

Visual Information Retrieval through Interactive Multimedia Queries (Ph.D. thesis)

Aurhor: José Manuel Torres {torres@comp.lancs.ac.uk}
Supervisor: Prof. David Hutchison {dh@comp.lancs.ac.uk}
Computing Department, Lancaster University, United Kingdom, 2005

Abstract

Supported by technological evolution, there is an exponential growth of the amount of digital visual information available and being used nowadays. Users wish to search that information by means of descriptive information, although clearly the so-called "Semantic Gap" continues to persist between users and machines. In an ideal scenario, image retrieval systems would reflect the best characteristics of humans with their capability in image understanding and computers with their tremendous processing power and tireless logic. As part of a general research trend in image retrieval, this thesis focuses on the development of a new approach to image retrieval that aims to get closer to the "Intelligent Image Retrieval" paradigm often discussed in the literature.

This thesis starts off with a general introduction into the above-mentioned research area and a description of the typical requirements of image users. A critical examination of the existing content-based image retrieval technology is provided. The core of the thesis presents the design and implementation of the Visual Object Information Retrieval (VOIR) prototype. The model implemented in VOIR combines two layers, the conceptual and the feature-based. It works at the region or visual object level and uses a textual thesaurus to explore associations between text and regions of images and to enable concept-level queries. VOIR implements a relevance feedback mechanism that improves associations between text and regions of images: the more the system is used, the more text-image associations are learned.

The evaluation of the prototype VOIR is carried out using a relatively small collection of 300 pre-segmented and hand-labelled images. These experiments serve to validate the use of the concept-based queries, the relevance feedback mechanism and the scheme used to learn associations between text and regions of images. The results indicate the very strong potential of the tested model.