UDC:004.65
Original scientific paper

IMPROVING RETRIEVAL EFFECTIVENESS BY INTEGRATED QUERYING OF IMAGES AND TEXT

***Zoran Stejic**

Yasufumi Takama

Kaoru Hirota

Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology 4259 Nagatsuta, Midori-ward, Yokohama 226-8502, Japan

stejic@hrt.dis.titech.ac.jp

Abstract: *An integrated retrieval system is developed, enabling simultaneous querying of the image and the text content of a document — a feature not supported by the existing retrieval systems. An integrated query is expressed using the natural language expressions. Image content is queried either based on the color-spatial properties or based on the induced sensations, derived from the color arrangements. The vagueness of the natural language expressions used to query the image content is modeled employing the fuzzy set theory. Integrated retrieval experiment, performed on a database of 1,022 Web pages containing images and text, shows that, when using the integrated queries, the number of retrieved documents in average reduces to 14% of the number obtained by the image or text queries. Image retrieval experiment, involving a database of 1,100 color images and evaluated using a questionnaire involving 10 subjects, shows 70% agreement between the system's and the user's ranking of images. It follows that the proposed integrated querying approach, compared to the conventional image and text querying approaches, enables a more effective retrieval, essential for the large-scale databases like the Internet.*

Keywords: *integrated querying, image and text retrieval, color-spatial properties, sensations.*

1. INTRODUCTION

Presently, more than 1.7 billion of Web pages are directly accessible through the Internet [6]. Around 75% of the data those pages contain is visual [7] — images of various sorts and videos. However, in a Web page, images almost always appear with the accompanying text. While a number of image retrieval [7, 4, 2, 10] and text retrieval [14, 1] systems are available, none of the existing retrieval systems considers the images and the text in an integrated manner, even though the two media together constitute the content of a multimedia document.

This fact limits the precision of the existing retrieval systems, since the user looking for the information given partly in visual and partly in textual form, cannot completely express the desired information through the query. In the case of retrieval from a large-scale database like the Internet, this can lead to a huge number of retrieved documents, even in the case when the documents user is actually looking for are not numerous.

The proposed system addresses this issue, enabling the *integrated querying* of images and text. Integrated querying enables more precise queries and a more complete access to the document contents, since the information carried by the image complements that carried by the text, and vice versa. For example, in a page from an online car catalog, while the photo is simply showing a "red racing car", the accompanying text reveals that the car is "Toyota Corolla, produced in 1995." Since the same catalog might contain hundreds of photos of red racing cars, and the accompanying data about the whole series of Toyota models, efficient retrieval is possible only if both the image contents and the text contents are used in an integrated manner.

An integrated query consists of the *image sub-query* and the *text sub-query*, both expressed using natural language (NL) keywords and phrases. Image sub-query addresses the image content: either the *color-spatial properties* (e.g., "*upper half* of the image is *mainly* covered by *light blue*"), or the *induced sensations* (e.g., "image *atmosphere* is *cold*"), derived from the spatial arrangement of colors. Text sub-query includes *all*, *any* or *phrase* type of queries, checking whether the document text contains a given set of keywords, one of the several keywords, or the exact sequence of keywords, respectively.

In the case of image sub-queries, fuzzy set theory is employed to model the inherent vagueness of the query keywords referring to the image content (e.g., "light blue" or "cold atmosphere").

In the document indexing stage, relevant image and text features are automatically extracted from the documents and stored in a form suitable for the comparison with the query keywords. Relevant image features are the color-spatial properties and the induced-sensations, while the relevant text features refer to the words appearing in the documents (keyword-based text indexing).

In the retrieval stage, documents are ranked with respect to the integrated query by first considering the text content, involving a crisp relevance judgment, and then performing the ranking of the resulting relevant documents based on the image content similarity.

Experiment using the integrated queries to retrieve the Web pages has shown that the number of retrieved documents in average reduces to 14% of the number obtained by the image or text queries. Furthermore, results of the image sub-queries show 70% agreement between the system's and the user's ranking of images.

The proposed integrated querying approach — not supported by any existing image or text retrieval system — enables a more precise and complete access to the multimedia document contents, resulting in a more effective retrieval, essential for the large-scale databases like the Internet. In addition, the natural language expressions-based querying gives user a uniform access to both the image and the text contents, making the retrieval more intuitive.

In Section 2. the idea of the integrated querying is elaborated and the proposed integrated retrieval system model is presented. Section 3. describes the image and text indexing and retrieval techniques proposed to realize the integrated retrieval. Finally, Section 4. contains the analysis of the system performance, based on the retrieval experiment.

2. INTEGRATED QUERYING AND MULTIMEDIA IR SYSTEM

2.1. Integrated querying concept

The multimedia documents targeted by the proposed integrated retrieval system contain the images accompanied by the text, and include typical Web pages.

To *accurately* and *completely* query the contents of such multimedia documents, both the image and the text content must be accessible through the query. Moreover, the contents of the two media must be accessible in the same way.

We propose an integrated querying approach, where both the image and the text content can be queried using the natural language keywords and phrases. Keyword-based querying, as a common feature of the text retrieval systems [14], as well as of some image retrieval ones [7], is intuitive and therefore suitable even for the non-experienced users. Even for querying the image content, keyword-based approach proves to be more usable than the more widespread query-by-example approach. Namely, except in the trivial cases, finding the query image by browsing a large database, or drawing the query image, is difficult compared to entering a couple of keywords addressing the image content. In addition, expressing the contents of images and text through the common medium — natural langauge — enables the "fusion" of these contents, and discovering the information not given explicitely through either images or text.

Querying the text content by the NL keywords in the simplest case doesn't require the transformation of the text, since the query keywords can be directly compared with the keywords appearing in the text. On the contrary, querying the image content requires that the image properties are extracted and converted to a form suitable for the comparison with

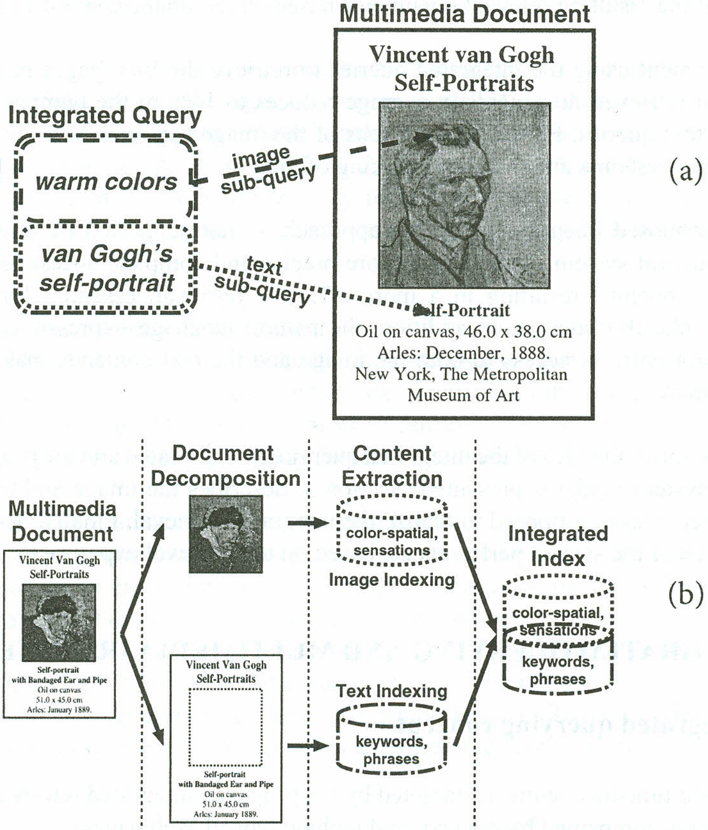


Figure 1: (a) Integrated query and multimedia document; (b) Integrated indexing of a document

the query keywords. Consequently, for the realization of the integrated retrieval, the most important part of the system is the image indexing. Choice and significance of the image properties for retrieval are elaborated in Section 3.1.

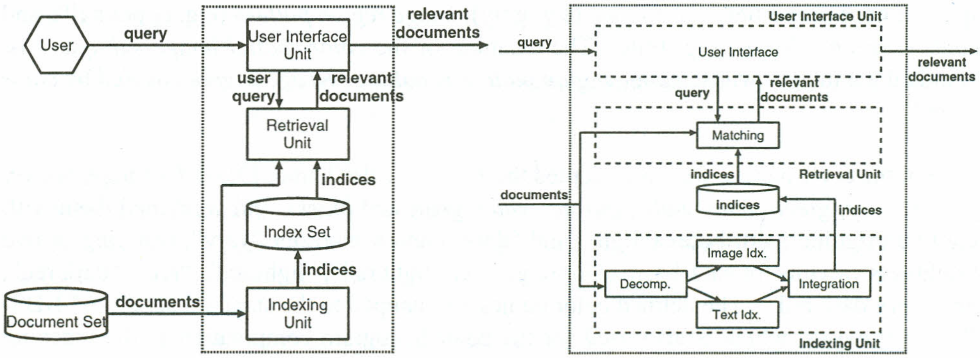
An integrated query (Figure 1-a) consists of the *image sub-query* referring to the image, and the *text sub-query*, referring to the text part of a document. The system, using the integrated document index (Figure 1-b), separately evaluates the similarity of the document's image and text content to the respective parts of the query, finally outputting the integrated document-query similarity degree. Similarity degrees assigned to the database documents define the ranking of the documents with respect to the query, from more to less relevant ones (Section 3.3.).

2.2. MULTIMEDIA IR SYSTEM MODEL

To realize the integrated retrieval of images and text, the proposed system incorporates the functionality of both the image and the text retrieval systems, accompanied by the functions realizing the integration of the two.

The proposed system consists of the indexing unit, the retrieval unit, and the user interface unit (Figure 2-a). The indexing unit produces a compact representation of the image and text contents — an *integrated document index*. The retrieval unit, using the integrated document indices, matches the documents to the user's query (*query evaluation*), and retrieves the relevant documents from the database. Finally, the user interface unit handles the interaction with the user — entering the query, displaying the relevant documents, etc.

The indexing unit (Figure 2-b) is made up of the decomposition, the image indexing, the text indexing, and the integration components. The decomposition component decomposes the document into the images and the text. The image and the text indexing components extract the contents of the corresponding media and produce the image and text indices, respectively. The integration component merges the two media indices into an integrated document index.



(a) (b)

Figure 2. (a) The structure of the proposed integrated retrieval system; (b) System structure on the components level

The retrieval unit (Figure 2-b) consists of the matching component. Matching component matches the integrated document indices against the user query (computing the document-query similarity degrees), ranks the documents accordingly, and retrieves the documents the most similar to the query (the *relevant* documents).

3. MULTIMEDIA DOCUMENT INDEXING AND RETRIEVAL

3.1. Image indexing and retrieval

The purpose of the image indexing is to extract the image content of a document and produce the image index. At the retrieval time, the image index is compared to the keywords in the image sub-query, resulting in the *image-query similarity degree* (e [0,1]), which defines the relevant images.

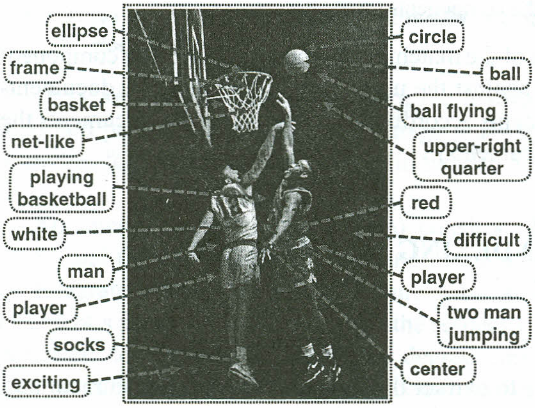
Since the image content is queried using NL expressions, only the image properties expressible in the NL are considered. Figure 3 summarizes the image properties [3], giving some examples related to the photo of the two basketball players in action. Considering the expressibility in the NL, and the computational cost of extracting the image properties [11], the properties chosen for the indexing are: the *color-spatial properties*, and the *color-induced sensations*. In the remainder of the section, the proposed image indexing and retrieval techniques are elaborated.

3.1.1. Color-spatial Properties Indexing and Retrieval

The purpose of the color-spatial indexing is to extract the color-spatial information from the image in a form suitable for the keyword-based querying and the retrieval.

Querying. Color-spatial information answers to the questions like: "What color is the upper half of the image?" or "To what extent is the image covered by gray color?". These two example questions contain the three types of terms that can be used to query the color-spatial image properties: *color name* (e.g. gray), *image region pattern* (e.g. upper half), and *color coverage degree* (e.g. little). The meaning of the *color-spatial image sub-query* is: "Find all the images whose *image region pattern* is *color coverage degree* covered by *color name*".

For the color names, we have adopted the 13 basic color names [9] *(red, orange, brown, yellow, olive, green, blue, violet, purple, white, gray,* and *black),* and combined them with the two brightness attributes ("light" and "dark") and a modifier "very", resulting in five brightness variants for each basic color (e.g., "very light red", "light red", "red", "dark red", and "very dark red"). The defined color names are mapped to the standard set of 267 NBS-ISCC [9] colors, which is also used for the color histogram computation in the indexing stage.



(a)

Image Property	*Example*	*NL*	*EE*
perceptual properties (low-level)			
color	"white", "red"	√	√
texture	"net-like"	–	–
shape	"ellipse", "circle"	√	–
spatial relationships	"upper-right quarter", "center"	√	√
semantic primitives (high-level)			
objects roles scenes	"basket", "man", "socks", "ball" "player" "playing basketball", "ball flying"	√	–
induced sensations (high-level)			
impressions emotions	"difficult" "exciting"	√	√*
Legend	*NL*: expressible in natural lang. *EE*: *easily* extractable		

(b)

Figure 3. (a) Photo illustrating the image properties expressed in the NL; (b) Image properties with examples, expressibility in natural language, NL, and ease of extraction, EE (* color-induced sensations, a special case of the induced sensations solely derived from the color arrangements, can be extracted at low computational cost, see Section 3.1.2.)

Regarding the image region patterns, we partitioned the image area at the different levels of detail and assigned intuitive names to the obtained image region patterns. The result of image partitioning are the 9 *basic region patterns* (Figure 4-a), along with the names: "image", "center", "border", "half, "diagonal", "quarter", "belt", "edge", and "corner". *Region patterns* are obtained by combining the basic region patterns with the the position (e.g., *"upper-left* corner"), and width (e.g., "narrow center", "wide center") modifiers (Figure 4-b).

Coverage degree of an image region by a color can be "little", "half, or "mainly", with the obvious meanings. These three expressions are modeled as the fuzzy sets.

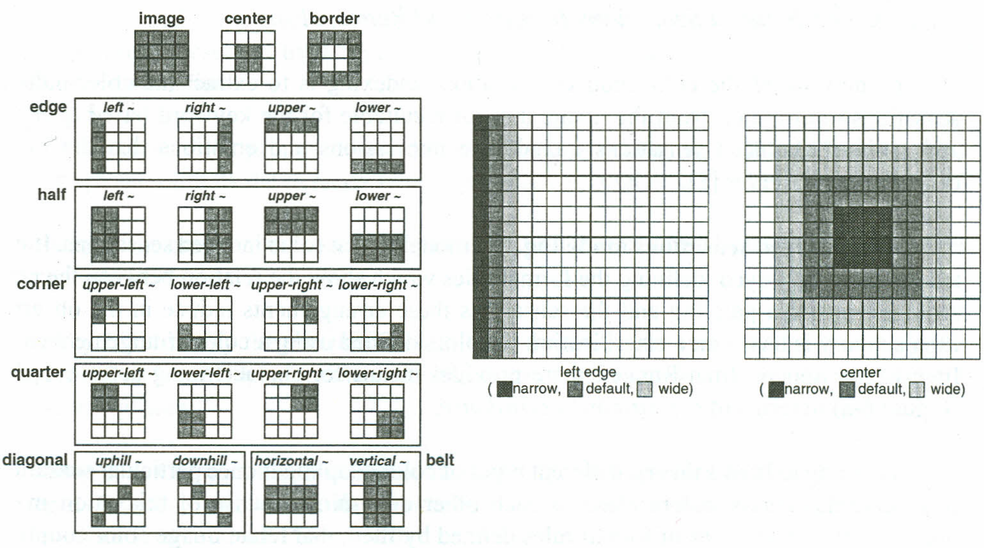
Indexing and retrieval. For indexing the color information, color histogram [12] is used, showing the distribution of the colors appearing in the image. The main deficiency of the color histogram — lack of the spatial information [3] — is compensated by the proposed algorithm, as elaborated in the following.

The outline of the proposed **color-spatial properties indexing algorithm,** which processes the image and outputs the *color-spatial properties index,* is (Figure 5-a):

Step 1-1. Uniformly partition the image into 16x16 blocks.

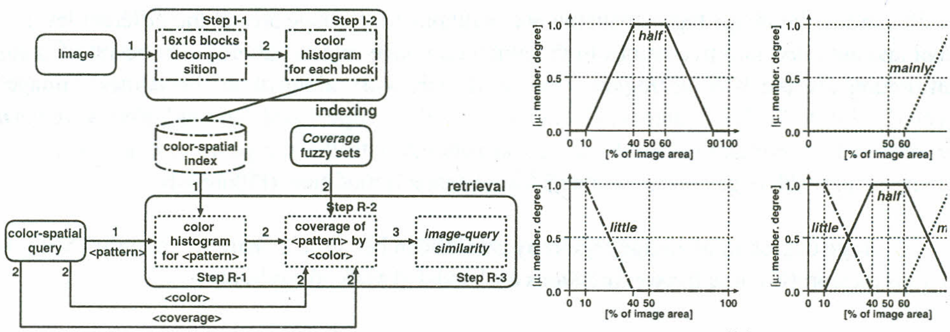
Step 1-2. Compute a color histogram for each obtained block. Color histogram is computed using the perceptually uniform MTM color space [8], being the closest to the human perception among the known color spaces [3].

The resulting 256 color histograms are used as the color-spatial properties index.



(a) (b)

Figure 4. (a) Image region patterns and NL names; (b) Image region patterns on 16 x 16 grid



(a) (b)

Figure 5. (a) The proposed algorithm for color-spatial properties indexing and retrie\
(b) Fuzzy sets for "coverage"

The outline of the proposed **color-spatial properties retrieval algorithm,** which uses image index to compute the *image-query similarity degree* (e [0,1]), is (Figure 5-a):

Step R-1. Compute the color histogram of the queried image region pattern, using block histograms stored in the image index.

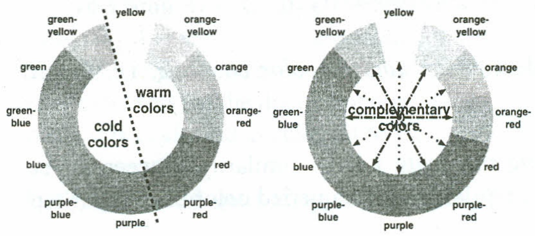
Step R-2. Based on the queried color and coverage degree, compute the image-qu similarity by evaluating the obtained color histogram using the fuzzy sets for "covera; (Figure 5-b).

3.1.2. Color-induced Sensations Indexing and Retrieval

The purpose of the color-induced sensations indexing is to extract the color-indu sensations information from the image in a form suitable for the keyword-based query and retrieval. Induced sensations include the impressions and emotions that the im: causes in the observer [3].

Color-induced sensations modeling. For modeling the color-induced sensations, Itte theory [5] is employed, defining the formal rules which show the relation between the cc arrangements in a painting and the sensations these arrangements induce in the obser Itten's theory is based on a classification of colors defined using a color arrangement cal Itten-Runge sphere. Itten-Runge sphere provides means for characterizing *color coupl* (Figure 6-a) in terms of *contrast* and *accordance.*

According to Itten's theory, different types of color coupling induce particular sensati (e.g., complementary colors close to each other determine a sense of relaxation in observer, etc.). Examples of formal rules defined by Itten, that relate image color coupli to the sensations, are shown in Figure 6-b.



(a)

Sensation	Color coupling
warmth	from yellow to purple-red
coldness	from purple to green-yellow
relaxation	complementary colors close to each other
uneasiness	no contrast (reinforced by dark yellow and purple)
liveliness	harmonic accordance

(b)

Figure 6: (a) The core of the Itten-Runge sphere with *warm-cold* and *complementary* color couplings; (b) Itten's rules showing the sensations induced by the color couplings

Itten's theory has been employed for the image retrieval by Corridoni [2], whose system is oriented towards the experts in the field of art, and the queries, expressed in a formal language, refer to the high level color properties like the complement, harmonic accordance, etc. On the contrary, we use Itten theory to enable querying of the induced sensations by the simple NL keywords.

Querying. In the *induced sensations image sub-query,* the terms used for querying the color-induced sensations are directly adopted from the Itten's theory. The queriable sensations referring to the *image atmosphere* are: "warm", "cold", "relaxing", "uneasy", and "lively".

Indexing and retrieval. In order to extract the induced sensations from an image, Itten's rules are mapped to the color-spatial image properties. This makes it possible to identify the color coupling present in the image and, based on it, to compare the image sensation with the queried sensation.

The outline of the proposed **induced sensations indexing algorithm,** which processes the image and outputs the *induced sensations index,* is (Figure 7-a):

Step 1-1. Compute the color histogram of an image using the 12 pure colors of the Itten-Runge sphere.

Step 1-2. Identify the *dominant colors* from the histogram (Step 1-1). Dominant colors are defined as (at most) 5 colors with the highest proportion, each covering not less than 1/12 of the image area.

Step 1-3. Compute the percentage of the image area covered by the dominant colors (Step 1-2) and, based on it, the *dominance degree* of the dominant colors (using the appropriate fuzzy set, see Figure 7-b).

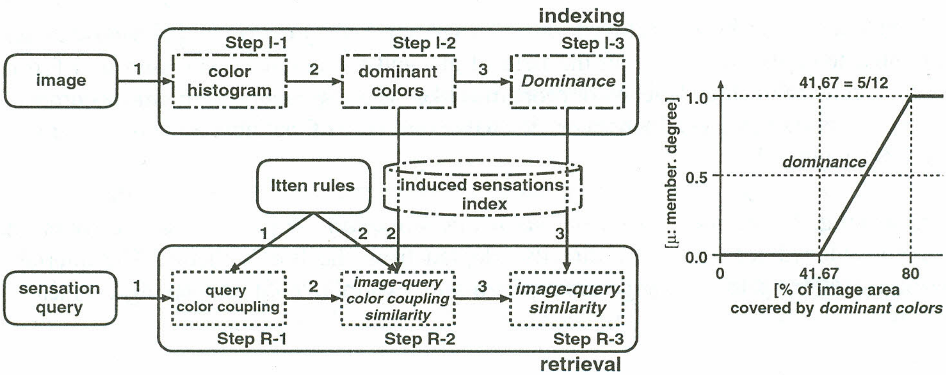
The set of dominant colors and their dominance degree make up the induced sensations index.

The outline of the proposed **induced sensations retrieval algorithm,** which uses the image index to compute the *image-query similarity degree* (G [0,1]), is (Figure 7-a):

Step R-l. Using the Itten's rules, identify the *queried color coupling,* i.e. the cc coupling inducing the queried sensation.

Step R-2. Compute the *color coupling similarity,* i.e. the similarity between the cc coupling present in the dominant colors (Step 1-2) and the queried color coupling (Step l).

Step R-3. Finally, to obtain the image-query similarity degree, the color coupl similarity (Step R-2) is weighted by the dominance degree (Step 1-3).



(a) (a)

Figure 7. (a) The proposed algorithm for induced sensations indexing and retrieval; (b) The domina degree of the dominant image colors

3. 2. Text indexing and retrieval

The purpose of the text indexing is to extract the text content of a document and produce the text index. At the retrieval time, the text index is compared to the keywords the text sub-query, resulting in the *text-query similarity degree* (G {0,1}), which defines relevant documents.

Querying. The text sub-query consists of the keywords that are expected to app in the text of a document. The keyword-based queries are chosen since they are intuiti widespread among the existing text retrieval systems, and can be easily implemented. T meaning of a query is limited to checking the existence of a keyword(s) or a phrase wit a document — i.e. *exact match* — not considering the number of occurrences of a keyw or any other information [1]. Query keywords can be connected either by AND or by Boolean operator, with the obvious meanings. In the case of a "phrase" query, the seque of keywords must appear in the document text exactly as in the query.

Indexing and retrieval. As the text indexing technique, the inverted file (or inver index) is adopted, being currently the most efficient indexing mechanism for the keywo based searching [1]. Inverted file stores, for each word appearing in the text (of all the documents), the list of all the documents containing that word. In the retrieval stage, the text-query similarity degree is evaluated in a crisp manner, based on the presence of the query keyword(s) in the document text. If the query keywords are present in a document, the document is evaluated as relevant to the query, otherwise the document is not relevant.

3.3. Integrated querying and retrieval

An integrated query (Section 2.1.) is a combination of the image sub-query (Section 3.1.) and the text sub-query (Section 3.2.), referring to the image and text part of a document, respectively. The image sub-query includes the color-spatial (Section 3.1.1.) and the induced sensations (Section 3.1.2.) types of image queries.

The *integrated similarity* (relevance) degree of multimedia document with respect to an integrated query is computed as a combination of: *image similarity* — similarity between the document image and the image sub-query (Section 3.1.); and *text similarity* — similarity between the document text and the text sub-query (Section 3.2.).

As pointed out earlier in this section, the image similarity is a real number in range between 0 and 1, while the text similarity is a binary value, 0 or 1. Based on these two values, the integrated similarity degree is defined as:

$$\text{integrated-similarity} = \begin{cases} \text{image-similarity} & \textit{if } \text{text-similarity} = 1 \\ 0 & \textit{otherwise} \end{cases}$$

In other words, the document ranking is done based on the image similarity, while the text similarity determines which database documents will be ranked. Crisp matching of the text-sub query serves as a filter which reduces the total number of documents that need to be ranked with respect to the image sub-query, leading to the efficient retrieval of the documents containing images and text.

4. EXPERIMENTS AND SYSTEM PERFORMANCE EVALUATION

The proposed integrated retrieval system is implemented in UNIX/X environment, using the GTK+ GUI toolkit.

The retrieval experiment is performed with two goals in mind: first, to **evaluate the integrated retrieval performance** (Section 4.1.), by comparing the integrated queries with the conventional image and text queries; and second, to **evaluate the image retrieval performance** (Section 4.2.), as a key component in realizing the keyword-based integrated retrieval of images and text (as explained in Section 2.1.).

4.1. Integrated retrieval experiment

The purpose of the integrated retrieval experiment was to evaluate the proposed integrated querying approach through the comparison with the conventional image and text querying approaches.

Integrated retrieval evaluation. In order to evaluate the integrated querying approach, we have compared the number of documents retrieved by the integrated queries, to the number of documents retrieved by the corresponding image and text queries. The purpose was to demonstrate the decrease in the number of the retrieved documents when using the integrated queries, which is especially valuable in the case of the retrieval from the large databases.

The experiment involved 13 integrated queries (Figure 9-a) executed on a test set of 1,022 Web-pages from the "The Vincent van Gogh Gallery" [13]. Each test page contains an image of Van Gogh's oil painting or a watercolor. Each painting is accompanied by the text revealing its title, origin, etc. As an illustration, integrated retrieval is demonstrated in Figure 8.

The results of the experiment (Figure 9-b) show that, when using the integrated queries, the number of retrieved documents is in average reduced to 14% (i.e. more than 7 times less) of the number obtained by the image or text queries. In the case when the information the user is looking for is a combination of the images and text, and when the retrieval from a large-scale database is considered, this means a significant increase in the retrieval effectiveness, compared to the conventional image or text retrieval.

4.2. Image retrieval experiment

The purpose of the image retrieval experiment was to evaluate the image retrieval subsystem, which processes the image sub-queries: color-spatial queries and induced sensations queries (Section 3.1.).

Color-spatial image sub-queries. A color-spatial image sub-query contains *color, coverage,* and *region pattern* expressions (Section 3.1.), describing how much of a given image region is covered by a given color. Accordingly, the experiments comparatively emphasizing each of the three expressions are performed.

In the experiment involving 27 image sub-queries, totally 1,100 color images (11 sets of 100 photos each) from the *Corel Corporation'?,* "Corel Gallery 3.0" image database are used. All images have strong color properties and include the photos of the skies, landscapes, flowers, fruits and vegetables, textiles, and various objects.

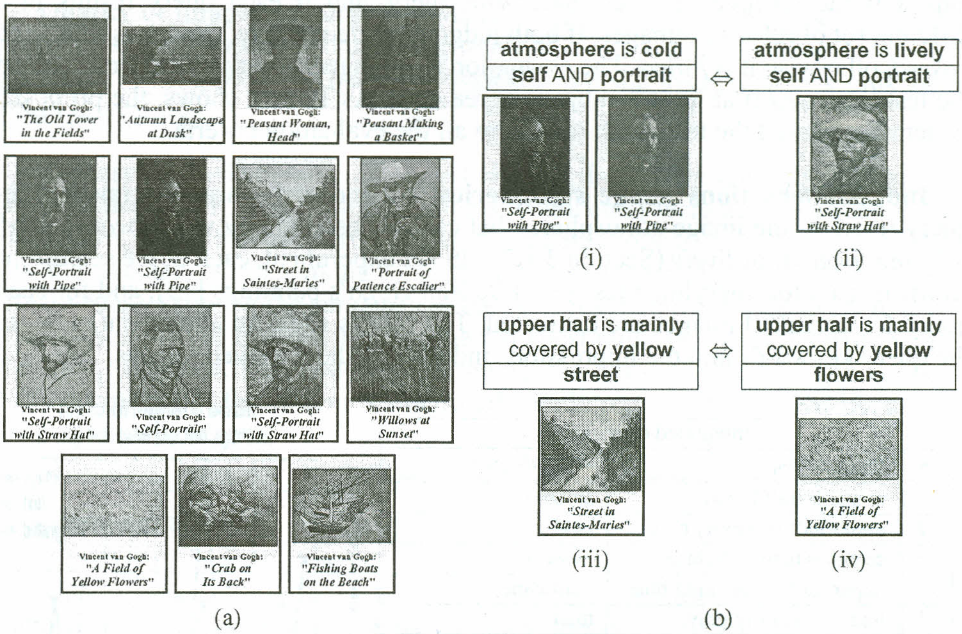


Figure 8. (a) Set of documents (corresponding to the actual Web pages) used for the integrated retrieval demonstration; (b) Integrated queries and the retrieved documents (upper half of the integrated query is the image sub-query, lower half is the text sub-query)

Evaluation Criterion	*# Images*	*# Queries*	*Passed*
basic colors	100	8	√
color brightness	100	4	√
color coverage	200	4	√
image region patterns	400	5	√
image region pattern width	300	6	√
Total	1,100	27	5/5

Table 1: Results of the image retrieval experiment using color-spatial image sub-queries

Experiments focusing on the different parts of the color-spatial image sub-query are summarized in Table 1. *Evaluation Criterion* refers to the image property that was under the focus in each experiment. The purpose of the experiments is to show whether (and how well) the system is able to distinguish the variations of the different queriable image properties. For each evaluation criterion, a set of queries are issued, each query expected to result in a different set of retrieved images. E.g. 8 *basic color* queries involved: *red, purple, orange, violet, green, blue, olive,* and *brown.*

The retrieval results are evaluated through a questionnaire involving 5 subjects (university students). For each set of queries (i.e. evaluation criterion), a subject judges if the retrieved images match the query, and if each query results in a different set of retrieved images. If both judgments are positive, the test is said to be *passed,* otherwise it *is failed.* The evaluation is averaged over the 5 subjects, so that the test is *passed* if at least 3 subjects agree on it. As Table 1 shows, the proposed system has passed the tests with respect to all the evaluation criteria.

Induced-sensations image sub-queries. An induced-sensations image sub-query refers to the image atmosphere, that can be described by words *warm, cold, relaxing, uneasy,* or *lively* (Section 3.1.2.). In the experiment, each of the sensations words is used for querying a test set of 30 Van Gogh's paintings [13], and for each query 6 top-ranked images are evaluated. The results show an agreement between the query keywords and the impressions induced by the retrieved images.

Integrated queries		
#	*Image sub-query*	*Text sub-query*
1	image is half brown	autumn
2	upper-half is mainly blue	field
3	center is mainly light brown	flowers
4	upper-half is half light blue	landscape
5	image is mainly gray	man
6	image is half yellow	peasant
7	image is half yellow	peasant AND woman
8	center is little white	roses
9	upper-half is half orange	self AND portrait
10	image is half red	still life
11	lower-half is mainly green	trees
12	upper-half is half dark red	vase
13	image is half light green	woman

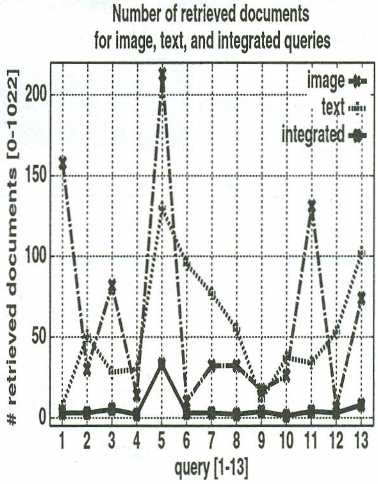
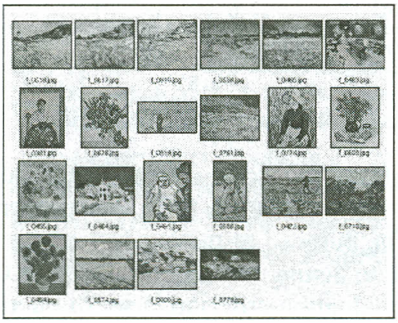


Figure 9. (a) The 13 integrated queries used in the experiment; (b) The mber of documents retrieved, for each query, by the image, text, and integrated querying approach



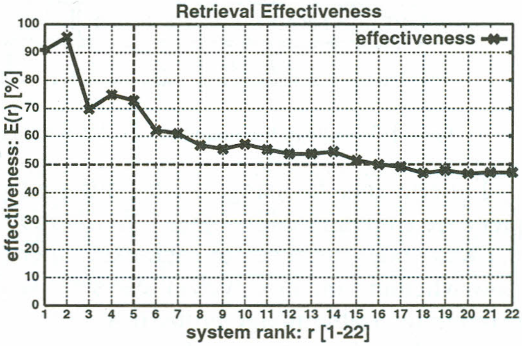


Figure 10: (a) Results of an image sub-query "image is mainly covered by yellow", ranked in left-right, up-down order. This set of images is used as a questionnaire.; (b) Retrieval effectiveness of the proposed system

Quality of image ranking. The quality of the image ranking produced by the proposed system is evaluated using the retrieval effectiveness [3], a measure which expresses the similarity between the system's and the average user's ranking of images.

The test set used contains 22 images (Figure 10-a), from the "The Vincent van Gogh Gallery" [13]. Ten subjects (university students) were asked to rank the 22 test images, using the same criterion as the system (coverage of the yellow color), and the average user's ranking is computed based on their ranks. The retrieval effectiveness, showing the similarity of the system's ranking and the average user's ranking is shown in Figure 10-b. For the 5 highest ranks (the answers that the user sees first, therefore, the most valuable ones), the retrieval effectiveness is above 70%, meaning that the system is able to efficiently retrieve most of the documents the user actually wants.

5. CONCLUSION

While a typical Web-page is a mixture of images and text, none of the existing information retrieval systems enables the integrated querying of both the image and the text contents of a document. Consequently, the user looking for the information given party in visual and partly in textual form, cannot completely express the desired information through the query, which limits the precision of the existing retrieval systems.

The developed integrated retrieval system enables the integrated querying of images and text. Integrated query is expressed through the natural language keywords and phrases, and consists of the image sub-query and the text sub-query.

Integrated retrieval experiment, performed on a database of 1,022 Web pages containing images and text, shows that, when using the integrated queries, the number of retrieved documents in average reduces to 14% of the number obtained by the image or text queries. Image retrieval experiment, using a database of 1,100 color images and based on a questionnaire involving 10 subjects, shows 70% agreement between the system's and the user's ranking of images.

The proposed integrated querying approach — not supported by any existing image or text retrieval system — enables a more precise and complete access to the multimedia document contents, resulting in a more effective retrieval, essential for the large-scale databases like the Internet. In addition, the natural language expressions-based querying gives user a uniform access to both the image and the text contents, making the retrieval more intuitive.

REFERENCES

[I] Baeza-Yates, R.; Ribeiro-Neto, B. (1999): *Modern Information Retrieval.* Addison Wesley Longman, Reading, MA.

[2] Corridoni J. M.; Del Bimbo A.; Pala P. (1999): *Image Retrieval by Color Semantics.* ACM Multimedia System Journal, Vol. 7, No. 3, pp. 175-183.

[3] Del Bimbo, A. (1999): *Visual information retrieval,* Morgan Kaufmann Publishers, Inc., San Francisco, California.

[4] Flickner, M. et. al. (1995): *Query by Image and Video Content: The QBIC System,* IEEE Computer, Vol. 28, No. 9, pp. 23-32.

[5] Itten, J. (1961): *The art of color,* Otto Maier Verlag, Ravensburg, Germany.

[6] Lawrence, S.; Giles, C. L. (1999): *Accessibility of Information on the Web,* Nature, Vol. 400, pp. 107-109.

[7] Lew, M. S. (2000): *Next-Generation Web Searches for Visual Content,* IEEE Computer, Vol. 33, No. 11, pp. 46-53.

[8] Miyahara, M.; Yoshida, Y (1988): *Mathematical transform of(R, G, B) color data to Munsell (H, V, C) color data,* In SPIE 1001, Visual Communications and Image Processing, pp. 650-657.

[9] Mundie, D. A. (1995): *The NBS/ISCC color system,* Polymath Systems, Pittsburgh, PA.

[10] Smith, J. R.; Chang S. F. (1996): *VisualSEEk: A fully automated content based image query system.* ACM Multimedia, Boston, MA.

II1] Stejic, Z. (2001): *Integrated retrieval of images and text using natural language "kansei" expres sions,* M.Sc. thesis, Dept. of Comp. Intelligence and Sys. Science, Tokyo Institute of Technology, Yokohama, Japan.

[12] Swain, M. J.; Ballard, D. (1991): *Color indexing,* International Journal of Computer Vision,Vol. 7, No. l,pp. 11-32.

[13] *The Vincent Van Gogh Gallery* (2000), http: / /www. vangoghgallery. com [14] Ziviani, N.; de Moura, E. S.; Navarro, G.; Baeza-Yates, R. (2000): *Compression: A Key for Next-Generation Text Retrieval Systems,* IEEE Computer, Vol. 33, No. 11, pp. 374.

Received: 15 December 2001
Accepted: 1 July 2003

Zoran Stejic
Yasufumi Takama
Kaoru Hirota

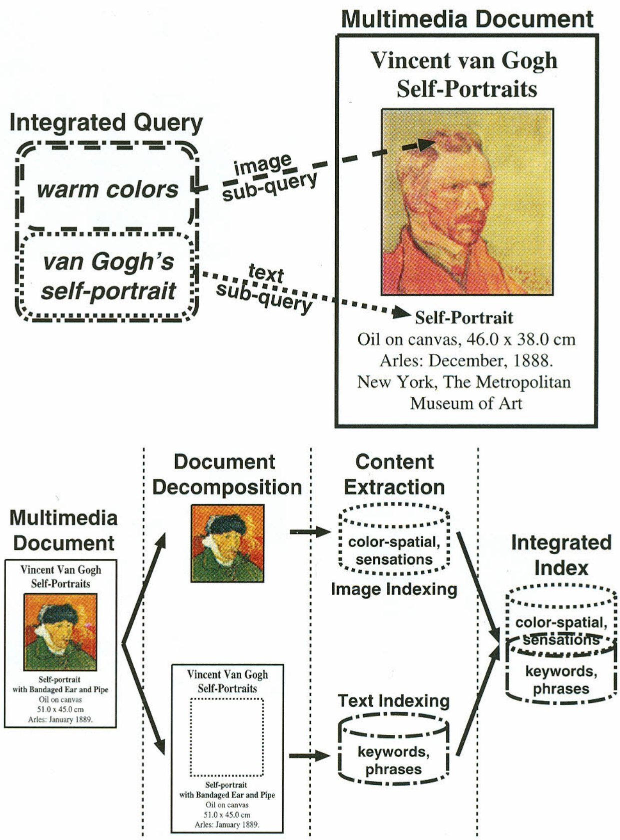
POBOLJŠANJE UČINKOVITOSTI PRETRAŽIVANJA DOKUMENATA INTEGRIRANIM UPITIMA U SLIKE I TEKST

Sažetak

Razvijen je integrirani sustav pretraživanja koji omogućuje simultano postavljanje upita o slikovnom i tekstualnom sadržaju dokumenta — što je osobina koja postoječim sustavima pretraživanja nedostaje. Integrirani upit oblikuje se izrazima prirodnog jezika. Slikovni sadržaj propituje se ili temeljem svojstava rasporeda boja u prostoru slike ili induciranim osjetima, izvedenima iz uzoraka rasporeda boja. Neopreciznost upita u prirodnom jeziku modelira se uz pomoć teorije fuzzy skupova. Pomoću anketnog upitnika s 10 pitanja proveden je eksperiment na uzorku od 1,100 slika. On je pokazao 70% suglasje između sustava i korisnika u pogledu rangiranja slika. To pokazuje da je predloženi sustav za postavljanje integriranih upita učinkovitiji u usporedbi s klasičnim pristupima pretraživanju teksta i slika, što je od velike važnosti za velike baze podataka kakvom se može smatrati i Internet.

Ključne riječi: *integirano postavljanje upita, pretraživanje teksta i slika, svojstva prostornog rasporeda boja*

APPENDIX (coloured figures from the paper *Z. Stejic, Y. Takama, K. Hirota: Improving retrieval effectiveness by integrated querying of images and text*)



(a)

(b)

Figure 1: (a) Integrated query and multimedia document; (b) Integrated indexing of a document

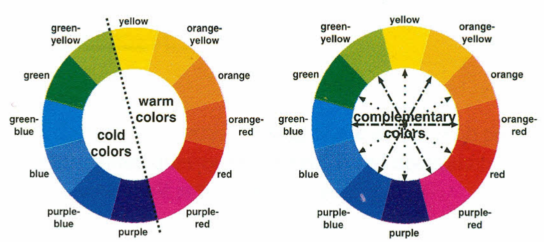


Figure 6: (a) The core of the Itten-Runge sphere with *warm-cold* and *complementary* color couplings

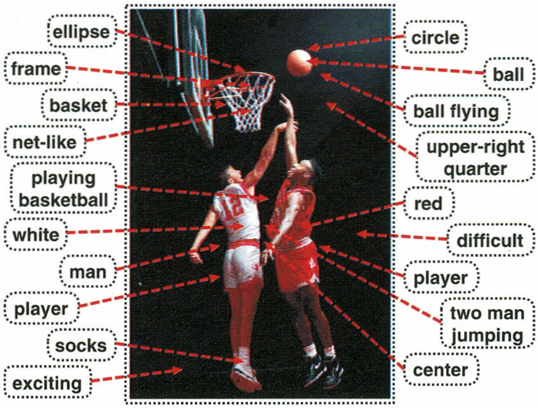


Figure 3. (a) Photo illustrating the image properties expressed in the NL

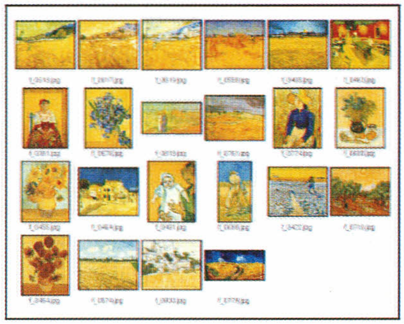


Figure 10: (a) Results of an image sub-query "image is mainly covered by yellow", ranked in left-right, up-down order. This set of images is used as a questionnaire.

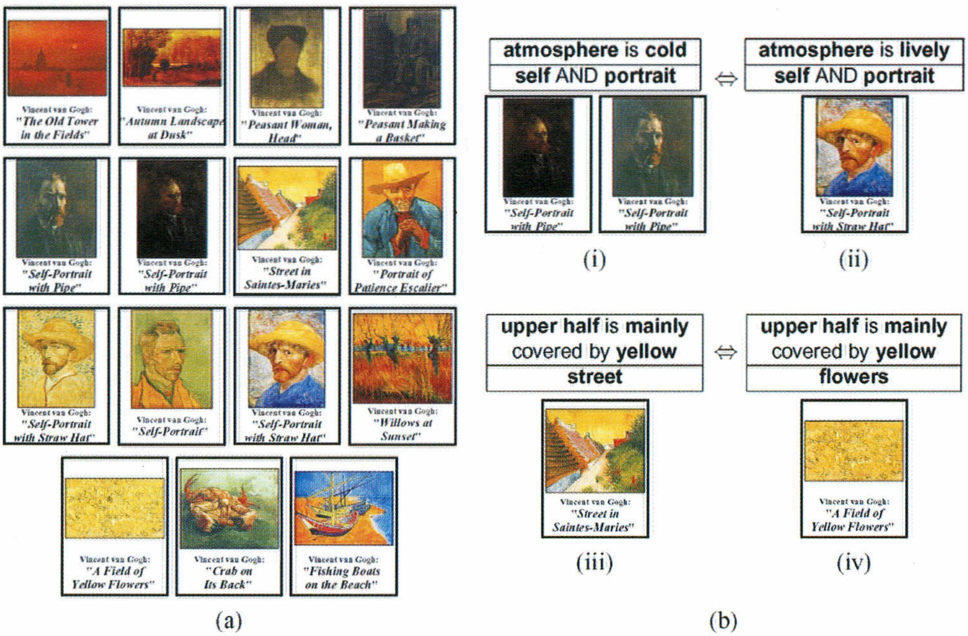


Figure 8. (a) Set of documents (corresponding to the actual Web pages) used for the integrated retrieval demonstration; (b) Integrated queries and the retrieved documents (upper half of the integrated query is the image sub-query, lower half is the text sub-query)