

Chapter XII

Personalized Content–Based Image Retrieval

Iker Gondra

St. Francis Xavier University, Canada

ABSTRACT

In content-based image retrieval (CBIR), a set of low-level features are extracted from an image to represent its visual content. Retrieval is performed by image example where a query image is given as input by the user and an appropriate similarity measure is used to find the best matches in the corresponding feature space. This approach suffers from the fact that there is a large discrepancy between the low-level visual features that one can extract from an image and the semantic interpretation of the image's content that a particular user may have in a given situation. That is, users seek semantic similarity, but we can only provide similarity based on low-level visual features extracted from the raw pixel data, a situation known as the semantic gap. The selection of an appropriate similarity measure is thus an important problem. Since visual content can be represented by different attributes, the combination and importance of each set of features varies according to the user's semantic intent. Thus, the retrieval strategy should be adaptive so that it can accommodate the preferences of different users. Relevance feedback (RF) learning has been proposed as a technique aimed at reducing the semantic gap. It works by gathering semantic information from user interaction. Based on the user's feedback on the retrieval results, the retrieval scheme is adjusted. By providing an image similarity measure under human perception, RF learning can be seen as a form of supervised learning that finds relations between high-level semantic interpretations and low-level visual properties. That is, the feedback obtained within a single query session is used to personalize the retrieval strategy and thus enhance retrieval performance. In this chapter we present an overview of CBIR and related work on RF learning. We also present our own previous work on a RF learning-based probabilistic region relevance learning algorithm for automatically estimating the importance of each region in an image based on the user's semantic intent.

INTRODUCTION

In recent years, the rapid development of information technologies and the advent of the Web have accelerated the growth of digital media and, in particular, image collections. As a result and in order to realize the full potential of these technologies, the need for effective mechanisms to search large image collections becomes evident. The management of text information has been studied thoroughly, and there have been many successful approaches for handling text databases (see Salton, 1986). However, the progress in research and development of multimedia database systems has been slow due to the difficulties and challenges of the problem.

The development of concise representations of images that can capture the essence of their visual content is an important task. However, as the saying “a picture is worth a thousand words” suggests, representing visual content is a very difficult task. The human ability to extract semantics from an image by using knowledge of the world is remarkable, though probably difficult to emulate.

At present, the most common way to represent the visual content of an image is to assign a set of descriptive keywords to it. Then, image retrieval is performed by matching the query text with the stored keywords (Rui & Huang, 1998). However, there are many problems associated with this simple keyword matching approach. First, it is usually the case that all the information contained in an image cannot be captured by a few keywords. Furthermore, a large amount of effort is needed to do keyword assignments in a large image database. Also, because different people may have different interpretations of an image’s content, there will be inconsistencies (Rui & Huang, 1998). Consider the image in Figure 1. One might describe it as “mountains,” “trees,” and/or “lake.” However, that particular description would not be able to respond to user queries for “water,” “landscape,” “peaceful,” or “water reflection.”

In order to alleviate some of the problems associated with text-based approaches, content-based image retrieval (CBIR) was proposed (see Faloutsos et al., 1993, for examples of early approaches). The idea is to search on the images directly. A set of low-level features (such as color, texture, and shape) are extracted from the image to characterize its visual content. In traditional approaches (Faloutsos et al., 1993; Gupta & Jain, 1997; Hara, Hirata, Takano, & Kawasaki, 1997; Kelly, Cannon, & Hush,

Figure 1. Sample image



25 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/personalized-content-based-image-retrieval/4160

Related Content

An Efficient System for Blocking Pornography Websites

Tarek M. Mahmoud, Tarek Abd-El-Hafeez and Ahmed Omar (2014). *Computer Vision and Image Processing in Intelligent Systems and Multimedia Technologies* (pp. 161-176).

www.irma-international.org/chapter/an-efficient-system-for-blocking-pornography-websites/108414

Fall Detection Depth-Based Using Tilt Angle and Shape Deformation

Fairouz Merrouche and Nadia Baha (2018). *International Journal of Computer Vision and Image Processing* (pp. 26-40).

www.irma-international.org/article/fall-detection-depth-based-using-tilt-angle-and-shape-deformation/214072

Moving Vehicles Detection in Traffic Video Using Modified SXCS-LBP Texture Descriptor

Arun Kumar H. D. and Prabhakar C. J. (2015). *International Journal of Computer Vision and Image Processing* (pp. 14-34).

www.irma-international.org/article/moving-vehicles-detection-in-traffic-video-using-modified-sxcs-lbp-texture-descriptor/159805

Surgeon Assistive Augmented Reality Model with the use of Endoscopic Camera for Line of Vision Calculation

Anastasia Daskalaki, Kostas Giokas and Dimitris Koutsouris (2013). *Image Processing: Concepts, Methodologies, Tools, and Applications* (pp. 658-674).

www.irma-international.org/chapter/surgeon-assistive-augmented-reality-model/77568

Clinical Machine Learning in Action: CAD System Design, Development, Tuning, and Long-Term Experience

Yoshitaka Masutani, Mitsutaka Nemoto, Yukihiro Nomura and Naoto Hayashi (2013). *Image Processing: Concepts, Methodologies, Tools, and Applications* (pp. 621-638).

www.irma-international.org/chapter/clinical-machine-learning-action/77566