Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 6, Issue, 01, pp.4750-4753, January, 2014

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

MULTIMODAL MOBILE VISUAL SEARCH USING REGION-BASED MATCHING ALGORITHM

***Malathi, S., Sivakumar, P. and Prakash, M.**

Department of Computer Science, KSR College of Engineering, India

ARTICLE INFO

Article History:
Received 05th October, 2013
Received in revised form
24th November, 2013
Accepted 19th December, 2013
Published online 31st January, 2014

Key words:

Multimodal, Visual search,
Exemplary image, Image Clustering,
Image Composition.

ABSTRACT

Searching is becoming pervasive and one of the most popular applications on mobile devices. The image search is based on the input query. Based on the query images is retrieval from database through mobile devices. Text, voice, sketches are used as query to search images on mobile devices. Search performance is the main constraint for any kind of search techniques. To improve the searching performance we need to create new technique. For that purpose, some of the searching techniques were studied below. From all that techniques, we decide to develop a new multimodal interactive image search on mobile device to visual search.

Copyright © 2014 Malathi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Image search is a hot topic in both computer vision and information retrieval with many applications. More consumers use phones or other mobile devices as their personal concierges surfing on the Internet. Along this trend, searching is becoming pervasive and one of the most popular applications on mobile devices. The bursting of mobile users puts forward the new requests for image retrieval (Neven ?). The images are searched based on the query given by the user. Text, Voice, Sketches, Photo and Content of the images are used as query to search images on mobile devices. In the text-based search, the user can type an entity name to find the images. The photo-to-search (Jia *et al.,* 2006) is becoming pervasive as the development of the computer vision. This enables the user to capture photos using the in-built camera on the phone and then initiate search queries about objects in visual proximity to the user. Then voices are used as query to search images. Another kind of image search is sketch-based image search (Tao Chen *et al.,* 2009). It uses hand-drawn sketches to search for satisfied images. The formation / construction model of visual search results based on text, voice, photo, sketch query is represented in the Figure 1.

Text-Based search

Image search is a hot topic in both computer vision and information retrieval with many applications. The traditional desktop image search systems with text queries (Tsai *et al.,* 2011) have dominated the user behavior for a quite long period. The user can either type an entity name or look up on an online local map to find the target Figure 2. shows the method of text to search

A small screen limits the presentation of searching results, which requires the top results to be more relevant while on the phone. However, using only text as search query can hardly meet this end. The surrounding texts of the web images are not always correct. Even the tags of the some human-labeled datasets such as Flickr images are unreliable. Moreover, on the one hand, the user must know the exact terms the annotator used in order to be able to retrieve the images he wants. On the other hand, textual annotations are also language-dependent. Actually, there are more images which have no text information on the web repository. All this deficiency can ruin a good user experience of text-based image search system on the mobile phone.

Voice to search

Voice queries are available on some devices, there are still many cases that semantic and visual intent can hardly be expressed by these descriptions for search. For example, in a common image search task, the user might have already conceived of the general idea of expected pictures such as color configurations and compositions. However, the users usually have to pick up ideal images amidst much more irrelevant results. In such cases where irrelevant images spoil the results and ruin the user experience, visual-aided tools can largely boost the relevance of search results and the user experience. Figure 3. shows the method of voice to search

Corresponding author: Malathi, S.,
Department of Computer Science, KSR College of Engineering, India.

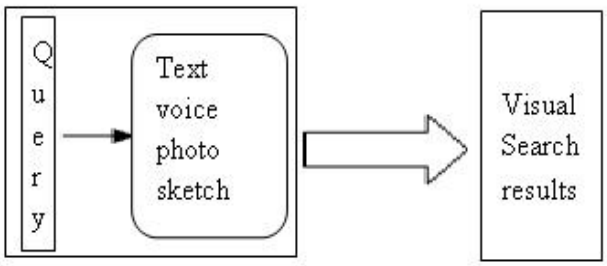

Figure 1. Types of query to search images

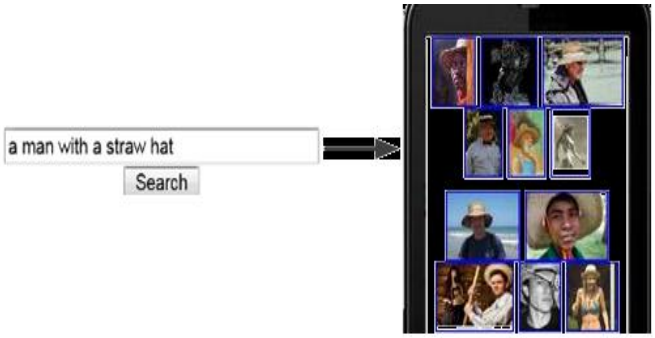

Figure 2. Text-to-search

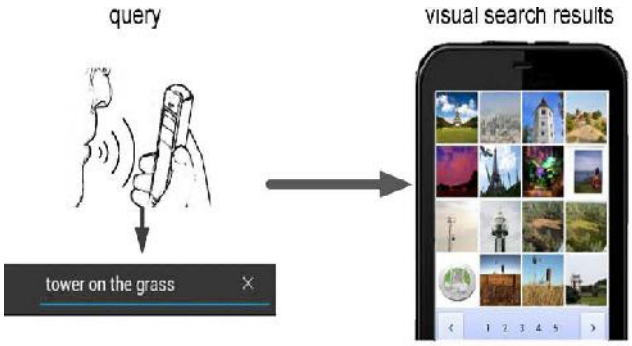

Figure 3. Voice to search

For example: Consider the user has no idea of the name of a restaurant but can only describe its particular appearance, such as "a restaurant with red door, two stone lions, and many red pillars in front;" or even in another totally different situation where the user wants to find "an oil paint of a man with straw hat." The common thing shared in both situations is that only with a scene or general picture in the user's mind, the user doesn't have the title or name of the target. Such kind of searches is not easy under present text-based search condition. But with the help of visual aids, which can search for images based on not only text but also image content, these tasks can be much easier. As a result, a powerful image search system with visual aids is desired.

Photo-to-search

Photo-to-search (Jia *et al.*, 2006) is becoming pervasive as the development of the computer vision and content-based image retrieval. This enables the user to capture photos using the in-built camera on the phone and then initiate search queries about objects in visual proximity to the user. Such applications enable users to search for what they see by taking a photo on the go which is shown in Figure 4. These applications search for the exact partial duplicate images in their database and provide the users with related information of the query images. However, the search is only available for some vertical domains, such as products, landmarks, CD covers, and etc., where the partial duplicate images of the query image have been indexed in their database.

Figure 4. Photo-to-search

Sketch to search

Outline sketches are typically easier and faster to generate than a complete color description of the scene. And they can be generated for arbitrary desired images, while example images may or may not be at hand when searching. In addition, input devices change in favor of sketching as touch-enabled devices become more common. In other words, sketch-based image retrieval (SBIR) (Eitz *et al.*, 2011) is a relevant means of querying large image databases. Several approaches for SBIR have been suggested. Figure 5 shows a sketch to search method.

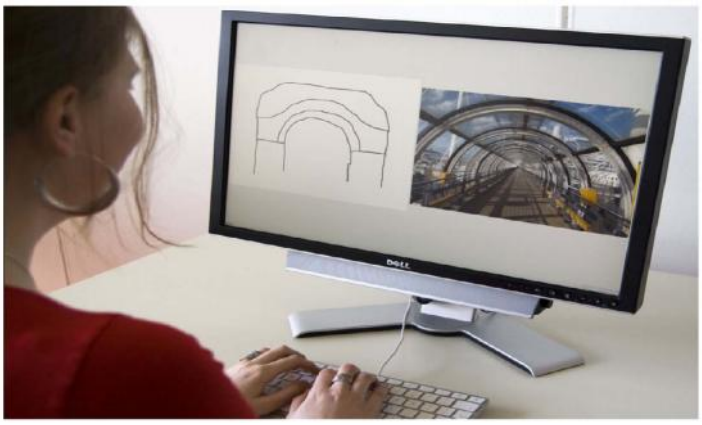

Figure 5. Sketch to search

However, to achieve interactive query response, it is impossible to compare the sketch to all images in the database directly. Instead, descriptors are extracted in preprocess and stored in a data structure for fast access. Very commonly, the descriptors are interpreted as points in a high-dimensional space and finding close matches means searching for nearest neighbors in this space. Sketch-based search allows users to express their visual intent in some way, it can hardly develop

complex meanings and is difficult to use for users without drawing experience.

Sketch to Photo search

A picture is said to be worth a thousand words. Very often, people compose pictures to convey ideas. A common approach is to sketch a line drawing by hand, which is flexible and intuitive. An informative sketch (Tao Chen *et al.*, 2009) requires some artistic skill to draw, and line drawings typically have limited realism. Figure.6. shows a simple freehand sketch is automatically converted into a photo-realistic picture by seamlessly composing multiple images discovered online.

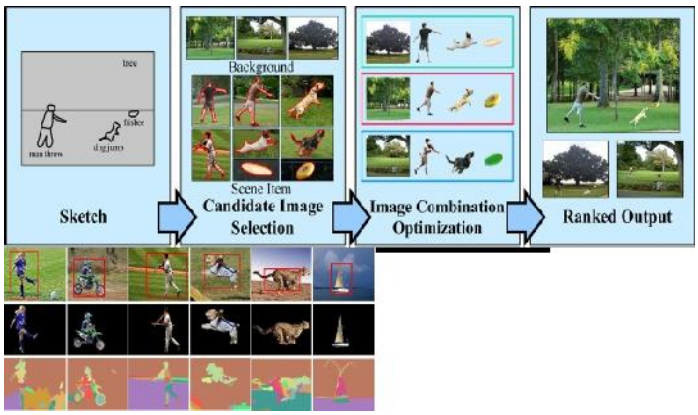

Figure 6. Sketch to photo search

The user provides a simple freehand sketch, where each scene item is tagged with a text label. Our goal is to convert this sketch into a photorealistic image. To achieve this, we search on-line for each scene item, and the background, using the text label. The results are filtered to exclude undesirable images. During filtering, each image is segmented to find scene elements matching items in the sketch. We then optimize the combination of the filtered images to seamlessly compose them, using a novel image blending technique. Several compositions are automatically generated and ranked.

Jigsaw-Visual search

The visual search on mobile devices was taking full advantage of multi-modal and multi-touch interactions on mobile devices. As shown in Fig. 7 users can easily formulate a composite image as their search intent by naturally interacting with the phone through voice and multi-touch. JIGSAW (Wang *et al.*, 2011) is an interactive mobile visual search application that enables users to naturally formulate their search intent in an interactive way and combines different visual descriptors for visual search. Figure 8 shows the framework of JIGSAW. On the client-side, a user first speaks a natural sentence to initiate a voice query. On the cloud side, the system employs speech recognition (SR) to transfer the speech to a piece of text, and then extracts entities from the text. Those entities are represented as exemplary images.

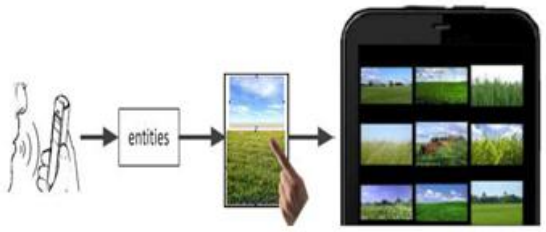

Figure 7.Visual search result

Further specify search intent by touching the screen and dragging their preferred exemplary images, and then formulating a composite visual query. Those exemplary images are automatically generated using a clustering process according to the extracted entities. Finally, we exploit both the text and the composite visual query to search for relevant images, by considering the position and the size of the exemplary images. In the next sections, we will describe the details of each component.

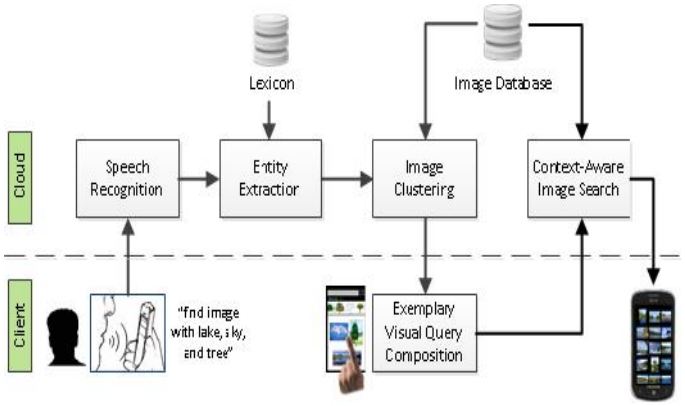

Figure 8. JIGSAW

Map snapper

The Map Snapper (Jonathon ?) project aimed to develop a system for robust matching of low-quality images of a paper map taken from a mobile phone against a high quality digital raster representation of the same map. It represents a methodology for performing content-based image retrieval and object recognition from query images. The Map Snapper, aimed to explore how computer vision techniques could be exploited for the matching of low-resolution digital photographs of Survey paper map products to digital versions of the same map. In particular, the aim was to ascertain of which part of the map the photograph was taken. The motivation for this comes from a desire to exploit the current ubiquity of mobile information devices that incorporate digital cameras, such as mobile phones and personal digital assistants, and combine these devices with Survey paper map products. The vision was of a product that would allow users to query a remote information system based on photos of a paper map taken with the device. The information system could then return useful information to the user via the device. For example, the returned information could include such things as events, facilities, opening times, and accommodation in the geographical region depicted by the query. This paper has

described the systematic approach taken to design a suitable algorithm for matching poor quality query images taken from a mobile phone against a high quality digital representation of a map. The design methodology was heavily evaluation-driven and involved many stages of incremental improvements in order to reach the final design. The outcome of this research is a fast, robust algorithm that meets the design criteria. The matching algorithm combines a number of computer vision techniques, including interest point extraction and local descriptor generation with multidimensional indexing. Geometric constraints were applied to ascertain whether the interest-point matches are consistent.

Conclusion

Text-based search engines are still available on mobile devices. But it is neither user-friendly on phone, nor machine-friendly for search engine. Voice queries must need general idea of expected pictures such as color configurations and compositions. Sketch-based search is difficult to use for users without drawing experience. Photo-to-search needs exact partial duplicate images in their database for search similar images. Thus user's search experience on mobile device is significantly improved by interactive mobile visual search system compare to all other techniques, which allow the users to formulate their search through multimodal interactions with mobile devices. The visual query generated by the user can be effectively used to retrieve similar images. Mobile visual search (Ystad and Sweden 2011) takes the advantages of multimodal and multi-touch functionalities on the phone.

REFERENCES

Eitz, M. Fak. 2011. IV-Elektrotechnik & Inf., Tech. Univ. Berlin, Berlin, Germany Hildebrand, K., Boubekeur, T., Alexa, M, "Sketch-based image retrieval: Benchmark and bag-of-features descriptors".

Jia M., X. Fan, X. Xie, M. Li, and W. Ma, 2006. "Photo-to-search: Using camera phones to inquire of the surrounding world," in *Proc. Mobile Data Manag.*,

Jonathon S. Harea, Paul H. Lewisa, Layla Gordonb and Glen Hartb, "Map Snapper: Engineering an Efficient Algorithm for Matching Images of Maps from Mobile Phones" a School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK; Ordnance Survey, Romsey Road, Southampton, SO16 4GU, UK

Neven Sr., Hartmut (Aachen, DE) Neven, Hartmut (Malibu, CA, US) United States Patent 7565139, "Image-based search engine for mobile phones with camera", www.freepatents online .com

Tao Chen, Ming-Ming Cheng, Shi-Min Hu 2009. Tsinghua University, Ping Tan - National University of Singapore, Ariel Shamir,"Sketch2photo: Internet image montage", The Interdisciplinary Center.

Tsai S. S., D. Chen, H. Chen, C.-H. Hsu, K.-H. Kim, J. P. Singh, and B. Girod, 2011. "Combining image and text features: a hybrid approach to mobile book spine recognition," in *Proc. ACM Multimedia*, pp.

Wang Y., T. Mei, J. Wang, H. Li, and S. Li, 2011. "JIGSAW: Interactive mobile visual search with multimodal queries," in *Proc. ACMMultimedia*, pp.73–82.

Xu H., J. Wang, X. Hua, and S. Li, 2010. "Image search by concept map," in *Proc. ACM SIGIR*, pp. 275–282.

Ystad, Sweden, "Mobile Visual Search from Dynamic Image Databases", 17th Scandinavian Conference, SCIA2011, May2011.

****