Query-by-Appearance: Visual Query Expansion to Support Domain-Specific Retrieval of e-Books

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

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

***Abstract*—This paper proposes a visual query expansion method called *Query-by-Appearance*, which utilizes techniques developed in the field of editorial design. Our system expands a simple query into one that is informative and well-designed one, allowing users to find their desired e-book intuitively. To simplify the selection of target media data on modern touch screen devices, our system visualizes the retrieved results in a two-dimensional plot that incorporates both layout-based and color-based relevance rankings. We conduct two experiments to evaluate the effectiveness of our Query-by-Appearance method, comparing our automatically generated queries and conventional keyword-based queries in order to determine the extent to which the editorial design templates improve media retrieval. The results show that our visual query expansion method successfully increases retrieval performance for display-oriented media data.**

Keywords - e-Book, Search Engine, Editorial Design, Query Expansion

I. INTRODUCTION

This paper describes a *Query-by-Appearance* system [1] and its implementation framework for the EPUB3 e-book data format. Our system utilizes modern editorial design techniques, which enrich the overall appearance of books and magazines, to interpret users' initial queries. By leveraging this editorial design knowledge, we develop an intuitive query editing method to find the desired e-book from the viewpoint of its visual appearance.

Books are traditionally regarded as highly organized media for storing the inherited knowledge of humankind for many years. With the rapid progress of mobile computing technologies, many books are becoming digitized and stored in online libraries and on personal devices. Due to their proliferation, portable and personal devices dominated by a large display, such as tablet computers and smartphones [2],[3], are commonly used to view e-books and web pages. E-books, in particular, are fast becoming popular; the Association of American Publishers reported that total e-book sales in January 2012 constituted 26% of all book sales in the United States [4]. Such proliferation and diversity of digital media increases the demand for an effective retrieval system. By enhancing the retrieval capability for e-books, we enable a wider scope for the sharing of human knowledge and more opportunities for its practical use.

Generic retrieval methods, such as content-based image retrieval systems [5],[6] and keyword-based image search engines [7], are insufficient for retrieving e-books because users require high domain-specific search quality for frequently accessed media data. For example, a fashion magazine requires a fashion-domain-specific query for its retrieval, whereas very different query is required to retrieve an outdoor and nature magazine. Most approaches to image retrieval on the Internet adopt document search techniques that index images files by associating them with the surrounding text and HTML tag attributes [7]. Due to ambiguity and poor quality, both in the user-supplied keyword query and the textual metadata used to describe the media, web image retrieval systems find it difficult to incorporate visual features of the target media itself in the media retrieval process.

Visual-oriented search mechanisms, which do not use text-based search methods, are promising because users often memorize e-book contents by association with their visual appearance. The overall visual appearance of the page layout is essential to e-book searches because the visual and spatial characteristics of graphical layouts are an important factor in recognizing the contents, as surveyed in [8]. For example, Woodruff et al. [9] indicated that the combination of textual information and thumbnails enhances the usability of Web search engines. However, because conventional book search systems utilize text information, such as tables of contents and indexes [10], there are no studies into layout-based search methods for e-books, in spite of the layout being an important factor for bookbinding.

Thus, this paper proposes a Query-by-Appearance system that improves the query input process by exploiting some knowledge base of book design. Figure 1 illustrates the concept of our Query-by-Appearance system. Query-by-Appearance system adopts editorial design as a template to restructure and modify rough queries entered as input by a user. When the system receives a query consisting of simple lines, it retrieves editorial design templates that are similar to the initial query, and uses them to enrich and improve the query. The system then retrieves actual e-book media data similar to those generated from the query.

This system enables users to find the desired e-book by merely inputting just a simple, intuitive query, because the system ranks e-books according to the similarity of their overall composition, layout, colors, and overview, rather than detailed and trivial differences between them. The system is applicable to the following areas: 1) searching for vaguely remembered visual content, such as the cover of a

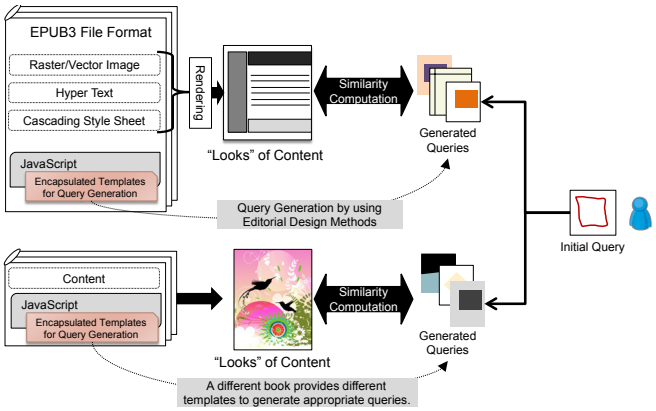

Figure 1. Conceptual view of the Query-by-Appearance System for Style-Oriented e-Book Retrieval Using Encapsulated Editorial Design Templates for Query Generation.

book or a web page, and 2) searching for DVD and CD covers/packaging by preparing appropriate templates.

We implement a prototype system that can be embedded into media data by using JavaScript, which is supported in EPUB3 and HTML5. This prototype system applies a self-search mechanism to the media data according to the internal page images and web page images in the browser history. The system encapsulates the templates derived from the editorial design methods into each e-book according to its style, such as that of a fashion magazine or mystery novel. This design principle makes it possible to interpret user queries by reflecting the characteristics of the target book itself.

The remainder of the paper is structured as follows. In Section II, we discuss several related studies, and Section III introduces several motivating examples to show the advantages of our Query-by-Appearance concepts. Section IV presents an architectural overview of our system. We describe the four fundamental data structures and three core functions of our system in Sections V and VI, respectively. Section VII introduces the prototype system. In Section VIII., we compare conventional keyword-based search method and our visual query-expansion method to show the advantages of editorial design. Finally, in Section IX, we offer our concluding remarks and some ideas for future work.

II. Related Work

In this paper, we present a system architecture that aims to improve the effectiveness of e-book retrieval approaches by exploiting editorial design knowledge as a visual query-expansion source. To cover all of these aspects, we divide our review of related work into three sections: 1) visual similarity analysis of media data, 2) document structure information extraction, and 3) query expansion for e-books.

A. Visual Similarity Analysis of Media Data

Several studies have investigated the development of a search engine that can consider the visual appearance of media. Content-Based Image Retrieval (CBIR) [5],[6] systems have been developed that allow users to search for images by visual similarity. Such systems support Query-by-Example, which requires a sample image as the query input. However, CBIR methods are unsuitable for finding complex multimedia data such as websites and e-books because they ignore the semantic information associated with the images, such as text and annotations. An alternative method [11] has been proposed that calculates the similarity between the visual components of web pages according to their overall visual appearance. The system proposed in [12] analyzes the visual link structure created by assigning numerical weights to each image. This technique incorporates visual signals into text-based search engines in order to improve the accuracy of conventional searches. Although conventional approaches are effective in retrieving image data when a detailed query, such as a sketch or sample image, is submitted, it is difficult for novices to input such queries, especially in the domain of well-designed e-books. Thus, a new approach to assist users to create domain-specific queries for multimedia data is required.

B. Document Structure Information Extraction

In the field of document database systems, the extraction of document structure information is recognized as an important topic [13]. Recently, Gao et al. [14] proposed a method of analyzing the physical and logical structure of PDF documents, including global typographies, reading order, logical elements, chapter/section hierarchy, and metadata. In a book, the Table of Contents is an important resource for extracting the logical structure [15],[16]. Spatial and semantic information, such as the structure of page numbers, headers, footers, headlines, figures, and body text, can be analyzed to recognize the logical structure of a book. Another method for layout recognition was proposed by Igarashi et al. [17], who used adaptive analysis to determine the implicit structures of card-handling editors. Regardless of these efforts, the development of a general structure analysis method for all document types is nearly impossible, a point highlighted in [18]. It is therefore essential to provide a novel retrieval method that is aware of domain-specific visual features for each type of e-book.

C. Query Expansion for e-books

Traditionally, query expansion is an effective way to help users by modifying their initial query [19]. There is a query expansion method targeting e-books [20] that adopts Wikipedia as the entry point for the book search. This method exploits the link graphs found in Wikipedia to define the relevance metric between books. To apply the query-expansion method to multimedia retrieval, concept-based query expansion [21], which provides high-level suggestions for expanding the original user query, is effective. For example, Hoque et al. [22] applied concept-based query expansion to the retrieval of web images by extracting a list of concepts that are related to the query from Wikipedia. Natsev et al. [23] proposed query expansion methods that integrate text-based keyword relevance analysis and content-based visual example analysis to identify the most relevant visual concepts for a given query. One example of a query expansion method applied to a domain-specific area is that of

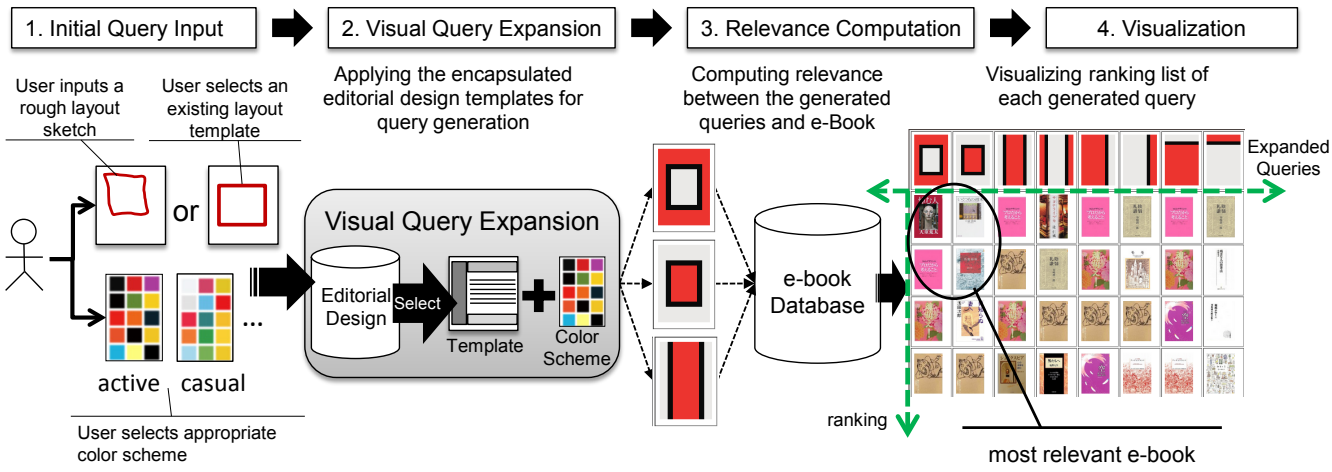

Figure 2. System Architecture of the Query-by-Appearance System.

Joly et al. [24], which describes visual query expansion for detecting specific logos in photographs.

Our Query-by-Appearance system differs from these existing query expansion methods in the following two aspects. First, as shown in Figure 1, Query-by-Appearance provides well-designed examples as alternatives to the initial query of a user. Second, in contrast to keyword-based approaches, Query-by-Appearance first retrieves the relevant editorial design templates, and then uses them to generate various query suggestions. This enables the system to provide well-designed visual queries for retrieving a suitable e-book. Moreover, although previous work on query expansion has mainly focused on the precision of the expanded queries, the main objective of this study is to demonstrate the effect of visual query expansion presentation methods by displaying the retrieval results in a two-dimensional ranking plot that incorporates layout-based and color-based relevance rankings.

III. Motivating Example

Our Query-by-Appearance method is used in following situation. For example, a user has 400 books in iPad News Stand. User tries to find book which title, author, published date is uncertain, but remember cover image vaguely. In this case, by using our system, user could retrieve desired book thorough inputting rough visual appearance. When inputting layout and color, it does not have to be correct and specific, because our system generates inputted query into nearest template. On the other hand, our system is not suitable for retrieving books which title and author is clear. In such case, user prefers to use keyword-based search method implemented in the conventional e-book stores and readers.

Another example is dedicated for a user experience aspect. A user draws a rough sketch on a touchscreen of smart devices, such as iPad and Android, according to his/her ambiguous memory. Then the system generates the well-designed e-book image data to retrieve search results. This approach is highly effective in web-based touch screen interfaces because it simplifies the query-input task. The touch screen interface visualizes both the layout similarity and color similarity in a two-dimensional matrix, enabling users to retrieve the desired e-book according to their preferences (i.e., the weights of the layout similarity and color similarity) by selecting a sub-matrix in the visualized two-dimensional matrix.

IV. System Architecture

Figure 2 illustrates the architectural overview of the proposed system. Query-by-Appearance provides an indirect e-book retrieval mechanism that is leveraged by a knowledge base of editorial design. Here "indirect" means that the system selects the relevant templates for query generation according to an initial query which consists of rough sketch and several colors. And then, the system generates actual queries by using the selected templates, which are well-designed and have coordinated colors. The system retrieves e-book images that are similar to the generated queries inspired by the initial rough query, and visualizes the results in a two-dimensional ranking. This query-assistance function enables users to input a layout and color scheme by using a web-based touch interface, allowing the search to be conducted intuitively through rough visual appearances, rather than technical knowledge.

The query templates consist of layouts and color schemes. The layout can be specified by a rough sketch or by selecting an appropriate template, and a well-designed color scheme, derived from a color-design book [25], is provided.

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

The Query-by-Appearance system for e-books performs the following six steps:

1. The user inputs a rough layout sketch on the HTML5 Canvas system.

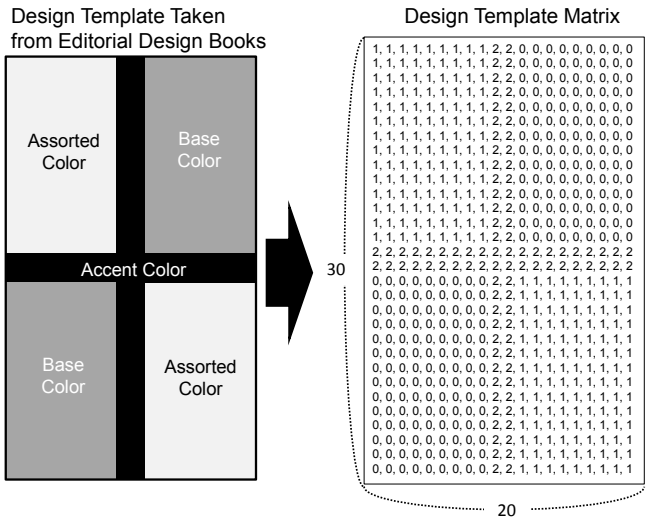

Figure 3. Structure of the Design Template Matrix.

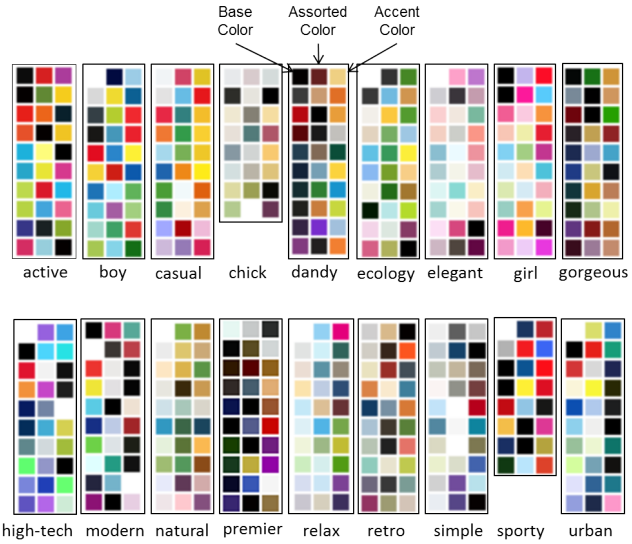

Figure 4. Color Schemes taken from an Editorial Design Book [25].

2. The system converts the rough layout query input into a matrix, which it compares with 30 templates stored in the knowledge base.
3. The system generates colored templates by adding color to the selected layout using the 20 color schemes in the knowledge base. The sets of colored templates are automatically expanded to image matrices that include the layout structure and the color scheme.
4. The system calculates the 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. E-book cover images that contain similar editorial design characteristics are visualized in a two-dimensional ranking plot with a generated query axis and a ranking axis.

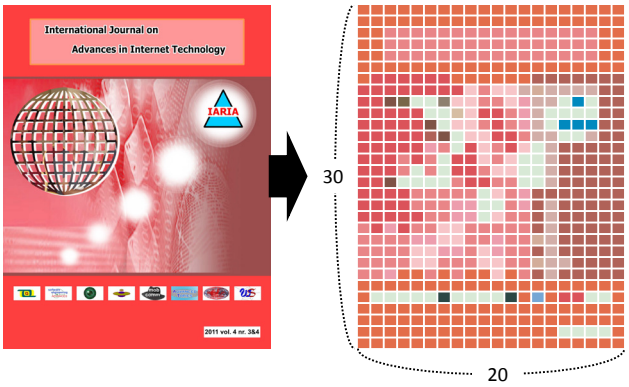

Figure 5. e-Book Matrix Generation.

6. Further retrieval is conducted when the user selects a color on a sub-matrix in the principal two-dimensional matrix.

V. Data Structures

This section shows data structures used in our visual query expansion method. Our system contains four fundamental data structures: A) Query Template Knowledge Base from Editorial Design, B) Query Matrix, C) e-book Database, and D) Visualization.

A. Query Template Knowledge Base from Editorial Design

Our editorial design knowledge base for visual query expansion consists of the layout template, and the color scheme. The left side of Figure 3 shows an example layout. The layout template matrix is the set of matrices. The right side of Figure 3 shows an example of a design template matrix. Layout information consists of a base color, an assorted color, and an accent color, which are assigned the numerical identifiers 0, 1, and 2, respectively. Each of color types defines the role of a color. the detailed and formal definition of those three colors are given later in this section. The layout template matrix T_L is defined as follows:

$$T_L := \begin{bmatrix} t_{[0,0]} & \cdots & t_{[0,m]} \\ \vdots & \ddots & \vdots \\ t_{[n,0]} & \cdots & t_{[n,m]} \end{bmatrix} \quad (1)$$

where $t_{[i,j]}$ indicates the color type at location $[i, j]$ in the template matrix. The system colors each column according to this number. It is important to mention that our Query-by-Appearance model does not assign specific colors to the layout matrix statically, but provides a dynamic color assignment mechanism to generate the query according to user demands. Thus, the layout template T_L only specifies the role of each color as the base color, assorted color, or accent color. According to these roles, the system dynamically assigns the actual colors defined in the color scheme template.

The color schemes are combinations of colors suitable for joint use. Figure 4 shows several examples of color schemas taken from an editorial design book [25]. Each color scheme consists of a base color, an assorted color, and an accent color. We have assigned an adjective to these color schemes in order to group them; the adjectives represent human color

Figure 6. Screenshot of the Prototype System for Visualizing e-book Search Results. In this screenshot, a user selected the layout template.

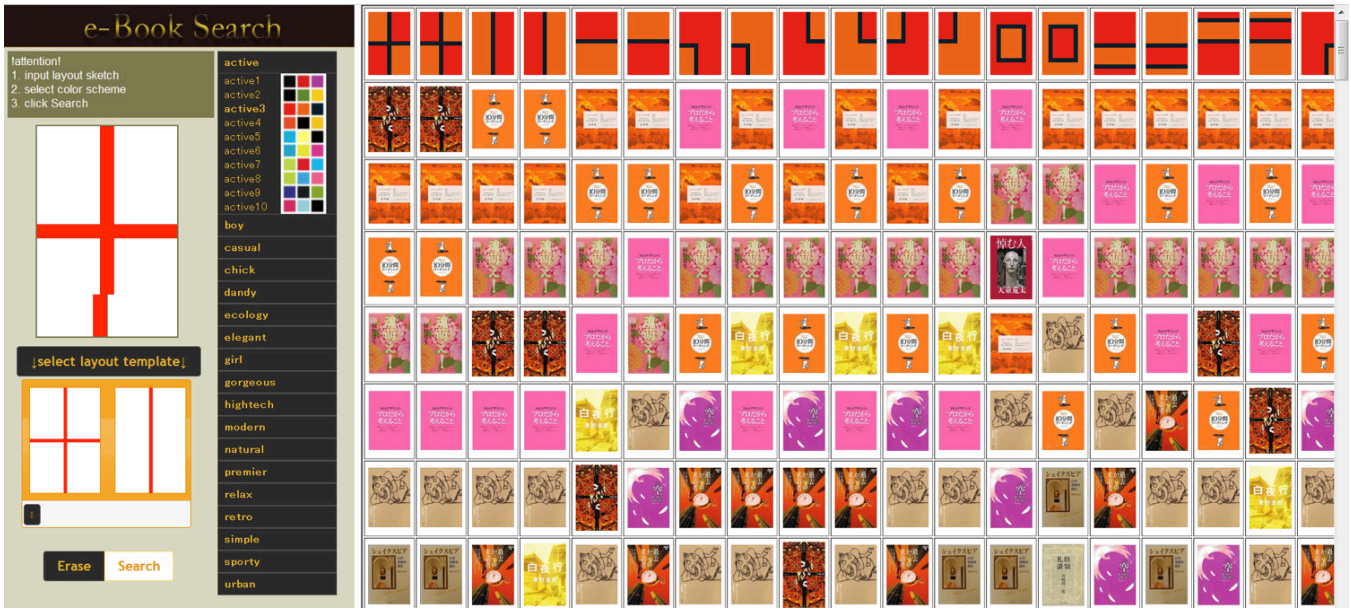

Figure 7. Screenshot of the Prototype System for Visualizing e-book Search Results. In this screenshot, a user sketched the layout to retrieve design templates.

perceptions. In total, we have defined 18 groups of color schemes by using a color scheme definition dictionary [25].

The color scheme is a triplet of colors, which is a combination of 102 primitive colors. As shown in Figure 4, each color scheme is assigned a specific adjective. Our system draws colors in each column according to this scheme, thus expanding the queries in order to generate e-book image data. The color scheme *CS* is defined as follows:

$$CS := \langle c_{base}, c_{assorted}, c_{accent} \rangle$$

where c_{base}, $c_{assorted}$, and c_{accent} denotes colors modeled in HSV color space. Our model represents colors using the HSV (hue, saturation, and value) color space. HSV is widely adopted in image and video retrieval because it describes perceptual and scalable color relationships. In order to select appropriate colors, we have selected 102 Munsell basic colors [26]. The details of those colors are as follows:

- A base color is the dominant color in the layout. It is used as the background or main theme color. This color occupies over 70% of the surface.
- An assorted color is the secondary dominant color. It has strong relationship with the base color. This color supports and enhances the impression of the base color. The assorted color occupies 25% of the surface.
- An accent color is a salient color that occupies less than 5% of the surface. It determines the structure and texture of the image. This color is very important because it is recognized as forming the basic shape of a figure drawn on the material.

B. Query Matrix

Our system converts each query into a matrix that represents the rough input of the user, such as simple lines drawn on HTML5 Canvas. We call this the query matrix, which is important for calculating the similarity between user input and layout templates in the knowledge base. The query matrix Q is defined as follows:

$$Q := \begin{bmatrix} w_{[0,0]} & \cdots & w_{[0,m]} \\ \vdots & \ddots & \vdots \\ w_{[n,0]} & \cdots & w_{[n,m]} \end{bmatrix} \quad (2)$$

where $w_{[i,j]}$ indicates whether a line is present at $[i, j]$ in the query. The system models an initial query in pixel level granularity because the templates are designed in a raster format. This is reasonable for dealing with many editorial design templates utilizing various design techniques such as blur effect and shadowing effect. Weights are assigned to each column as follows: if a line passes through a column, we assign a weight of 1.0 to the column; regions surrounding columns with a weight of 1.0 are assigned a weight of 0.5; if no line passes through a column, we assign a weight of 0.0. These weights are used to find the inner product of the user query matrix and the layout template matrix defined in the previous subsection.

C. e-Book Database

Our system converts JPEG-format image data into a matrix structure by clustering the image into the 102 system colors. We call the resultant data structure an e-book matrix. This is a 30 × 20 data matrix in which each cell represents the color at the corresponding point in the image, as shown in Figure 5. The e-book matrix is used to calculate the similarity with the expanded query matrix obtained from the user input and template knowledge base. The e-book database matrix D is defined as follows:

$$D := \begin{bmatrix} d_{[0,0]} & \cdots & d_{[0,m]} \\ \vdots & \ddots & \vdots \\ d_{[n,0]} & \cdots & d_{[n,m]} \end{bmatrix} \quad (3)$$

where $d_{[i,j]}$ represents each element of the matrix, and $[i, j]$ represent the rows and columns, respectively, of the image data in the matrix.

VI. Core Functions

Query-by-Appearance system consist of three core functions as follows: A) the visual query expansion function, B) the query-by-appearance function, and C) the visualization function. When the system receives an initial query, which is a rough sketch, the system applies the visual query expansion function to generate several queries by making the initial query well-formed according to the editorial design templates. And then the system applies the query-by-appearance function to calculate similarity scores for each generated query and each e-book in the database. Finally the system visualizes the retrieval result by ranking e-books in the descending order of scores.

A. Visual Query Expansion

Step-1: 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 knowledge base. The system uses the selected templates to generate actual queries. We call this "*query assistance* function". The inner product of the transformed user input matrix, which consists of sketched line data, and the template matrices stored in the template knowledge base is calculated in order to analyze the degree of similarity. Higher values of the inner product indicate greater similarity.

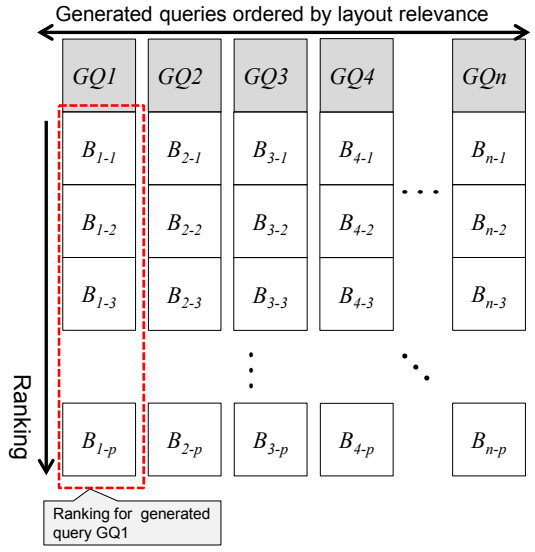

Figure 8. Visualization Matrix Structure.

The function is defined as follows:

$$f_{layout}(q, t) := \sum_{i=1}^{w} \sum_{j=1}^{h} q_{[i,j]} t_{[i,j]} \quad (4)$$

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 $[i, j]$ in the template matrix. This calculation measures the line similarity between the query matrix and the template matrix by checking all the pixels. This similarity measure is used for only choosing the relevant template, and is not used for actual e-book retrieval phase.

Step-2: Assigning Colors to the Template. Our system provides two strategies for assigning colors to a template. The first strategy allocates a base color to the dominant area of the template and allocates an assorted color to the secondary dominant area of the template. This strategy is called "straight color assignment". The straight color assignment strategy keeps original roles of the base color and the assorted color. In contrast, the second strategy allocates a base color to the secondary dominant area of the template and allocates an assorted color to the primary dominant area of the template. This exchange of the base color the assorted color is a well-known technique for creating variations of color scheme. The color assignment function is defined as follows:

$$f_{assign}(T_L, \mathrm{CS}) \rightarrow \mathrm{Q} := \begin{bmatrix} hsv_{[0,0]} & \cdots & hsv_{[0,m]} \\ \vdots & \ddots & \vdots \\ hsv_{[n,0]} & \cdots & hsv_{[n,m]} \end{bmatrix} \quad (5)$$

where T_L indicates the template data, and CS denotes a color scheme containing three HSV colors.

Step-3: Color Space Conversion. The system converts RGB color values of each image to HSV at the pixel level. We adopt the following well-known RGB to HSV conversion equation.

$$\mathrm{V} = \max(\mathrm{R}, \mathrm{G}, \mathrm{B}) \tag{6}$$

$$\mathrm{S} = 255 \times \frac{\max(\mathrm{R}, \mathrm{G}, \mathrm{B}) - \min(\mathrm{R}, \mathrm{G}, \mathrm{B})}{\max(\mathrm{R}, \mathrm{G}, \mathrm{B})} \tag{7}$$

$$\mathrm{H} = \begin{cases} 60\dfrac{B-G}{\max(\mathrm{R},\mathrm{G},\mathrm{B}) - \min(\mathrm{R},\mathrm{G},\mathrm{B})} & R == \max(\mathrm{R},\mathrm{G},\mathrm{B}) \\ 60\left(2 + \dfrac{R-B}{\max(\mathrm{R},\mathrm{G},\mathrm{B}) - \min(\mathrm{R},\mathrm{G},\mathrm{B})}\right) & G == \max(\mathrm{R},\mathrm{G},\mathrm{B}) \\ 60\left(4 + \dfrac{G-R}{\max(\mathrm{R},\mathrm{G},\mathrm{B}) - \min(\mathrm{R},\mathrm{G},\mathrm{B})}\right) & B == \max(\mathrm{R},\mathrm{G},\mathrm{B}) \end{cases} \tag{8}$$

B. 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:

$$f_{qba}(Q, B) \coloneqq \sum_{i=1}^{w} \sum_{j=1}^{h} \Delta_{godlove}\left(Q_{[i,j]}, B_{[i,j]}\right), \tag{9}$$

$$\Delta_{godlove} \coloneqq \frac{2S_1S_2\left(1 - \cos\left(2\pi\frac{|H_1 - H_2|}{100}\right)\right) + (|S_1 - S_2|)^2 + (4|V_1 - V_2|)^2}{2} \tag{10}$$

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 [27] to calculate the distance between two HSV colors. To improve the precision of color distance, the CIEDE2000 delta equation [28] can be applied. We adopt Godlove's delta equation here due to its faster computational performance.

Our system uses a two-dimensional ranking mechanism to present the results of the e-book image search. This visualization format enables users to select books on the basis of their layout and color similarity. The system samples template data from a wider field than the user's input in order to display the retrieval results interactively on the basis of the layout and color choices.

The two-dimensional ranking procedure employs the following steps: 1) The generated queries are shown on the horizontal axis. Each column contains a particular query template and color. 2) The ranking of e-book image data is shown on the vertical axis. 3) The system stacks the e-book image data according to the similarity score of queries and e-book metadata in each cell.

C. Visualization

The system visualizes the layout and color similarity using a two-dimensional ranking, as shown in Figure 6 and Figure 7. Results with similar editorial design characteristics are visualized in a grid layout, where the horizontal axis corresponds to the variation of expanded queries and the vertical axis corresponds to the relevance ranking.

The visualization model places the most relevant e-book in the upper-left position because this placement implies a high relevance scores for both the layout and color scheme. This ranking visualization scheme consists of two relevance scores for each cell. The ranking array $Cell_x$ is defined as follows:

$$Cell_x \coloneqq \langle L_i, C_j, \{B_{x-1} \cdots B_{x-p}\}\rangle \tag{11}$$

where L_i is the layout ID, C_j is the color ID, and $\{B_{x-1} \cdots B_{x-p}\}$ determine the ranking order in a column. The order of books in a certain column depends on their color relevance. Figure 8 shows the basic layout of this visualization model. Figure 6 and Figure 7 show screenshots of prototype system implementation that visualize the above model. In this model, each cell represents a specific "rank" of the corresponding e-book. Our system generates the ranking according to the relevance score, which is calculated with inner product, by sorting them in a descending order.

VII. System Implementation

We have implemented a prototype of the Query-by-Appearance system that calculates the similarity between generated queries and editorial design templates. The screenshots of the prototype, which uses HTML5 Canvas and JavaScript, is shown in Figure 6 and Figure 7. In Figure 6, a user selected the layout template, and the system generated the query according to the selected layout and color schemes. In Figure 7, a user sketched the layout to retrieve design templates, and the system listed the ranking of the layout templates and assigned the colors for the retrieval results. An important feature of this implementation is that it uses standard web technologies, which are also used by the current EPUB3 specification. Thus, we can apply our method to EPUB3-based e-books without significant modification.

Figure 9 shows the detailed architecture of our prototype system, which specifically includes the modern HTML5 technologies Canvas API, Web Storage API, and Web Workers API. The system consists of the following three modules: a query editor, a visual query expansion system, and a search result visualization engine. We now describe these components in detail.

The main user interface is the query editor, which uses HTML5 Canvas to produce two-dimensional interactive vector graphics. Users sketch the layout query on the query editor, and select a color scheme. When they have finished drawing, the query editor encodes the figure into JavaScript Object Notation (JSON) format and passes it to the visual query expansion system. This procedure allows our system to share the JSON-encoded figure among multiple web workers to parallelize the execution of visual query expansion.

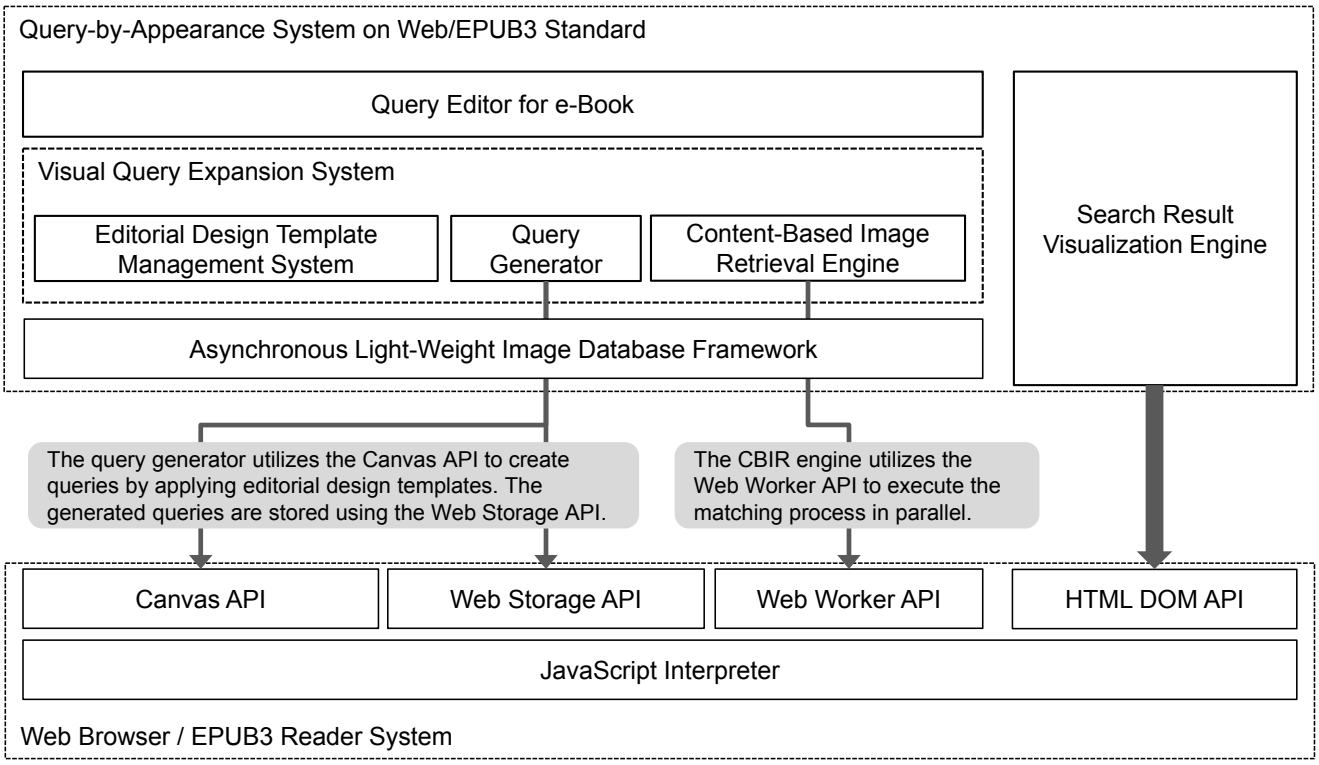

Figure 9. Prototype System Architecture of the Query-by-Appearance System Implemented Using Modern HTML5 Technologies.

The visual query expansion system is a core component of the proposed system. It expands the input query into well-designed layout and color information by retrieving relevant editorial design templates and assigning colors to the retrieved templates. The visual query expansion system consists of the following sub-modules: an editorial design template management system, a query generator, and a content-based image retrieval engine. The editorial design templates are embedded in an HTML file as plain text, and the editorial design template management system parses this file to reconstruct the editorial design template objects. By storing them in HTML5, we can deploy various templates according to the e-book content. The query generator applies the retrieved templates to generate actual queries, and stores the generated queries using the Web Storage API in order to reduce the communication overhead between web browsers and the server. Finally, the content-based image retrieval engine executes the generated queries over the database contents. This retrieval process is parallelized by the Web Workers API, and the retrieved e-books are presented to the user by the search result visualization engine.

The most important feature of this architecture is that the query expansion process and the data retrieval process can be executed on the client side. This client-side execution framework is beneficial as it permits our method to retrieve contents that are inside a book. In addition, this system spawns real OS-level threads from the Web Workers API to parallelize the e-book retrieval process. Modern HTML5 technologies enable us to implement complex processes in web browsers.

VIII. EVALUATION

This section evaluates the effectiveness of our Query-by-Appearance system when applied to the existing books. The experimental data comprise 130 book cover images from Amazon.co.jp. We perform two evaluation experiments: Experiment 1 is an evaluation of retrieval precision, and Experiment 2 is an evaluation of the effectiveness of our system by measuring the time consumed for retrieving the desired book by using our system and the conventional book search system.

A. Experiment-1: Outline of Experimental Studies

In this section, we evaluate the effectiveness of our system by examining its retrieval precision. The aim of this experiment is to clarify the effectiveness of retrieving book cover images by utilizing the editorial design methodology. We compare the retrieval precision using two search methods: 1) queries are generated using only layout templates and 2) queries are generated by integrating layout templates and color scheme templates. We show that the integration of a layout template and a color template makes a significant contribution to the e-book retrieval result.

For this experiment, we prepared: 1) 130 book cover images from Amazon.co.jp, 2) five queries, and 3) five answer data sets for each query. The queries and answer data sets were formed in consideration of the basic structures of editorial design: a symmetrical layout, a diagonal layout, and a layout with a gravity point in the center. Thus, the following five queries were chosen (see Figure 10):

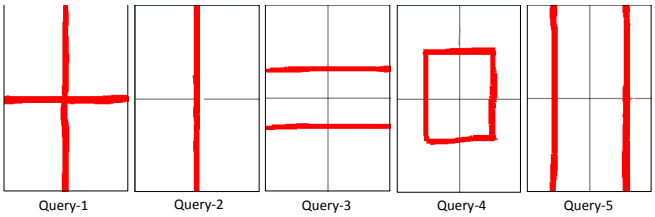

Figure 10. Five Experimental Queries drawn on the Canvas.

Figure 11. Top 20 Retrieved Images for the Combination of Query 3 and "cool" (left), and Query 1 and "sporty" (right).

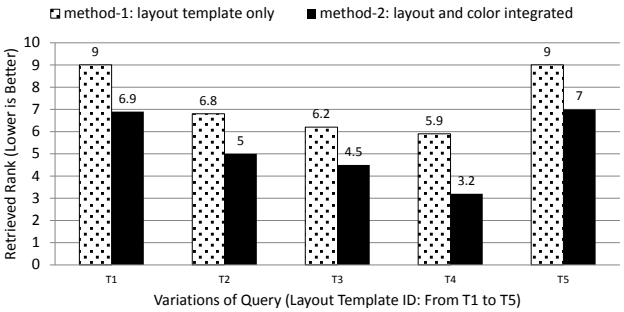

Figure 12. Retrieved Rank Scores Using Search Method 1 (left-hand bars) and Search Method 2 (right-hand bars).

Query-1. Draw a cross shape to divide the canvas into four sections.

Query-2. Draw a vertical line along the center of the 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 three colors (as shown in Figure 3 and 4). The system calculates the similarity between the queries and the generated e-book image data.

Figure 13. Ten Books to be Retrieved in Experiment-2 (Images are Shaded Off for Protecting Copyrights).

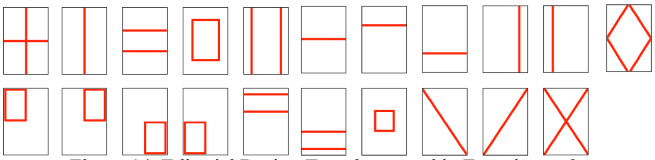

Figure 14. Editorial Design Templates used in Experiment-2.

B. *Experiment-1: Experimental Results*

In this section, we evaluate the retrieval precision of the five queries specified in the previous subsection. Figure 11 shows some retrieval results for a combination of a query and color scheme. Our approach assigns two scores for the retrieved rank. The first is the average number of correct answers in the top 20 that were retrieved using only template relevance. The second score is the average rank of correct answer data that has the chosen color scheme. Our approach assigns a score based on the retrieved ranking. Thus, we obtain one score for the average correct ranking using layout relevance alone (i.e., from search method 1) and one for the average rank of correct answer data when color scheme templates were integrated (i.e., search method 2).

Figure 12 shows the average scores of the top 20 ranking for each query. This result shows that search method 2 (integrating both layout and color templates) achieves better retrieval precision than method 1 (using the layout template only), as a superior (lower) rank has been assigned to results that correspond to the query. This shows that our system effectively retrieves e-book images using layout templates and color schemes.

C. *Experiment-2: Outline of Experimental Studies*

In this section, we evaluate the effectiveness of our system by measuring the time consumed for retrieving the desired book by using our system and the conventional book search system. The aim of this experiment is to clarify the effectiveness of our visual query expansion method by comparing it with the exiting book retrieval system. We have compared our system and the keyword-based search engine implemented in Amazon.co.jp. It is important to mention that this experiment assumes that a user forgets the details about the target book and cannot submit the book's title correctly. In such a situation, our visual query expansion method

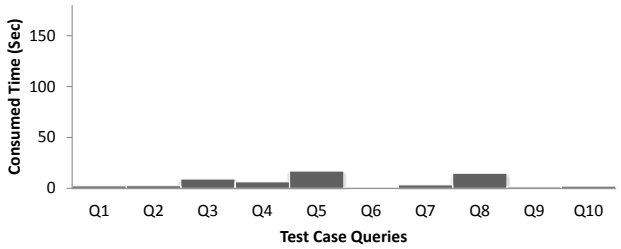

Figure 15. Time Consumed to Retrieve the Correct Result by using Query-by-Appearance (Method-1).

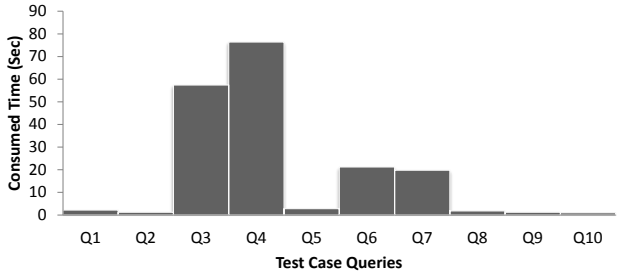

Figure 16. Time Consumed to Retrieve the Correct Result by Submitting an Author's Name as a Query for Amazon.co.jp (Method-2).

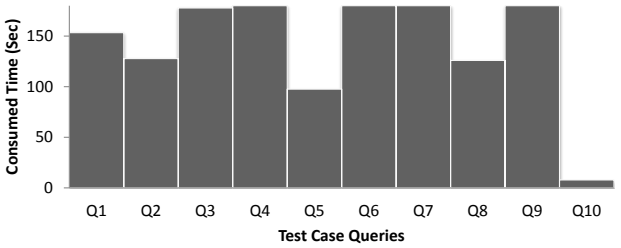

Figure 17. Time Consumed to Retrieve the Correct Result by Submitting a Publisher's Name as a Query for Amazon.co.jp (Method-3).

Figure 18. Time Consumed to Retrieve the Correct Result by Submitting a Publication Date as a Query for Amazon.co.jp (Method-4).

effectively contributes to user's retrieval process. We compare the time consumed to find the correct result by using four search methods: 1) queries are generated using Query-by-Appearance system, 2) publisher's name is submitted for Amazon.co.jp, 3) author's name is submitted for Amazon.co.jp, and 4) publication date is submitted for Amazon.co.jp. We have asked five casual people for five test subjects to find ten books by using those four methods. Images of ten books are shown in Figure 13. Figure 14 shows the editorial design templates used in Experiment-2. We have measured the time (second) until each test subject find the correct result. We show that our visual query expansion method and visualization method make a significant contribution to reduce the time required to find out the correct result.

D. Experiment-2: Experimental Results

In this section, we evaluate the retrieval time of the ten queries by using four methods specified in the previous subsection. We plot the averaged time of five test subjects for each query. When a test subject could not find the correct result, or gave up finding the book, we considered it as consumption of 180 seconds. Figure 15 shows the time consumed to find the correct result by using Query-by-Appearance system. Figure 16 shows the time consumed to find the correct result by submitting author's name into Amazon.co.jp. Figure 17 shows the time consumed to find the correct result by submitting publisher's name into Amazon.co.jp. Figure 18 shows the time consumed to find the correct result by submitting publication date into Amazon.co.jp.

The average time consumed for searches carried out using our Query-by-Appearance system (Method-1) was 6 seconds; the average time consumed for searches carried out using author's name (Method-2) was 19 seconds; the average time consumed for searches carried out using publisher's name (Method-3) was 104 seconds; the average time consumed for searches carried out using publication date (Method-4) was 141 seconds. This confirms that our visual query expansion method and visualization method significantly improve the e-book retrieval efficiency when it is compared with the existing search engine.

IX. CONCLUSIONS AND FUTURE WORK

This paper proposed the *Query-by-Appearance* system for e-books. The system provides an intuitive and effective query input and visual retrieval method based on the similarity of the overall layout and color scheme. A unique feature of this system is its visual query expansion function, which exploits structural design and analysis from the field of editorial design. This assistant mechanism is intuitive because it allows users to find a desired e-book by submitting a simple query. Our visual query expansion function is a design-based template-matching function that compares a user's rough query with the embedded editorial design templates. We have performed an evaluation of the effects of our visual query expansion method on realistic search tasks. Experimental results showed that visually expanded queries perform a better and more consistent performance than manually conducted queries. In future work, we plan to develop a prototype system that supports the full specifications of EPUB3, and to perform a feasibility study by evaluating the scalability and effectiveness of our approach when applied to existing e-books.

ACKNOWLEDGMENT

This research was 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] Koike, Y. and Kurabayashi, S., "Query-by-Appearance System for Style-Oriented Media Retrieval," In Proceedings of the Seventh International Conference on Internet and Web Applications and Services (ICIW 2012), pp.290–295, Stuttgart, Germany, May 27-June 1, 2012.

[2] 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.

[3] 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.

[4] American Association of Publishers, "Publishing Industry Has Strong January Revenue Growth In Print Books And Ebooks For All Audiences," 2012, http://www.publishers.org/press/62/, (accessed 2012-8-6).

[5] Datta, R., Joshi, D., Li, J., Wang, J.Z., "Image retrieval: Ideas, influences, and trends of the new age," ACM Comput. Surv. 40, pp. 1–60, 2008.

[6] Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain, R., "Content-based image retrieval at the end of the early years," IEEE Transactions on Pattern Analysis and Machine Intelligence, 22, pp.1349–1380, 2000.

[7] Kherfi, M.L., Ziou, D., Bernardi, A., "Image Retrieval from the World Wide Web: Issues, Techniques, and Systems," ACM Computing Surveys, 36, pp. 35–67, 2004.

[8] Shipman, F.M., Marshall, C.C., Moran, T.P., "Finding and using implicit structure in human-organized spatial layouts of information," In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI'95. pp. 346–353. ACM, 1995.

[9] Woodruff, A., Faulring, A., Rosenholtz, R., Morrsion, J., Pirolli, P., "Using thumbnails to search the Web," In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI'01. pp. 198–205. ACM, 2001.

[10] Abdullah, N., Gibb, F., "Using a Task-Based Approach in Evaluating the Usability of BoBIs in an E-book Environment," In Proceedings of the 30th European Conference on Information Retrieval, pp. 246–257. Springer-Verlag, 2008.

[11] Kudělka, M., Takama, Y., Snášel, V., Klos, K., Pokorný, J., "Visual Similarity of Web Pages," In Proceedings of the 6th Atlantic Web Intelligence Conference - AWIC 2009, pp. 135–146. Springer-Verlag, 2010.

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

[13] Tang, Y.Y., Yan, C.D., and Suen, C.Y., "Document Processing for Automatic Knowledge Acquisition," IEEE Trans. on Knowledge and Data Engineering, 6(1), pp. 3–21, 1994.

[14] Gao, L., Tang, Z., Lin, X., Liu, Y., Qiu, R., and Wang, Y., "Structure extraction from PDF-based book documents," In Proceeding of the 11th Annual International ACM/IEEE Joint Conference on Digital Libraries - JCDL'11, pp. 11–20, 2011.

[15] He, F., Ding, X., and Peng, L., "Hierarchical Logical Structure Extraction of Book Documents by Analyzing Tables of Contents," In Proc. of the International Conference on Document Recognition and Retrieval XI, pp. 6–13, 2004.

[16] Lin, C., Niwa, Y., and Narita, S., "Logical Structure Analysis of Book Document Images Using Contents Information," In Proc. of the 4th International Conference on Document Analysis and Recognition, ICDAR'97, pp.1048–1054, 1997.

[17] Igarashi, T., Matsuoka, S., Masui, T., "Adaptive recognition of implicit structures in human-organized layouts," Proceedings of the 11th IEEE International Symposium on Visual Languages, pp. 258–266, 1995.

[18] Lee, K.H., Choy, Y.C. and Cho, S.B., "Logical Structure Analysis and Generation for Structured Documents: A Syntactic Approach," IEEE Transaction on Knowledge and Data Engineering, 15(5), pp.1277–1294, 2003.

[19] Harman, D., "Relevance feedback revisited," In Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 1–10, 1992.

[20] Koolen, M., Kazai, G., Craswell, N., "Wikipedia pages as entry points for book search," In Proceedings of the Second ACM International Conference on Web Search and Data Mining - WSDM '09. pp. 44-53. ACM Press, 2009.

[21] Fonseca, B.M., Golgher, P., Pôssas, B., Ribeiro-Neto, B., Ziviani, N., "Concept-based interactive query expansion," In Proceedings of the 14th ACM International Conference on Information and Knowledge Management - CIKM'05, pp. 696, 2005.

[22] Hoque, E., Strong, G., Hoeber, O., Gong, M., "Conceptual Query Expansion and Visual Search Results Exploration for Web Image Retrieval," In Proceedings of the 7th Atlantic Web Intelligence Conference, AWIC 2011, pp. 73–82. Springer-Verlag, 2011.

[23] Natsev, A. (Paul), Haubold, A., Tešić, J., Xie, L., Yan, R., "Semantic concept-based query expansion and re-ranking for multimedia retrieval," In Proceedings of the 15th ACM International Conference on Multimedia, pp. 991–1000, 2007.

[24] Joly, A., Buisson, O., "Logo retrieval with a contrario visual query expansion," In Proceedings of the 17th ACM International Conference on Multimedia, pp. 581–584, 2009.

[25] Totogawa. Y., "Color Scheme & Color Design," ISBN: 978-4797359237, Softbank Creative, 2012 (In Japanese).

[26] Newhall, S.M., Nickerson, D., and Judd, D.B., "Final Report of the O.S.A. Subcommittee on the Spacing of the Munsell Colors," Journal of the Optical Society of America, 33(7), pp. 385-411, 1943.

[27] Godlove, I.H. "Improved Color-Difference Formula, with Applications to the Perceptibility and Acceptability of Fadings," Journal of the Optical Society of America, Vol. 41, NO. 11, pp.760-770, 1951.

[28] Luo, M.R., Cui, G., and Rigg, B., "The development of the CIE 2000 colour-difference formula: CIEDE2000," Color Research & Application, 26(5), pp. 340-350, 2001.