Research Paper Engineering

Narrowing Semantic Gap in Content-based Image Retrieval

* Naveen kumar Joshi ** Ritesh Khedekar

* Student Scholar, IES-IPS,Indore (M.P.)

* Student Scholar, IES-IPS,Indore (M.P.)

ABSTRACT

Due to the low-level image features it utilizes, the semantic gap problem is hard to bridge and performance of CBIR systems is still far away from users' expectation. Image annotation, region-based image retrieval and relevance feedback are three main approaches for narrowing the "semantic gap". In this paper, recent development in these fields is reviewed and some future directions are proposed in the end.

Keywords: Content-based Image Retrieval; Region-based Image, Retrieval; Image Annotation; Relevance Feedback

I INTRODUCTION

Continuing advances in digital image capture and storage are resulting in a proliferation of imagery and associated problems of information overload in image domains. Searching interested images based on visual properties or contents of images is a challenging problem and it has received much attention from researchers in the last decades. Content-based image retrieval (CBIR) is any technology that in principle helps to organize digital picture archives by their visual content. [1], therefore, anything ranging from an image similarity function to a robust image annotation engine falls under the purview of CBIR, such as image understanding, computer vision, machine learning, information retrieval, human-computer interaction, database systems, Web and data mining, information theory, statistics, and psychology contributing are becoming part of the CBIR community [2].

CBIR is the retrieval of interested images from image collections to match the query based on visual properties of images themselves. The visual properties used are often low-level features, such as colors, textures, shapes etc. extracted from the images. CBIR is often accomplished by comparing these low-level features based on an assumption that similarity of low-level features of images can reflect the perceptual similarity among images [3, 4]. Some significant CBIR research works have been reported in [5, 6]. In one of the earliest CBIR work by IBM, the IBM's QBIC system, two kinds of low-level features, color and shape, are used. In [5], another well known early work on CBIR, the Berkeley Chabot system, both text-based image description and low-level visual features of images are used in retrieving images from a collection of photographs of California Department of Water Resources [5]. Most of the early work in CBIR considered color features, especially the color histogram, as one type of their low-level visual features. Generally speaking, CBIR can be categorized into two kinds: Annotation-Based Image Retrieval (ABIR) approaches [7, 8] and Query-By-Example (QBE) approaches [9]. For ABIR approaches, they are of cross-medium type since queries proposed by the user are often in the form of text and the search targets are images (e.g. the user wants to find images in a database containing horses, in which the word "horses" the user inputs to the search system is in the form of text). On the other hand, QBE approaches are of mono-medium type because both the user's queries and the search targets are images (e.g. the user wants to find images in a database containing horses in which queries the user provides are images depicting horses) [10,11].

Although great interest and a large number of new techniques and systems are emerged in content-based image retrieval in the last decades, the gap between low-level visual features and high-level semantic understanding of images, which is also known as the semantic gap problem[12], the gap between the object in the world and the information in a computational description derived from a recording of that scene is the bottleneck to further improvement of the performance of a content-based image retrieval system. Therefore, in order to solve the problems and improve CBIR performance, image annotation, Region-Based Image Retrieval (RBIR) approaches and relevance feedback have been received more attention in the recently years.

In this paper, we review three main approaches in CBIR research. The rest of this paper is arranged as follows. In Section 2, some work in image annotation which narrowing the semantic gap between a user's retrieval request and the actual information or features used to perform retrieval are introduced. Some key methods and techniques of RBIR approaches which based on the fact that high-level semantic understanding of images can be better reflected by local features of images, rather than global features are presented in Section 3. In Section 4, some work of relevance feedback methods are discussed. In Section 5, we give the conclusion of this review and propose future directions of CBIR.

II. IMAGE ANNOTATION

The difficulty of narrowing "semantic gap" lies in estimating the context of the search request or the usage context for retrieved images [8]. For example, in content-based approaches a user query (usually in the form of a sample image), is matched against low-level features (e.g. colors, shapes), extracted from an image, thereby providing no insight into why the user wants to retrieve the specified imagery or of how the retrieved imagery will actually be used. In order to reveal image semantics at a higher level, one of the approaches is use different models and machine learning methods to find the relation between image visual features and semantics, then label the image with keywords [13-15], this is image annotation.

III. REGION-BASED IMAGE RETRIEVAL

In traditional CBIR, features are often global features, which are extracted from the entire images. However, global features can hardly reflect the semantic understanding of images as the interpretation of image is often by objects in images.

These objects are often the targets users want to search from the retrieval system. Therefore, more and more researchers nowadays have begun to pay more attention to “local features”, which is believed to better reflect the notion of objects and narrow down the semantic gap between low-level features and high-level semantic understanding of images.

RBIR is an image retrieval approach which focuses on contents from regions of images, not the content from the entire image in early CBIR. For RBIR, it first segments images into a number of regions, and extracts a set of features, which are known as “local features”, from segmented regions. A similarity (or distance) measure determining the similarity between target regions in the query and a set of segmented regions from other images is utilized later to determine relevant images to the query based on local regional features. The motivation of RBIR approaches are based on the fact that high-level semantic understanding of images can be better reflected by local features of images, rather than global features. Since users are often more concerned with interested objects in images when using an image retrieval system, considering images based on regions (RBIR) allows users to pay more attention to regional properties that may better characterize objects which are also made up of local regions. This strategy is able to better reflect the characteristics of images from the perspective of image regions and objects, and in turn further improve the image retrieval performance.

A step prior to constructing local features of images is to segment images into several regions, which may possibly retain their own semantic meaning. However, region segmentation, which is based on low-level features of images, is often inconsistent on images of the same scene captured at different time and it is still an open issue arousing many controversies and research efforts in computer vision nowadays. Therefore, some researchers proposed to simplify or even avoid this hard problem. A method of image retrieval based on the information provided by histogram analysis of the intensity or grayscale values of images is proposed in [. In Marques et al. segmented out regions that match salient Region of Interest (ROI) defined by studying human perception. Also, Wang et al. applied the same ROI idea in their RBIR work [36]. The way they construct their ROIs is mainly based on the wavelet decomposition of an image and they used the lowest frequency sub-band as an approximation of the original image. After forming the original ROI on the approximation representation of images, they selected centroid of the ROI as a bench-mark reference and found the maximum distance from each pixel to the bench-mark for subregion segmentation. In Fan et al. proposed a new framework to achieve a middle-level understanding of contents of images. The method they developed is to find salient objects based on some pre-defined basic vocabularies. One of the earliest RBIR work, the Blobworld system [38], utilized four kinds of local features, color, texture, location and shape, to represent blobs (regions) selected by the user from the query. Many subsequent RBIR work continued with this local region-based representation. In [39], Chen and Wang first segmented an image into several regions and then represented these regions using multi-dimensional fuzzy sets as their local feature representation. Other researchers also incorporated this wavelet-based representation in their RBIR system. Based on an assumption that the decomposition by wavelet transform mimics the one in human visual systems, Sun and Ozawa also proposed a hierarchical approach for RBIR .Applied multiple features

IV. RELEVANCE FEEDBACK

Relevance feedback (RF) is an interactive supervised learning technique that has been proposed to bridge the Semantic gap between the low-level image features used and the semantic content of the images and, thus, to improve the retrieval results. In particular, RF attempts to insert the subjective human perception of image similarity into a CBIR system. In order for this to be accomplished, the user is required to assess, in each RF round, the retrieved images as relevant or irrelevant to the initial query and to submit his/her assessment as a feedback to the system. Then, the system takes into account this feedback and updates in an appropriate way the image ranking criterion.

V. CONCLUSION AND FUTURE DIRECTIONS

Due to the low-level image features it utilizes, the semantic gap problem is hard to bridge and performance of CBIR systems is still far away from users’ expectation. Image annotation, region-based image retrieval and relevance feedback are three main approaches for narrowing the “semantic gap”. In this paper, we have presented the early years of image retrieval with progress in these fields in the current decade.

Although many promising results have been reported from reducing semantic gap research, there are still many problems need to be tackled. (1) How to choose learning samples. Relevance feedback has been the most popular way for learning semantic of an image from the user. However, it is a tedious work for users labeling relative or irrelative repetitiously. Long-term learning based on historic data such as search logs is a good way to expand the set of labeled samples. More efforts need to be put on long-term learning to alleviate the “small sample” problem in short term learning from relevance feedback. (2) How to choose similarity measures. Similarity measures determining which images should be considered to be relevant to the query. Improper choice or definition of them would likely to return user images irrelevant to his/her query. (3) How to structure metadata. Image annotation tags some keywords to a specific area of an image. It helps users in organizing and searching image content. However, how to structure metadata to enable users to describe, extract and search information based on images in a more accurate and efficient way is a challenging question.

REFERENCES

[1] Ritendra Data, Dhiraj Joshi, etc. “Image Retrieval: Ideas, Influences, and Trends of the New Age”, ACM Computing Surveys, 40(2), Article 5, 2008. | [2] J. Z. Wang, N. Boujemaa, etc. “Diversity in multimedia information retrieval research.” In Proceedings of the ACM SIGMM International Workshop on Multimedia Information Retrieval (MIR) 2006. | [3] Niblack, C., Barber, R., Equitz, W., Flickner, M., Glasman, E., Petkovic, D., et al. “The QBIC project: Querying images by content using color, texture, and shape.” In Proceedings of SPIE, pp. 173-187, 1993. | [4] Smith, J. “Integrated spatial and feature image systems: Retrieval, analysis, and compression”. PhD Thesis, Columbia University, 1997. | [5] Ogle, V., & Stonebraker, M. “Chabot: Retrieval from a relational database of images.” IEEE Computer, 28(9), pp. 40-48. 1995. | [6] Smith, J. R., Chang, S.-F. “VisualSEEk: A fully automated content-based image query system.” In Proceedings of ACM international conference on multimedia, pp. 87-98, 1996. | [7] Inoue, M.. “On the need for annotation-based image retrieval”. In Proceedings of workshop on information retrieval in context, 2004 | [8] Dympna O'Sullivan ,David C. Wilson ,Michela Bertolotto. “Task-based annotation and retrieval for image information management”. Multimed Tools Appl, Vol 54: pp. 473-497, 2011. | [9] Xiang Sean Zhou ,Thomas S. Huang. “Edge-based structural features for content-based image retrieval”, Pattern Recognition Letters, 22(5): pp. 457-468, 2001. | [10] Gevers, T., & Smeulders, A. “PicToSeek: A content-based image search engine for the WWW”. In Proceedings of international conference on visual information systems. pp. 93-100,1997. | [11] Sclaroff, S., Taycher, L., & Cascia, M.. “ImageRover: A content-based image browser for the world wide web”. In Proceedings of IEEE international workshop on content-based access of image and video libraries, pp. 2-9. 1997 | [12] A. Smeulders, et al., “Content-based image retrieval at the end of the early years”, IEEE Trans. PAMI, 22(12): pp. 1349-1380, 2000. | [13] Peng Huang, Jiajun Bu. “Improve Image Annotation by Combining Multiple Models”, In Third International IEEE Conference on Signal-Image Technologies and Internet-Based System, pp. 3-9, 2008; | [14] Taha Osman, et al, “An Integrative Semantic Framework for Image Annotation and Retrieval”, In 2007 IEEE/WIC/ACM International Conference on Web Intelligence, pp. 366-373, 2007 | [15] J. J., L. V., and M. R., “Automatic image annotation and retrieval using cross-media relevance models.” In SIGIR'03 Conference, pp. 119-126, 2003. | [16] Chih-Fong Tsai, Chihli Hung, “Automatically Annotating Images with Keywords: A Review of Image Annotation Systems”, Recent Patents on Computer Science, vol. 1, pp. 55-68, 2008.