Query-by-Appearance System for Style-Oriented Media Retrieval

Yuka Koike
Faculty of Environment and Information Studies
Keio University
5322 Endo, Fujisawa, Kanagawa, 252-0882, Japan
t09334yk@sfc.keio.ac.jp

Shuichi Kurabayashi
Faculty of Environment and Information Studies
Keio University
5322 Endo, Fujisawa, Kanagawa, 252-0882, Japan
kurabaya@sfc.keio.ac.jp

***Abstract*— This paper proposes the *Query-by-Appearance* system that provides an intuitive and effective query-input and visual retrieval method for media data, especially e-books, based on similarity of content's layout and color. The unique feature of this system is a query-assistance function that helps users to input their requirements by utilizing the knowledge which have been developed in the field of "Editorial Design". Editorial design is an essential methodology to enrich overall appearance of books and magazines. The query input assistant retrieves editorial design templates that are similar to the input query, and generates images by combining these templates with the chosen color scheme. The system then retrieves actual e-book images that are similar to those generated from the query. Finally, the system visualizes the retrieval results, which consist of two relevance scores (layout and color) in a two-dimensional ranking plot. This assistant mechanism allows users to find the desired e-book by submitting a query, which consists of simple lines, intuitively.**

Keywords*— *e-Book; Search Engine; Editorial Design.

I. INTRODUCTION

The proliferation of portable, personal devices dominated by a large display, such as tablet computers and smartphones [1][2] has made it common for e-books and Web pages to viewed on such devices. E-books, in particular, are fast spreading; Association of American Publishers reported that the total e-book sale from January to October of 2010 constituted 8.7% of all book sales in the United States [3]. Such proliferation and diversity of digital media data increase the demand of a system for retrieving them. Generic retrieval methods, such as keyword-based search engines and content-based image retrieval systems, are not sufficient for retrieving them, because users require highly-domain-specific search quality in such daily-life media data. For example, a fashion magazine requires a fashion-domain specific query for retrieving it, and an outdoor and nature magazine requires different query for retrieving them. It is essential to provide a query which is tailored to each type of e-books.

However, common users have difficulty in learning a new and specific search method for each type and genre of media data. It is important to develop a novel and intuitive search engine for those daily-life multimedia data. Visual-oriented search mechanisms, which do not use text-based search methods, are promising because users often memorize e-book contents by associating them with their visual appearances. We suppose that the overall visual appearance of page layout is essential to e-book search. In spite of layout being an important factor for bookbinding, there are no studies about layout searching method for e-book, because conventional systems utilize CBIR (Content-Based Image Retrieval) mechanisms for retrieving e-book.

Toward the above objective, this paper proposes a "*Query-by-Appearance*" system that improves the query input process by exploiting the knowledge of book design. This system provides a query-generation mechanism that enhances and decorates a user's rough and simple query by generating candidates of queries inspired by the initial input. This system enables users to find a desired e-book just inputting intuitive and simple query alone because the system ranks e-books according to the similarity of overall compositions, layouts, colors, and overviews, rather than detailed and trivial differences between them. We call this visually rich search method "*Query-by-Appearance*".

The unique feature of this system is a book style-oriented query generation, which is enhanced by the knowledge of "editorial design", for helping users to input a complex and sophisticated query by generating candidates of queries. Editorial design, developed in the book industry, is an essential methodology to enrich and to beautify books and magazines by configuring the overall appearance, including the visual layout, such as photographs and illustrations. Editorial design techniques provide sophisticated way to construct and to lay out visual objects' compositions. Our "*Query-by-Appearance*" system utilizes the editorial design methods as query templates. The query templates consist of layouts and color schemes. When the system receives a query, which consists of simple lines, the system applies the query templates to enrich and to improve the query, and generates the well-organized and colored image query by combining the layout templates and the color scheme templates.

We design "*Query-by-Appearance*" system as an embedded system in an e-book format, such as EPUB3. The system encapsulates the templates, which are derived from the editorial design methods, into each e-book according to the style of books such as a fashion magazine and a mystery novel. This design principle makes it possible to interpret a user's query with reflecting the target books' own characteristics. A core function of this system is a design-based template-matching function that compares a user's rough query input with the embedded editorial design templates.

 ISBN: 978-1-61208-200-4

The system is applicable to the following area: 1) searching dimly remembered visual contents, such as the title of book, web pages, and 2) searching jackets of DVD, CD, and packaging of a product by preparing appropriate templates. We show a prototype system implementation that is embeddable into a media data, by using JavaScript supported in EPUB3 and HTML5. This prototype system implements self-search mechanism in those media data according to the internal page images and images of web pages in the browser history.

The remainder of the paper is structured by follows. In section two, we discuss several related studies. In section three, we show an architectural overview of our system. In sections four and five, we describe the four fundamental components and three core functions of our system. In section six, we show the prototype of our system. In section seven, we evaluate the effectiveness of our system. In section eight, we give concluding remarks.

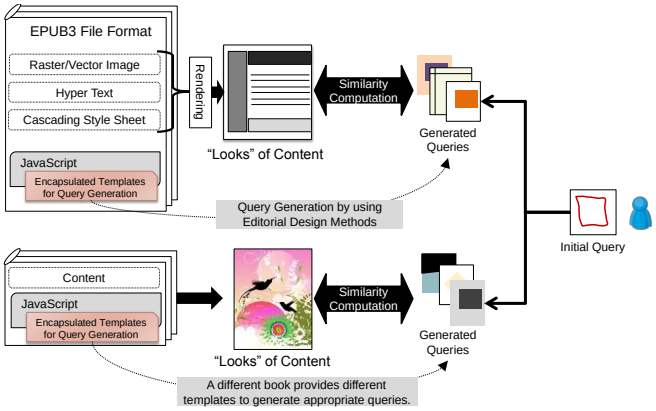

Figure 1. A Concept of Query-by-Appearance System for Style-Oriented e-Book Retrieval by Using Encapsulated Editorial Design Templates for Query Generation.

II. RELATED WORKS

There are several studies on developing a search engine that considers the visual appearance of media. Recently, CBIR [4][5] systems have been developed that allow users to search for images by visual similarity. Such systems support "*Query-by-Example*," which requires an example image as the query input. However, basically, CBIR methods are not suitable to find complex multimedia data such as Web sites and e-books because these methods ignore the semantic information associated with the images, such as text and annotations. Alternatively, a method [6] has been proposed to calculate similarity among the visual components of Web pages by their overall visual appearance. The system proposed in [7] analyzes the visual link structure created by assigning numerical weights to each image. This system incorporates visual signals into text-based search engines in order to improve the accuracy of conventional search engine. Those conventional approaches are effective to retrieve image data by submitting the detailed query, such as a sketch, an example image, and a combination of them. However, it is difficult for novices to input such as detailed query, especially in the domain of well-designed e-books. Thus, a new approach to assist users to create the domain-specific query for multimedia data is required.

III. SYSTEM ARCHITECTURE

Figure 2 illustrates an architectural overview of this system. Our "*Query-by-Appearance*" system provides an indirect e-book retrieval mechanism that is leveraged by the knowledge base of editorial design. This system retrieves the e-book images that are similar to editorial design characteristics inspired by the query. The results are visualized in two-dimensional ranking, such as generated query axis and ranking axis. This query-assistance function enables users to input layout sketch and color by using Web-based touch UI, and search e-books through rough visual appearances intuitively, by accepting a simple query consisting of layouts and colors, rather than technical knowledge.

The system uses the generated e-book image data for retrieving an e-book by calculating similarity based on the color and structure of appearance. This approach is highly effective to develop a Web-based touch UI because this system simplifies the query-input task. The touch UI visualizes both the layout similarity and the color similarity in a two-dimensional matrix. This visualization enables users to retrieve the desired e-book according to their preferences, which means weights of the layout similarity and the color similarity, by selecting sub-matrix in the visualized two-dimensional matrix.

The most important function of this system is the query-assistance function that utilizes both layout and color to enrich and extend the input query. As described, this system employs a design template database which stores layouts and color scheme data. The templates are defined by sets of matrices; so, the system is able to select the relevant templates by calculating the similarity in layout structure. In addition, 20 color schemes are provided by the combination of 102 colors.

This system performs the *Query-by-Appearance* for e-books by involving the following six steps:

1. The user inputs a rough layout sketch on the Canvas system.
2. Second, the system converts the rough layout query input into a matrix, and compares this matrix with 30 templates that are stored in the knowledgebase.
3. The system generates colored templates by adding color to the selected layout templates using the 20 color schemes stored in the knowledgebase. The sets of colored templates are expanded automatically to sets of image matrices, which include the layout structure and the color scheme.
4. The system calculates similarity between the expanded query matrices and e-book image data stored in an e-book database. The e-book image data is preliminarily clustered into 102 colors.
5. The e-book cover images that contain similar editorial design characteristics are visualized in a two-dimensional ranking with generated query axis and ranking axis.

 ISBN: 978-1-61208-200-4

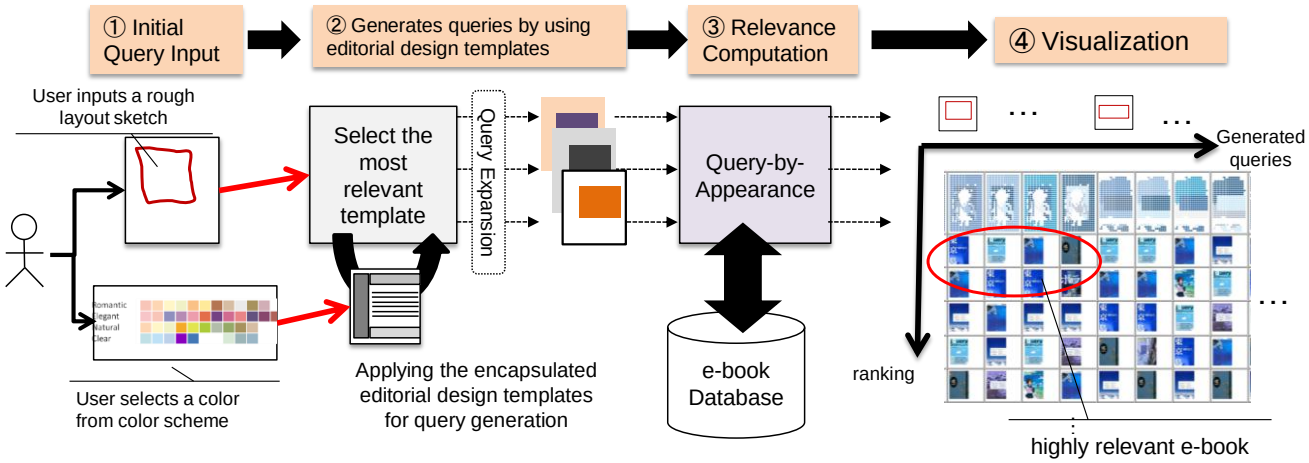

Figure 2. System Architecture of Query-by-Appearance System.

6. Further retrieval is done by selecting a color on a sub-matrix in the visualized two-dimensional matrix.

Our visualization mechanisms allow users to retrieve the desired e-book according to similarity in layout and color scheme.

IV. DATA STRUCTURES

Our system contains four fundamental components: A) Query Matrix, B) Template Knowledgebase, C) e-book Database, and D) Visualization.

A. *Query Matrix*

Our system converts each query into a matrix that represents the user`s rough input, such as simple lines drawn on HTML5 Canvas. We call this matrix a query matrix. The query matrix is important for calculating the similarity between the user input and the layout templates in the knowledgebase. We assign weights to each column as follows: if a line passes through a column, 1.0 is assigned to the column; if no line passes through a column, 0.0 is assigned to it. This weight is assigned to get the inner product of the user`s query matrix and layout template matrix, as shown in the *Template Knowledgebase*.

The query matrix Q is defined as follows:

$$\mathrm{Q} \coloneqq \begin{bmatrix} w_{[0,0]} & \cdots & w_{[0,m]} \\ \vdots & \ddots & \vdots \\ w_{[n,0]} & \cdots & w_{[n,m]} \end{bmatrix}$$

where $w_{[n,m]}$ indicates whether a line is present at $[n,m]$ in the query.

B. *Template Knowledgebase*

The system provides matrix structures representing: 1) the layout template 2) the assigned-color template, and 3) the color scheme. We call the data structures a design template structure. All of the above are stored in the Template Knowledgebase, which assists in expanding the input queries.

1) The layout template matrix is the set of matrices obtained by exploiting the method of editorial design. We assign weights to each column as follows: if there is most obvious layout border passes through a column, 1.0 is assigned to that column; the regions surrounding 1.0, 0.5 is assigned to it; if there is no important layout structure to the column, 0.0 is assigned to it.

The layout template matrix $\mathrm{T_L}$ is defined as follows:

$$\mathrm{T_L} \coloneqq \begin{bmatrix} t_{[0,0]} & \cdots & t_{[0,m]} \\ \vdots & \ddots & \vdots \\ t_{[n,0]} & \cdots & t_{[n,m]} \end{bmatrix}$$

where $t_{[n,m]}$ indicates whether a line is present at the location $[n,m]$ in the template matrix.

2) The assigned-color template matrix is a set of matrices containing numbers from 1 through 5 in each column. The system draws color in each column according to this number. 1 represents the color that appears most often in the image, whereas 5 represent the color that appears least often.

3) The color scheme is a set of color data obtained by a combination of 102 colors. This system draws colors in each column by using this color scheme, to perform the expanded queries in order to generate e-book image data.

The color scheme $\mathrm{T_C}$ is defined as follows:

$$\mathrm{T_C} \coloneqq \langle \mathrm{c}_1, \mathrm{c}_2, \cdots, \mathrm{c}_k \rangle$$

where c_k denotes each color used in the color scheme, and k is the total number of colors in the color scheme. In the examples in Figure 3, each color scheme is assigned a specific adjective.

20 color schemes are made using a dictionary that defines adjectives for color schemes, such as “Romantic”, or “Elegant” [8].

C. *e-book Database*

Our system converts JPEG-format image data into the matrix structure by clustering the image into 102 colors. We call the data structure an e-book database matrix. This e-book database matrix contains clustered image data that is used to calculate the similarity with the expanded query matrix obtained from the user input and template knowledgebase.

The e-book database matrix D is defined as follows:

 ISBN: 978-1-61208-200-4

$$D := \begin{bmatrix} d_{[0,0]} & \cdots & d_{[0,m]} \\ \vdots & \ddots & \vdots \\ d_{[n,0]} & \cdots & d_{[n,m]} \end{bmatrix}$$

where d represents each element of the matrix, and $[n, m]$ represent the rows and columns of the image data, respectively, in the matrix.

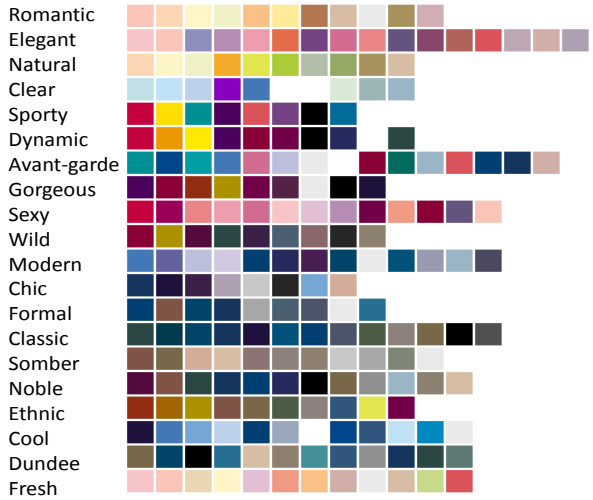

Figure 3. Sets of color schemes.

D. Visualization

The system visualizes the layout and color similarity in a two-dimensional ranking. This ranking visualization consists of two relevance scores for each cell. The ranking array Cell_x is defined as follows:

$$\text{Cell}_x := \langle L_i, C_j, R_L, R_C, \{B_1 \cdots B_p\} \rangle$$
$$\vdots$$
$$\text{Cell}_{x+n} := \langle L_i, C_j, R_L, R_C, \{B_1 \cdots B_p\} \rangle$$

where Cell_x to Cell_{x+n} comprise L_i for Layout id, C_j for color id, R_i for layout ranking, R_c for color ranking, and $\{B_1 \cdots B_p\}$ for ranking order on a column, when more than one image data has the same layout template and color scheme. The order of books in a certain column is decided by color relevancy, where p represents the order of the books.

V. CORE FUNCTIONS

The system contains three fundamental functions: A) Selecting the most relevant template, B) Assigning-colors to the template C) Query-by-Appearance, and D) Two-dimensional ranking.

A. Selecting the most relevant template

The system selects the most relevant template by comparing the user input with the template matrices stored in the template knowledgebase. We call this function query assistance. The inner product of the transformed user input matrix, which consists of line sketched data, and the template matrices that are stored in the template knowledgebase, is calculated in order to analyze the similarities. A higher value for this inner product indicates greater similarity.

The function is defined as follows:

$$f_{layout}(q,t) := \sum_{i=1}^{w}\sum_{j=1}^{h} q_{[i,j]} t_{[i,j]}$$

where $q_{[i,j]}$ indicates whether a line is present at position [i,j] in the query matrix, and $t_{[i,j]}$ indicates whether a line is present at the position [i,j] in the template matrix. This calculation measures the line similarity between the query matrix and the template matrix by checking all the pixels.

B. Assigning-colors to the template

The system assigns colors 1 through 5 to the template, according to the users selected color scheme and number of colors. This function is important for generating a colored template, which contains template, keyword, and number information. The function is defined as follows:

$$f_{assign}(T_L, \text{keyword}, \text{number}) \rightarrow Q := \begin{bmatrix} hsv_{[0,0]} & \cdots & hsv_{[0,m]} \\ \vdots & \ddots & \vdots \\ hsv_{[n,0]} & \cdots & hsv_{[n,m]} \end{bmatrix}$$

where template T_L indicates the template data that contains numbers 1 to 5 for drawing, a keyword indicates the name of the color scheme, and number indicates the number of colors in color scheme to draw on the template.

C. Query-by-Appearance

Our system calculates the correlation between the generated query matrix and e-book image data matrices in order to retrieve the relevant e-book image. The Query-by-Appearance operation is defined as follows:

$$f_{qba}(Q,B) := \sum_{i=1}^{w}\sum_{j=1}^{h} \Delta_{godlove}\left(Q_{[i,j]}, B_{[i,j]}\right),$$

$$\Delta_{godlove} := \frac{2S_1S_2\left(1-\cos\left(2\pi\frac{|H_1-H_2|}{100}\right)\right) + (|S_1-S_2|)^2 + (4|V_1-V_2|)^2}{2}$$

where $Q_{[i,j]}$ denotes a color at the specific point in the generated query matrix, and consists of HSV components H_1, S_1, and V_1. $B_{[i,j]}$ denotes a color at a specific point in the e-book image data, and consists of H_2, S_2, and V_2. This calculation measures the color similarity between the expanded queries and the e-book image data for each block. We employ Godlove's delta equation [9] to calculate the distance between two HSV colors.

D. Two-dimensional ranking

Our system provides a two-dimensional visualization mechanism for presenting the results of the e-book image search. This visualization enables users to select desired books on the basis of the layout and the color similarity. The system takes a wider sample of template data compared to the user`s input, in order to display the retrieval results interactively on the basis of the choice of layout and color.

This system employs the following three steps:

Step-1. The generated queries are shown on the horizontal axis. Each column contains a query with template and color.

Step-2. The ranking of e-book image data is shown on the vertical axis.

 ISBN: 978-1-61208-200-4

Step-3. The system stacks the e-book image data according to the similarity score of queries and e-book meta data in each cell.

VI. System Implementation

We have implemented a prototype system for *Query-by-Appearance* that calculates the similarity between generated queries and editorial design templates. The prototype system is implemented using HTML5 Canvas and JavaScript as shown in Figure 4. The important feature of this implementation is that it uses standard web technologies, which are also used by the current EPUB3 specification. So, we can apply our method to EPUB3-based e-books without significant modification. The implemented system consists of the following three modules:

- View: The user sketches the layout query on HTML5 Canvas and selects a color scheme. The system calculates the similarity between the queries and template knowledgebase.
- Controller: The retrieval is done by clicking on the search button.
- Model: The retrieval target database and two templates (layout matrix, color scheme) are encoded in the JSON format and embedded in the HTML5.

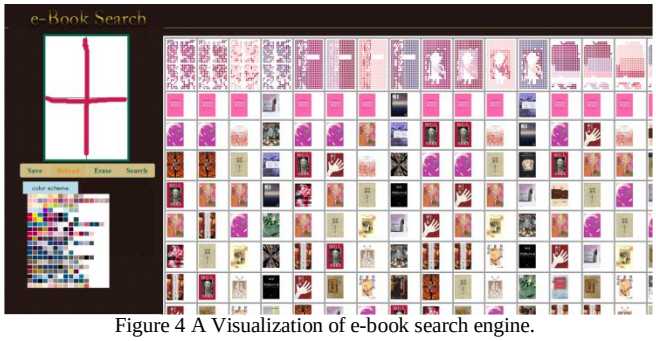

Figure 4 A Visualization of e-book search engine.

VII. Evaluation

A. Outline of experimental studies

In this section, we evaluate the effectiveness of our system by examining the retrieval precision for our generated queries. The task of this experiment is to clarify the effectiveness of retrieving book cover image by utilizing the knowledge of "editorial design". We compare the retrieval precision using two search methods as follows: method-1) uses queries generated using only layout templates, and method-2) uses queries generated by integrating layout templates and color scheme templates. We show that the integration of a layout template and a color template has significant contribution to the e-book retrieval result.

In this experiment, we have prepared: 1) 300 book cover images from Amazon.co.jp, 2) five queries, 3) five answer data sets for each query. The queries and answer data sets are set up considering the basic structure of "editorial design", as specified as follows: 1) the symmetrical layout, 2) the diagonal layout, 3) the layout with gravity point on center. We chose the following five queries (Figure 5):

Query-1. Draw the shape of a cross to divide the canvas into four sections.
Query-2. Draw a vertical line along the center of canvas to divide it into two sections.
Query-3. Draw two horizontal lines on the canvas to divide it into three sections.
Query-4. Draw a rectangle in the center of the canvas to divide it into two sections.
Query-5. Draw two vertical lines on the canvas to divide it into three sections.

We have used color schemes that consist of four colors (as shown in Figure 3). The system calculates the similarity between those queries and generated e-book image data.

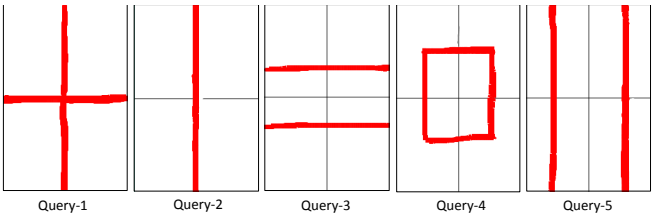

Figure 5. The Experimental Queries drawn on the Canvas.

B. Experimental Results

In this section, we have evaluated the retrieval precision of generated queries in order to clarify the effectiveness of our approach. Figure 6 shows the results of the retrieval. Our approach gives two average scores for retrieved rank. The first score is the average of correct answer data shown in the top 20, which has retrieved only by template relevance. The second score is the average rank of correct answer data that has the chosen color scheme.

Figure 7 shows the resulting scores of this experiment. The vertical axis is the retrieved score. The horizontal axis is the variations of the query that is the template of correct answer sets. We calculate the average scores of each query in the top 20 ranking as shown in Figure 6. This result shows that method-2 (integrating both layout and color templates) achieves better retrieval precision than method-1 (using the layout template only), since a superior rank has been assigned to the result corresponding to the query. This result shows that the system effectively retrieves e-book images by using layout template and color scheme.

VIII. Conclusion and Future Works

This paper proposes a "*Query-by-Appearance*" system for e-books, which provides an intuitive and effective query input and visual retrieval method base on the similarity in overall layout and color scheme. The unique feature of this system is the query-assistance function that exploits structural design and analysis knowledge from the field of "editorial design". Editorial design is an essential methodology to enrich the appearance of books and magazines. The query input assistant retrieves the editorial design templates that are similar to the input query, and generates actual e-book images by combining the color templates and the layout templates. Then, the system retrieves e-book images that are similar to the generated e-

 ISBN: 978-1-61208-200-4

book image. Finally, the system visualizes the retrieval results, which consist of two relevance scores (layout, color) in a two-dimensional ranking; generated query axis and ranking axis. This assistant mechanism is intuitive because it allows users to find a desired e-book by submitting a query that consists of simple lines.

As a future work we are planning to develop a prototype system that supports full-spec EPUB3, and to perform a feasibility study by evaluating scalability and effectiveness of our approach applied to the existing e-books.

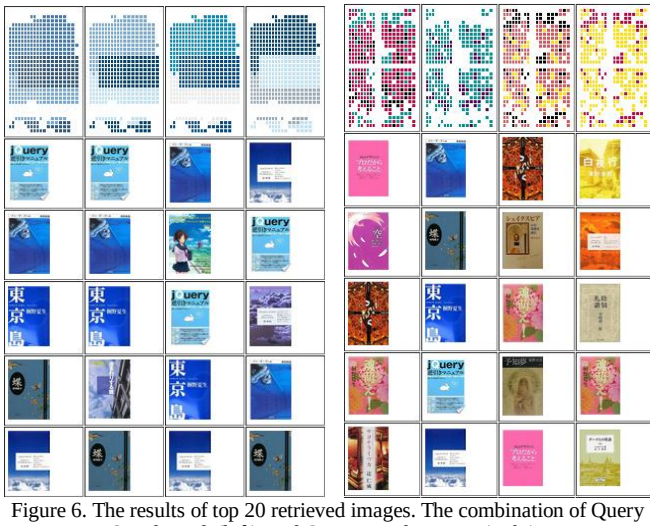

Figure 6. The results of top 20 retrieved images. The combination of Query 3 and "cool"(left), and Query 1 and "sporty"(right).

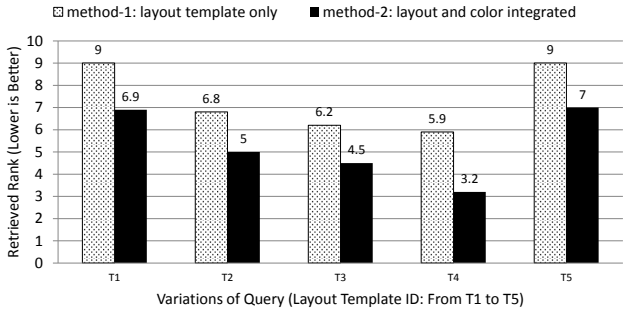

Figure 7. The result of retrieved rank of "only layout template" and "colored template".

ACKNOWLEDGMENT

This research is supported by SCOPE: Strategic Information and Communications R&D Promotion Programme of the Ministry of Internal Affairs and Communications, Japan: "Kansei Time-series Media Hub Mechanism for Impression Analysis / Visualization / Delivery Intended for Video / Audio Media".

REFERENCES

[1] Gartner, "Gartner Says Worldwide PC Shipments in First Quarter of 2011 Suffer First Year-Over-Year Decline in Six Quarters", 13 April 2011, http://www.gartner.com/it/page.jsp?id=1632414.

[2] IDC, "Nearly 18 Million Media Tablets Shipped in 2010 with Apple Capturing 83% Share; eReader Shipments Quadrupled to More Than 12 Million, According to IDC", 10 March 2011, http://www.idc.com/about/viewpressrelease.jsp?containerId=prUS22737611.

[3] American Association of Publishers, "AAP Reports October Book Sales", 2010, http://www.publishers.org/main/PressCenter/Archicves/2010_Dec/AAPReportsOctoberBookSales.htm, (accessed 2010-9-10).

[4] Baeza-Yates, R. and Ribeiro-Neto, B. "Modern Information Retrieval: The Concepts and Technology behind Search (2nd Edition)," Chapter 14, Addison-Wesley Professional, 2011.

[5] Datta, R., Li, J., and Wang, J. Z. "Content-Based Image Retrieval - Approaches and Trends of the New Age," in Proceedings of the 7th ACM SIGMM International Workshop on Multimedia Information Retrieval, ACM: New York, NY, USA, 2005.

[6] Furusawa, T., Watai, Y., Yamasaki, T., Aizawa, K. "A Visual Similarity Metric for View-based Web Page Retrieval," The Journal of the Institute of Image Information and Television Engineers, Vol. 62, No.2, 2008, pp.209-215, DOI: 10.3169/itej.62.209 (In Japanese).

[7] Jing, Y., and Baluja, S. "PageRank for Product Image Search," In Proceedings of the 17th International Conference on World Wide Web (2008), pp. 307-316, April 2008.

[8] Nagumo, H. "Color Scheme Imaging", ISBN-13: 978-4766111705, Graphics-sha, 2000 (In Japanese).

[9] Godlove, I. H. "Improved Color-Difference Formula, with Applications to the Perceptibility and Acceptability of Fadings". J. Opt. Soc. Am., Vol. 41, NO. 11, pp.760-770, 1951.

 ISBN: 978-1-61208-200-4