based on hierarchical structure of figures. Therefore, this search function can not fully deal with various similarity measures in user 's memory.

DRIP system [7] was developed as a system with a presentation creation support function. This system is unsuitable for reuse of slides including images and graphic data because only text data can be reused when a presentation is created.

In our research, we will focus on reusing accumulated presentation slides and supporting the creation of new presentations. In order to comprehensively search candidate slides, we develop the function which can search similar slides using the slide itself as a query based on the various criteria: shape, color, and text.When it is difficult to express the search query as a character string, we develop a sketch image search that makes it easy to intuitively search the desired slide based on the handwritten sketch of the slide. Since images in slides and presentation files that are not in OpenXML format such as ppt format are also search targets, our system handles slides as images and text, and does not analyze OpenXML.

2. Proposed presentation slide retrieval system

Figure 1 shows the overview of this system, which consists of one metadata repository and three search functions. The following sections describe the details of each function.

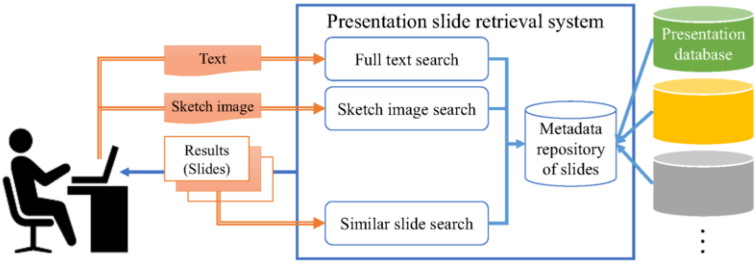

Figure 1. Overview of presentation slide retrieval system.

2.1. Metadata repository of slides

In order to save time opening and closing the presentation files and searching different repository, slide unit search is realized by decomposing the presentation files into slide units. URIs (Uniform Resource Identifier) are assigned to all the slides accumulated in several different presentation databases. URIs and metadata of slides are managed in one metadata repository which is based on RDF (Resource Description Framework).

2.2. Similar slide search

In creating a presentation, it is very important to confirm better candidate slides. This makes it possible for users to select the slide which is more suitable for the purpose, reducing the work of editing and correcting the slide. Therefore, when discovering the slide that the user thinks is beneficial, this system makes it possible to easily search similar slides. A presentation slide itself can be regarded as an image and most of slides includes character strings. Therefore, the similarity is calculated based on the shape and color derived from the slide image, and the text included in the slide.

2.2.1. Feature vector for shape : Bag of visual words [8]

The slide image is converted into feature vector based on a set of SURF feature quantities which are local feature quantities that are invariant to scale and rotation, robust to light and dark. First, all the SURF feature quantities are extracted from all the slide image set and categorized into K clusters using K-means clustering method. Then, Centrobaric vectors for each cluster (visual words) are obtained. K is set to 500 in this system. Next, SURF features extracted from each slide image are assined to the closest visual words. Finally, 500-dimensional feature vector for each slide image is obtained as a normalized histogram on visual words.

2.2.2. Feature vector for color : Color histogram [9]

In the case of 256 gradations of RGB, 16,777,216 colors can be expressed. However, users can neighter memorize nor identify all the exact color of slide image. Therefore, RGB is reduced to 4 gradations each and the color of the pixel of each slide image is assigned to the reduced colors. 64-dimensional feature vector is obtained as a normalized histogram on colors.

2.2.3. Feature vector for text : Bag of words [10]

The character strings are extracted from each AutoShape included in the slide and unnecessary spaces included for the appearance are removed. All the words are extracted from all the slides using morphological analysis. Among them, the upper 1,000 words with the highest occurrence frequency except for the stop word are selected. The frequency of occurrence of the 1,000 words for each slide is calculated to obtain 1,000-dimensional feature vector as a normalized histogram on text.

2.2.4. Similarity based on Bhattacharyya coefficient [11]

For the two histograms **I** and **M,** the similarity is calculated using the following equation:

$$BC(\boldsymbol{I},\boldsymbol{M}) = \sum_{j=1}^{n} \sqrt{I_j M_j} \tag{1}$$

where n is the number of dimensions of histograms. Threshold values are defined for each of the three features, and the similar slides exceeding the threshold value are associated with each other from in terms of three features: shapes, colors and text.

2.3. Skech image search

2.3.1. Feature quantity of contour line

Contour lines are obtained from a query sketch image and slides by canny edge and contour line tracking. Next, the obtained contour lines are described by P type Fourier descriptor [12], and amplitude spectra are extracted as feature quantities. The P-type

Fourier descriptor is one of methods for expressing a curve line on the plane, and this amplitude spectrum is invariant with regard to the translation, scale, and rotation of the contour line.

2.3.2. Similarity calculation

The Euclidean distances $d(\boldsymbol{x}, \boldsymbol{y})$ between the extracted amplitude spectra are calculated. The contour lines are similar as the distance is smaller [13]. When N amplitude spectra are obtained from sketch image K and M amplitude spectra are obtained from slide image S, their distances are defined by the following equation:

$$Dist(K,S) = \sum_{i=1}^{N} \min_{j=1,2,3,\ldots,M} d(\boldsymbol{a}_i, \boldsymbol{a}_j) \tag{2}$$

where $\boldsymbol{a}_i$ is the amplitude spectrum. The smaller $Dist(K, S)$ is, the more similar the two images are.

2.4. User interface

Figure 2 shows the user interface of presentation slides retrieval system. Users can search slides by entering queries on text forms in the area A. Search results is shown in the area B. When a useful slide is found, you can store the slide in each item of the table of contents of the new presentation prepared in advance in the area C. Finally, we can merge and download the slides stored in the table of contents as a pptx file. Our system has a user interface for sketch image search shown in Figure 3a. Users can draw a sketch as a search query in the sketch area using tablet PC in Figure 3b, and then obtain slides displayed in descending order of similarity in the results area.

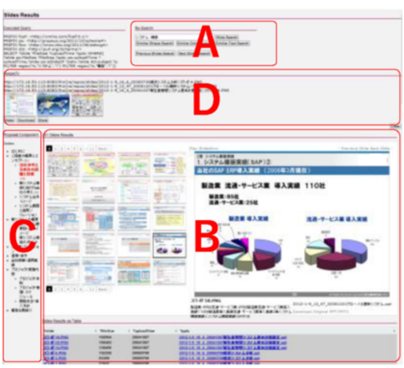

Figure 2. User interface of the slides retrieval system.

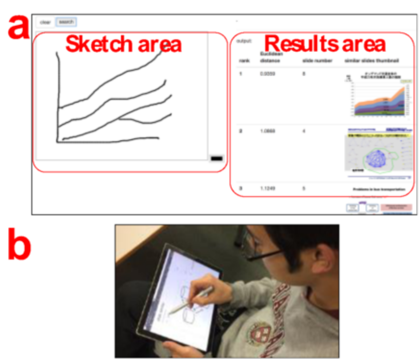

Figure 3. User interface of sketch image search.

3. Case study

3.1. Evaluation of similar slide search

The performance of similar slide search function is evaluated by comparing it with full text search function. The evaluation criterion is the number of search steps until all the search target slides are found from the accumulated slides.

In this evaluation, 17 existing presentation files including a total of 1632 slides are used. The search target slides are six slides including a figure of a V-shaped process which illusrates a system development standard process called SDEM [14] as shown in Figure 4.

Figure 4. Search target slides.

Search results with only full-text search, and search results with full text search and similar slide search are shown in Table 1 and Table 2, respectively.

In order to find all the target slides with full text search function, it was necessary to include a string “development” in the query without a string “standard”. Because search results by full text search depend on the input search query, it takes time to find an appropriate search query. On the other hand, all the target slides could be found with fewer steps by similar slide search based on the slides (Slide ID = 1, 2, 3, 4) obtained by full text search for the query string “SDEM”. Specifically, text similar slide search (Query ID = 3, 6, 12 in Table 2) made it possible to find all the six target slides. In addition, those precisions was larger than the precision of full text search (Query ID = 7 in Table 1), which could find all the search target slides.

Table 1. Search results by fullt text search. Query ID, query string, the number of search results, search target slide IDs contained in the search results, precision, recall , and F value.

ID	Query string	Num of results	Correct slide ID	P	R	F
1	SDEM	9	1,2,3,4	0.44	0.67	0.53
2	"body of standard process"	4	3,4	0.50	0.33	0.40
3	"standard process"	10	3,4	0.20	0.33	0.25
4	standard process	29	1,2,3,4	0.14	0.67	0.23
5	"development standard"	10	2,3	0.20	0.33	0.25
6	"development process"	8	1,2,4,5,6	0.63	0.83	0.71
7	development process	56	1,2,3,4,5,6	0.11	**1.00**	0.19

Table 2. Search results by similar slide search. Query ID, slide ID as search query, type of similalrity, the rest which are the same as those in Table 1.

ID	Query slide ID	Type	Num of results	Correct slide ID	P	R	F
1	1	shape	10	1,6	0.20	0.33	0.25
2	1	color	10	1	0.10	0.17	0.13
3	1	text	23	1,2,3,4,5,6	0.26	**1.00**	0.41
4	2	shape	10	2,4,5	0.30	0.50	0.38
5	2	color	10	2	0.10	0.17	0.13
6	2	text	35	1,2,3,4,5,6	0.17	**1.00**	0.29
7	3	shape	10	3,4	0.20	0.33	0.25
8	3	color	10	3	0.10	0.17	0.13
9	3	text	10	1,2,3,4	0.40	0.67	0.50
10	4	shape	101	2,3,4,5,6	0.05	0.83	0.09
11	4	color	10	4	0.10	0.17	0.13
12	4	text	34	1,2,3,4,5,6	0.18	**1.00**	0.31

3.2. Evaluation of sketch image search

Next experiment verifies that the target slide can be searched by developed sketch image search function. In this experiment, a user watches the target slide and draws a sketch for the target slide. It is evaluated that it is possible to obtain the target slide from the registered 10 slides. Figure 5 shows 10 slides accumulated in the system. Figure 6 shows the sketch image for slide with ID 8 drawn by the user.

Table 3 shows the distances calculated by the equation (2) between the drawn sketch image in Figure 6 and the 10 slides in Figure 5. Result showed that the most similar slide to the input image was the slide with ID 8. Therefore, it was confirmed that sketch image search function works well even for the simpler input image than the original slide image.

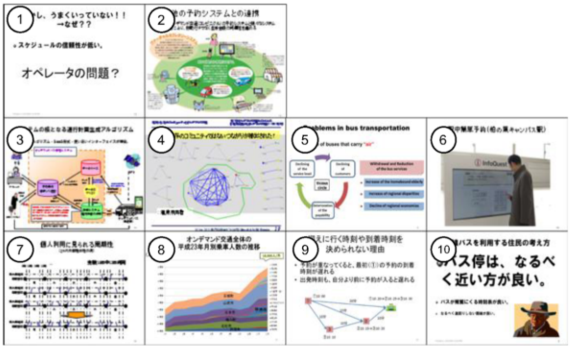

Figure 5. Accumulated slides in the slide retrieval system.

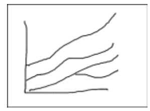

Figure 6. Drawn sketch image for the slide ID 8.

Table 3. Distances between sketch image and slide images.

Slide ID	Dist
1	2.80
2	1.97
3	1.97
4	1.09
5	1.13
6	1.35
7	1.97
8	**0.94**
9	1.35
10	1.94

3.3. Evaluation of effectiveness in practical work

In order to evaluate effectiveness in practical work, a draft of presentation material including about 40 slides is created by one user on developed system. This user has 16 yeas of experience in sales. The data set is the same as 3.1, and the time limit is set to 2 hours. We compare the current method using only Windows Explorer and Microsoft PowerPoint and the method of using this system in terms of required time. In order to reduce bias of prior knowledge, the interval between the two experiments is set to one week.

Table 4 shows the difference of creation time. The creation time using this system was 19.7% shorter than the current method because this system reduced time to open and close files, to browse irrelevant slides in the same file, and to reassemble extracted slides. Table 5 shows search efficiency of similar slide search, which was calculated based on the system usage log. From the viewpoint of the number of extracted slides per search, text similar slide search function is more than five times as efficient as full text search function, while the absolute value of the number of slides extracted by full text search function is the largest.

Table 4. Comparison of required time (minutes).

Current method		Using this system	
Preprocess	5	Preprocess	0.5
Search & Extraction	55	Search & Extraction & Reassembly	66
Reassembly	21		
Modification	28	Modification	21
Total	109	Total	87.5

Table 5. Search efficiency of similar slide search.

Type of search		Num of searches	Num of stored slides	Search efficiency
Full Text Search		92	33	0.36
Similar Slide Search	Shape	8	2	0.25
	Color	4	2	0.50
	Text	14	29	2.07

4. Conclustion

In our research, we developed a slide retrieval system which can comprehensively search slides necessary for presentation creation in a short time. Slide unit search is realized by decomposing presentation files into slide units and managing them by URI and metadata in one metadata repository. In order to comprehensively and efficiently search reusable slide candidates, similar slide search function was developed which can search similar slides by using the specified slide itself as a search query. In order to realize slide retrieval based only on visual memory when a user cannot remember the character string as a search query, a retrieval function based on handwritten graphic information was developed.

In case study, the performance of similar slide search and sketch image search was evaluated. The presentation creation time using this slide retrieval sytem was reduced by 19.7% compared with the current method. Through this case study, we found that the search efficiency of text similar slide search was the highest among the search functions.

References

[1] H. Yokota, T. Kobayashi, T. Muraki, and S. Naoi, UPRISE: Unified Presentation Slide Retrieval by Impression Search Engine, *IEICE Transactions on Information and Systems*, Vol. 87, No. 2, 2004, pp. 397–406.

[2] A. Behera, D. Lalanne, and R. Ingold, DocMIR: An Automatic Document-based Indexing System for Meeting Retrieval, *Multimedia Tools and Applications*, Vol. 37, no. 2, 2008, pp. 135–167.

[3] A. Kushki, M.Ajmal and N.Plataniotis, Hierarchical Fuzzy Feature Similarity Combination for Presentation Slide Retrieval, *EURASIP Journal on Advances in Signal Processing*, 2008, pp. 1–19.

[4] G. M. Liew, M. Y. Kan, Slide Image Retrieval: A Preliminary Study, *Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries*, 2008, pp. 359-362.

[5] S. Tanaka, T. Tezuka, A. Aoyama, F. Kimura and A. Maeda, Slide Retrieval Technique Using Features of Figures, *Proceedings of the International MultiConference of Engineers and Computer Scientists*, Vol. 1, 2013. pp. 424-429.

[6] Y. Sakuragi, A. Aoyama, F. Kimura, and A. Maeda, A Method for Estimating Meanings for Groups of Shapes in Presentation Slides, *International Journal of Computer Theory and Engineering*, Vol. 8, No. 1, 2016, pp. 74-79.

[7] T. Tsuchida, S. Ohira and K. Nagao, A Presentation Slide Authoring System Based on Discussion Content, IEICE technical report, *Office Information Systems*, Vol. 108, No. 53, 2008, pp. 85–90.

[8] J. Sivic and A. Zisserman : Video Google: A Text Retrieval Approach to Object Matching in Videos, *Proceedings of 9th IEEE International Conference on Computer Vision*, Vol. 2, 2003, pp. 1470–1477.

[9] M. J. Swain and D. H. Ballard, Color Indexing, *International Journal of Computer Vision*, Vol. 7, No. 1, 1991, pp. 11–32.

[10] J. M. Fishbein and C. Eliasmith, Integrating Structure and Meaning: A New Method for Encoding Structure for Text Classification, *Proceedings of the IR research, 30th European conference on Advances in information retrieval*, 2008, pp. 514-521.

[11] T. Kailath, The Divergence and Bhattacharyya Distance Measures in Signal Selection, *IEEE Transactions on Communication Technology*, Vol. 15, No. 1, 1967, pp. 52–60.

[12] Y. Uesaka, A New Fourier Descriptor Applicable to Open Curves, *Electronics and Communications in Japan (Part I: Communications)*, Vol. 67, No. 8, 1984, pp. 1–10.

[13] I. Hattori and Y. Kumagai and G. Ohashi, Query-by-Sketch Image Retrieval for Partial Retrieving Using Fourier Descriptor, *The IEICE transactions on information and systems (Japanese edition)*, Vol. 93, No. 12, 2010, pp. 2678-2862.

[14] T. Oshima, M. Kashiwagi and H. Fukao, Fujitsu's System Development Methodology: SDAS, *FUJITSU Scientific & Technical Journal*, Vol. 42, No. 3, 2006, pp. 277-285.