Mobile Searching Based On Image,Speech and Text Queries

Divya Agrawal[1], Kaweri Pawar[2,] Snehal More[3] & Prof.K.D.Tamhane[4]

[1] *Divya Agrawal, Computer Department, Pravara Rural Engineering College Loni, Maharahtra, India*

[2] *kaweri Pawar, Computer Department, Pravara Rural Engineering College Loni , Maharahtra, India*

[3] *Snehal More, Computer Department, Pravara Rural Engineering College Loni, Maharahtra, India*

[4] *Prof.K.D.Tamhane, Computer Department, Pravara Rural Engineering College Loni, Maharahtra, India*

ABSTRACT

This paper describes a interactive multimodal image retrieval on mobile (smartphone) devices. For full advantage of the user can easily search few natural multimodal interactions with his/her mobile devices.the multimodal input and natural user interactions of mobile devices, the system, the Joint search with ImaGe, Speech, And Word Plus (JIGSAW). Users have no precise descriptions or names to address them, it is designed for users who already have pictures in their minds. The user can easily find the desired images through a few naturalmultimodal interactions with his/her mobile device, by describing it using speech and then refining the recognized query by interactively composing a visual query using exemplary images. Compared with our previous work JIGSAW, the algorithm has been significantly improved in three aspects: 1) to remove the artificial block partitions; segmentation-based image representation is adopted 2) inverted index is constructed instead of brute force matching. To achieve 5% gain in terms of search performance and is ten times faster, the proposed JIGSAW is able.

Keyword:- *Interactive search, Mobile device, Mobile visual search, Multimodal search,*

---

1. Introduction

IMAGE retrieval is a hot topic in both computer vision and information retrieval with many applications. The traditional desktop image retrieval systems with text queries have dominated the user behavior for a quite long period. As their personal concierges surfing on the Internet, However, while on the go, more and more consumers use phones or other mobile devices. It is the superlative popular applications on mobile devices, along this trend, searching is becoming pervasive. It is reported that one-third of the Internet search difficulties will come from phones by 2014. The new requests for image retrieval, the bursting of mobile users puts forward. First, there is a huge gap in user interface between desktop and mobile devices especially for the input methods. While recent mobile devices always provide multimodal input methods including cameras, GPS, microphones, and multitouch screens, on desktop, keyboards and mice are the main input devices. On some mobile devices which are hardly user-friendly, there are sometimes also tiny keyboards.

A difficult task, this difference makes typing on mobile devices. on the mobile device also differs from that on desktop, it is also observed that the users' search interest. Which are largely related to local information such as local spots and local business like cinemas and restaurants, the searches initiated by smart phones always have strong purposes. For mobile users include text-based search and local map search, existing search alternatives. To find the target, the user can either type an entity name or look up on a online local map. As the development of the

computer vision and content-based image retrieval, moreover, photo-to-search is becoming pervasive. This enabled the user to capture images using the in-built camera on the phone and then initiate search find about objects in visual proximity to the user. Identifying products, comparison shopping, finding information about buildings, movies, compact CDs, real estate, print media, artworks, etc, this advance offers various applications. Which requires the top results to be more relevant while on the phone, a small screen limits the presentation of searching results. Search query can hardly meet this end, using only text. The web images are not always correct, the surrounding texts.

As Flickr images are unreliable, even the tags of the some human-labeled datasets. The exact terms the annotator used in order to be able to retrieve the images he wants, on the one hand, the usermust know. Textual annotations are also language-dependent, On the other hand. Actually, there are more images which have no text information on theweb repository. Image search system on the mobile phone, all this deficiency can ruin a good user experience of text-based.

Test	Purpose	Expected Result
Speech to text	**Text recognized or not from the given input speech**	**Text is recognized from the input speech.**
Text to Search	**If text is properly Entered or not**	**Input text is properly given to Edit text**
Image to Image	**Whether image is uploaded properly or not**	**Image is uploaded properly to server**
Information Search	**Whether information is properly given or not**	**Proper information is given as input**

Table : Validation Test Cases

Though picture search has become a common tool on the PC since 10 years ago, with which the user can input text query to retrieve relevant images, Compared with text search, map search, and photo-to-search, visual search is still not very popular on the phone. on mobile device is that the existing image search applications do not perfectly accommodate to the smartphones and local oriented user intent, A main reason why such image search applications are not popular. The user experience on the phone is not always enjoyable, the search results are rarely useful. No matter whether a tiny keyboard or a touch screen is used, first of all, typing is a tedious job on the phone. Even though voice queries are available on some phones, there are still many cases that semantic and visual intent can hardly be expressed by these descriptions for search. The user might have already conceived of the general idea of expected pictures such as color configurations and compositions, in a common image search task. However, the users usually have to pick up ideal images amidst much more irrelevant results. Visual-aided tools can largely boost the relevance of search results and the user experience, in such cases where irrelevant images spoil the results and ruin the user experience. Such as 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 subject or name of the target. Such kind of searches are not easy under present text-based search condition, let's further consider such a scenario in which the input user has no idea of the name of a restaurant but can only describe its particular appearance. these tasks can be much easier, but with the help of visual aids, which can retrieved for images based on not only text but also image content. A powerful image search system with visual aids is desired. Regarding content-based image search, one kind of popular products, including Google Image, TinEye on PC, and Google Goggles on smart phone, can accept single images as search queries, and return to the user similar pictures or even with information mined from their databases. These engines are able to achieve impressive results, with very large databases. However, to initiate such a visual search, the user must have an exists pictures on hand as a query .Moreover, it needs partially duplicate pictures or exact the

same thing exists in the database. Another kind of image retrieval engines designed for desktop, including GazoPa and some other sketch-based image search researches like, use hand-drawn sketches to find for satisfied images

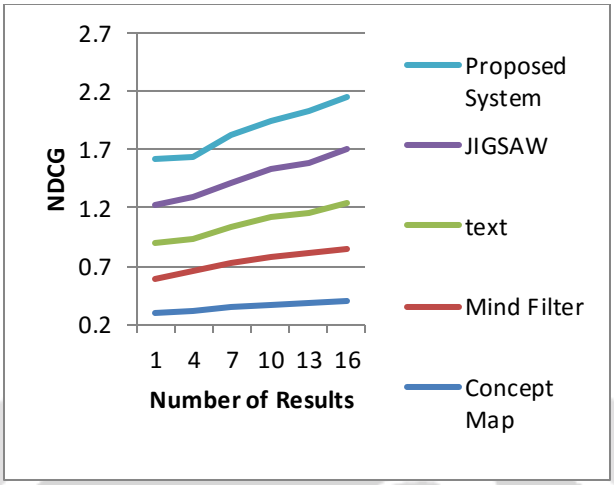

Fig. 1 The curve of five different search methods for the "similar image search" task.

It can hardly develop complicated meanings and is difficult to use for users without drawing experience, though sketch- based search allows users to specific their visual intent in some way. MindFinder and Concept Map also provide visual aids to search for images. By multiple exemplary image patches, visually and semantically similar images are search. By the engine according to the lexicons, the user offers lexicons and then composes a visual query using multiple image patches given.In these works, images are unnaturally divided into blocks in which features are then extracted. To selections and positions of exemplars, the performance is very sensitive. For further information please refer to the papers. Interestingly, in the authors build a Sketch2Photo system that uses simple text-annoter line sketch to automatically synthesize accurate images. They also employ text and sketch to retrived for templates which are then stitched on a background to created a montage. However, their work focuses on pictures composing instead of image search. Inspired by these works, in this paper, we design a multimodal mobile search system to do visual search.

2. Architecture Diagram

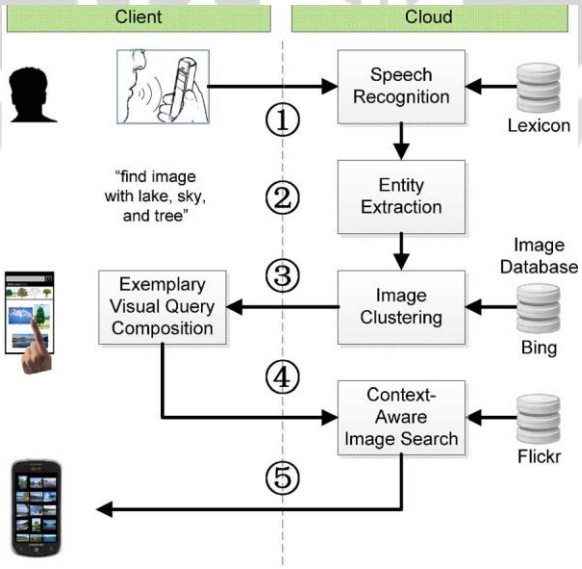

Fig. 2 Architecture Diagram for Proposed System

3. CONCLUSION

We have introduced an interactive mobile visual search system which allows the users to formulate their search intent through natural multimodal interactions with smart phones devices. The system represents the first study on smartphones visual search by taking the benifits of multimodal and multitouch functionalities on the phone. The proposed system provides a game-like interactive pictures search scheme with composition of multiple exemplars. The visual query generated by the user can be effectively used to search similar images by the proposed method. Subjective experiment shows that JIGSAW is an effective complementary tool to existing mobile search applications, especially in cases where users have only partial visual clues in their minds. We significantly improved the algorithm adopted by many similar image search frameworks, i.e., block based matching, and introduced segmentation-based image representation into the previous JIGSAW system. We also indexed the features as visual words into a inverted index as well as proposed a relative position checking scheme. The intend method outperforms all the other methods in three objective experiments. Compared to text-based search system the performance of the proposed system is boosted. The user's search experience on mobile device is thus significantly to make better by this game-like image search system.

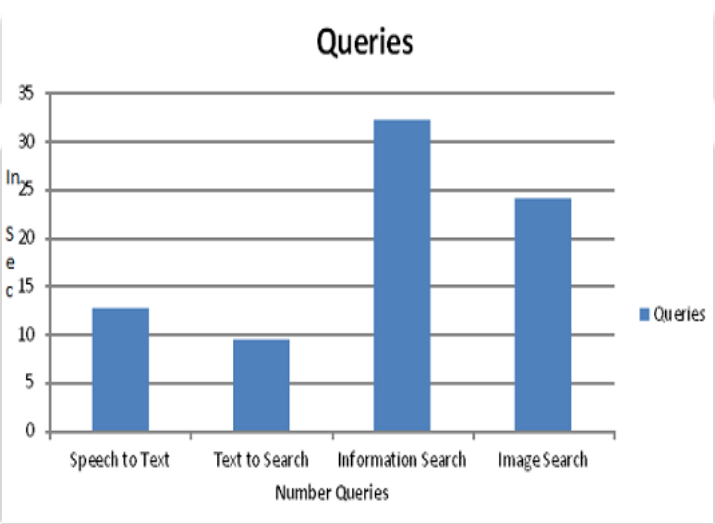

Fig. 3 Query Result

4. FUTURE WORK

Our future works consists the following issues. First, we will try to use the graphical image structure within each exemplar, which may further improve the similar image retrieval results. Second, we will further develop the usability of our system and improve the user experience. For example, we may deploy the visual search system on other mobile devices with larger screen such as tablets. Thus more powerful interactions and functions can be introduced into the system. Third, we will focus on is combining low-level features into mid-level features. Because a relatively small vocabulary size degrades the searching speed and large vocabulary size is too sensitive to feature variances, multiple low-level features can be combined into more robust and discriminative visual words.

5. REFERENCES

[1] [Online]. Available: http://www.pwc.com/gx/en/communications/review/ features/mobile-data.jhtml
[2] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded-up robust features," in Proc. ECCV, 2008, vol. 110, no. 3, pp. 346–359.
[3] N. Ben-Haim, B. Babendo, and S. Belongie, "Improving web-based image search via content based clustering," in Proc. IEEE Conf. Comput. Vis. Pattern Recogn. Workshop, 2006, p. 106.

[4] Y. Cao, H. Wang, C. Wang, Z. Li, L. Zhang, and L. Zhang, "MindFinder: Interactive sketch-based image search on millions of images," in Proc. ACM Multimedia, 2010, pp. 1605–1608.
[5] V. Chandrasekhar, D. M. Chen, A. Lin, G. Takacs, S. S. Tsai, N. M. Cheung, Y. Reznik, R.Grzeszczuk, and B.Girod, "Comparison of local feature descriptors for mobile visual search," in Proc. IEEE Int. Conf. Image Process., 2010, pp. 3885–3888.
[6] V. Chandrasekhar, G. Takacs, D. Chen, S. Tsai, R. Grzeszczuk, and B. Girod, "CHoG: Compressed histogram of gradients a low bit-rate feature descriptor," in Proc. IEEE Conf. Comput. Vis. Pattern Recogn., Jun. 2009, pp. 2504–2511.
[7] T. Chen, M.-M. Cheng, P. Tan, A. Shamir, and S.-M. Hu, "Sketch2photo: Internet image montage," ACM Trans. Graph., vol. 28, no. 5, pp. 124:1–124:10, Dec. 2009.
[8] K. Church, B. Smyth, P. Cotter, and K. Bradley, "Mobile information access: A study of emerging search behavior on the mobile internet," ACM Trans. Web, vol. 1, no. 1, May 2007.
[9] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in Proc. IEEE Conf. Comput. Vis. Pattern Recogn., Jun. 2009, pp. 248–255.
[10] Digimarc Discover [Online]. Available: https://www.digimarc.com/ discover/