UDC 621.397.3+681.3.06+002.5

Development of "TRIPIT", the Associative Image Search Engine for Tagged Images

Hajime TERAYOKO*, Yosuke OHASHI*, Tetsuya SAWANO*, and Norihisa HANEDA*

Abstract

We have developed a new image retrieval engine called "TRIPIT" to search relevant images from a keyword that will satisfy more than 70% of users with the search results.

In this paper, we introduce the structure and procedure of development of the technology behind "TRIPIT".

1. Introduction

We have developed a relevant image search engine called TRIPIT that can search for relevant images using an ambiguous idea as a start point. TRIPIT is a technology that identifies the relevance between tags attached to images and can search for images related to the input tag. The TRIPIT relevant image search engine is a step towards a world where anyone can search for the information they want without influence from the clarity, or lack thereof, of the search term or the user's search ability.

In this paper, we discuss the development background, issue and hypothesis formulation, technical summary, experiment and evaluation methods, results, and considerations related to TRIPIT.

2. TRIPIT Development Background

In recent years, along with the popularization of the Internet and digital consumer electronics, large images and other large size digital contents (hereafter collectively referred to as contents) have come to be in wide circulation. At the same time, searching for these contents has become part of daily life for not only users knowledgeable about information technology, but for men and women of all ages. The prevailing method for searching for contents is the keyword search. The user inputs keywords, and contents containing matching keywords are returned. However, in order to search for contents with this method, the user must specify a clear keyword that specifies the item to search for.

On the other hand, if we look at the real world, it is not always the case that searching is done with a clear target from the start. Users often begin searching from an unclear idea. Let's take shopping as a typical example. When people go shopping, it is often the case that even if they do not have something clearly in mind that they want to buy, they find something that they potentially want while browsing through all the items lined up in the store. For example, people may go shopping for winter clothing with the plan to purchase a trench coat or a down jacket, only to return with a sport jacket that piqued their interest. Cases such as this are hardly rare.

However, in searching for contents, there is no decisive search method that relies on feeling or the senses. For users unaccustomed to searching, this is a large problem.

3. Issue and Hypothesis Formulation

To deal with this situation, we made it our issue to find a way to make it possible for users to start with an ambiguous idea and search for the image that they potentially want. We then hypothesized that it is possible to reach the image that the user potentially wants by repeatedly following connections between related images. We developed the relative image search engine TRIPIT to prove this hypothesis.

There are many methods to accomplish this search engine. One typical method is to search for similar images. But, the problem with this search method is its slow search speed, because it requires image processing. Therefore, we focused our attention on tagged images on online photo sharing sites such as "Flickr" (http://www.flickr.com/). (The number of such images is rapidly increasing.) Tags are keywords attached to an image. Keywords that indicate the photographic subject or the circumstances around the time the photograph was taken are added by users as tags. We chose the approach of searching for related images through the use of the linguistic relationship between image tags.

Fig. 1 shows a conceptual diagram of TRIPIT. The user inputs the tag "flower," and TRIPIT outputs images such as those shown below that contain tags related to "flower."

1) Images that contain the higher-category tag "vegetation."
2) Images that contain the lower-category tag "tulip" or "hydrangea."

Original paper (Received December 17, 2007)

* Internet Business Development Division
New Business Development Division
FUJIFILM Corporation
Senzui, Asaka, Saitama 351-8585, Japan

3) Images that contain the parallel-category tag “gardening” or “bird.”

As shown above, TRIPIT is a relative image search engine that allows the user to search for related images based on the relationship between image tags. The goal of TRIPIT is to allow users to reach the image that they potentially want by having the user first specify an ambiguous tag and then repeatedly following the connections between related images.

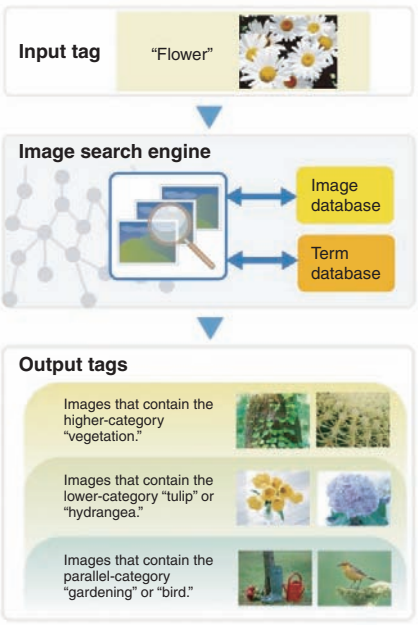


Fig. 1 Conceptual diagram of “TRIPIT”.

4. Structure of TRIPIT

4.1 Overall Structure

Fig. 2 shows the structure of the TRIPIT system.

TRIPIT uses an existing set of tagged image information as the search target and then presents the results of the relative search based on the image information. The TRIPIT system has two major functions.

The first major function is registering contents. In contents registration, the image information from an existing image database (DB) that you want to use as the basis of the search is transferred to TRIPIT via https and registered. Image information refers to image IDs, tags, image URLs, and other such information.

The second major function is searching. TRIPIT’s relative search function is packaged together with a REST Web API. When the user specifies a tag as the parameter and invokes the Web API, the image information relating to the specified tag is output in XML format. You can display the related images based on the acquired related image information.

The point to take note of is that the exchanged information does not include the images themselves, only information such as the image IDs and tags is exchanged. The major merit of TRIPIT is that because it carries out the relative image search based on the tags attached to the images, it does not exchange the images themselves and therefore creates a minimum amount of network traffic. Another merit worth mentioning is that because you can add the relative search function as an add-on to an existing image DB, you can incorporate TRIPIT with minimum system changes. We will next explain TRIPIT’s internal system structure and the thesaurus, a key component of TRIPIT.

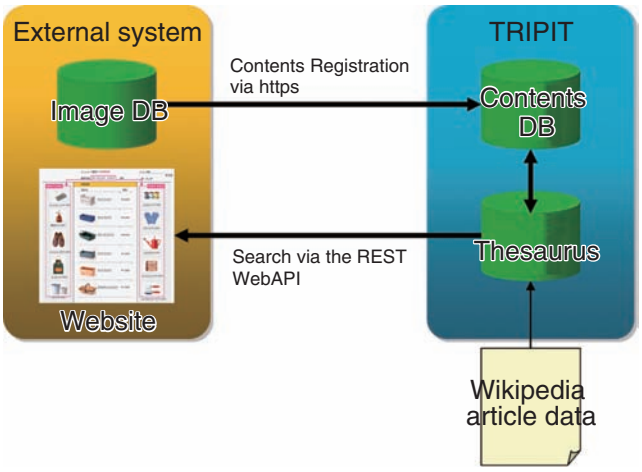


Fig. 2 Structure of “TRIPIT”.

4.2 Internal Structure

There are two DBs within the TRIPIT system.

The first DB is the thesaurus. TRIPIT carries out image searches based on the tags related to the tag input by the user. TRIPIT uses the thesaurus to derive the related tags during the search. A vast amount of tag relationship information is stored within the thesaurus. The second DB is the contents database. This DB records information regarding the images sent from external image databases that was registered during contents registration.

When a relative search is carried out via the Web API, the search is carried out by using the two DBs in the process shown below.

1) Search for tags related to the input tag in the thesaurus.
2) Search for image information in the contents DB that contains the related tag found in step 1.
3) Output the image information found in step 2 in XML format.

4.3 Thesaurus Construction

First, we developed a framework for constructing the thesaurus founded on the category information and the link information between items for data in a digital dictionary. Then, we used the online encyclopedia “Wikipedia” (http://ja.wikipedia.org/) as the first subject for the thesaurus. When we made the aforementioned framework applicable to Wikipedia, we jointly conducted development with Yoichi Hara from Server Domain Inc., knowledgeable in the fields of language processing and DB engineering.

Wikipedia is an online encyclopedia. Its greatest merit is that users can freely edit articles; as such it is an encyclopedia that is built on user participation. A version of Wikipedia has been released for each language in the world. As of the time of this writing, December 2007, approximately 440,000 articles have been contributed to the Japanese version. We will next discuss the method used to build the thesaurus from Wikipedia.

Wikipedia has a tree-shaped category structure, and every article belongs to one or more categories. In addition, there is a link structure with links to other articles within article text. The construction of TRIPIT's thesaurus is based on Wikipedia's category structure and link structure.

Fig. 3 shows an outline of Wikipedia's category and link structures as well as the method used to build the thesaurus. We will next outline the method used to extract higher-category words, lower-category words, and parallel-category words in relation to the reference word.

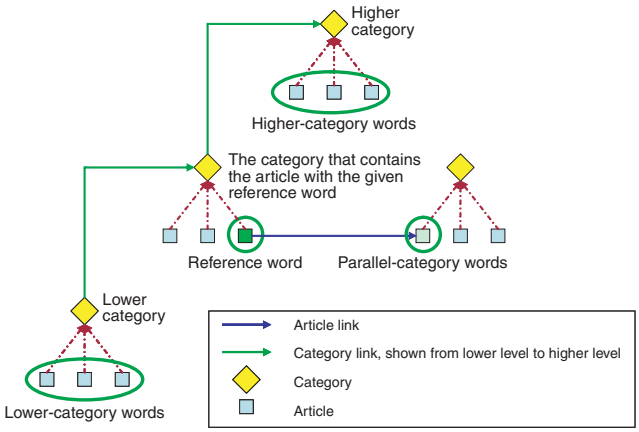


Fig. 3 How to make a thesaurus using Wikipedia.

All of the headings in the articles that belong to the category that is at the level above the category that contains the article with the given reference word are used as higher-category words. Only the category above the category that contains the article with the given reference word is used for extracting higher-category words. The process does not recursively proceed to subsequent higher level categories. Similar to finding the higher-category words, all of the headings in the articles that belong to the category that is at the level below the category that contains the article with the given reference word are used as lower-category words. The headings in the articles linked from the article that contains the given reference word are used as parallel-category words.

The thesaurus stores the type of associated word as well as the association rating that indicates the strength of the association. The simplified formula used to calculate the score, that is, the association rating, between one term W1 and its associated word W2 is as follows:

$$score = seed - \log(LinksToW1) \times \log(LinksFromW2)$$

The seed is the initial value of the score and differs based on factors such as the type of association and whether or not there are reciprocal links between W1 and W2. LinksToW1 is the number of articles that link to W1. Therefore, a word in the heading of an article that is linked to by many other articles has a low association rating with all of its associated words. LinksFromW2 is the number of articles that are linked from W2. Therefore, a word in the heading of an article that has many links to other articles has a low association rating.

5. Evaluation

We developed TRIPIT as outlined in section 4, and to confirm its results, we installed the TRIPIT relevant search function into a live website and evaluated our hypothesis.

5.1 Selected Website

To perform the evaluation, we installed TRIPIT into the "photolibrary" website (http://www.photolibrary.jp/), which is administered by Photolibrary Co. Ltd.. "photolibrary" is a website where users can purchase photographs contributed by other users. It is a site for selling photographs that is built on user participation. From July 25, 2007 to September 30, 2007, we carried out the experimental installation of TRIPIT on this website in a form that we will discuss later. We selected "photolibrary" for our experiment for the following two reasons. First, the items sold are images, and the website has a clear need for image searching. Second, all images have already been tagged by the website's users.

We will next explain the relevant search function provided by TRIPIT after installing it into the "photolibrary" website.

5.2 Related Tag Display

The related tag display is a function that displays tags related to the input keyword when the results from the keyword search are displayed. The section outlined in red in Fig. 4 is the related tag display.

Fig. 4 Screenshot of related tags.

If the user clicks one of the displayed related tags, the system performs a further search based on the related tag. By clicking on related tags, the user can search from one related image to another without inputting keywords.

5.3 Related Image Display

The related image display is a function that acquires and displays a group of images related to the current image that the user is viewing the details of. The section outlined in red in Fig. 5 is the related image display. If the user clicks one of the related images, the details page for the clicked image is displayed. This feature allows the user to browse from one related image to another simply by clicking the mouse.

5.4 Flash Image Search

Fig. 6 shows a screenshot of the Flash image search. This is one form of the TRIPIT installation. The Flash image search uses an Adobe Flash program to provide interactive relative image search functionality. This was a new page on the website created for the TRIPIT installation.

The images related to the image located in the center of the screenshot are displayed around the central image. If the user clicks one of the related images, that image moves to the center, and related images are displayed around it. If the user clicks the central image, the image magnifies, and the user can move to the details page for the image or a page that displays a list of other images that contain the same tags as this image.

Fig. 5 Screenshot of related images.

Fig. 6 Screenshot of flash image search.

6. Results and Considerations

To evaluate our hypothesis, we analyzed user access to the "photolibrary" website, and we held a survey for its users.

6.1 Results of Site Access Analysis

We analyzed 20,123,796 page views taken from 80 days prior to installing TRIPIT and 68 days after installing it.

The probability that a user clicked related tags from the keyword search page was 55.1%, showing that more than half of the users searched repeatedly on related tags when they performed a keyword search. Fig. 7 shows the percentage of pages accessed after a flash image search. From this result, in 95% of the cases when the user executed a flash image search, they next either performed a further search, such as a keyword search or an image search, or browsed, such as viewing the image details or viewing the magnified image. In only 5% of the evaluated cases did users leave the website after performing a Flash image search.

Based on the above results, our hypothesis that users would follow connections between related images was proven by the fact that many users did so in practice when TRIPIT was installed.

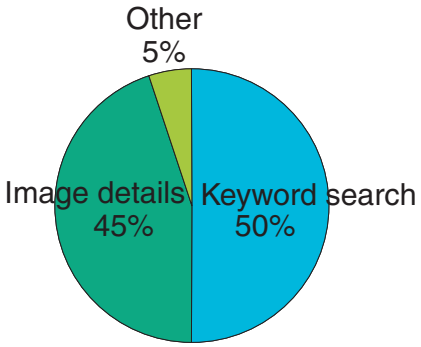
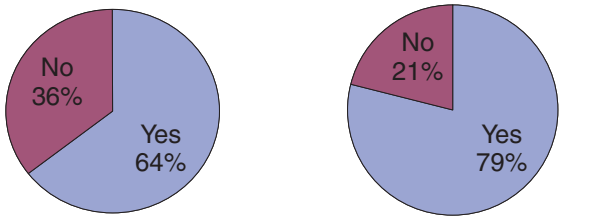


Fig. 7 Pages accessed after flash image search.

6.2 Results of the User Survey

We received completed surveys from 343 users. Fig. 8 shows the results of the survey. First, to the question "Could you search for useful images with TRIPIT?", 64% of the surveyed users answered "Yes." Second, to the question "Could you find images that you did not originally plan to search for?", approximately 80% of the surveyed users answered "Yes." The results of this survey show that TRIPIT was a beneficial search method for many users.



Could you search for useful images with TRIPIT?

Could you find images that you did not originally plan to search for?

Fig. 8 Questionnaire result of "TRIPIT".

6.3 Considerations

From the results given above, we have confirmed our hypothesis for TRIPIT that it is possible to reach the image that the user potentially wants by repeatedly following connections between related images, and we have proven the technical validity of TRIPIT.

On the other hand, from the series of evaluations performed by Photolibrary Co. Ltd., there have been requests for improvement to areas including the following: the validity of related words, support for unknown words that do not exist in the thesaurus, and the processing speed of the Flash image search. In addition, these requests have clarified some of the technical issues.

7. Conclusion

Based on the goal of searching for an image that the user potentially wants from an ambiguous idea as a start point, we have developed the relevant image search engine TRIPIT that searches for related images through the linguistic relationship between tags. From our experience of installing TRIPIT on a website for selling photographs, we have proven that TRIPIT's relative image search contributes to allowing users to search for the image that they potentially want.

TRIPIT is not limited just to images; it is a generic technology that can be used for any tagged contents. In the future, we plan to expand TRIPIT's scope to include searching for movies, recommending merchandise on e-commerce websites, and other applications.

Future issues especially include improving the accuracy of the thesaurus that serves as TRIPIT's fundamental technology and adding support for unknown words. Along with investigating the technical issues given above, we also plan to investigate converting TRIPIT's relative image search functionality into a commercial product.

("Flickr," "Wikipedia," and "Adobe Flash" in this paper are the registered trademarks of Yahoo Inc., Wikimedia Foundation Inc., and Adobe Systems Inc. in the United States, respectively. And "TRIPIT" is applied for trademark registration by FUJIFILM Corporation.)