Information Sciences 399 (2017) 154–182

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Image retrieval using the extended salient region

Jing Zhang, Shengwei Feng, Da Li, Yongwei Gao, Zhihua Chen, Yubo Yuan*

Department of Computer Science and Engineering, East China University of Science and Technology, Shanghai, 200237, China

ARTICLE INFO

Article history:
Received 8 March 2016
Revised 11 January 2017
Accepted 3 March 2017
Available online 7 March 2017

Keywords:
Extended salient region
Region matching
MLAP
Image retrieval
Image processing

ABSTRACT

The salient region is the most important part of an image. The salient portion in images also attracts the most attention when people search for images in large-scale datasets. However, to improve image retrieval accuracy, considering only the most salient object in an image is insufficient because the background also influences the accuracy of image retrieval. To address this issue, this paper proposes a novel concept called the extended salient region (ESR). First, the salient region of an input image is detected using a Region Contrast (RC) algorithm. Then, a polar coordinate system is constructed; the centroid of the salient region is set as the pole. Next, the regions surrounding the salient region are determined by the neighboring regions, moving in a counterclockwise direction. The resulting combination of the salient region and its surrounding regions is defined as the ESR. We extract the visual content from the ESR using the well-known Bag of Words (BoW) model based on Gabor, SIFT and HSVH features and propose a graph model of the visual content nodes to represent the input image. Then, we design a novel algorithm to perform matching between two images. We also define a new similarity measure by combining the similarities of the salient region and the surrounding regions using weights. Finally, to better evaluate the image retrieval accuracy, an improved measure called the mean label average precision (MLAP) is proposed. The results of experiments on three benchmark datasets (Corel, TU Darmstadt, and Caltech 101) demonstrate that our proposed ESR model and region-matching algorithm are highly effective at image retrieval, and can achieve more accurate query results than current state-of-the-art methods.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

1.1. Motivations

Because of the steady growth in computing power, the declining cost of storage, the rapid development of the Internet, and advanced digital image-capturing technologies, the number of stored images has exploded in recent years. Consequently, effective indexing and searching of large-scale image databases has become a necessary task and an urgent problem. Due to the enormous workload, the cost of manual labeling and the subjectivity of human perception, text-based labeling methods have become impractical. Therefore, content-based image retrieval (CBIR), which is based on automatically extracted visual features, has attracted increasing attention [1,9,28,29].

Classical CBIR takes an entire query image as the input and then retrieves similar images. However, this reliance on a global view of images fails to capture the important properties of individual objects [9,28]. Images can nearly always be

* Corresponding author.
E-mail address: ybyuan@ecust.edu.cn (Y. Yuan).

http://dx.doi.org/10.1016/j.ins.2017.03.005
0020-0255/© 2017 Elsevier Inc. All rights reserved.

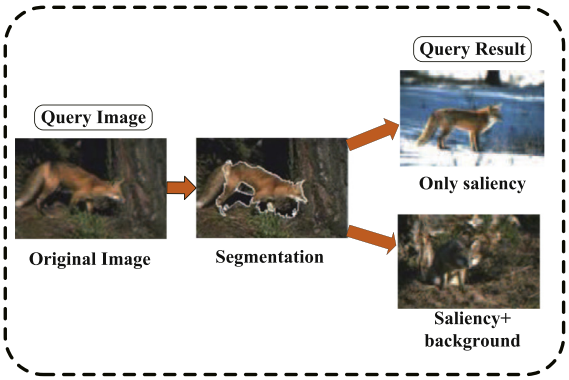

Fig. 1. Comparison of image retrieval by considering only the salient region or the salient region plus the background.

segmented into several regions in which each region contains different low-level features and semantic information. A global image content representation ignores the semantic and feature differences of these image regions, often causing a query to fail. To narrow the semantic gap and improve image retrieval performance, region-based image retrieval (RBIR) was proposed [14,16,23]. RBIR overcomes the drawback of considering only global features by representing images at an object level, which is closer to way the human visual system perceives images. RBIR algorithms segment the images into several regions and represent each region individually, which improves the accuracy of the overall representation. In addition, region matching replaces whole-image matching by focusing on the semantic information and correlation of the individual regions. Hence RBIR has become a research hotspot in the image retrieval field.

However, not all the parts of an image are equally important; usually, one region tends to attract more attention. Such regions are called salient objects [5]. A salient object contains the most important semantic content of an image and, thus, largely determines the accuracy of query results. Hence, there are two main types of algorithms in RBIR. The first type considers only the salient region and ignores other regions [37,38]. However, by considering only the salient region and neglecting other regions, the query results will be insufficiently precise. Fig. 1 illustrates the query results obtained by either considering or ignoring the non-salient regions. Obviously, the query results that also consider the non-salient regions are more accurate and, therefore, can better satisfy the query requirements.

The second type of RBIR algorithm treats all regions equally [18,23]. However, because not all the parts of an image are equally important—-most are background—if we treat them equally, the important objects will be outweighed and the query accuracy will be reduced. Hence, to achieve more precise image retrieval, we not only need to consider the importance of the salient object but also the similarity of background regions.

1.2. Contributions and structure

Considering the background is important in achieving accurate image retrieval. We propose a novel concept called an extended salient region (ESR) and define a new model of image retrieval based on this concept. The original contributions made in this paper are as follows.

- We propose a new concept called the extended salient region. A salient region is typically the most important area in an image and contains the main semantic information. In the ESR model, we not only fully consider the importance of the salient region but also consider the areas surrounding the salient region. This approach improves the accuracy of our image content representation but avoids the interference caused by representing and searching irrelevant background content.
- We adopt a graph model to represent the ESR that is both concise and powerful. The image content representation by the ESR combined with the graph model is more precise and, to some extent, better describes the semantic construction of an image.
- We propose a new image retrieval algorithm based on the ESR that finds the most similar images using an effective region-matching algorithm. Our algorithm not only considers the visual similarity between regions but also the relative positions of the regions. Experiments demonstrate that our method can achieve more accurate image retrieval results than existing methods.
- To evaluate image retrieval query results more effectively, we propose a new evaluation measure called Mean Label Average Precision (MLAP), which employs all the labels in an image to evaluate the image retrieval performance.

The remainder of this paper is organized as follows. Section 2 reviews related works. The notations and operators used in this study are defined in Section 3. In Section 4, we introduce the ESR concept. The method of image representation based on ESR is introduced in Section 5, and its visual content representation using the Bag of Words approach is demonstrated in Section 6. Section 7 provides an image retrieval method based on ESR. The algorithms underlying this method are presented

Download English Version:

https://daneshyari.com/en/article/4944498

Download Persian Version:

https://daneshyari.com/article/4944498

Daneshyari.com