

Chapter VI

Content-Based Image Retrieval: From the Object Detection/Recognition Point of View

Ming Zhang

University of Calgary, Canada

Reda Alhaji

University of Calgary, Canada

ABSTRACT

Content-Based Image Retrieval (CBIR) aims to search images that are perceptually similar to the query-based on visual content of the images without the help of annotations. The current CBIR systems use global features (e.g., color, texture, and shape) as image descriptors, or use features extracted from segmented regions (called region-based descriptors). In the former case, descriptors are not discriminative enough at the object level and are sensitive to object occlusion or background clutter, thus fail to give satisfactory result. In the latter case, the features are sensitive to the image segmentation, which is a difficult task in its own right. In addition, the region-based descriptors are still not invariant to varying imaging conditions. In this chapter, we look at the CBIR from the object detection/recognition point of view and introduce the local feature-based image representation methods recently developed in object detection/recognition area. These local descriptors are highly distinctive and robust to imaging condition change. In addition to image representation, we also introduce the other two key issues of CBIR: similarity measurement for image descriptor comparison and the index structure for similarity search.

1. INTRODUCTION

The explosive growth of digital images in our lives requires efficient image data management systems for image storage and retrieval. The early image retrieval systems are based on manually annotated descriptions and have the following drawbacks (Chang and Hsu, 1992): 1. manually annotating is too expensive for large database; 2. the annotation is subjective and context-dependent. From the early 1990s, **content-based image retrieval (CBIR)** became an active and fast developing research area. Simply speaking, **CBIR** is a method of image retrieval that searches images (from the database) that are similar to the query image based on visual content (by “appearance” according to human perception). More formally, we may define the CBIR as follows:

Definition of content-based image retrieval: Given a large image database U , an image representation method based on image primitives (e.g., pixel intensities) and a dissimilarity measure $D(p,q)$ defined on the image representation, find (using certain index) the M images $p \in U$ with the lowest dissimilarity to the query image q , the resulting M images are ranked by ascending dissimilarity.

According to the above definition, the **CBIR** is query by example. Our definition is narrower than the general case **CBIR** as query by example is the most common form of **CBIR**. Our chapter is based on this definition.

As the ultimate goal of image retrieval system is to find images that the users are interested in and the result images are determined by the content of the query image, the first problem we have to deal with is: Can the query image always express clearly what the users are interested in?

In a traditional database, a query is a formally phrased information request that clearly expresses the user’s information needs. Put it in the context of CBIR, the user’s information request should be unambiguously determined by the visual content of the query image. If a user is looking for a specific object or the objects that are very similar to a specific object by appearance, for example, his lost dog or the dogs that look very much like his dog, he may use a photo of his dog as the query image. In this case, the photo is better than any words to express clearly the user’s interest. However, users who are only interested in a generic category of objects, for instance, “animal”, can never be able to express this information need by submitting an image of a specific dog. They will not get satisfactory result by doing so. Some researchers refer to this as the semantic gap problem. To bridge the semantic gap is to make the system generalize semantic concept from a single specific image, which is, in our opinion, impossible, because a high level concept can only be generalized from a large number of instances using machine learning techniques.

In this chapter, we will not consider the semantic gap problem. We assume that the CBIR is used when the users’ information need is better expressed by image than by words. Specifically, query by example is preferred when the user is looking for a specific object/scene or a narrow category of objects/scenes that the instances belonging to this category are similar by appearance, in both cases the user’s information need cannot be expressed unambiguously by simple words. From this perspective, the content-based image retrieval can be roughly regarded as an object detection/recognition problem. Therefore, we may use some techniques originally developed for object detection/recognition in CBIR system.

In this chapter, we discuss three key issues of the CBIR:

28 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/content-based-image-retrieval/4154

Related Content

Automatic Learning Improves Human-Robot Interaction in Productive Environments: A Review

Mauricio Andres Zamora Hernandez, Eldon Caldwell Marin, Jose Garcia-Rodriguez, Jorge Azorin-Lopez and Miguel Cazorla (2017). *International Journal of Computer Vision and Image Processing* (pp. 65-75).

www.irma-international.org/article/automatic-learning-improves-human-robot-interaction-in-productive-environments/188762

A Review of Registration Methods on Mobile Robots

Vicente Morell-Gimenez, Sergio Orts-Escolano, José García-Rodríguez, Miguel Cazorla and Diego Viejo (2013). *Robotic Vision: Technologies for Machine Learning and Vision Applications* (pp. 140-153).

www.irma-international.org/chapter/review-registration-methods-mobile-robots/73189

Rotation Invariant Texture Image Retrieval with Orthogonal Polynomials Model

R. Krishnamoorthi and S. Sathya Devi (2011). *International Journal of Computer Vision and Image Processing* (pp. 27-49).

www.irma-international.org/article/rotation-invariant-texture-image-retrieval/64184

Recent Trends of Addressing COVID-19 Disease by AI/ML

Shawni Dutta, Utsab Mukherjee and Digvijay Pandey (2023). *Handbook of Research on Thrust Technologies' Effect on Image Processing* (pp. 395-419).

www.irma-international.org/chapter/recent-trends-of-addressing-covid-19-disease-by-aiml/328045

Construction of 3D Triangles on Dupin Cyclides

Bertrand Belbis, Lionel Garnier and Sebti Foufou (2011). *International Journal of Computer Vision and Image Processing* (pp. 42-57).

www.irma-international.org/article/construction-triangles-dupin-cyclides/55099