

Users' Seeking Behavior and Multilingual Image Tags

Miguel E. Ruiz and Pok Chin

University of North Texas, Department of Library and Information Sciences
1155 Union Circle 311068, Denton, Texas 76203-1068
Miguel.Ruiz@unt.edu Pok.Chin@unt.edu

ABSTRACT

Sharing and tagging images is an activity that is very popular among web users. This has led to the creation of Internet services such as Flickr that allow users to post their own pictures and tag them with the purpose of sharing them with friends and displaying their work in a public forum. This poster presents the results of a study conducted in the framework of the Interactive Image Retrieval (iCLEF) task, which is part of the Cross Language Evaluation Forum (CLEF). The main goal of the study is to identify the challenges that users face when they search for images annotated in languages other than English. Our results indicate that the users find the task to be very hard due to the difficulty of selecting terms that match those assigned by the creators of the images. The results also suggest that the increasingly popular tagging of images has limitations similar to those encountered with controlled vocabularies, specifically vocabulary mismatch among searchers and creators of content.

Keywords

Information retrieval, information seeking behavior, user studies, social tagging, image retrieval.

BACKGROUND

Many user studies on image searching have investigated image users' behavior by analyzing metadata, logs, and queries. For example, Matusiak (2006) found that user-provided image metadata varied in terms of description, accuracy, and consistency. These metadata used users' own words to reflect their perception about images. Jørgensen and Jørgensen (2005) found that image users used more descriptive and thematic queries, and less unique term to search for images. Rorissa and Iyer (2008) found that labels used for image category were usually generic and interpretive. Spink and Jansen (2006) found that image users used two to three terms per query to conduct image

search. Queries used for web, news, and audio were short and took longer sessions to complete. Despite different findings from these studies, they all have a common ground: Different image users have different needs and backgrounds that affect the way they interpret images. According to Goodrum (2000), an efficient image retrieval system had to take into account various users' needs and types of images available.

METHODOLOGY

Our study is based on the interactive image retrieval (iCLEF) track of the Cross Language Evaluation Forum (CLEF) (Gonzalo, Peinado, Clough, & Karlgren, 2009). The main goal of iCLEF 2009 was to collect as much information as possible from the interaction of the users with the system. To accomplish this goal the organizers made available a system called Flickling which implements a multilingual interface for accessing Flickr images. It supports 6 European languages (English, Dutch, French, German, Italian, and Spanish) and allows users to submit queries in any of the supported languages and retrieve images tagged in all the supported languages. The interface itself has been localized in all the 6 supported languages and can be set by the user when they start their search section. For this study we used the Flickling interface and recruited users to be observed as they use the system to find images assigned in the game. The users were asked to complete this game in two scenarios: one in which they could use all the features of the game (including asking for hints) and another in which they could not ask for hints. This second condition is similar to the traditional known-item search task.

RESULTS AND ANALYSIS

The results presented here are based on the data collected from 6 users recruited from the College of Information at UNT. Despite the small sample, the information collected from the retrieval task and the face to face interviews generated a rich set of data that included searches for 35 images (which we call here image search sessions), 303 queries, 1022 search terms with 289 unique terms, and the users' opinions of the level of difficulty of each image search task. For the search tasks we analyze each session to determine the search strategy being used by the user. We classify these strategies according to three types:

- General to specific: the user starts with general term and adds more terms that narrows the set of results.

- Specific to general: The user starts with a very specific query and changes it towards a more general set of terms.
- Parallel: the user tends to change the terms in her/his initial query for other equivalent or synonym terms.

Of the 35 search sections that were conducted by our users, 7 used a general to specific search strategy, 5 used a specific to general, and 23 used parallel queries. This seems to be consistent with the nature of the task since known item search usually requires the user to try different ways to describe the contents of the image until they either find it or give up. The strategies of using general to specific correlate with queries where the user claims that there were too many results returned, while the parallel strategy tends to correlate with queries where the user had a hard time describing the image.

A total of 303 queries were submitted by the users with an average query length of 3.37 terms. We also categorized each of the terms used in the query according to the 7 basic attributes ascribed to images as proposed by Greisdorf and O'Connor (2002): color, shape, texture, object, action, location, and affect. Since several images in our experiments included some text in the image that could be used as a clue by the user we decided to add this to the set of attributes ascribed to images. Our results are in line with the findings presented by Greisdorf and O'Connor since most of the words used by users tend to describe the objects present in the image (65%), followed by color (14%), action (8%), location (6%) and affect (6%). Textual clues occurred in 5 of the images and they were used by the users as part of their queries. Although the percentage is too small, these textual clues were important for finding the correct images in 4 out of the 5 cases where they appeared.

	Success	Failure
No Hints	2 (12%)	15 (88%)
Hints	14 (78%)	4 (22%)

Table 1 Success rates in the experimental conditions

The success rates related to the two conditions of our experiments are presented in Table 1. This table reveals that the known item search using images annotated with multilingual tags is extremely hard (our users were able to find the given image in 12% of the cases). As expected when the users are allowed to use the hints from Flickling the success rate increases significantly to 78%.

During the interviews most users reported having a very hard time describing the contents of the images. There were also problems with the translations provided by the system and many users resort to finding appropriate translations using Google and adding them to their searches. The users also expressed that they wish the system had a spell checking mechanism.

CONCLUSION

Our findings confirm research results reported in previous studies such as the fact that finding images is still quite challenging for users, even when the images have been tagged by a community such as Flickr. The major challenge reported by the users is trying to accurately describe with words the contents of an image. The multilingual aspect of the tags also adds an extra level of complexity that makes this task extremely challenging for users who are predominantly English speakers. We also suspect that vocabulary mismatch between the user's query and the tags assigned to the image plays a role but we were not able to measure the degree of this problem in our experiments. Our results also confirm the findings reported by Greisdorf and O'Connor (2002) regarding the distribution of terms ascribed to images with objects being the preferred type of characteristics followed by color/shape/texture. We recognize that this study has limitations related to the small amount of users as well as the convenience sample that we used to recruit participants.

REFERENCES

- Champ, H. (2009, October 10th). *Flickr Blog*. Retrieved May 30, 2010, from Flickr: <http://blog.flickr.net/en/2009/10/12/4000000000/>
- Gonzalo, J., Clough, P., & Karlgren, J. (2008). Overview of iCLEF 2008: Search Log Analysis for Multilingual Image Retrieval. *Working Notes for the CLEF 2008 Workshop*. Aarhus, Denmark.
- Gonzalo, J., Peinado, V., Clough, P., & Karlgren, J. (2009). Overview of iCLEF 2009: Exploring Search Behaviour in a Multilingual Folksonomy Environment. *Working Notes for the CLEF 2009 Workshop*. Corfu, Greece.
- Goodrum, A. (2000). Image information retrieval: An overview of current research. *informing Science*, 3 (2), 63-67.
- Greisdorf, H., & O'Connor, B. (2002). Modeling what users see when they look at images: A cognitive viewpoint. *Journal of Documentation*, 58 (1), 6-29.
- Jørgensen, C., & Jørgensen, P. (2005). Image querying by image professionals. *Journal of the American Society for Information Science and Technology*, 56 (12), 1346-1359.
- Matusiak, K. K. (2006). Towards user-centered indexing in digital image collections. *OCLC Systems & Services: International Digital Library Perspectives*, 22 (4), 283-298.
- Rorissa, A., & Iyer, H. (2008). Theories of cognition and image categorization: What category labels reveal about basic level theory. *Journal of the American Society for Information Science and Technology*, 59 (9), 1383-1392.
- Spink, A., & Jansen, B. J. (2006). Searching multimedia federated content web collections. *Online Information Review*, 30 (5), 485-495.