

“INFORMATION RETRIEVAL FROM THE INTERNET AND WWW, USING SEARCH BY IMAGE: A TUTORIAL”

By Prof. Paul Nieuwenhuysen,

Vrije Universiteit Brussel, Belgium

Abstract

This tutorial workshop is based on a continuing investigation of the power, applicability, usefulness and limitations of search by image through the Internet. In this relatively new method for information retrieval, a query does not consist of text, but of an image file. The search results lead to images on the WWW and also to related texts. Other terms used for this method are

- Search(ing) by example
 - Reverse image search(ing)
 - Reverse image lookup = RIL
 - Backwards image search(ing)
 - Inside search(ing)
- **Some of our findings:**
 1. Several online services are available free of charge to search by image.
 2. Differences among these services are substantial.
 3. The search service offered by Google performs relatively well.
 4. Google can reveal images present on the Internet, which are copies of the query/source image; however, the success is quite variable from case to case.
 5. This recall performance is strongly correlated with the performance of a more classical Google search by text to find copies of the query/source image file on the Internet.
 6. Even images that are modified versions of the query/source image can be revealed by Google; more specifically, such modified versions can differ from the source image in size and in colours; even fragments of the source image that are present on the internet can be revealed.
 7. Our tests have demonstrated that since 2014 search by image can not only find images that are visually similar to the query/source image, but can even retrieve images that are semantically similar/related to the query/source image, even when visual similarity is not obvious. The search results may also include a description of the subject on the image, and this can of course be interesting if the user has not yet much knowledge about the subject, so that using a specific text query becomes possible. Furthermore, other information related to the image and relevant links may also be included in the search results.
 8. The performance of search by image to find images that are semantically similar to the query/source image is improving.
 - **Various applications can be shown:**
 - Starting from your own image, you may find copies or even modified versions on the WWW.
 - Starting from some interesting image, that you have not created, but that you consider as interesting, and that is perhaps not the original version, you may find other and better versions.
 - Starting from some interesting source image, you may find images with a subject that is related to the subject of that source image.

Keywords

reverse image search, Google, information discovery, information retrieval, semantic gap