Feature Extraction in Content-Based Image Retrieval

Jacob John Foley
School of Science and Technology, University of New England, Australia

Paul Kwan
School of Science and Technology, University of New England, Australia

INTRODUCTION

In recent decades, the increased usage and availability of digital cameras has created a vast amount of new information captured in the form of digital images. These images have been given an unprecedented level of accessibility through the Internet and sharing in social media. It is difficult to represent these images using text descriptions due to the amount of labour required to annotate large collections and the occurrence of inconsistencies in annotations caused by the differing perceptions of the individual annotators. This makes searching images using text-based methods ineffective (Rui, Huang & Chang, 1999).

New techniques in Content Based Image Retrieval (CBIR) are being developed to accommodate indexing and searching images using Feature Extraction. Feature extraction algorithms use the content of digital images to produce Feature Vectors, which represent the important details of an image in a concise form and allow for complex analysis of the source image.

Feature vectors can be compared using Similarity Measures that quantitatively describe the difference between two sets of feature vectors. A high similarity between feature vectors corresponds to a high likelihood that the two vectors are being used to represent the same or a similar object. Techniques can be applied to further enhance the results obtained from similarity measures, such as relevance feedback, context mining, supervised machine learning, object ontologies and semantic templates (Liu, Zhang, Lu & Ma, 2007).

CBIR has applications in fields such as medicine, defence, weather forecasting, security, personal photo collections, social photo sharing and digital cultural heritage. Any system which benefits from assisted photo organisation and identification is a potential application.

Datta, Joshi, Li & Wang (2008) classify searches as being made via association, where there is no clear goal initially but refined searches lead to the target image, via an aimed search, where a clear goal is known, or via a category search, where the user wishes to find images within a category with no clear goal. CBIR provides an effective means of performing all three kinds of search.

In this article, we examine the most common features for indexing and searching images. This provides an introduction to the concepts used in Feature Extraction. For further information on specific techniques and their implementations, refer to the Additional Reading section.

BACKGROUND

While the field of CBIR is being developed to address the issue of reliable and efficient description of images, it faces certain challenges. One of the greatest challenges to CBIR is the issue of the semantic gap.

In an example provided by Datta, Joshi, Li & Wang (2008), the reader is asked to consider what a "perfect" picture of a subject might be in terms of its features. Not only will this vary on the individual, but it will also be difficult to describe in terms of a set of features that correspond to the semantic concept of perfection. Even descriptions such as "find images of a cheerful crowd" have sufficient ambiguity that a computer program would have difficulty associating a set of feature vectors with the desired result.

An additional challenge relates to the selection of feature extraction algorithms. There is not a single correct approach to feature extraction for general purpose CBIR, nor a single type of feature which is demonstrably

DOI: 10.4018/978-1-4666-5888-2.ch583

better across all applications than other features. This grants significant leeway for creative and exploratory approaches in the field. The similarity measures used to compare feature vectors also vary, with popular approaches using distances such as Euclidean distance, Minkowsky distance (Kaur & Jyoti, 2013), or trees using their edit distance (Yang, Kalnis & Tung, 2005).

Veltkamp & Tanase (2000) describe alternatives to CBIR. These include browsing an image database until the target image is located, specifying the target image in terms of a keyword or image query, or providing a sketch of the desired image and using relevance feedback to improve the results. These approaches are labor intensive and have less potential for computer automation than CBIR.

In the rest of this article, five common features for which feature extraction algorithms have been developed will be described. A significant amount of research conducted in Colour and Saliency in computer vision has been inspired by biological vision systems. Signal processing techniques can be used to extract and analyse Shape and Texture in an image. Object Recognition and Region Matching are refinements of these signal processing approaches that use an additional level of abstraction to enhance the sophistication of the results produced by these extraction algorithms. Spatial Location is a feature which can increase the effectiveness of other methods.

COMMON IMAGE FEATURES

Colour

The use of colour in image processing is inspired by biological vision systems, which use the dominant colour and saliency of an image to provide excellent results in scene identification and navigation. Colour and intensity information is readily available in digital images, providing the raw data for the colour feature extraction algorithm. Determining the dominant colour is relatively inexpensive in terms of computation when compared to methods relying on texture, shape or object recognition, hence its widespread use.

Humans are able to find the general composition of an image with a relatively short exposure time using the biological computation of an image's gist, even down to exposure times as low as 30 milliseconds. This is possible as the spatial distribution and colour information of an image is used to identify a scene in biological systems prior to the extraction and recognition of individual objects and their relationships. This kind of holistic information is resistant to distortion such as blurriness and visual noise, making it an excellent feature of a visual system (Oliva & Torralba, 2006).

Context-based scene recognition takes inspiration from biological systems and identifies a scene by calculating the image's gist using saliency and colour information. For example, identifying a scene as being a beach makes it more likely that a large blue area corresponds to the ocean rather than ice. This contrasts with object-based scene recognition, which aims to detect and identify objects and infer the scene based on their relationships, such as ice and snow indicating an arctic scene (Siagian & Itti, 2007).

Oliva & Torralba (2006) provide evidence to support that the human visual system processes low and high-frequency information at different spatial scales, and that the visual system may be selective in which frequency band is used first. When asked to identify a scene or the emotion expressed on a face, viewers used low-frequency information relating to colour as opposed to high-frequency information relating to shape, and were able to answer questions asked about images with exposures of only 30 milliseconds. High-frequency information was obtained at exposure times of 150 milliseconds. However, when the viewers were asked to identify the gender of a face, high-frequency information was favoured within the 30 millisecond exposure time. This supports the approach of using combinations of features such as colour and shape when the CBIR approach is modelled on biological systems.

Techniques that utilise colour as a feature vary primarily based on whether global information over the entire image is used, or whether the spatial distribution of local areas of colour is taken into account. The colour spaces used can also vary, with perceptually uniform colour spaces like CIELAB being generally preferable as the distance between colours can be calculated using Euclidean distance, and these colour spaces correspond more closely to human perception of colour. While there are a wide range of techniques used for extracting colour feature vectors, a common approach is to create a colour histogram of the dominant colours in an image and compare histograms to determine the similarity between images (Lei, Fuzon & Bo, 1999).

Related Content