

An XML fragment retrieval method with image and text using textual information retrieval techniques (Extended Abstract)

Yu Suzuki¹, Masahiro Mitsukawa¹, Kenji Hatano²,
Toshiyuki Shimizu³, Jun Miyazaki⁴, and Hiroko Kinutani⁵

¹ Ritsumeikan University, 1-1-1 Noji-Higashi, Kusatsu, Shiga 5258577, Japan

² Doshisha University, 1-3 Tatara Miyakodani, Kyotanabe, 6100394, Japan

³ Kyoto University, Yoshida, Sakyo, Kyoto 6068501, Japan

⁴ Nara Institute of Science and Technology 8916-5 Takayama, Ikoma, Nara 6300192, Japan

⁵ University of Tokyo, 4-6-1 Komaba, Meguro, Tokyo 1538505, Japan

Abstract. In this paper, we propose a retrieval method using textual information retrieval techniques, such as vector space model, for XML documents which consist of image and text. Recently, many image retrieval systems are proposed. However, these systems are mainly based on pattern recognition techniques. Therefore, the features of images are also based on these recognition techniques, such as color histogram, and shape of the object in images. Generally, these systems do not consider weight of features, which means how important these features are, which are generally used in textual information retrieval systems. In this paper, we propose a method considering weight, such as TFIDF, to identify the importance degree of features. Using our proposed method, the system can retrieve intuitively similar retrieval target images to user's query images.

1 Introduction

In this paper, we propose a method for retrieving images using textual information retrieval method, such as tfidf-based method.

Recently, there are many documents which consist of not only textual information but also images, movies and music information, especially on the Internet. Therefore, many researchers have developed many information retrieval systems which deal with data which includes textual information, images, and other media. Then, these retrieval systems are widely used for many purposes. In this paper, we focus on these multimedia document retrieval system, especially multimedia XML document retrieval system.

When we develop XML document retrieval system, we should develop image retrieval system for retrieving images which is placed on retrieval target XML documents. Most image retrieval systems are based on the recognition techniques of images. This means that, the system try to recognize objects, scene, color, texture from images for constructing Content-based Image Retrieval (CBIR) systems.

On the other hand, when we develop textual information retrieval system, we do not deal with simple n-gram or basic boolean-like method. Instead of these naive methods, we deal with vector space model[1], probabilistic model, or several modern information

retrieval model to discover appropriate retrieval result documents. One of the reason about this retrieval model selection is that users do not need exactly same terms as query terms. Users should retrieve intuitively relevant documents to the user's queries.

We believe that if we merge textual information retrieval method with image information retrieval method, we can retrieve intuitively relevant images to the user's query images. We think that the aims of the pattern recognition techniques based retrieval methods are similar to simple n-gram and basic boolean-like methods of textual information retrieval systems.

We develop multimedia XML document retrieval systems which consist of textual XML document retrieval system and the image retrieval system described in this paper.

2 Overview of our Multimedia XML retrieval system

In this section, we describe an overview of our proposed system. Our system can divide into three parts.

1. **The system retrieves XML fragments using textual XML retrieval system.**
Using textual XML fragment retrieval method described in paper [2], we retrieve XML fragments.
2. **The system retrieves images using image XML retrieval system.**
Using image retrieval method described in section 3, we retrieve retrieval target images.
3. **The system merge two result lists.**
Using two retrieval results from textual XML fragments retrieval system and image retrieval system, the system generate one retrieval result by calculating the integrated retrieval status values of each XML fragment.

In this paper, we describe the second part, such as a method for retrieving images.

3 Image retrieval method using textual information retrieval techniques

In this section, we describe a method for retrieving images using textual information retrieval techniques.

In this method, we process the following four methods.

1. **The system extracts features from retrieval target images.**
From retrieval target image retrieval system, the system extracts feature values. We use CIELab color space as feature values.
2. **The system extracts features from user's query image.**
From user's query image, the system also extract feature values using same method as the method for retrieval target images.
3. **The system calculates tfidf-based weights from image features.**
General image retrieval simply deals with image features in straight. In our method, we calculate weights of feature values for emphasizing intuitively characteristic features of images.

4. **The system calculates retrieval status values using Earth Mover's Distance.**
We deals with Earth Mover's Distance [3] to calculate the retrieval status values between the retrieval target image and the user's query image.

4 Conclusion

In this paper, we proposed a multimedia XML fragment retrieval system using image retrieval method with textual information retrieval techniques. In our proposed method, we deal with textual information retrieval method for retrieving images.

Our implemented system can only process queries which specify exact images, then our system cannot process concept clauses. Therefore, we should consider how to calculate retrieval status values using concept clauses.

References

1. Salton, G., McGill, M.: Introduction to Modern Information Retrieval. McGraw-Hill Book Company (1983)
2. Hatano, K., Shimizu, T., Miyazaki, J., Suzuki, Y., Kinutani, H., Yoshikawa, M.: Ranking and displaying search results based on content-and-structure conditions of xml documents and queries. In: Informal Proceedings of INitiative for the Evaluation of XML Retrieval. (2007)
3. Rubner, Y., Tomasi, C., Guibas, L.J.: The earth mover's distance as a metric for image retrieval. International Journal of Computer Vision **40**(2) (2004) 99 – 121