Multilingual interactive experiments with Flickr

Paul Clough
Department of
Information Studies
University of Sheffield
Sheffield, UK
`p.d.clough@sheffield.ac.uk`

Julio Gonzalo
Departamento de Lenguajes
y Sistemas Informáticos
UNED
Madrid, Spain
`julio@lsi.uned.es`

Jussi Karlgren
Swedish Institute of
Computer Science
Stockholm
Sweden
`jussi@sics.se`

Abstract

This paper presents a proposal for iCLEF 2006, the interactive track of the CLEF cross-language evaluation campaign. In the past, iCLEF has addressed applications such as information retrieval and question answering. However, for 2006 the focus has turned to text-based image retrieval from Flickr. We describe Flickr, the challenges this kind of collection presents to cross-language researchers, and suggest initial iCLEF tasks.

1 Information Retrieval Evaluation by User Experiment

Information retrieval systems, especially text retrieval systems, have benefited greatly from a fairly strict and straight-laced evaluation scheme, which enables system designers to run tests on versions of their system using a test collection of pre-assessed data. These relevance-oriented experiments shed light on comparative system performance and enable both introduction of new algorithms and incremental optimization. However, batch-oriented system evaluation based on large amounts of data, abstracted away from situational constraints, variation in usage, and interactiveness issues only addresses some of the bottlenecks to build a successful system.

The CLEF[1] Interactive Track (iCLEF[2]) is devoted to the comparative study of user inclusive cross-language search strategies. Over the past 5 years, iCLEF has studied three cross-language search tasks: retrieval of documents, answers and annotated images (Gonzalo and Oard, 2002; Gonzalo et al., 2005). All tasks involve the user interacting with information systems in a language different from that of the document collection. Although iCLEF experiments continue producing interesting research results, which may have a substantial impact on the way effective cross-language search assistants are built, participation in this track has remained low across the five years of existence of the track. Interactive studies, however, remain as a recognized necessity in most CLEF tracks.

Therefore, to encourage greater participation in 2006 our focus has turned to FLICKR[3], a large-scale, web-based image database with the potential for offering both challenging and realistic multilingual search tasks for interactive experiments. Our aim in selecting a primarily non-textual target to study textual retrieval is based on some of the multi-lingual and dynamic characteristics of FLICKR. We will outline them below.

2 The Flickr system

The majority of Web image search is text-based and the success of such approaches often depends on reliably identifying relevant text associated with a particular image. FLICKR is an online tool for managing and sharing personal photographs and currently contains over five million freely accessible images. These are available via the web, updated daily by a large number of users and available to all web users (users can access FLICKR for free, although limited to the upload of 20MB of photos per month).

---

[1]http://www.clef-campaign.org/
[2]http://nlp.uned.es/iCLEF/

[3]http://www.flickr.com/

2.1 Photographs in the collection

It is estimated that the complete FLICKR database contains 37 million photos with approximately 200,000 images added daily by 1.2 million members[4]. FLICKR provides both private and public image storage, and photos which are shared (around 5 million) can be protected under a Creative Commons (CC) licensing[5] agreement (an alternative to full copyright). Images from a wide variety of topics can be accessed through FLICKR, including people, places, landscapes, objects, animals and events. This makes the collection a rich resource for image retrieval research.

2.2 Annotations

In FLICKR, photos are annotated by authors with freely chosen keywords in a naturally multilingual manner: most authors use keywords in their native language; some combine more than one language. In addition, photographs have titles, descriptions, collaborative annotations, and comments in many languages. Figure 5 provides an example photo with multilingual annotations; Figure 5 shows what the query "cats" retrieves from the database, compared with what the query "chats" retrieves.

Annotations are used by the authors to organize their images, and by any user to search on. Keywords assigned to the images can include place names and subject matter, and photos can also be submitted to online discussion groups. This provides additional metadata to the image which can also be used for retrieval. An explore utility provided by FLICKR makes use of this user-generated data (plus other information such as Clickthroughs) to define an "interestingness" view of images[6].

3 Flickr at iCLEF 2006

Many images are accompanied by text, enabling the use of both text and visual features for image retrieval and its evaluation (Müller et al., 2006, see e.g.). Images are naturally language independent and often successfully retrieved with associated texts. This has been explored as part of ImageCLEF (Clough et al., 2005) for areas such as information access to medical images and historic photographs. The way in which users search

for images provides an interesting application for user-centered design and evaluation. As an iCLEF task, searching for images from FLICKR presents a new multilingual challenge which, to date, has not been explored. Challenges include:

- Different types of associated text, e.g. keywords, titles, comments and description fields.

- Collective classification and annotation using freely selected keywords (known as folksonomies) resulting in non-uniform and subjective categorization of images.

- Annotations in multiple languages.

Given the multilingual nature of the FLICKR annotations, translating the user's search request would provide the opportunity of increasing the number of images found and make more of the collection accessible to a wider range of users regardless of their language skills. The aim of iCLEF using FLICKR will be to determine how cross-language technologies could enhance access, and explore the user interaction resulting from this.

4 Proposed tasks

For iCLEF, participants to this evaluation campaign will be provided with the following:

- A subset of the Flickr collection including annotations and photographs[7].

- Example (realistic) search tasks. Ideally these search tasks will reflect real user needs which could be derived from log files, studies or similar retrieval tasks.

- A framework in which to run an evaluation.

5 Summary

Flickr will allow us to create an extremely interesting interactive task based on truly heterogeneous annotations (that will in turn hopefully attract more participants). Using images from within a Web environment is a realistic and contemporary search challenge and allows many important research questions to be addressed from

---

[4]These figures are accurate as of October 2005: http://www.wired.com/news/ebiz/0,1272,68654,00.html

[5]http://creativecommons.org/image/flickr, http://flickr.com/creativecommons/

[6]http://www.flickr.com/explore/interesting

[7]We are currently in negotiations with Yahoo! (owners of Flickr) and Flickr to provide researchers with legitimate access to a subset of the collection.

a quickly developing field. User-centered studies are required within both text and image retrieval, but are often neglected as they require more effort and time from participating groups than a system-centered comparison that can often be run without human intervention. Still, user-centered evaluation cannot be replaced and the influence of the user on the results is in general stronger than the influence of the system itself.

References

Paul Clough, Henning Müller, and Mark Sanderson. 2005. The clef 2004 cross language image retrieval track. In Carol Peters, Paul Clough, Julio Gonzalo, Gareth Jones, Michael Kluck, and Bernardo Magnini, editors, *Multilingual Information Access for Text, Speech and Images: Results of the Fifth CLEF Evaluation Campaign*, number 3491/2005 in Lecture Notes in Computer Science, pages 597–613. Springer, Heidelberg, Germany.

Julio Gonzalo and Doug Oard. 2002. The clef 2002 interactive track. In *Advances in Cross-Language Information Retrieval*, number 2785 in Lecture Notes in Computer Science. Springer-Verlag, Berlin-Heidelberg-New York.

Julio Gonzalo, Paul Clough, and A Vallin. 2005. Overview of the clef 2005 interactive track. In *Working notes of the CLEF workshop*, Vienna, Austria, September.

Henning Müller, Paul Clough, William Hersh, Thomas Deselaers, Thomas Lehmann, and Antoine Geissbuhler. 2006. Using heterogeneous annotation and visual information for the benchmarking of image retrieval systems. In *SPIE conference Photonics West, Electronic Imaging, special session on benchmarking image retrieval systems*, San Diego, February.
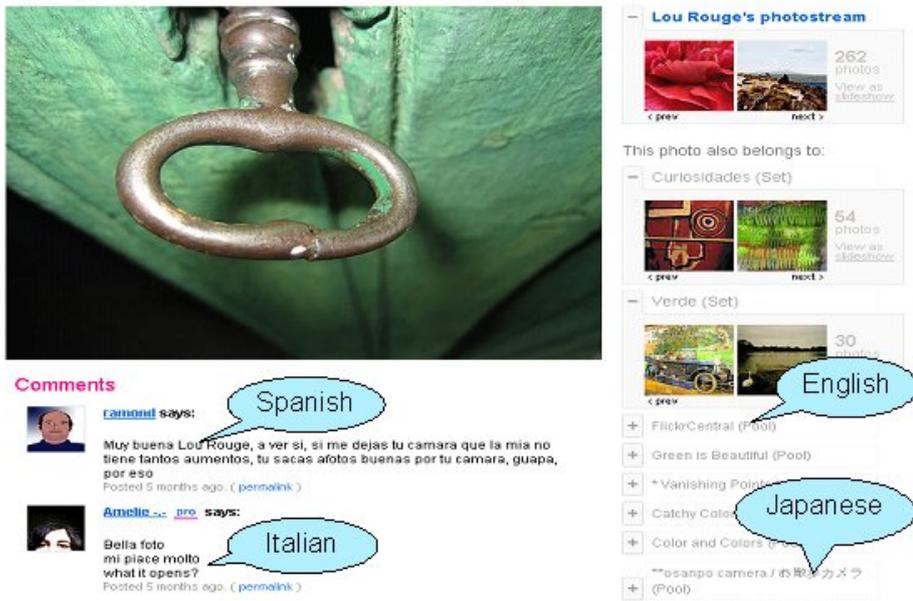
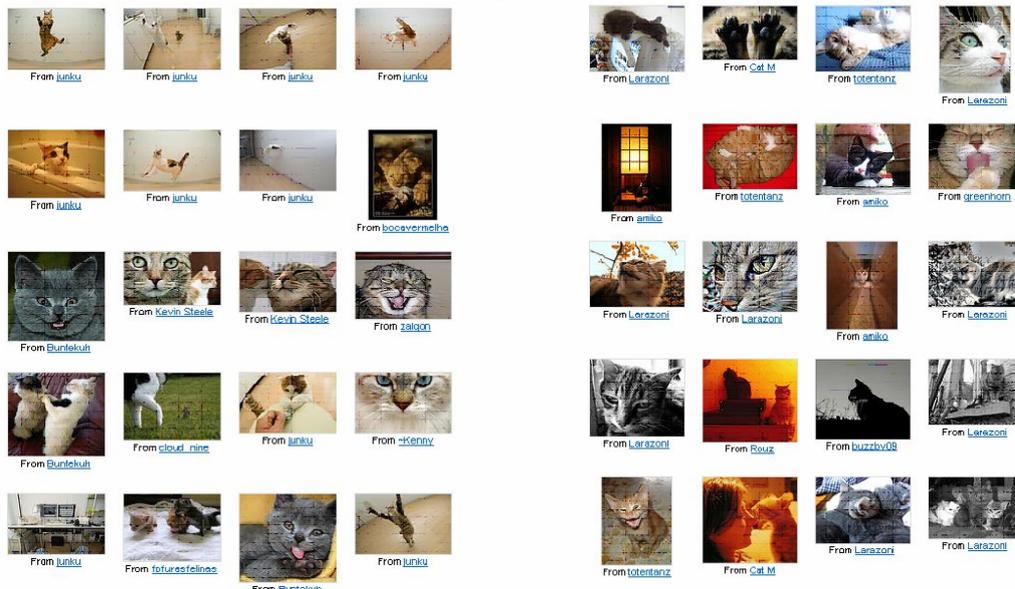Figure 1: Example multilingual annotations in Flickr.

Figure 2: Retrieval of "cats" (left) and "chats" (right).