

Presenting Mathematical Expression Images on Web to Support Mathematics Understanding

Kuniko Yamada¹, Hiroshi Ueda², Harumi Murakami¹, and Ikuo Oka¹

¹Osaka City University,
3-3-138, Sugimoto, Sumiyoshi, Osaka 558-8585 Japan
k16uz90z10@st.osaka-cu.ac.jp

²Stroly Inc.
7-5-1, Seikadai, Seika, Soraku, Kyoto 619-0294 Japan

Abstract. People cannot use a text search to find mathematical expressions because expressions cannot be replaced with words. Our research uses an ordinary text search and presents appropriate mathematical expression images (hereinafter called math-images) for input keywords. First we classify a set of the top ranking images from all the images in HTML files by scoring them. We focus on three viewpoints that are unique to mathematical expression images and mark the images by using these viewpoints. Then by adding bonus points to these marked images, the best three images are chosen from the set and presented with an explanation of the keyword and the surrounding information in the HTML files. We conducted two experiments to optimize the parameters of the expression giving the mark and to evaluate the effect of the bonus points. The rate of the average correct images of the best three was 79.5%.

Keywords: mathematical expression image, mathematics understanding, mathematical information retrieval, web search, SVM, Wikipedia.

1 Introduction

For such fields as science, technology, or economics, searching for mathematical expressions is necessary because these concepts are explained using expressions. However, mathematical information retrieval is difficult. Since expressions are not annotated by names, we cannot search for them by names; we cannot use a text search.

Our research uses an ordinary text search given a mathematical term as a keyword and presents appropriate math-images for an input keyword. We focus on the viewpoints where important math-images are on a separate line in a document, and many of these images have unique feature quantities. Adding another viewpoint where these important images often have a keyword around them to the viewpoints, we obtain a set of the top ranking math-images. Focusing on one more viewpoint where important information often appears in the first part of a document, we obtain the best three math-images from a set of the top rankings. Our experiment result showed that at least two of the best three

math-images were correct images. To support mathematics understanding, the best three images are shown on a screen display with keyword explanations and the surrounding information in HTML files.

2 Approach

2.1 Mathematical Expression

General ways of writing mathematical expressions on the web use pdfs, images embedded in a HTML file, or MathML based on XML. The following is the writing style for mathematical expressions in a document. (a) A variable, a sign, or an additional expression is on a line and is expressed with math notation instead of characters because the mathematical font is special. (b) An important theorem or a formula is on a separate line, even though it is in a sentence. (c) Connected by equal signs, a mathematical expression is too long from which a theorem is derived or a calculation example is shown.

2.2 Method

In the following steps, we explain Fig.1.

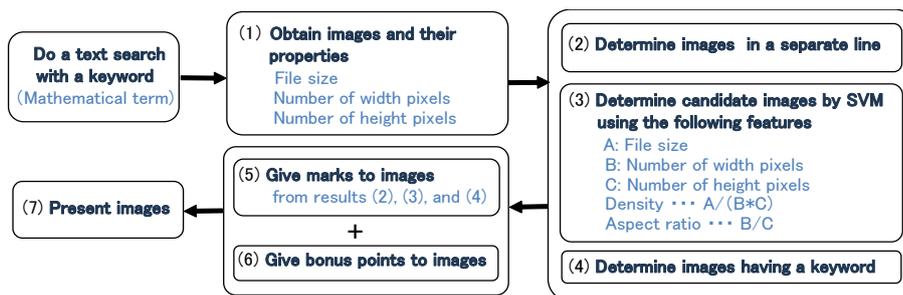


Fig. 1. Overview

(1) **Obtain images:** When the HTML files of a mathematical term are given, we obtain the images and extract the following image information: the file size, the number of width pixels, and the number of height pixels. When there is no text between the images, they are concatenated and considered as one.

(2) **Determine images on a separate line:** When an image satisfies the following two conditions, it is defined as being on a separate line: (a) An image tag to which the image belongs has no text before or after it. (b) The image tag has tags that denote a new line before and after, such as `
`, `<p>`, `<tr>`, `</br>`, `</p>`, `</tr>`.

(3) **Determine candidate images by SVM:** We determine whether an image is a candidate by Support Vector Machines (SVM). The features given to

the SVM are file size, the number of width pixels, the number of height pixels, the density, and the aspect ratio.

(4) **Determine images having a keyword:** We search for the closest image to a keyword as the window size that is set from -80 to +80 Japanese characters surrounding the keyword. If it exists, the image is defined as having a keyword. However, if an image and a keyword appear in the same sentence, it is preferentially defined as having a keyword.

(5) **Give marks to images:** Each image is given a mark using (2) to (4) and ranked in descending order of points. Fig. 2 shows an example of the top ranking images when the keyword is “spherical harmonics”. The first two images are incorrect and the others are correct.

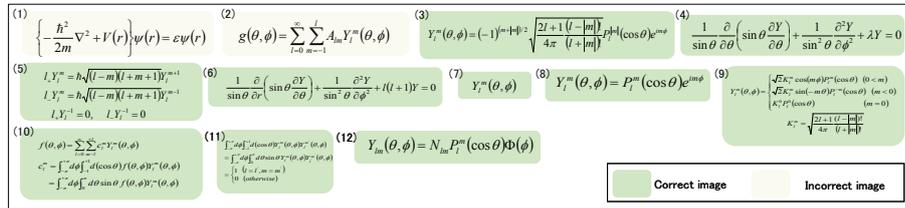


Fig. 2. Example of top ranking images for keyword: spherical harmonics

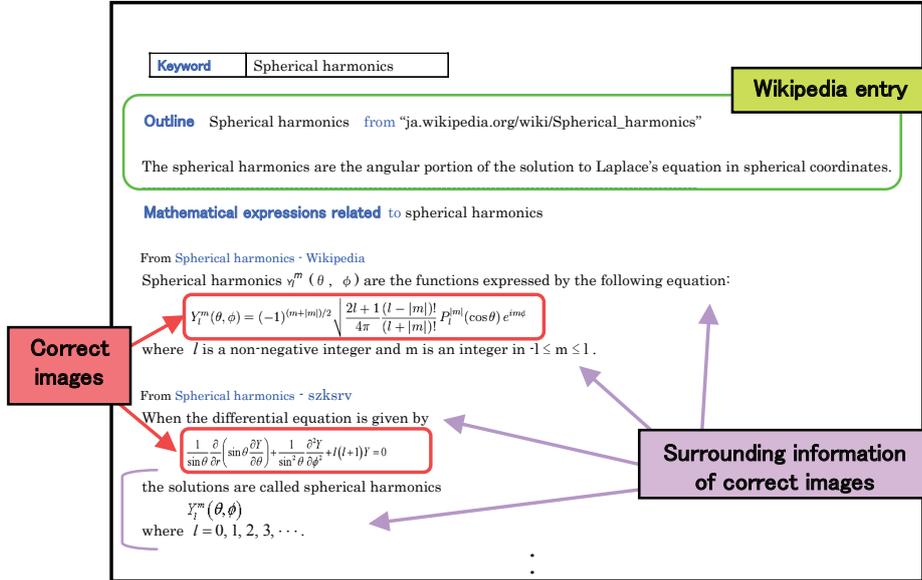


Fig. 3. Screen display example of keyword: spherical harmonics

(6) **Give bonus points to images:** In Fig. 2, the first two happen to be incorrect, because their order is affected by the ranking of the original web pages. This order is not useful when we choose the best three or five from the top ranked images. To solve this problem, we propose a bonus point system where the important information on a keyword often appears in the first part of a document. This system gives an additional point to an image when it obtains the highest marks in the order of the appearance in the HTML file. Then the possibility increases that the actually correct image in each HTML file rises to the higher position in the set of the top rankings.

(7) **Present image:** Fig. 3 shows a screen display example that has all correct images (The third image is omitted). At the top, the display shows the outline extracted from the first paragraph of Wikipedia that matches the keyword. Below it, the display shows the best three images with the information around them in the HTML files. Adding the surrounding information provides clearer explanation about the images.

3 Experiment

3.1 Dataset

We extracted 16 keywords from the science and engineering web syllabi of Osaka City University and obtained the top five HTML files for each keyword using a text search in Experiment 1. We added five keywords in Experiment 2, for a total of 105 web pages and 3,262 images. These images were judged manually and SVM used eight keyword search results as training data. In Experiment 1, the remaining eight keyword search results were evaluation data. In Experiment 2, 13 keyword search results including additional five keyword search results were evaluation data.

3.2 Experiment 1

To give marks to the images, we developed Expression (1), where *points* is the acquisition mark, *image* is the math-image, and α_k is the weights. When $k = 1$, an image in a separate line is correct; when $k = 2$, an image determined to be correct by SVM is correct; when $k = 3$, an image with a keyword is correct. The aim of this experiment is to optimize the values of α_k and $\delta_3(\textit{image})$ when an image is incorrect.

$$\textit{points} = \sum_{i=1}^3 \alpha_k \delta_k(\textit{image}) \quad (1)$$

When $k = 1, 2$

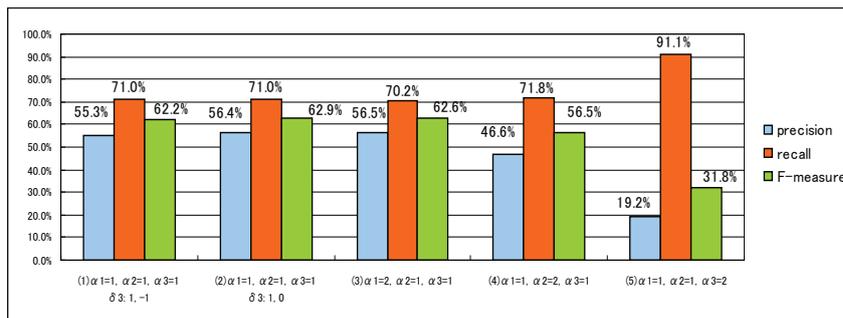
$$\delta_k(\textit{image}) = \begin{cases} 1 & (\textit{image} : \textit{correct}) \\ -1 & (\textit{image} : \textit{incorrect}) \end{cases}$$

When $k = 3$

$$\delta_k(\textit{image}) = \begin{cases} 1 & (\textit{image} : \textit{correct}) \\ -1 \textit{ or } 0 & (\textit{image} : \textit{incorrect}) \end{cases}$$

First, to determine -1 or 0 of $\delta_3(\textit{image})$, two sets were created by fixing each $\alpha_k = 1$. One had -1 as its δ_3 value and the other had 0, and then the top ranking

images were compared using precision, recall, and F-measure. From Figs. 4 (1) and (2), we adopted 0. Next, to determine the α_k values, four sets were created. The first had $\alpha_1 = \alpha_2 = \alpha_3 = 1$ as its condition and the second to the fourth had conditions where only one value of each α_k was 2 and the others were 1. Then we compared these sets using F-measure. In Figs. 4 (2) to (5), since (2) is the highest, we adopted $\alpha_1 = \alpha_2 = \alpha_3 = 1$.



(1) and (2): Comparison of applying 0 and -1 to images without a keyword
 (2) to (5): Comparison when changing weights

Fig. 4. Results of Experiment 1

3.3 Experiment 2

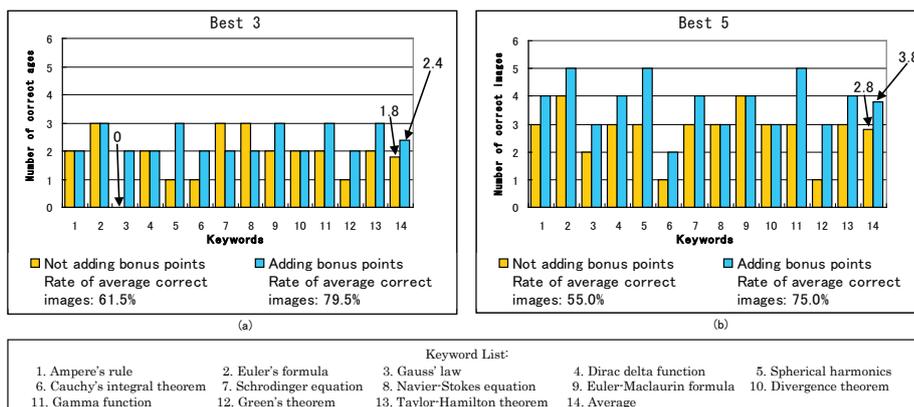


Fig. 5. Comparison of adding and not adding bonus points

Expression (1) was fixed to $points = \sum_{i=1}^3 \delta_k(image)$ and $\delta_3(image) = 0$ when the image is incorrect. Using this expression, we examined the validity of our bonus point system. In Fig. 5, adding bonus points was basically better than not adding them. Especially in Fig. 5 (a), the number of correct images of the third keyword, “Gauss’ law”, rose to 2 from 0. The rates of the average correct images showed that the bonus point system was effective, and since the rate of the average correct images of the best three exceeded the rate of the best five, we adopted the best three to which bonus points were added.

4 Related work

Even though various methods have addressed mathematical information retrieval [1], they are searching for math notations with similar structure or semantics for a given expression. Among those methods, content-based image retrieval (CBIR) inputs a digital image to search for similar images [2], and instead of using images, others use MathML [3] or LaTeX [4].

This paper’s approach is quite different. Focusing on the unique features of mathematical expressions and math-images led us to use text information. Even people who are unfamiliar with handling MathML or LaTeX can use our method.

5 Conclusions

We proposed a method to present math-images with the information around them in the HTML files. Our method uses an ordinary text search that inputs a mathematical term as a keyword. To obtain the best three math-images, we developed Expression (1) and optimized its parameters. Our results showed that at least two of the best three were correct, and the rate of the average correct images of the best three was 79.5%. Then using these three images, we showed a screen display with an explanation of the keyword and information about the images to support mathematics understanding.

References

1. Zanibbi, R., Blostein D.: Recognition and Retrieval of Mathematical Expressions. *International Journal on Document Analysis and Recognition* 15(4), 331-357 (2012)
2. Shirmenbaatar, M., Koga, H., Watanabe, T.: Content-Based Mathematical Formula Image Retrieval System. In: *The 4th forum on Data engineering and Information Management (DEIM2012)* (2012)
3. Yokoi, K., Nghiem, M.-Q., Matsubayashi, Y., Aizawa, A.: Contextual Analysis of Mathematical Expressions for Advanced Mathematical Search. *Research journal on Computer science and computer engineering with applications* 43, 81-86, Polibits (2011)
4. Zanibbi, R., Yuan, B.: Keyword and image-based retrieval for mathematical expressions. In: *Proceedings of Document Recognition and Retrieval XVIII, SPIE*, pp. 011-019, (2011)